**Response to Referee’s Report #2  
‘Sustainable Factor Investing: Where Doing Well Meets Doing Good’**IREF\_2019\_983R1

# INTRODUCTION

First of all, we would like to thank the Editor, Professor Brian Lucey, and the Associate Editor, Professor Larisa Yarovaya, for forwarding the constructive feedback of the Referee and for providing us with the opportunity to submit a revised version of our paper. We would also like to thank the Referee for the feedback and the insightful comments, which we believe have allowed us to significantly improve the paper. We have amended the manuscript taking all remarks carefully into consideration.

In Section 2, we re-iterate the comments of the referee (*in italics*) and explain the revisions made in the paper to address each of these comments.

# RESPONSE TO COMMENTS

**Comment 1:**

*This study explores forming portfolios by combining traditional investing strategies with ESG scores.  This idea is not original, as similar work has been published in other venues (Nagy, 2013; 2015; Bender, et al., (2017)).**[1] For instance, the work by Bender, et al., (2017) shows that highly rated ESG stocks have historically been high value, negative momentum, high quality, low volatility and larger size.  The authors also explore integrating ESG ratings and factor exposures and provide simple integration* *strategies. [2] Compared with this manuscript, Bender et al (2017) present a less rigorous analysis and they analyze a different market.  Bender et al (2017) propose a different methodology to integrate ESG, using an optimizer, instead of building portfolios by sorting stocks terciles/quintiles of the factors.  Similar idea is also in Keiser (2018), a paper cited in this manuscript, that analyzes the US and European markets.  These studies introduce the same idea reflected in this manuscript, thus, diminishing the relevance of its contributions.*

The comment is well-taken. We thank the referee for pointing this out.

We acknowledge the similarities between our work under review and the works of Bender et al. (2017) and Kaiser (2018). However, we understand the concern of the referee, thus we have taken additional measures to ensure that the contribution of our work is clearer.

We highlight the novelty of our work as follows. First, both studies cited above do not investigate ESG integration within the Australian equity market. The uniqueness of the Australian market, including the effects of dividend imputation system (Cannavan, Finn & Gray, 2004; Cummings & Frino, 2008), mandatory superannuation scheme (Huynh, Mallik & Hettihewa, 2006), and high retail ownership (Henker & Henker, 2010; ASX, 2015: p49) are well-documented in the literature. Thus, the lack of Australian ESG studies, and more specifically, studies that integrate ESG with known risk factors, presents an important knowledge gap our paper attempts to address. Our findings can also be viewed as an out-of-sample test for studies that focus on other developed markets.

Second, our paper also differs from the above-mentioned studies in terms of portfolio construction and factor selection. Kaiser (2018)restricts the combination of ESG with value, growth and momentum to long-only portfolios. We investigate the construction of zero-investment (long/short) portfolios similar to Asness et al. (2013). Furthermore, we employ quantile/tercile portfolio breakpoints, following the rationale of Asness et al. (2013). This approach also differs from the optimization-based portfolios in Bender et al. (2017).

To address the concern of the referee, we highlight the above discussions by inserting footnote 1 on page 2: paragraphs 3 & footnote 3 on page 3: paragraph 3 in the revised manuscript.

**Comment 2:**

*This manuscript is also related to the work by Lee, Fan, and Wong (2018).  Lee, Fan, and Wong (2018) also analyze the consequences of ESG based investing in firms listed in Australia during the 2006-2016 period.  They find that high ESG rated portfolios do not significantly outperform or underperform the market, whereas low ESG rated portfolios underperform the market.  The manuscript object of this report deals with more complex investment strategies, but the work by Lee, Fan, and Wong (2018) also diminishes its marginal contribution.*

*While reading these two papers, I noticed that they use different sources for ESG scores, even though this paper explains why Bloomberg is superior to ASSET4. Do the results differ depending on the ESG database?*

The comment is well-taken. We wish to thank the referee for their constructive feedback.

Our work significantly differs from Lee, Fan, and Wong (2018) in two ways. First, Lee et al. (2018)only investigate high and low ESG rated portfolios (termed “tilt portfolios”), where only ESG scores are considered in the portfolio formation process. The paper under review studies the integration of ESG with five important fundamental factors in Australia. Second, we would like to stress the fact that Bloomberg ESG scores are employed in our study since most asset managers have access already to Bloomberg, whilst services such as Datastream, Regnan and RobecoSAM will likely entail additional costs. Meanwhile, we failed to identify any papers in the Australian context that uses the Bloomberg ESG scores.

Another reason why we have not used more than one provider of ESG data is due to the issues of ratings divergence. Halbritter and Dorfleitner (2015) contend that the magnitude and direction of ESG and firm performance is largely dependent on ratings provider, company sample and time period studied. This divergence (or convergence) problem is confirmed with varying results from utilising different ESG scores, sample size and time period by Halbritter and Dorfleitner (2015), Lee et al. (2018) and Limkriangkrai et al. (2017). Nevertheless, to address the concern of the referee, we re-evaluate our results using the Asset4 dataset employed by Lee et al. (2018). We acknowledge the support of the authors for generously providing their dataset. Using Asset4 ESG scores, we replicate the results of Table 2 in the manuscript (see below Table I). These results are, in general, consistent with the Bloomberg ESG scores. Specifically, quality and momentum strategies report similar performance. For this reason, we have not included these results in the revised version of the manuscript. We are more than happy to include these should the referee insists.

**Table I ESG integrated factors with ASSET4 data**

This table reports the performance of ESG integrated factors. The quality signal is represented by the ROE ratio. The minimum volatility signal is computed through a rolling 60-month standard deviation. The momentum signal is based on the past 12-month return, skipping the most recent month. The size is based on the market-capitalization and value is based on the inverse of price-to-book ratio. For standard factors (S), we sort stocks into quartiles based on the respective factor signal at the end of each month. For quality, momentum and value, we take long (short) positions in stocks within the highest (lowest) quartile. For minimum volatility and size, we take long (short) positions in stocks within the lowest (highest) quartile. Each long-short portfolio is market value-weighted and rebalanced monthly. The sample period covers January 2006 through December 2016. ESG scores are not integrated within the S portfolios. M1, M2 and M3 represent methods of ESG integration. For M1, we first screen non-ESG rated firms from the “All” sample, and then construct the factor portfolio as discussed above using the remaining stocks. For M2, we first form factor portfolios in the “All” sample, and then screen non-ESG rated firms off the long and short portfolios. For M3, we first screen non-ESG rated firms from the “All” sample, and then create a new sort signal by assigning a 50/50 weighting to the factor signal and ESG scores. Portfolios are then formed on this combined signal. The market wide benchmarks consist of passive long positions in “All” stocks, “Rated” stocks, S&P/ASX200 Index and the All Ordinaries Index. “All” represents the sample of stocks with top 75% market-capitalization, “Rated” refers to the sample of ASSET4 ESG rated stocks. Average monthly returns, Newey-West adjusted t-statistics, volatility, skewness, kurtosis and maximum drawdown (MaxDD) are reported. Sharpe ratios are annualized.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Quality | | | |  | Minimum Volatility | | | |  | Momentum | | | |
|  | S | M1 | M2 | M3 |  | S | M1 | M2 | M3 |  | S | M1 | M2 | M3 |
| Returns | 1.68% | 1.55% | 2.05% | 1.23% |  | 1.21% | 1.40% | 1.89% | 0.79% |  | 1.08% | 1.08% | 1.26% | 1.60% |
| *t-*statistics | 3.17 | 3.18 | 3.18 | 3.21 |  | 1.20 | 1.47 | 1.83 | 1.40 |  | 1.69 | 1.50 | 2.03 | 3.41 |
| Volatility | 5.56% | 5.18% | 6.59% | 4.31% |  | 7.78% | 7.80% | 8.29% | 4.50% |  | 6.10% | 7.00% | 6.62% | 4.73% |
| Sharpe | 1.15 | 1.13 | 1.21 | 1.05 |  | 0.58 | 0.67 | 0.88 | 0.63 |  | 0.65 | 0.57 | 0.71 | 1.28 |
| Skewness | 0.66 | 0.49 | 0.20 | 0.07 |  | 0.12 | 0.24 | 0.13 | -0.31 |  | 0.04 | -0.67 | 0.23 | 0.11 |
| Kurtosis | 4.85 | 4.02 | 4.14 | 4.81 |  | 3.54 | 3.44 | 3.15 | 4.18 |  | 3.40 | 6.56 | 4.19 | 3.82 |
| MaxDD | -0.30 | -0.29 | -0.37 | -0.15 |  | -0.45 | -0.41 | -0.40 | -0.28 |  | -0.54 | -0.64 | -0.48 | -0.36 |
|  | Size | | | |  | Value | | | |  | Market | | | |
|  | S | M1 | M2 | M3 |  | S | M1 | M2 | M3 |  | All | Rated | ASX200 | All Ords |
| Returns | 0.33% | -0.50% | 1.34% | -0.55% |  | -0.89% | -1.13% | -0.97% | 0.00% |  | 0.55% | 0.28% | 0.20% | 0.20% |
| *t-*statistics | 0.76 | -0.89 | 0.96 | -1.30 |  | -2.17 | -2.52 | -2.20 | -0.01 |  | 1.39 | 0.73 | 0.58 | 0.57 |
| Volatility | 4.47% | 5.40% | 11.96% | 4.65% |  | 4.33% | 4.83% | 4.68% | 3.56% |  | 3.76% | 3.74% | 4.05% | 4.12% |
| Sharpe | 0.26 | -0.31 | 0.42 | -0.40 |  | -0.68 | -0.77 | -0.68 | 0.00 |  | 0.25 | -0.04 | -0.10 | -0.09 |
| Skewness | -0.24 | 0.70 | 2.47 | 0.33 |  | 0.26 | -0.29 | -0.03 | 0.29 |  | -0.76 | 0.33 | -0.66 | -0.75 |
| Kurtosis | 3.46 | 6.86 | 13.47 | 4.24 |  | 3.38 | 4.76 | 4.00 | 2.90 |  | 3.53 | 3.59 | 3.14 | 3.52 |
| MaxDD | -0.38 | -0.78 | -0.82 | -0.67 |  | -0.82 | -0.88 | -0.86 | -0.36 |  | -0.45 | -0.31 | -0.50 | -0.51 |

**Comment 3:**

*A main difficulty to conduct this study is to decide how to integrate or combine existing investment strategies.  The potential number of combinations is unlimited, and there is not a clear reason to choose one specific investment strategy other than it worked well in the past.  This manuscript proposes three methods of integration, that occasionally yield different results (see Table 2 and 4).  Furthermore, in the process the authors discard alternatives because they do not work:*

*·            In foot note 9: “A fourth method of integration, ESG momentum (i.e. changes in ESG ratings), similar to the second method outlined in Melas, Nagy and Kulkarni (2018) was also investigated. We fail to find significant returns after integrating ESG momentum into quality, minimum volatility, momentum, value and size strategies”.*

*·            In page 14 “Since minimum volatility and value strategies do not report statistically significant returns before or after integration, we focus on quality, momentum and size for the remainder of the paper”.*

*The authors also acknowledge, in footnote 10, that the choices of specific benchmarks is arbitrary, and mention that they conduct robustness checks using alternative benchmarks. Of special relevance, given the good performance of the third method, is the construction of M3.  The authors combine signal for each stock i as: [0.5 × ESG\_scorei + 0.5 × Signal\_scorei].  The authors acknowledge the subjectivity in choosing this 50-50 weighting, and they explain that “we assign alternative weightings (0.1 – 0.9) to ESG scores relative to factor signals under M3. The results remain qualitatively similar, however, the performance is generally weaker when higher weightings are assigned to ESG scores.”*

*An interested reader may wonder, why did the authors chose to report the results using 50-50 weighs?, how weak are the results when using different weights?, why do the results become weaker?, Is there a non-subjective criteria to choose these weights?*

*These comments have a common root: This type of research lacks a theoretical motivation.  Not only there is not a good reason for building portfolios based on terciles, instead of quintiles, etc, but there is also not a good justification for not using a combination of the factors, or changing these strategies overtime.**[3]  The authors present several methods of integration, some of them work better than others in the sense that they produce better returns than other strategies using historical information.  This line of research involves a process of choosing the strategy that works better ex-post, and does not tell us much about the drivers of these returns, and thus about the functioning of stock markets.*

This comment is well-taken. Once again, we wish to think the referee for giving us the opportunity to think deeply again about “sustainable factor investing”. In summary, the referee has raised three issues including *i*) the weighting between factor and ESG scores; *ii*) multi-factor ESG integration; *iii*) theoretical motivation. We address each issue separately as follows.

First, we believe that a 50/50 weighting scheme is the least arbitrary (or most non-subjective) scheme that can be applied to the integration of ESG with a factor signal. This is because unlike any other weighting combinations, an equally weighted signal implies that one does not take any stand on the information being weighted. Another reason why 50/50 is preferred is that previous studies have found that a naïve 1/N strategy always outperforms mean-variance optimization and its extensions in terms of Sharpe ratio, certainty-equivalent return, or turnover (DeMiguel, Garlappi & Uppal, 2007). However, we have computed the results based on alternative weights as a robustness test (not reported in the manuscript). For the referee’s information, we plot these results as below in Figure I, where Panels A, B and C represent the returns, *t*-statistics and Sharpe ratios respectively, when ESG scores are weighted as *w*={0.1, 0.2…0.9} versus the factor signal. The returns for quality and momentum range between 1-2% per month across weighting schemes, with *t*-stats and returns increasing (decreasing) for momentum (quality) when higher convicted to ESG. For precision, we also updated footnote 13 on page 12 to reflect these discussions.

Second, the comment on multi-factor integration is well-taken. Following the suggestion of the referee, we conduct additional tests by integrating ESG scores with multiple factors simultaneously. Table II below reports these results. We first consider the combination of quality and momentum factors (Panel A) since they report the strongest results post-ESG integration in our main tests. Subsequently, we add the size factor to the mix in Panel B due to its significance (albeit being weaker). Finally, all factors are combined in one framework including value and minimum volatility (Panel C). The results in Table II reveal that non-ESG integrated multi-factor strategies consistently outperform various ESG integrated methods, except for the combination of all five factors as it reports almost identical Sharpe ratios compared to the non-ESG integrated strategy. Interestingly, regardless of factor combination, a higher weighting towards ESG in the integration process consistently reduces the volatility of the strategy. However, this does not translate to a higher Sharpe ratio compared to non-integrated multi-factor strategies. These results suggest that in a multi-factor setting (as in Panel C), ESG scores are best treated as a signal like other fundamental variables (ESG1/N). As discussed previously, the ESG1/N approach should be considered as the least subjective method. This is also consistent with DeMiguel et al. (2007).

**Table II ESG Integrated Multi-factors**

This table reports the performance of standard (S), equal weighted factor and ESG integration (ESG1/N), 50% ESG weighting and 50% equal weighted factors (ESG1/2) and 75% ESG weighted and 25% equal weighted factors. For instance, in ESG1/N integrated quality and momentum, 33.33% weighting is applied to each signal. Furthermore, in ESG1/2 integrated quality, momentum and size, 50% weighting is applied to ESG and 16.67% weighting applied to remaining factors. The sample period covers 2006-2016. Average monthly returns, Newey-West adjusted *t-*statistics, volatility and annualized Sharpe ratios are reported.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | S |  | ESG(1/N) |  | ESG(1/2) |  | ESG(3/4) |
| A. Quality, Momentum | | | | | | |
| Returns | 2.06% |  | 1.65% |  | 1.72% |  | 1.49% |
| *t*-statistics | 3.71 |  | 3.19 |  | 3.64 |  | 3.18 |
| Volatility | 5.89% |  | 5.76% |  | 5.31% |  | 4.82% |
| Sharpe | 1.36 |  | 1.09 |  | 1.23 |  | 1.16 |
|  | B. Quality, Momentum, Size | | | | | | |
| Returns | 1.65% |  | 1.29% |  | 1.31% |  | 1.21% |
| *t*-statistics | 3.29 |  | 2.53 |  | 3.45 |  | 3.18 |
| Volatility | 5.01% |  | 5.73% |  | 4.90% |  | 4.37% |
| Sharpe | 1.25 |  | 0.84 |  | 1.00 |  | 1.02 |
|  | C. Quality, Momentum, Size, Value, Minimum Volatility | | | | | | |
| Returns | 1.74% |  | 2.05% |  | 1.73% |  | 1.26% |
| *t*-statistics | 3.68 |  | 4.57 |  | 4.14 |  | 2.87 |
| Volatility | 4.52% |  | 5.36% |  | 4.84% |  | 4.67% |
| Sharpe | 1.47 |  | 1.49 |  | 1.36 |  | 1.00 |

We have fought long and hard about whether the above findings should be presented in the manuscript. Ultimately, we believe that multi-factor ESG integration goes beyond the scope of our current work. This is mainly because investors’ preference towards factors are completely arbitrary, i.e. one can choose to invest in any single factor(s) or combinations of factors. Our main goal is assessing the impact of ESG integration on single-factor exposures, *not* obtaining the “optimal” or “efficient” factor exposure within an investable universe as discussed in Clark, De Silva & Thorley (2016), Ghayur, Heaney & Platt (2018) and Lester (2019). For this reason, we have opted to not report the above table in the manuscript, as it considerably shifts the focus of the paper. However, a new footnote (#15) on page 15 has been inserted to highlight the key results discussed.

Third, we thank the referee for pointing out “the lack of theoretical motivation”. While we share the concern of the referee, like most other studies in the ESG investing literature, our motivation is to examine the empirical link between firm’s non-financial (i.e., ESG) and financial performance (i.e., stock returns). Our particular interest is to understand how ESG scores *interact* with firm characteristics that are fundamental to the pricing of Australian stocks. As noted by the referee, there is no universal method to build ESG integrated portfolios. For instance, the optimization-based approach in Bender *et al.* (2017), portfolio sorts in Kaiser (2018) and more recently an ESG-efficient frontier in Pedersen, Fitzgibbons & Pomorski (2019), to name a few. Our work proposes and tests one of many ways to integrate ESG into the investment process, in a previously untested, developed stock market with unique features. Such knowledge would aid the design of better investment strategies and help accelerate the “mainstreaming of sustainable investing” globally (Grei, 2018).

**Comment 4:**

*Using 5 observations in 2006 (September) is troublesome, and the authors may consider not using this year in this analysis.  It is not only difficult to make any generalization of the results using 5 observations, but the authors use only 1 firm after sorting into quartiles.  Because of the lack of observations, it is difficult to conduct this analysis during the 2007-2009 financial crisis.  Related to the sample construction, close to 30% of the sample consists of financial firms that are usually excluded in analyses of the post financial crisis period because they receive substantial support during the crisis (e.g. Lins et al., 2017).*

We thank the referee for this constructive feedback.

Indeed, we have neglected the effects of financials in the previous version. We also take note of the comment on the size of the cross-section. While the number of “Rated” stocks is quite small in 2006, between 2007-2009, there are on average 112 stocks in the cross-section, suggesting that our long-short factor portfolios should be sufficiently diversified.

Nevertheless, following the advice of the referee, we re-evaluate the performance of ESG integrated versus non-ESG integrated factors by excluding financial stocks (SF vs SNF and M3F vs M3NF) in Table III below. In addition to this, we repeat the same analysis eight times from 2007 to 2014, by varying the starting period. Interestingly, during the crisis period, quality and size (momentum) without financials in fact outperform (underperforms) the same strategies that include financials. Nonetheless, if one considers the results from 2007-2009 crisis period troublesome due to the size of the cross-section, the results from 2010-2016 solidify the effectiveness of ESG integration in quality and momentum strategies. Overall, following Lins et al. (2017), our results are robust to the exclusion of financial firms.

To include this robustness results, we have inserted a new paragraph on page 16: paragraph 2 that summarizes these results.

**Table III Effects of Financials and Size of Cross-section**

This table reports the performance of standard (S) and ESG integrated (M3) quality, momentum and size strategies across various time periods. Standard inclusive of the Financials GICS sector is defined as SF, whilst S excluding Financials is defined as SNF. Further, M3 with Financials included is abbreviated to M3F, while M3NF indicates the exclusion of Financials. The performance displayed is across Januray of the indicated year until December 2016. For example, 2007 indicates the performance from January 2007 until December 2016 etc.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | Quality | |  |  |  | Momentum | |  |  |  | Size | |  |
|  |  | SF | SNF | M3F | M3NF | SF | SNF | M3F | M3NF | SF | SNF | M3F | M3NF |
| 2007 | Returns | 1.43% | 1.51% | 2.03% | 2.10% |  | 0.77% | 1.16% | 1.55% | 1.34% |  | 1.33% | 1.58% | 1.48% | 1.73% |
|  | *t-*statistics | 3.41 | 2.84 | 4.03 | 3.90 |  | 1.35 | 1.69 | 2.73 | 2.25 |  | 3.24 | 3.54 | 2.06 | 2.34 |
|  | Volatility | 4.52% | 5.62% | 5.75% | 5.84% |  | 5.56% | 6.23% | 6.14% | 6.28% |  | 3.45% | 3.77% | 5.79% | 6.10% |
|  | Sharpe | 1.19 | 1.01 | 1.37 | 1.40 |  | 0.50 | 0.69 | 0.96 | 0.80 |  | 1.44 | 1.59 | 0.96 | 1.08 |
| 2008 | Returns | 1.41% | 1.45% | 2.03% | 2.19% |  | 0.86% | 1.18% | 1.64% | 1.43% |  | 1.36% | 1.57% | 1.60% | 1.88% |
|  | *t-*statistics | 3.12 | 2.53 | 3.70 | 3.70 |  | 1.38 | 1.56 | 2.70 | 2.23 |  | 2.97 | 3.16 | 2.03 | 2.33 |
|  | Volatility | 4.65% | 5.76% | 5.99% | 6.03% |  | 5.74% | 6.41% | 6.09% | 6.19% |  | 3.50% | 3.80% | 6.01% | 6.32% |
|  | Sharpe | 1.14 | 0.95 | 1.31 | 1.42 |  | 0.54 | 0.68 | 1.02 | 0.86 |  | 1.45 | 1.57 | 1.01 | 1.15 |
| 2009 | Returns | 1.06% | 1.21% | 1.81% | 2.05% |  | 0.70% | 1.08% | 1.76% | 1.51% |  | 1.61% | 1.83% | 2.05% | 2.40% |
|  | *t-*statistics | 2.58 | 2.11 | 3.36 | 3.29 |  | 1.05 | 1.34 | 2.71 | 2.38 |  | 3.27 | 3.42 | 2.44 | 2.89 |
|  | Volatility | 4.29% | 5.41% | 5.24% | 5.53% |  | 5.14% | 5.74% | 5.65% | 5.55% |  | 3.40% | 3.75% | 6.07% | 6.27% |
|  | Sharpe | 0.91 | 0.83 | 1.32 | 1.44 |  | 0.49 | 0.69 | 1.19 | 1.03 |  | 1.79 | 1.87 | 1.31 | 1.52 |
| 2010 | Returns | 1.34% | 1.72% | 2.07% | 2.51% |  | 1.39% | 1.95% | 2.21% | 2.07% |  | 0.98% | 1.15% | 0.96% | 1.33% |
|  | *t-*statistics | 3.81 | 3.39 | 3.99 | 4.41 |  | 2.58 | 3.18 | 3.49 | 4.04 |  | 2.73 | 2.90 | 1.73 | 2.42 |
|  | Volatility | 3.72% | 5.00% | 4.68% | 5.21% |  | 4.72% | 5.06% | 5.59% | 5.28% |  | 2.84% | 3.11% | 5.04% | 5.29% |
|  | Sharpe | 1.34 | 1.31 | 1.72 | 1.92 |  | 1.10 | 1.49 | 1.55 | 1.52 |  | 1.27 | 1.37 | 0.69 | 0.94 |
| 2011 | Returns | 1.47% | 1.87% | 2.58% | 3.06% |  | 1.64% | 2.34% | 2.51% | 2.30% |  | 0.82% | 0.97% | 0.60% | 0.99% |
|  | *t-*statistics | 3.79 | 3.29 | 5.32 | 5.60 |  | 2.76 | 3.55 | 3.54 | 4.02 |  | 2.32 | 2.52 | 1.06 | 1.76 |
|  | Volatility | 3.80% | 5.20% | 4.66% | 5.11% |  | 4.97% | 5.29% | 5.85% | 5.55% |  | 2.77% | 2.95% | 5.09% | 5.32% |
|  | Sharpe | 1.45 | 1.38 | 2.22 | 2.46 |  | 1.25 | 1.74 | 1.71 | 1.63 |  | 1.08 | 1.20 | 0.42 | 0.68 |
| 2012 | Returns | 1.60% | 2.08% | 2.45% | 2.93% |  | 1.65% | 2.33% | 2.37% | 2.18% |  | 0.80% | 0.92% | 0.41% | 0.78% |
|  | *t-*statistics | 3.57 | 3.13 | 4.22 | 4.53 |  | 2.37 | 3.04 | 2.75 | 3.13 |  | 1.96 | 2.08 | 0.63 | 1.20 |
|  | Volatility | 4.10% | 5.59% | 4.85% | 5.18% |  | 5.15% | 5.52% | 6.30% | 5.88% |  | 2.85% | 3.03% | 5.32% | 5.58% |
|  | Sharpe | 1.48 | 1.44 | 2.00 | 2.31 |  | 1.22 | 1.67 | 1.49 | 1.45 |  | 1.01 | 1.10 | 0.27 | 0.50 |
| 2013 | Returns | 1.64% | 2.08% | 2.34% | 2.66% |  | 2.05% | 2.36% | 2.34% | 2.16% |  | 0.90% | 0.99% | 0.41% | 0.77% |
|  | *t-*statistics | 2.99 | 2.63 | 3.55 | 3.65 |  | 2.60 | 2.58 | 2.30 | 2.54 |  | 1.99 | 2.02 | 0.50 | 0.94 |
|  | Volatility | 3.93% | 5.11% | 5.00% | 5.27% |  | 4.86% | 5.53% | 6.37% | 5.99% |  | 2.91% | 3.12% | 5.77% | 6.07% |
|  | Sharpe | 1.58 | 1.58 | 1.84 | 2.03 |  | 1.64 | 1.68 | 1.45 | 1.41 |  | 1.13 | 1.16 | 0.25 | 0.46 |
| 2014 | Returns | 1.48% | 1.74% | 1.84% | 2.36% |  | 1.24% | 1.51% | 1.57% | 1.43% |  | 1.12% | 1.29% | 0.64% | 0.92% |
|  | *t-*statistics | 2.10 | 1.71 | 2.41 | 2.75 |  | 1.69 | 1.57 | 1.64 | 2.01 |  | 2.59 | 2.85 | 0.88 | 1.27 |
|  | Volatility | 3.98% | 5.32% | 4.61% | 5.02% |  | 3.74% | 4.76% | 5.02% | 4.53% |  | 2.57% | 2.72% | 4.84% | 5.12% |
|  | Sharpe | 1.39 | 1.25 | 1.53 | 1.85 |  | 1.23 | 1.20 | 1.19 | 1.18 |  | 1.60 | 1.76 | 0.47 | 0.65 |

**Comment 5:**

*The analysis of the impact of business cycles and market conditions are also difficult to justify in Australia, a country without a recession in almost 30 years.  The average inflation in Australia (used in Table 3 to assess market conditions) has been 2.51%, during the 2006-2016 period, with only 2008 being above 4% (4.35%).  Thus Australia did not experience drastic changes in market conditions similar to that analyzed in prior work cited by the authors (e.g. Nofsinger and Varma (2014); see also Lins et al., 2017).  Thus, I wonder about the need for this analysis and the about the interpretation of the findings in Table 3.*

This comment is well-taken. We understand the concerns of the referee on business cycles.

Unfortunately, our sample period is dictated by the availability of Bloomberg ESG ratings data, thus we are unable to include a longer time-series that incorporates additional variations in the business cycle. Thus, instead of focusing on “recessions”, we have used the business cycle conditions provided by OECD. In footnote 14, we acknowledged the term “Recession” within Table 3 may be inappropriate considering that Australia did not experience a technical recession in the last 25 years. Therefore, to be more precise, the terminology of “Low growth” (“High growth”) is utilised instead of “Recession” (“Expansion”) throughout the revised manuscript except for when the OECD data is described on page 16 and in Table 3.

Furthermore, we also plotted the OECD business cycles for both the Australian (in red) and the US (in black) economy in the figure below. From 1990 to 2017, although Australia has not experienced a technical recession, the OECD business cycle counts exhibit considerable variations. Even in our sample period from 2006, there are noticeable non-overlapping high versus low growth periods with the US. Noticeably, the US recession counts do reflect the NBER business cycle variations.

Source: Australia: <https://fred.stlouisfed.org/series/AUSRECM>; USA: <https://fred.stlouisfed.org/series/USARECDM#0>

As for the comment on inflation, although Australia has not experienced hyperinflation within our sample period, we would like to stress the fact that our inflation analysis is a relative comparison. The Reserve Bank of Australia (RBA) targets an inflation rate of 2-3% over the medium-term. The following plot illustrates the Australian and US inflation rates from the 1990s. The median inflation rate in our sample (from 2006) is almost identical to the median over the last 25 years and is also consistent with the US median. Thus, we believe our inflation analysis does convey meaningful findings.

Source: Australia: <https://www.rba.gov.au/inflation/measures-cpi.html>   
 USA: <https://www.usinflationcalculator.com/inflation/consumer-price-index-and-annual-percent-changes-from-1913-to-2008/>

For these reasons, we would like to keep the business cycle and inflation analysis in the revised manuscript. In our opinion, Table 3 is particularly informative because it clearly highlights the benefit of ESG integration in factor strategies over non-integrated factors in different economic and market conditions. However, we are happy to reconsider should the referee insist.

**Comment 6:**

*Finally the authors mention the following as a contribution from this study: “Moreover, since forcing allocations across industries leads to inferior factor performance while ESG integration no longer delivers improvements, managers should clearly communicate the opportunity costs arising from mandated sector diversification”.  However, I do not think that the empirical analysis conducted in this study (consisting on selecting the best and worst firm in each sector based on each factor signal and ESG scores) is sufficient to justify this statement, and it is not enough to make any normative statement.*

We thank the referee for pointing out this confusion. The referee’s concern is well-taken.

The statement stems from the fact that ESG integrated Best-of-Sector (BOS) strategy will likely underperform a non-restricted BOS strategy.

To address this ambiguity however, we have removed the following from the abstract:

“Moreover, since forcing allocations across industries leads to inferior factor performance while ESG integration no longer delivers improvements, managers should clearly communicate the opportunity costs arising from mandated sector diversification.”

The following statement on page 5: paragraph 2 has been toned down from:

“Since ESG integration increases portfolio tilts to higher scoring industries, managers should clearly communicate the opportunity costs arising from mandated sector diversification.”

To:

“Since ESG integration increases portfolio tilts to higher scoring industries, managers may wish to communicate the potential opportunity cost arising from mandated sector diversification.”

# FINAL COMMENT

We are indebted to the referee for helping us improve the quality of the paper. We would like to sincerely thank the anonymous referee again for her/his constructive feedback. We have made our best efforts to address these comments, we hope that our revisions meet your expectations. We have no doubts that the paper has benefited significantly from this exercise.

**REFERENCE**

ASX. (2015). The Australian share ownership study. Retrieved on January 27, 2020 from <https://www.asx.com.au/documents/resources/australian-share-ownership-study-2014.pdf>

Asness, C. S., Moskowitz, T. J., & Pedersen, L. H. (2013). Value and momentum everywhere. *The Journal of Finance*, *68*(3), 929-985.

Bender, J., Sun, X., & Wang, T. (2017). *Thematic indexing, meet smart beta! merging ESG into factor portfolios*. Working Paper. State Street Global Advisors.

Cannavan, D., Finn, F., & Gray, S. (2004). The value of dividend imputation tax credits in Australia. *Journal of Financial Economics*, *73*(1), 167-197

Clarke, R., De Silva, H., & Thorley, S. (2016). Fundamentals of Efficient Factor Investing (corrected May 2017). *Financial Analysts Journal*, *72*(6), 9-26

Cummings, J. R., & Frino, A. (2008). Tax effects on the pricing of Australian stock index futures. *Australian Journal of Management*, *33*(2), 391-406.

DeMiguel, V., Garlappi, L., & Uppal, R. (2007). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy?. *The Review of Financial studies*, *22*(5), 1915-1953.

Ghayur, K., Heaney, R., & Platt, S. (2018). Constructing Long-Only Multifactor Strategies: Portfolio Blending vs. Signal Blending. *Financial Analysts Journal*, *74*(3), 70-85.

Greis, M. (2018). *Mainstreaming Sustainable Investing*. CFA Institute Research Foundation.

Halbritter, G., & Dorfleitner, G. (2015). The wages of social responsibility—where are they? A critical review of ESG investing. *Review of Financial Economics*, *26*(1), 25-35.

Henker, J., & Henker, T. (2010). Are retail investors the culprits? Evidence from Australian individual stock price bubbles. *The European Journal of Finance*, *16*(4), 281-304

Huynh, W., Mallik, G., & Hettihewa, S. (2006). The impact of macroeconomic variables, demographic structure and compulsory superannuation on share prices: the case of Australia. *Journal of International Business Studies*, *37*(5), 687-698.

Kaiser, L. (2018). *ESG integration: Value, growth and momentum*. Working Paper. University of Liechtenstein.

Lee, D. D., Fan, J. H., & Wong, V. (2018). *No more excuses! Performance of ESG integrated portfolios in Australia*. Working Paper. Griffith University.

Lester, A. (2019). On the Theory and Practice of Multifactor Portfolios. *The Journal of Portfolio Management*, *45*(3), 87-100.

Limkriangkrai, M., Koh, S., & Durand, R. B. (2017). Environmental, social, and governance (ESG) profiles, stock returns, and financial policy: Australian evidence. *International Review of Finance, 17*(3), 461-471.

Lins, K. V., Servaes, H., & Tamayo, A. (2017). Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis. *The Journal of Finance, 72(4), 1785-1824.*

Pedersen, L. H., Fitzgibbons, S., & Pomorski, L. (2019). *Responsible investing: The ESG-efficient frontier*. Working paper. AQR capital management.