Sustainable stock indices and long-term portfolio decisions

Robert J. Bianchi & Michael E. Drew

To cite this article: Robert J. Bianchi & Michael E. Drew (2012) Sustainable stock indices and long-term portfolio decisions, Journal of Sustainable Finance & Investment, 2:3-4, 303-317

To link to this article: https://doi.org/10.1080/20430795.2012.715577

Published online: 05 Sep 2012.

Submit your article to this journal

Article views: 303

View related articles

Citing articles: 2
Sustainable stock indices and long-term portfolio decisions
Robert J. Bianchi* and Michael E. Drew

Griffith Business School, Department of Accounting, Finance and Economics, Griffith University, 170 Kessels Road, Nathan, Brisbane, 4111, Queensland, Australia

(Received 11 April 2012; final version received 22 July 2012)

What is the long-term behaviour of sustainable stock index returns and the accretive benefits to portfolio diversification? We consider these issues through the prism of a long-term investor by replicating the risk and reward behaviour of sustainable stock indices from 1927 through 2010. We find that these indices exhibit long-term mean, variance and tail-risk characteristics that are commensurate with conventional U.S. stocks. We also reveal that recent performance appears worse than their performance over the long term. On the question of portfolio diversification, we find that only one of the three sustainable stock indices investigated dominates the efficient frontier. Our findings suggest that the stock screening process of these indices has important implications regarding the desirability of these investments for long-term investors.

Keywords: socially responsible investments (SRI); sustainability; asset pricing; portfolio analysis

1. Introduction

The rise of ‘sustainable’ investment practices over the last two decades has seen the proliferation of new stock indices designed to measure the return of a portfolio of companies pursuing a strategy of corporate sustainability and social responsibility. Although these new benchmarks play an important role regarding the performance of socially responsible investment (SRI) screens, one of the current challenges for plan sponsors is that these indices have relatively short histories. The relatively limited time series of SRI benchmark returns is especially problematic for pension funds, which are faced with the daunting task of constructing mean-variance efficient portfolios for members over long time horizons. From a practical perspective, the limited sample of SRI returns may heighten both model and estimation error in a portfolio analysis.

To address the disconnect between short-run data and long-term portfolio decisions, this study examines the characteristics of the three most popular sustainable stock indices, namely, the MSCI KLD 400 Social Index (KLD), Dow Jones Sustainability U.S. Index (DJSI) and the Calvert Social Index (CSI) (hereafter collectively known as the ‘sustainable stock indices’). We statistically map the relatively short empirical histories of these respective indices on the well-established Fama and French (1992, 1993) and Carhart (1997) U.S. market risk factors from 1927 through 2010. The motivation for the analysis is to provide investors with a long-term perspective of the risk and return characteristics of these sustainable stock indices.

Our findings reveal that the sustainable stock index returns exhibit market beta with negative risk exposures to the Fama and French (1992, 1993) small-minus-big (SMB) and high-minus-low
(HML) risk factors. Although these sustainable stock indices exhibit common risk factors, we reveal a significant variation in the returns of these sustainable indices over the long term. These findings suggest that the stock selection methods of each sustainable stock index can invoke a significant variation in their average returns while displaying common risk factors and risk characteristics. We employ mean-variance and mean-conditional value-at-risk portfolio analyses and we find that the KLD is the sustainable stock index with the most desirable return/risk characteristics for investors over the long term.

The remainder of this paper is organised as follows. Section 2 reviews the relevant literature for this study. Section 3 describes the data and Section 4 outlines the methodology employed in the analysis. Section 5 presents the empirical analysis and section 6 provides the conclusion and discussion of the results.

2. Related literature

Global sustainable finance initiatives, such as the United Nations Principles of Responsible Investments (UNPRI), have seen the emergence of responsible investing as part of the governance agenda of pension funds and the investment management industry. The UNPRI is one institutional initiative whereby public sector and private sector fund managers strive toward higher standards in environmental, social and governance (ESG) issues to enhance global investment practices. While there is an emerging shift in the global investment management industry towards ESG-based investment decision-making, the empirical evidence regarding the benefits to investors remains mixed.

A number of studies have demonstrated the empirical benefits of sustainable investment decision making. Gompers, Ishii, and Metrick (2003) document the presence of a corporate governance risk premium in stock returns. Other studies by Thomas (2001) and Derwall et al. (2005) show that excess returns can be earned through environmentally aware investment strategies. Overall, these studies suggest that investors can earn a risk premium from sustainable investments.

Another strand of literature offers evidence to suggest that sustainable investing returns are no different to conventional market returns. The early works by Diltz (1995), Goldreyer and Diltz (1999), Guerard (1997), Hamilton, Jo, and Statman (1993) and Sauer (1997) find no statistical differences in the performance between SRI/ethical funds versus conventional funds. These early findings have been further supported by more recent studies by Bauer, Derwall, and Otten (2007), Bauer, Koedijk, and Otten (2005), Benson, Brailsford, and Humphrey (2006), Cortez, Silva, and Areal (2009), Humphrey and Lee (2011) and Statman (2000).


It is clear that the empirical literature is mixed in terms of the performance of sustainability based investments. To better understand the behaviour of these sustainable stock indices, researchers have examined the common risk factors that drive these returns. Kurtz and diBartolomeo (1996) found that sustainable index returns can be explained by important macroeconomic variables such as the oil price and the level of industrial production. Statman (2000) shows that the KLD index exhibits similar returns to a large market capitalisation index. Lee and Faff (2009) find that the Fama and French (1992, 1993) three-factor model can explain the variation in returns of the Dow Jones family of sustainable indices.

While the literature has examined the performance of sustainable investments and the sources of their returns, little attention has been dedicated to their potential portfolio diversification benefits or
otherwise. From a Markowitz (1952) mean-variance perspective, the work of Rudd (1981) reminds us that portfolio performance suffers when portfolios are constrained to specific industries or sectors as this lowers diversification, resulting in a less efficient frontier. This suggests that portfolio diversification is reduced when stock indices are constructed from ‘sustainability-type’ stock selection screens and criteria. The argument of lower portfolio diversification has been supported in the sustainable finance literature by Bello (2005), Lee, Faff, and Langfield-Smith (2009) and Statman (2006).

It is our conjecture that the absence of portfolio-based studies is due to the short empirical history of sustainable stock indices, which poses a challenge for this type of analysis. The second issue that is yet to be fully considered in the debate relates to the tail-risk characteristics of sustainable investments. In a post-global financial crisis world, our understanding of tail-risk in sustainability-based investments is critical given that it is well recognized that empirical financial market returns exhibit unexpected large losses in the extreme left-tail of the distribution of returns. To address this issue, the portfolio selection literature by Krokhmal, Palmquist, and Uryasev (2000), Rockafellar and Uryasev (2000, 2002) and Uryasev (2000) has introduced the mean conditional-value-at-risk (M-CVaR) portfolio approach whereby this optimization procedure minimizes the tail-risk of a portfolio rather than the volatility of portfolio returns.

It is against this background that this study contributes to the sustainability debate by employing the methodology of Agarwal and Naik (2004) to construct the long-term monthly return series of the KLD, DJSI and CSI indices. Agarwal and Naik (2004) employ an in-sample/out-of-sample procedure to lengthen a decade of monthly hedge fund returns into a time series of many decades that can be employed in portfolio optimizations. This study will utilize the Agarwal and Naik (2004) technique to reconstruct and analyse the long-term return and risk characteristics of these sustainable stock indices. We will then employ these long-term returns in the mean-variance and M-CVaR portfolio frameworks to determine whether sustainable stock indices are accretive to portfolio diversification.

3. Data
This study employs the KLD, DJSI and CSI as they are the three major sustainable stock indices that are employed by practitioners and are extensively referenced in the literature (refer to Statman (2006) for a detailed review of the design and methodological construction of these sustainable stock indices). Panel A of Table 1 reports their descriptive statistics and shows that they are similar in all aspects with the exception of their monthly means and median returns. As stated in Statman (2006), these three sustainable stock indices exhibit slight variations in terms of their ESG emphasis and this explains the differences in their historical mean returns. Panel B of Table 1 reports the Fama and French (1992, 1993) risk factors (U.S. stock returns, SMB and HML) and the Carhart (1997) momentum risk factor, which are employed in this study to model and replicate the monthly returns of the sustainable stock indices over the long term. Panel C presents the summary statistics of the two conventional asset classes (i.e. stocks and bonds) to be employed in the portfolio analysis from 1970–2010. We employ the MSCI U.S. Standard Core Gross (Equity) Index as the proxy that genuinely represents U.S. stock returns earned by a pension fund over the long term. The U.S. bond proxy in panel C is constructed by splicing a number of datasets as a long term U.S. bond market proxy is unavailable. The U.S. bond proxy is formed by splicing the following returns: (i) the U.S. government long interest rate sourced from Robert Shiller from January 1970 to December 1972 converted into a 30-year bond monthly return; (ii) the Lehman Brothers U.S. Government Long Term Bond Index from January 1973 to December 1989; and (iii) the Citigroup U.S. Broad Investment Grade Index from January 1990 to October 2010.
Table 2 presents the ESG stock selection criteria of the three sustainable stock indices employed in this study. Please note that all three indices employ criteria, which are proprietary in nature, therefore, little disclosure and transparency is available in terms of their implementation. Table 2 reports that all three indices employ ESG principles, however, each stock selection criteria is unique with their emphasis on different forms of ESG factors.

4. Methodology
As discussed previously, the empirical challenges faced by researchers in analysing sustainable investments from a portfolio analysis perspective are many and varied. There are two specific

Table 2. ESG stock selection criteria.

<table>
<thead>
<tr>
<th>KLD</th>
<th>DJSI</th>
<th>CSI</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Environment</td>
<td>- Climate change strategies</td>
<td>- Products</td>
</tr>
<tr>
<td>- Community and society</td>
<td>- Energy consumption</td>
<td>- Environment</td>
</tr>
<tr>
<td>- Employees and supply chain</td>
<td>- Human resource development</td>
<td>- Workplace</td>
</tr>
<tr>
<td>- Customers</td>
<td>- Knowledge management</td>
<td>- Integrity</td>
</tr>
<tr>
<td>- Governance and ethics</td>
<td>- Stakeholder relations</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Corporate governance</td>
<td></td>
</tr>
</tbody>
</table>

This table presents the ESG criteria employed as the stock selection screens for the MSCI KLD 400 Social Index (KLD), the Dow Jones Sustainability U.S. Index (DJSI) and the Calvert Social Index (CSI).
problems in this study. First, the length of the time series of each sustainable stock index varies with differing commencement dates. This makes it difficult to compare returns, risk measures and covariation metrics of similar time horizons. The second problem is that the history of the time series is relatively short in comparison with other asset classes such as stocks and bonds. The data series for the KLD, DJSI and CSI commence in 1990, 1999 and 2000, respectively. From a long-term portfolio construction perspective, this empirical sample is very short and does not include the multiple macroeconomic growth and contraction periods of an economy, the varying periods of tight and loose monetary policy and the long-term dynamics of risky asset classes. The following section details the methodology employed in this study to construct the monthly returns of these sustainable stock indices so that a long-term portfolio analysis can be estimated.

4.1. Model development

While short empirical data samples abound in the field of empirical finance, perhaps nowhere is this more acute than in the study of alternative investments (particularly using hedge fund indices as a proxy for the return behaviour of alternative investments). This strand of literature, led by the seminal contribution of Agarwal and Naik (2004), faced similar data limitations considered in this study. Hedge fund indices, like their younger sustainable stock cousins, have considerably shorter return histories in comparison with traditional stock and bond benchmarks. In the spirit of Agarwal and Naik (2004), the following sections outline the methodology employed to reconstruct the long-term returns of the three sustainable stock indices in this study.

4.1.1. In-sample model

As the three sustainable stock indices are U.S. based, it is reasonable to hypothesize that they would exhibit risk factors similar to that of their home market. To quantify this, we estimate the Carhart (1997) four-factor model on the first half of the empirical sample of monthly returns (i.e. the in-sample period) for each sustainable stock index. The Carhart (1997) four-factor model is the Fama and French (1992, 1993) three-factor model with the additional momentum risk factor. We operationalize the Carhart (1997) four-factor model by estimating the following ordinary least squares (OLSs) regression on the in-sample data

\[ (R_t - R_{f,t}) = \alpha + \beta_1(R_{m,t} - R_{f,t}) + \beta_2(SMB_t) + \beta_3(HML_t) + \beta_4(UMD_t) + \varepsilon_t \]

where \( R_t \) represents the sustainable index return, \( R_{f,t} \) is the risk-free rate of return estimated from the U.S. government 1-month Treasury bill, \( \alpha \) is the intercept term, \( \beta_1 \) is the first regression coefficient, \( R_{m,t} \) is the U.S. market proxy, \( \beta_2 \) is the second regression coefficient, \( SMB_t \) is the Fama and French (1992, 1993) factor mimicking portfolio for size, \( \beta_3 \) is the third regression coefficient, \( HML_t \) is the Fama and French (1992, 1993) factor mimicking portfolio for book-to-market, \( \beta_4 \) is the fourth regression coefficient, \( UMD_t \) which is the Carhart (1997) factor mimicking portfolio for the 12-month return momentum, and \( \varepsilon_t \) is the regression error terms.

4.1.2. Out-of-sample verification

The OLS regression coefficient estimates from the in-sample model are then employed with the Carhart (1997) four risk factors to construct a set of out-of-sample model returns for each sustainable stock index. We then employ a variety of hypothesis tests to verify whether the out-of-sample model returns are statistically the same as the out-of-sample empirical returns of each sustainable stock index. To test whether the in-sample regression coefficients are robust at modelling the
out-of-sample returns, we employ a parametric *t*-test to test for differences in means, the non-parametric Wilcoxon signed rank test to test for differences in medians and the parametric *F*-test to examine the difference in the variance of returns.

4.1.3. Long-term returns of sustainable stock indices

The in-sample regression coefficients are then employed with the four-factor Carhart (1997) risk factors to construct the historical monthly returns of the three sustainable stock indices from 1927 to the start date of their empirical returns. We then examine the risk and return characteristics of the short empirical history of the three indices and compare them with their reconstructed long-term returns.

4.2. Portfolio selection

This study employs the Markowitz (1952) mean-variance (MV) analysis and the Rockafellar and Uryasev (2000) M-CVaR frameworks to analyse the returns of the three sustainable stock indices in a portfolio context. We impose a short selling constraint on all asset classes and we exclude portfolio leverage to reflect the practical limitations of asset allocation decisions with sustainable stock indices. We employ monthly returns for the January 1970 to October 2010 period as that is the available sample period with reliable U.S. stock and U.S. bond index returns.

The portfolio analysis in this study is viewed from the perspective of an U.S investor. There are two rationales for the U.S. approach to the portfolio analysis in this study. First, the three sustainable stock indices are composed of U.S. stocks and therefore, they are expected to exhibit U.S. stockmarket characteristics. This means that these sustainable stock indices will typically compete with conventional U.S. stocks as an asset class for inclusion in an investment portfolio. Modern portfolio theory implicitly suggests that sustainable stock indices will not dominate U.S. stocks as sustainable stock indices represent a specific market sector or segment and do not reflect the universe of highly diversified U.S. stock returns. The second rationale is the lack of available bond index returns outside of the U.S.; therefore, we develop the portfolio analysis from a U.S. investor perspective.

5. Results

5.1. In-sample multi-factor model

Table 3 presents the in-sample regression estimates. The results show that the U.S. market beta (i.e. Rm-Rf) is the most significant risk factor that explains the variation in returns of all three sustainable stock indices. This is expected given that the companies included in these sustainable indices come from larger U.S. broad-based stock indices. The DJSI reports a lower market beta coefficient of 0.8988, thereby representing lower systematic risk than the other two indices. Table 3 also reveals that all three indices report small but significantly negative regression coefficient estimates for the SMB factor, which reflects the design of these sustainable indices which are concentrated towards large market capitalization firms. Table 3 also reports small but significantly negative regression coefficient estimates for the HML factor, which suggests that all three indices are dominated by growth stocks rather than value stocks. These characteristics corroborate with the public information of exchange traded funds (ETFs) based on these indices (via Morningstar fund ratings announcements), which confirms that all three sustainable stock indices are classified within the ‘large market capitalization-growth’ investment style category. This ETF information supports the evidence of the negative SMB and HML risk factors. Table 3 also reports small negative regression coefficient estimates for momentum (up-minus-down
zero financing portfolio momentum risk factor, UMD) for all three indices that cannot be explained by this analysis. The intercept terms in Table 3 suggest that the KLD and DJSI earn 26 and 30 basis points per month of excess returns which are statistically significant over the 10 and 5 years in-sample period, respectively. Conversely, the CSI loses 14 basis points per month of excess returns over its 5 year in-sample period after controlling for the Carhart (1997) risk factors. Overall, the adjusted-\(R^2\) range from 0.90 to 0.98 suggests that the Carhart (1997) four-factor model is effective at explaining the variation of returns of these three sustainable stock indices during their in-sample periods.

### Table 3. In-sample regression estimates.

<table>
<thead>
<tr>
<th>Variables</th>
<th>(C)</th>
<th>Rm-Rf</th>
<th>SMB</th>
<th>HML</th>
<th>UMD</th>
<th>Adj (R^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient</td>
<td>0.0026</td>
<td>1.0102</td>
<td>-0.1564</td>
<td>-0.1578</td>
<td>-0.0824</td>
<td>0.9402</td>
</tr>
<tr>
<td>S.E.</td>
<td>0.0010</td>
<td>0.0319</td>
<td>0.0340</td>
<td>0.0529</td>
<td>0.0692</td>
<td></td>
</tr>
<tr>
<td>(t)-statistic</td>
<td>2.4581(^*)</td>
<td>31.7108(**)</td>
<td>-4.5991(**)</td>
<td>-2.9846(**)</td>
<td>-1.1913</td>
<td></td>
</tr>
<tr>
<td>(p)-value</td>
<td>0.0154</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0035</td>
<td>0.2359</td>
<td></td>
</tr>
<tr>
<td>Coefficient</td>
<td>0.0030</td>
<td>0.8988</td>
<td>-0.1701</td>
<td>-0.2193</td>
<td>-0.0805</td>
<td>0.9035</td>
</tr>
<tr>
<td>S.E.</td>
<td>0.0015</td>
<td>0.0390</td>
<td>0.0375</td>
<td>0.0653</td>
<td>0.0309</td>
<td></td>
</tr>
<tr>
<td>(t)-statistic</td>
<td>2.0869(^*)</td>
<td>23.0482(**)</td>
<td>-4.5375(**)</td>
<td>-3.3560(**)</td>
<td>-2.6034(^*)</td>
<td></td>
</tr>
<tr>
<td>(p)-value</td>
<td>0.0408</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0013</td>
<td>0.0114</td>
<td></td>
</tr>
<tr>
<td>Coefficient</td>
<td>-0.0014</td>
<td>1.0185</td>
<td>-0.0893</td>
<td>-0.1433</td>
<td>-0.0360</td>
<td>0.9779</td>
</tr>
<tr>
<td>S.E.</td>
<td>0.0010</td>
<td>0.0289</td>
<td>0.0407</td>
<td>0.0321</td>
<td>0.0149</td>
<td></td>
</tr>
<tr>
<td>(t)-statistic</td>
<td>-1.3796(**)</td>
<td>35.2485(**)</td>
<td>-2.1938(^*)</td>
<td>-4.4597(**)</td>
<td>-2.4126(**)</td>
<td></td>
</tr>
<tr>
<td>(p)-value</td>
<td>0.1729</td>
<td>0.0000</td>
<td>0.0322</td>
<td>0.0000</td>
<td>0.0190</td>
<td></td>
</tr>
</tbody>
</table>

This table presents the Carhart (1997) four-factor model regression results for the KLD, DJSI and CSI for the first half of their respective sample periods. Panel A presents the regression estimates for the KLD in-sample period from May 1990 to July 2000 consisting of 123 monthly observations. Panel B reports the DJSI in-sample period from January 1999 to November 2004 consisting of 71 monthly observations. Panel C shows the CSI in-sample period from May 2000 to July 2005 consisting of 63 monthly observations. The table presents the regression estimates with the intercept (C), Fama and French (1992, 1993) excess return (Rm-Rf), Small-Minus-Big zero financing portfolio risk factor (SMB), high-minus-low book value to equity ratio zero financing portfolio risk factor (HML), Carhart (1997) Up-minus-down zero financing portfolio momentum risk factor (UMD) and the respective adjusted \(R^2\). The table reports the regression coefficients, heteroscedasticity and autocorrelation-consistent standard errors, \(t\)-statistics and \(p\)-values. \(^*\) and \(**\) denote statistical significance at the 5% and 1% levels, respectively.

To ensure that the Carhart (1997) risk factors in the in-sample period genuinely reflect the true economic risks of the three sustainable indices over the long term, one would expect that the estimated regression coefficients in Table 3 to be robust in the out-of-sample period. We perform an out-of-sample analysis for every index by constructing a set of model returns by employing the Carhart (1997) factor loadings in Table 3. We then calculate the difference between the out-of-sample empirical returns of the KLD, DJSI and CSI and their respective model returns using the Carhart (1997) in-sample regression coefficients. To verify the robustness of the out-of-sample returns, we estimate whether the differences between the empirical and model returns are statistically significant. To achieve this, we calculate the parametric \(t\)-tests, non-parametric Wilcoxon signed-rank tests and \(F\)-test of the time series on the out-of-sample returns. For all three indices, the test statistics reveal that the differences in mean, median and variance are

### 5.2. Out-of-sample analysis

To ensure that the Carhart (1997) risk factors in the in-sample period genuinely reflect the true economic risks of the three sustainable indices over the long term, one would expect that the estimated regression coefficients in Table 3 to be robust in the out-of-sample period. We perform an out-of-sample analysis for every index by constructing a set of model returns by employing the Carhart (1997) factor loadings in Table 3. We then calculate the difference between the out-of-sample empirical returns of the KLD, DJSI and CSI and their respective model returns using the Carhart (1997) in-sample regression coefficients. To verify the robustness of the out-of-sample returns, we estimate whether the differences between the empirical and model returns are statistically significant. To achieve this, we calculate the parametric \(t\)-tests, non-parametric Wilcoxon signed-rank tests and \(F\)-test of the time series on the out-of-sample returns. For all three indices, the test statistics reveal that the differences in mean, median and variance are
statistically insignificant. These results reveal that the out-of-sample model returns are statistically similar to the out-of-sample empirical returns for all three sustainable stock indices. We can therefore conclude that the Carhart (1997) four-factor model estimated on the in-sample data in Table 3 is robust at explaining the empirical out-of-sample returns. These results provide us with statistical evidence to allow us to employ the Carhart (1997) in-sample regression coefficients in Table 3 to construct long-term monthly returns from January 1927.

5.3. Long-term performance

Panel A of Table 4 summarizes the return and risk statistics of the three sustainable stocks indices by comparing their short-term empirical performance with their long-term modelled history. To provide a direct comparison, Panel B of Table 4 presents the statistics for the broad-based U.S. Composite Stock Index returns. Overall, the findings in Table 4 provide a number of new insights in relation to the performance and risk of sustainable stock indices. Table 4 reveals that all three sustainable stock indices exhibit stronger performance (i.e. higher mean returns) over the long term than in the short term, regardless of the commencement date of the index. Second, Table 4 shows that the 99% VaR and 99% CVaR (i.e. the extreme tail-risk estimates) of sustainable stock indices in the short term are a magnitude smaller than what is likely to be experienced by investors over the long term. These results are consistent with the classical trade-off in finance.

Table 4. Summary statistics of systematic returns of the KLD, DJSI and Calvert indices.

<table>
<thead>
<tr>
<th>Panel A: Sustainable stock indices</th>
<th>Mean</th>
<th>SD</th>
<th>Median</th>
<th>Minimum</th>
<th>Maximum</th>
<th>VaR 95%</th>
<th>VaR 99%</th>
<th>CVaR 95%</th>
<th>CVaR 99%</th>
</tr>
</thead>
<tbody>
<tr>
<td>DJSI: 1/1927–12/1998</td>
<td>0.73</td>
<td>4.81</td>
<td>0.91</td>
<td>-25.16</td>
<td>30.95</td>
<td>-6.61</td>
<td>-13.16</td>
<td>-10.88</td>
<td>-18.60</td>
</tr>
<tr>
<td>DJSI: 1/1999–10/2010</td>
<td>0.24</td>
<td>4.94</td>
<td>0.65</td>
<td>-17.11</td>
<td>10.42</td>
<td>-8.00</td>
<td>-10.91</td>
<td>-10.38</td>
<td>-14.04</td>
</tr>
<tr>
<td>Calvert: 1/1927–4/2000</td>
<td>0.46</td>
<td>5.50</td>
<td>0.71</td>
<td>-29.13</td>
<td>36.26</td>
<td>-7.77</td>
<td>-15.26</td>
<td>-12.84</td>
<td>-21.51</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Broad-based U.S. stock index</th>
<th>Mean</th>
<th>SD</th>
<th>Median</th>
<th>Minimum</th>
<th>Maximum</th>
<th>VaR 95%</th>
<th>VaR 99%</th>
<th>CVaR 95%</th>
<th>CVaR 99%</th>
</tr>
</thead>
</table>

This table presents the mean returns, standard deviations, medians, minimum monthly returns, maximum monthly returns, empirical VaR and empirical CVaR (reported as a percentage) at the 95 and 99% confidence levels for the systematic returns of the three sustainable stock indices for the empirical sample period and the long-term model returns commencing in January 1927. Panel A reports the statistics for the KLD, DJSI and CSI. For comparative purposes, Panel B reports the statistics for the U.S. Composite Stock Index.
between risk and return. These findings of higher return and tail-risk over the long term must be evaluated against the broad based U.S. stockmarket index in panel B which also experienced lower returns and tail-risk since the 1990s. Put simply, the increase in return and tail-risk of these sustainable stock indices over the long-term is commensurate with the return and risk characteristics of the overall U.S. stockmarket.

5.4. Risk characteristics

To better understand the return-risk characteristics of the sustainable stock index returns over the long term, Table 5 presents a variety of absolute and relative risk statistics estimated by splicing the long-term modelled returns with the short-term empirical returns to form a complete monthly time series from 1927–2010. Panel A of Table 5 shows that the raw returns of the KLD outperformed U.S. stocks while the DJSI and CSI underperformed over the entire 84 year period. The

<table>
<thead>
<tr>
<th>Variable</th>
<th>KLD</th>
<th>DJSI</th>
<th>CSI</th>
<th>U.S. Stocks</th>
<th>U.S. T bills</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: Moments of the distribution</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.99%</td>
<td>0.66%</td>
<td>0.39%</td>
<td>0.91%</td>
<td>0.30%</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>5.38%</td>
<td>4.83%</td>
<td>5.47%</td>
<td>5.47%</td>
<td>0.25%</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.0789</td>
<td>-0.0617</td>
<td>0.0091</td>
<td>0.1323</td>
<td>1.0214</td>
</tr>
</tbody>
</table>

| Panel B: Standardised tail Z-scores |
| 1st quantile              | -2.8708| -2.8940| -2.9411| -2.9746     | -1.7717      |
| 5th quantile              | -1.6284| -1.6786| -1.6235| -1.6055     | -1.7118      |
| 97.5th quantile           | 1.7701 | 1.7157 | 1.7758 | 1.6719      | 1.8236       |
| 99th quantile             | 2.2200 | 2.1281 | 2.2265 | 2.3195      | 2.3030       |

| Panel C: Empirical tail risk measures |
| 95% VaR (empirical)        | -7.77% | -7.43% | -8.47% | -7.86%      | 0.01%        |
| 99% VaR (empirical)        | -14.30%| -12.34%| -14.81%| -14.95%     | -0.01%       |
| 95% CVaR (empirical)       | -11.85%| -10.95%| -12.72%| -12.62%     | 0.00%        |
| 99% CVaR (empirical)       | -19.63%| -17.89%| -20.49%| -26.58%     | -0.04%       |

| Panel D: Quantile regression CAPM betas against U.S. stocks |
| Q 1st                    | 0.9470 | 0.8389 | 0.8553 | –           | –            |
| Q 2.5th                  | 0.9439 | 0.8464 | 0.8722 | –           | –            |
| Q 5th                    | 0.9626 | 0.8532 | 0.8659 | –           | –            |
| Q 10th                   | 0.9735 | 0.8670 | 0.8747 | –           | –            |
| Q 25th                   | 0.9825 | 0.8878 | 0.8780 | –           | –            |
| Q 50th                   | 0.9925 | 0.8878 | 0.8817 | –           | –            |
| Q 75th                   | 0.9826 | 0.8785 | 0.8769 | –           | –            |
| Q 90th                   | 0.9783 | 0.8582 | 0.8688 | –           | –            |
| Q 95th                   | 0.9702 | 0.8457 | 0.8571 | –           | –            |
| Q 97.5th                 | 0.9636 | 0.8453 | 0.8615 | –           | –            |
| Q 99th                   | 1.0040 | 0.8859 | 0.9111 | –           | –            |

This table presents the long-term return behaviour of the KLD, DJSI, CSI, the Fama and French U.S. market raw return and the U.S. government 1 month Treasury-Bill returns for the full period from January 1927 to October 2010. Panel A reports the first four moments of the distribution of returns. Panel B presents the standardized returns of each investment. The 1%, 5%, 95% and 99% quantiles for a normal distribution are -2.3263, -1.6449, 1.6449 and 2.3263, respectively. Panel C reports the empirical VaR and CVaR estimates at the 95 and 99% confidence intervals, respectively. Panel D presents the quantile regression single-factor CAPM beta estimates of the excess returns of the three sustainable stock indices against U.S. excess returns at the 1st, 2.5th, 5th, 10th, 25th, 50th, 75th, 90th, 95th, 97.5th and 99th quantile levels.
standard deviations were relatively similar with the exception of the DJSI, which exhibits a marginally lower volatility in returns.

Panel B of Table 5 shows that the Z-scores of the first quantiles of the distribution of returns of the three sustainable stock indices is more negative than the losses expected from normally distributed returns. This suggests that the tail-risks of sustainable stock indices are non-normal; however, the left-tail is commensurate with the tail-risk estimate from the broad based U.S. stockmarket.

Panel C of Table 5 compares the empirical VaR and CVaR estimates over eight decades. Panel C shows that the three sustainable index returns exhibit similar tail-risks as the broad-based U.S. stockmarket with the exception of the 99% CVaR. The three sustainable index returns exhibit smaller 99% CVaRs than the overall U.S. stockmarket. This positive outcome in the tail-risk statistic motivates us to examine the risk-adjusted empirical betas and alternative portfolio selection frameworks that consider the extreme left-tail observations as an alternative measure of risk.

Motivated by the non-normality of returns in the left-tail of the distribution of returns, panel D of Table 5 presents quantile regression-based CAPM betas. The betas suggest that all three sustainable indices exhibit lower systematic risk than the overall market. The findings also suggest that the betas of the sustainable stock indices are sufficiently stable and consistent even when measuring systematic risk in the left tail of the distribution of returns. The DJSI exhibits the lowest systematic beta, which is consistent given that it also reports the lowest volatility of returns and the lowest tail-risk of the three sustainable stock indices.

5.5. **Portfolio selection**

This section of the study reports the historical mean-variance (MV) and mean-conditional value-at-risk (M-CVaR) portfolio analyses. The portfolio analysis examines the U.S. risk-free rate, U.S. stocks, U.S. bonds and the sustainable stock indices from January 1970 to October 2010. All tests in this section employ no short sales and no leverage to reflect the genuine investment governance and constraints of a typical pension fund.

5.5.1. **Mean-variance analysis**

Table 6 reports the summary statistics of the long-term returns employed in the mean-variance analyses. The striking feature from panel A is that the risk-free rate earned 0.45% per month (5.4% per annum) which was nearly double the return of the CSI which earned 0.25% per month (3.0% per annum). Another interesting feature of panel A is that U.S. bonds earned a higher return than the DJSI and the CSI. Panels A and B of Table 6 clearly show that the KLD, DJSI and CSI exhibit variations in their long-term mean returns even though they possess common Carhart (1997) risk factors, very high correlations with each other and similar volatility in returns. It is evident that the stock selection procedures of these sustainable stock indices cause sufficient variation in the first moment of the distribution of returns.

The historical inputs in Table 6 are employed in traditional MV analyses, which are reported in Table 7. The portfolio optimizations reveal that the KLD is the preferred sustainable stock index in a traditional MV analysis. Table 7 also shows that the MV analysis prefers U.S. stocks and U.S. bonds rather than the DJSI and CSI on the basis of their historical mean returns, standard deviations and correlations.

5.5.2. **Mean-conditional value at risk analysis**

To examine whether sustainable stock indices minimise portfolio tail-risk, we calculate the various M-CVaR portfolios at the 90, 95 and 99% CVaR confidence intervals. Table 8 presents
the results of the MV and M-CVaR portfolio optimisations with stocks, bonds and the KLD. The estimates show that the KLD is an undesirable investment for minimizing 90 and 95% CVaR; however, the KLD becomes more desirable when minimizing extreme tail-risk at the 99% CVaR level. This finding suggests that the KLD benefits long-term investors with low-to-medium target returns (i.e. low-to-medium risk); however, it is not effective at reducing tail-risk for portfolios targeting high monthly returns (and high levels of risk). We do not report the results of the DJSI and CSI as both were found to be undesirable in both MV and M-CVaR portfolio optimizations.

Table 6. Moments and correlations (1970 to 2010).

<table>
<thead>
<tr>
<th>U.S. Stocks</th>
<th>U.S. Bonds</th>
<th>KLD</th>
<th>DJSI</th>
<th>CSI</th>
<th>T Bills</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: First two moments</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.87</td>
<td>0.66</td>
<td>0.95</td>
<td>0.52</td>
<td>0.25</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>4.53</td>
<td>2.03</td>
<td>4.77</td>
<td>4.49</td>
<td>4.91</td>
</tr>
<tr>
<td><strong>Panel B: Correlation coefficients</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U.S. stocks</td>
<td>1.00</td>
<td>0.275**</td>
<td>0.987**</td>
<td>0.977**</td>
<td>0.980**</td>
</tr>
<tr>
<td>U.S. bonds</td>
<td>1.00</td>
<td>0.273**</td>
<td>0.251**</td>
<td>0.254**</td>
<td></td>
</tr>
<tr>
<td>KLD</td>
<td>1.00</td>
<td>0.982**</td>
<td>0.987**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DJSI</td>
<td>1.00</td>
<td>0.984**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calvert</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This table presents the first two moments of the distribution of returns and the correlations for the investments employed in the mean-variance analysis for the period January 1970 to October 2010. The assets employed in the mean variance analysis are the MSCI US Equity Index as the proxy for U.S. Stocks, the spliced bond return data series as the proxy for U.S. Bonds, the MSCI KLD 400 Social Index (KLD), the DJSI U.S. Index (DJSI), the Calvert Social Index (CSI) and the U.S. Government 1 month Treasury-Bill return as the risk-free rate (T-Bills). Panel A reports the first two moments of the distribution of returns for each investment. Panel B presents the correlation coefficients of each investment. * and ** denote statistical significance at the 5 and 1% levels, respectively.

Table 7. Mean variance analysis (1970 to 2010).

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>Ret.</th>
<th>SD</th>
<th>Sharpe</th>
<th>Skew.</th>
<th>Kurt.</th>
<th>w(%) Stocks</th>
<th>w(%) Bonds</th>
<th>w(%) Sust. Idx</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: KLD</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MVP</td>
<td>0.67</td>
<td>2.00</td>
<td>–</td>
<td>0.159</td>
<td>7.914</td>
<td>8.1</td>
<td>91.9</td>
<td>0.0%</td>
</tr>
<tr>
<td>ORP</td>
<td>0.75</td>
<td>2.31</td>
<td>0.13/0.44</td>
<td>0.103</td>
<td>4.805</td>
<td>0.0</td>
<td>68.3</td>
<td>31.7%</td>
</tr>
<tr>
<td><strong>Panel B: DJSI</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MVP</td>
<td>0.64</td>
<td>1.99</td>
<td>–</td>
<td>0.155</td>
<td>7.789</td>
<td>0.0</td>
<td>90.7</td>
<td>9.3%</td>
</tr>
<tr>
<td>ORP</td>
<td>0.71</td>
<td>2.17</td>
<td>0.12/0.41</td>
<td>0.082</td>
<td>5.360</td>
<td>27.5</td>
<td>72.5</td>
<td>0.0%</td>
</tr>
<tr>
<td><strong>Panel C: CSI</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MVP</td>
<td>0.67</td>
<td>2.00</td>
<td>–</td>
<td>0.159</td>
<td>7.914</td>
<td>8.1</td>
<td>91.9</td>
<td>0.0%</td>
</tr>
<tr>
<td>ORP</td>
<td>0.71</td>
<td>2.17</td>
<td>0.12/0.41</td>
<td>0.082</td>
<td>5.360</td>
<td>27.5</td>
<td>72.5</td>
<td>0.0%</td>
</tr>
</tbody>
</table>

This table presents the mean-variance portfolio (MV) optimization results for the minimum variance portfolio (MVP) and the optimal risky portfolio (ORP) based on the monthly returns for the January 1970 to October 2010 sample period. The investment universe of risky assets in these tests consists of U.S. Stocks, U.S. bonds and the respective sustainable stock index. Panel A reports the portfolio results that include the MSCI KLD 400 Social Index (KLD) as the sustainable stock index. Panel B reports the portfolio results that include the DJSI U.S. Index (DJSI) as the sustainable stock index. Panel C reports the portfolio results that include the Calvert Social Index (CSI) as the sustainable stock index. Ret denotes the portfolio monthly return. SD denotes the monthly standard deviation of portfolio returns. Sharpe denotes the monthly and annualized Sharpe ratio of portfolio returns. Skew denotes the skewness of portfolio returns. Kurt denotes the kurtosis of portfolio returns. The term w(%) denotes the optimal portfolio weight to each asset class.
To better understand the relationship between MV and M-CVaR portfolios, Table 9 presents the standard deviation and tail-risk estimates from the portfolio optimizations based on an investment universe of stocks, bonds and the KLD. Table 9 shows that the differences in the portfolio CVaR between the MV and M-CVaR optimizations grow in magnitude as we attempt to minimize CVaR towards the extreme left tail of the distribution of returns. Furthermore, Table 9 reveals that the largest differences in portfolio tail-risk between MV and M-CVaR portfolios occur in the low-risk portfolio combinations. This finding suggests that MV and M-CVaR investors can achieve similar risk outcomes unless they are constructing low-risk portfolios.

Overall, the portfolio analysis suggests that the KLD provides portfolio diversification benefits to both MV and M-CVaR investors over the long term. Conversely, the portfolio analysis demonstrates that the DJSI and the CSI are undesirable in a historical portfolio analysis due to their inadequate mean returns. The portfolio analysis suggests that the differences in the constituents of these sustainable stock indices causes sufficient variation in the mean returns to significantly alter the optimal portfolio weights of long-term investors even though all three indices exhibit similar risk factors and characteristics.

6. Conclusion

This study examined three popular sustainable stock indices and quantified their return and risk characteristics over the long term. To address the empirical challenge of their relatively short performance history, we employed the Agarwal and Naik (2004) procedure and Carhart (1997) four-factor model to replicate the returns of these sustainable stock indices from 1927 through 2010. This study found that the variation of returns for all three sustainable stock indices can be

Table 8. Optimal portfolio weights for the MSCI KLD 400 social index.

<table>
<thead>
<tr>
<th>Monthly return</th>
<th>Portfolio weight MV</th>
<th>Portfolio weight 90% M-CVaR</th>
<th>Portfolio weight 95% M-CVaR</th>
<th>Portfolio weight 99% M-CVaR</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.46</td>
<td>0.3</td>
<td>0.6</td>
<td>0.6</td>
<td>0.6</td>
</tr>
<tr>
<td>0.48</td>
<td>3.1</td>
<td>2.6</td>
<td>2.8</td>
<td>3.0</td>
</tr>
<tr>
<td>0.51</td>
<td>5.8</td>
<td>4.6</td>
<td>5.1</td>
<td>7.0</td>
</tr>
<tr>
<td>0.53</td>
<td>8.6</td>
<td>6.9</td>
<td>7.8</td>
<td>10.4</td>
</tr>
<tr>
<td>0.56</td>
<td>11.4</td>
<td>9.3</td>
<td>10.0</td>
<td>13.8</td>
</tr>
<tr>
<td>0.59</td>
<td>14.2</td>
<td>12.1</td>
<td>12.3</td>
<td>17.3</td>
</tr>
<tr>
<td>0.61</td>
<td>16.9</td>
<td>14.7</td>
<td>14.9</td>
<td>20.8</td>
</tr>
<tr>
<td>0.64</td>
<td>19.7</td>
<td>17.3</td>
<td>17.1</td>
<td>24.1</td>
</tr>
<tr>
<td>0.67</td>
<td>22.5</td>
<td>19.4</td>
<td>19.6</td>
<td>27.4</td>
</tr>
<tr>
<td>0.69</td>
<td>25.3</td>
<td>21.7</td>
<td>22.5</td>
<td>30.7</td>
</tr>
<tr>
<td>0.72</td>
<td>28.0</td>
<td>25.1</td>
<td>26.1</td>
<td>34.1</td>
</tr>
<tr>
<td>0.74</td>
<td>30.8</td>
<td>27.9</td>
<td>27.9</td>
<td>37.4</td>
</tr>
<tr>
<td>0.77</td>
<td>38.2</td>
<td>39.4</td>
<td>39.4</td>
<td>40.7</td>
</tr>
<tr>
<td>0.80</td>
<td>47.1</td>
<td>48.1</td>
<td>48.1</td>
<td>48.9</td>
</tr>
<tr>
<td>0.82</td>
<td>55.9</td>
<td>53.8</td>
<td>53.8</td>
<td>54.7</td>
</tr>
<tr>
<td>0.85</td>
<td>64.7</td>
<td>65.3</td>
<td>65.3</td>
<td>66.4</td>
</tr>
<tr>
<td>0.87</td>
<td>73.5</td>
<td>71.0</td>
<td>71.0</td>
<td>72.2</td>
</tr>
<tr>
<td>0.90</td>
<td>82.4</td>
<td>82.5</td>
<td>82.5</td>
<td>82.9</td>
</tr>
<tr>
<td>0.93</td>
<td>91.2</td>
<td>92.5</td>
<td>92.5</td>
<td>92.6</td>
</tr>
<tr>
<td>0.95</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>

This table presents the optimal portfolio weights from the mean-variance (MV) analysis and the M-CVaR portfolio frameworks. The risky assets in the investment universe includes U.S. stocks, U.S. bonds and the MSCI KLD 400 Social Index (KLD).
explained with market beta and negative exposures to the Fama and French (1992, 1993) SMB and HML risk factors.

Despite the identification of common risk factors, this study estimated a significant variation in the first moment of returns of these sustainable stock indices over the long term. This finding suggests that the stock selection screens of these indices are sufficiently different to cause a significant variation in the return outcomes for investors even though they all exhibit common market beta, large market capitalization and growth risk factors. The KLD was found to be the best performing index while the CSI was the worst. The DJSI exhibited the lowest risk profile among the three indices.

In terms of risk, this study found that the CVaR losses of sustainable stock indices over the long-term are worse than those experienced by investors in the recent past. Despite the higher levels of historical tail-risk, the CVaR estimates of these sustainable stocks indices were found to be commensurate with the tail-risk of the overall U.S. stockmarket.

Finally, we constructed optimal portfolios using the MV and M-CVaR frameworks. The portfolio analysis found that the KLD dominated both MV and M-CVaR portfolios due to its attractive long-term average rate of return. The DJSI and CSI indices were undesirable in a portfolio analysis as their long-term average returns were not high enough to offset any portfolio diversification or tail-risk benefits.

The goal of this study was to explore the long-term return behaviour of sustainable stock indices through the prism of a long-term investor. The findings highlight that the return outcomes from sustainable stock indices are highly influenced by the stock screening methods employed by
each index provider, resulting in divergent mean returns. While we found differences in return outcomes, the screening techniques of these indices tend to exhibit common risk factors, similar volatility and tail-risk characteristics. The portfolio analysis in this study suggests that these differences in the construction of sustainable index returns causes wide variations in return outcomes that have important implications for pension funds and long-term investors.

Acknowledgement
The authors gratefully acknowledge the support from the Griffith Business School New Research Initiatives Program (NRIP) in Superannuation and Funds Management. The authors are responsible for all errors.

Notes
1. Other studies including Lee and Faff (2009) and Lee, Faff, and Langfield-Smith (2009) examine the Dow Jones Sustainability Indexes Group by employing the Carhart (1997) multifactor model and include additional risk factors to include industry and country momentum risk factors. These studies incorporate these additional momentum risk factors as their study examines companies at the firm level and are able to construct these industry and country momentum effects.
2. The Fama and French (1992, 1993) Rm, SMB and HML risk factors and Carhart (1997) momentum risk factor are well recognized in the finance literature as the most important risk factors in explaining the variation in U.S. stock returns.
3. We gratefully acknowledge the resources of the Professor Kenneth French data library at: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html
4. CVaR is also referred to as expected shortfall (ES) and expected tail loss (ETL) in various studies in the finance literature.

References


