

# Responsible Investing: Delivering competitive performance

Amy O'Brien  
Head of Responsible Investment

Lei Liao, CFA  
Jim Campagna, CFA  
Quantitative Portfolio Managers  
Social Choice Equity Strategy

## RI indexes delivered market-like returns and risk, while pursuing social goals

### Executive summary

- Growing interest in Responsible Investing (RI)—also known as Socially Responsible Investing (SRI)—raises an important question: Does pursuing social goals—limiting the range of potential investment opportunities—require sacrificing performance?
- A TIAA analysis of leading RI equity indexes over the long term found no statistical difference in returns compared to broad market benchmarks, suggesting the absence of any systematic performance penalty.
- Moreover, incorporating environmental, social and governance (ESG) criteria in security selection did not entail additional risk. RI indexes and their broad market counterparts had similar risk profiles, based on Sharpe ratios and standard deviation measures.
- Although return patterns were *similar* over the long term, there were significant return and tracking error differences between RI indexes and broad market benchmarks over shorter periods. By narrowing the range of eligible investments, the RI process introduced biases that caused short-term index performance to deviate from broad market benchmarks, resulting in tracking error.
- RI index construction methodology is an important determinant of tracking error. Investors should consider specific ESG methodology and the relevant market benchmark when selecting an RI strategy.

### Growing interest in RI, but performance questions persist

Interest in Responsible Investing (RI) is increasing rapidly. From 2003 to 2014, RI assets in the U.S. more than tripled to reach \$6.57 trillion,<sup>1</sup> according to the Forum for Sustainable and Responsible Investment (US SIF Foundation). This represents nearly 18% of assets under professional investment management in the U.S. as tracked by Cerulli Associates.

RI strategies apply various environmental, social and governance (ESG) criteria in selecting public companies for inclusion in a portfolio. The process of incorporating nonfinancial criteria restricts the range of investment opportunities, potentially limiting returns. On the other hand, companies that wisely manage ESG risks and opportunities may also improve financial measures, potentially enhancing stock performance.

The key question for investors: Does investing in an RI strategy require sacrificing performance or taking on additional risk, compared to a broad market index?

Many studies on the performance of RI mutual funds versus non-RI funds have attempted to answer this question.<sup>2</sup> However, the range, variety and diversity of RI fund management strategies make apples-to-apples comparisons difficult. Instead, TIAA sought answers through a simpler comparison, analyzing the performance of several leading RI indexes versus broad market benchmarks. We focused on equity strategies because indexes with longer-term track records are readily available—and represent the majority of RI assets. (For performance of the TIAA-CREF Social Choice Equity Fund vs. the Russell 3000 Index, see Appendix 6 on page 15.) It is important to note that RI indexes themselves are not perfectly comparable due to differences in index construction and ESG evaluation processes. However, they provide a close proxy for RI as a strategy versus the broad market.

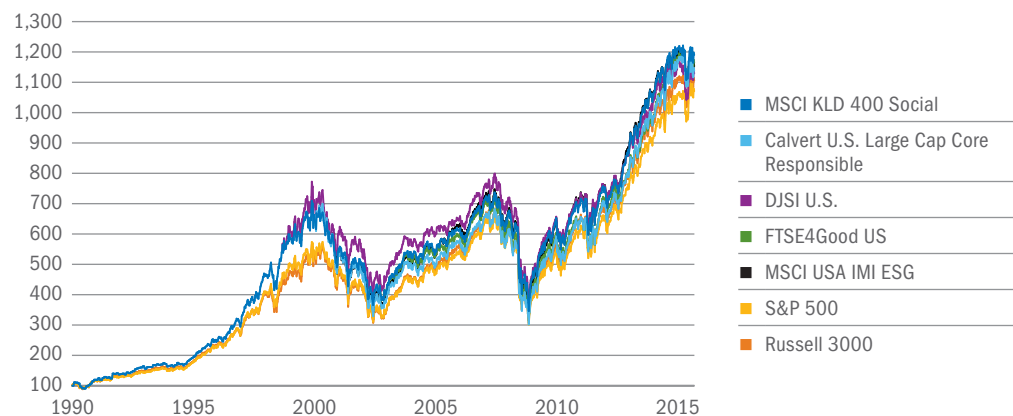
### How RI performed versus broad market indexes

We selected five widely known U.S. equity RI indexes with track records of at least 10 years: Calvert U.S. Large Cap Core Responsible Index, Dow Jones Sustainability U.S. Index (DJSI U.S.), FTSE4Good US Index, MSCI KLD 400 Social Index, and MSCI USA IMI ESG Index.<sup>3</sup> We compared returns for these indexes with two widely recognized U.S. equity-based indexes, the Russell 3000 and S&P 500 indexes.<sup>4</sup> We also examined volatility measures and calculated Sharpe ratios to understand risk-adjusted results. Finally, we compared index returns with respective benchmarks to determine tracking error rates. We also sought to determine whether differences in results were statistically significant or caused by random variation.

The result: our analysis found **no statistical difference** in RI index returns compared to the two broad market benchmarks. In other words, RI can achieve comparable performance over the long term without additional risk, despite using a smaller universe of securities meeting ESG criteria. Exhibit 1 illustrates the similarity of cumulative returns for RI indexes and broad market benchmarks over the long term.

#### Exhibit 1: Comparing returns of RI indexes and broad market indexes

Cumulative returns: Five U.S. RI indexes vs. S&P 500 and Russell 3000 indexes



Data through 12/31/15. Series indexed to 100, inception dates: S&P 500, Russell 3000, and MSCI KLD 400 Social, 5/4/1990; DJSI U.S., 1/1/1999; Calvert U.S. Large Cap Core Responsible, 4/28/2000; MSCI USA IMI ESG, 12/22/2000; and FTSE4Good US, 1/3/2003. MSCI indexes include aggregated, multisource histories prior to acquisition on 9/1/2010.

It is not possible to invest in an index. Performance for indices does not reflect investment fees or transactions costs.

Sources: FactSet Research Systems Inc., Morningstar, Inc., MSCI Inc., and TIAA

## RI returns were comparable to broad market indexes

Returns for the RI indexes were similar to each other and compared to the broad market. Ten-year average annual performance for the five U.S. RI indexes ranged from 6.05% to 7.49% versus 7.31% and 7.35% for the S&P 500 and Russell 3000 indexes, respectively. The gap between best and worst average annual performance spanned 144 basis points. (Exhibit 2)

### Exhibit 2: RI index returns were comparable to broad market indexes

Index average annual returns (as of 12/31/2015)		■ Top performer ■ Bottom performer			
	1 Yr	3 Yr	5 Yr	10 Yr	
MSCI USA IMI ESG	-1.56%	13.90%	11.11%	6.82%	
MSCI KLD 400 Social	0.94%	15.72%	12.26%	7.22%	
Calvert U.S. Large Cap Core Responsible	1.44%	16.47%	13.31%	7.49%	
FTSE4Good US	2.14%	16.61%	13.39%	7.37%	
DJSI U.S.	-1.17%	13.53%	10.60%	6.05%	
Range top/bottom performer	3.70%	3.08%	2.79%	1.44%	
S&P 500	1.38%	15.13%	12.57%	7.31%	
Russell 3000	0.48%	14.74%	12.18%	7.35%	

Based on daily returns for periods ending 12/31/15.

Sources: FactSet, Morningstar, MSCI, and TIAA

More importantly, statistical analysis showed no meaningful difference in returns when comparing RI indexes with relevant broad market indexes.<sup>5</sup> Any return variations appeared to be random and not systematic. For the analysis, performance was measured from the period when weekly returns first became available for each index. Track records ranged from 13 years for FTSE4Good US Index, to 21 years for MSCI KLD 400 Social Index. Time periods were long enough to ensure results were statistically valid.

### Volatility and risk-adjusted measures were also comparable:

Standard deviations for the RI indexes clustered fairly closely together and were similar to the S&P 500 and Russell 3000 indexes:

- Average annualized standard deviations for the RI indexes ranged from 15.69% to 17.13% over the past 10 years, compared to 16.53% and 17.05% for the S&P 500 and Russell 3000, respectively (Exhibit 3).
- The spreads between standard deviations for RI indexes and benchmarks averaged only 29 basis points for the 10-year period.
- Even though some standard deviations topped 40% during the 2008–2009 market collapse, the maximum spread between RI indexes and their benchmarks averaged only 1.78% for the 10-year period.

Meanwhile, *risk-adjusted returns* also showed little variation from broad market indexes. Sharpe ratios, or returns per unit of risk, also tracked fairly closely over various time periods, with average RI index Sharpe ratios mirroring the underlying market or lagging only slightly.

- For the 10-year period, RI index average annual Sharpe Ratios ranged between 0.78 and 0.89, compared with 0.87 and 0.88 for the Russell 3000 and S&P 500, respectively (Exhibit 3).

With standard deviations of returns and Sharpe ratios comparable between RI indexes and benchmarks, this suggests that incorporating ESG criteria in investment decisions doesn't require taking on additional risk relative to broad market benchmarks.

### Exhibit 3: Volatility measures and risk-adjusted returns were similar overall

Standard deviation and Sharpe ratios: RI indexes and broad market benchmarks

Index	Average Annualized Standard Deviation (%)	Spread*** vs. benchmark for 10-Yr period (%)		Average Annual Sharpe ratio
	10 Yr	Avg.	Max.	10 Yr
Calvert U.S. Large Cap Core Responsible*	17.13	0.08	1.61	0.87
DJSI U.S.**	15.69	0.84	2.87	0.78
FTSE4Good US**	16.77	0.24	2.50	0.89
MSCI KLD 400 Social**	16.27	0.26	1.13	0.86
MSCI USA IMI ESG*	17.01	0.04	0.78	0.84
RI Index average	16.57	0.29	1.78	0.85
S&P 500	16.53			0.88
Russell 3000	17.05			0.87

\* Benchmark: Russell 3000; \*\* Benchmark: S&P 500. Based on weekly returns for 10-year period through 12/31/15;

\*\*\* Spread equals index value minus benchmark value. Spread averages and maximums are absolute values.

Sources: FactSet, Morningstar, MSCI, and TIAA

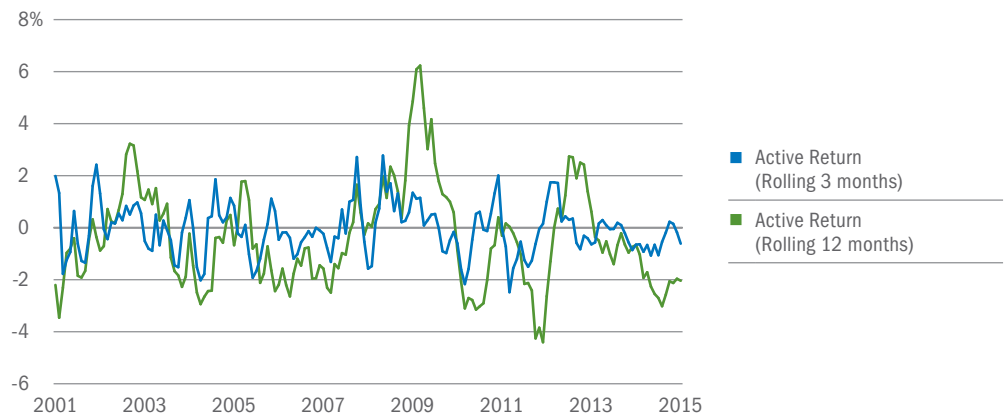
## RI indexes: Significant short-term performance variations

Although our results showed that index returns patterns were similar, they were not the same. In particular, performance variations increased significantly over short time periods, compared to broad market indexes. Short-term differences are to be expected because any strategy that does not replicate the index, such as the RI process, introduces portfolio biases causing performance to deviate from broad market indexes.

Exhibit 4 shows an example of how variable short-term performance can be. Measured on a 3- and 12-month basis, the rolling active return of the MSCI USA IMI ESG Index reveals considerable volatility, relative to an equivalent broad market benchmark, the Russell 3000 Index. This multicap RI index outperformed by as much as 6% and underperformed by more than 4% on a 12-month basis.

### Exhibit 4: RI indexes subject to greater short-term performance variations

Rolling active return – MSCI USA IMI ESG vs. Russell 3000 (2001-2015)



Rolling returns calculated for 3-month and 12-month periods for MSCI USA IMI ESG for period 12/29/2000 – 12/31/15 vs. Russell 3000 based on monthly returns.

Sources: FactSet, MSCI, and TIAA

## RI index tracking error rates varied measurably

All the RI indexes had performance that deviated from broad market indexes, as measured by their tracking error. An important question for investors is whether some RI indexes more closely matched the performance of broad market benchmarks than others. We performed statistical analysis to determine whether the tracking error rates were similar or different across the five RI indexes. Results showed that tracking error differences were statistically significant and, therefore, important for investors to consider.

We compared the MSCI USA IMI ESG Index and Calvert U.S. Large Cap Core Responsible Index to the Russell 3000 Index because these RI indexes include smaller-capitalization stocks. (MSCI actually benchmarks this RI index against its own MSCI USA IMI Index.) The remaining RI indexes were compared with the S&P 500 because they included primarily large-capitalization stocks. Among the indexes, the MSCI USA IMI ESG Index showed the lowest tracking error at 1.76%, and the DJSI U.S. had the highest at 3.43% (Exhibit 5). All tracking error rates were measured from the inception date for each index or first availability of weekly data through 12/31/15. Although time periods varied by index, all were sufficiently long to ensure statistical validity.

### Exhibit 5: Tracking error variations were significant across RI indexes

Tracking error rates since index inception through 12/31/15 (%)

RI Index	Min	Max	AVG	ST DEV	Benchmark
MSCI USA IMI ESG	0.79	3.96	1.76	0.61	Russell 3000
MSCI KLD 400 Social	1.23	5.51	2.59	1.00	S&P 500
Calvert U.S. Large Cap Core Responsible Index	1.24	5.46	2.71	1.07	Russell 3000
FTSE4Good US	1.68	6.02	2.89	0.98	S&P 500
DJSI U.S.	2.00	7.28	3.43	1.37	S&P 500

Weekly index total returns through 12/31/15. Beginning dates: MSCI KLD 400 Social, 11/11/1994; DJSI U.S., 1/8/1999; Calvert U.S. Large Cap Core Responsible Index, 5/5/2000; MSCI USA IMI ESG, 4/1/2001; and FTSE4Good US, 1/10/2003. MSCI indexes include aggregated, multisource histories prior to acquisition on 9/1/2010. Dates reflect first availability of weekly returns after index inception date.

Sources: FactSet, Morningstar, MSCI, and TIAA

Average tracking error for the MSCI USA IMI ESG Index was meaningfully lower than for Calvert U.S. Large Cap Core Responsible Index. For RI indexes benchmarked to the S&P 500, the MSCI KLD 400 Social tracking error was lower than DJSI U.S. and FTSE4Good US by a statistically significant margin. However, there appears to be no statistical difference between tracking errors for DJSI U.S. and FTSE4Good US.<sup>6</sup>

The consistency of returns versus a selected benchmark is an important consideration for investors in measuring performance and managing risk. Tracking error does not introduce absolute risk per se, but is a source of relative risk versus a benchmark. Low tracking error indicates the index's performance and risk characteristics closely match the benchmark's profile.

Investors considering RI strategies may be indifferent to the level of tracking error as long as long-term performance is comparable to the broad market. However, they should be aware of tracking error variations and their causes. Institutional investors, for example, may be constrained by client mandates to limit tracking error within specific ranges and against specific benchmarks.

## Index methodology drives short-term return variability

Variations in tracking error and short-term returns, relative to benchmark indexes, are by-products of the RI process. Some approaches for incorporating ESG criteria can eliminate or concentrate holdings in certain industries, resulting in portfolio characteristics that differ from the market.

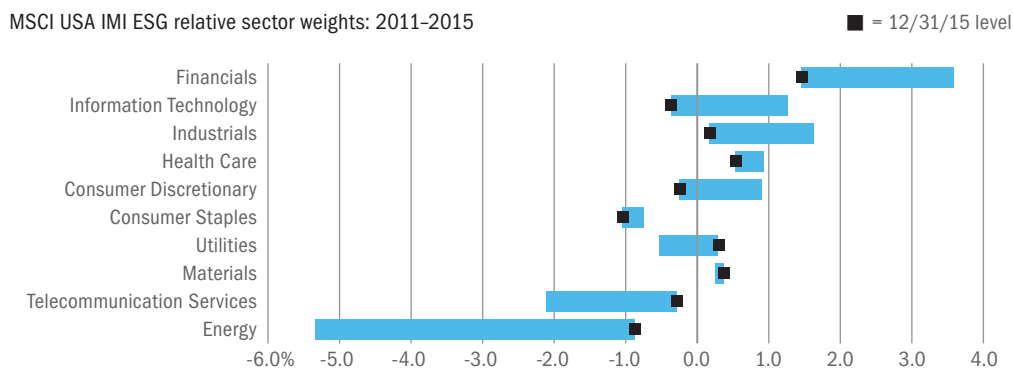
The five RI indexes use explicit ESG criteria to select a smaller subset of stocks from a universe of eligible companies. A particular strategy can involve excluding certain industries (such as gambling, tobacco and firearms), favoring companies that are leaders among their sector peers in managing relevant ESG risks and opportunities, or a combination of both. Decisions about how stocks are rated, selected and managed differentiate RI indexes from each other and the broad market.

The ESG evaluation and rating process itself can vary as indexes use different research approaches to select companies for inclusion in the index. Company assessments may differ depending on the ESG approach, the range of factors considered, and relative emphasis on the “E,” “S,” or “G” components. The potential impact on performance of different ESG research approaches was beyond the scope of this paper. However, an understanding of these differences may help investors select an RI index appropriate for their needs. (See Appendix 4 for detail on ESG rating and index methodologies.) Differences in index construction determine the biases and characteristics that drive tracking error and short-term return variability, including some of the following factors:

**Sector/industry weighting:** Certain industries or companies may tend to be excluded or have lower ESG ratings due to the nature of their business, such as gambling, tobacco or firearms. Conversely, some industries, like technology, may tend to receive higher ratings because they may face fewer ESG challenges than other industries. These variations can impact performance and alter the investment style versus the benchmark. Omitting mining and energy companies due to environmental concerns, for example, could potentially exclude some value-oriented companies and introduce a growth style bias to an index. Heavily capitalized and highly concentrated industries often include some of the largest companies, so including or excluding them could also skew the average size of companies in the index.

As an example, Exhibit 6 shows sector over/under weights for the MSCI USA IMI ESG Index compared to MSCI’s own broad market index over a five-year period and how these levels changed over time.

### Exhibit 6: Sector weight deviations can change over time



Sector weight ranges for MSCI USA IMI ESG relative to MSCI USA IMI index for period 12/31/2010 – 12/31/2015.

Sources: MSCI and TIAA

**Number of holdings:** In general, the more stocks in a portfolio, the lower the level of unsystematic risk. The fewer the names, the greater the odds individual companies or sector concentrations may drive returns, resulting in greater volatility. Choices about target market cap size can play a part in the number of names.

The MSCI USA IMI ESG Index broadly targets stocks of all capitalizations with higher ESG ratings, starting with a universe of over 2,400 securities. In contrast, the DJSI U.S. begins with the 600 largest-cap U.S. companies in the Dow Jones Sustainability North America Index and selects the most highly rated 20%. As a result, the MSCI index held over 1,000 mostly large- and mid-cap names as of 12/31/15, while the Dow Jones index held only 125 large-cap names. The Calvert U.S. Large Cap Core Responsible Index had 696 holdings, the MSCI KLD 400 Social Index held 400 issues, and the FTSE4Good US had 175 holdings.

Overall, we found that RI indexes with a larger number of stocks tended to have lower tracking error. The MSCI USA IMI ESG Index and the MSCI KLD 400 Social Index had the largest number of holdings among indexes tracking the Russell 3000 and S&P 500, respectively, and the lowest tracking error. The DJSI U.S. had the fewest names and the highest tracking error.

**% of stocks in the benchmark (by number and market cap):** We found the greater the percentage coverage of stocks in the benchmark, the lower the index's tracking error. The number of stocks in an index versus the benchmark is a very simple measure of coverage. In this case, holdings for the five RI indexes ranged from 23% to 80% of the number of stocks in their respective benchmarks (Russell 3000 or S&P 500). The two MSCI indexes had the largest coverage percentage and the lowest tracking error.

More relevant to assessing coverage and its likely impact on tracking error is the overlap of RI index holdings with the benchmark's actual constituents. Also important is how closely the market capitalization weighting of index holdings matches the benchmark constituents. However, examining coverage at that level of detail is beyond the scope of this paper (see Appendix 5 for a sample comparison of index characteristics: MSCI USA IMI Index vs. Russell 3000 Index).

**Efforts to address tracking error:** RI indexes have various procedures for adjusting position size and weights to help the index more closely and consistently track a broad market benchmark. These may include optimizing sector/industry weightings, limiting the size of individual holdings, periodic rebalancing, and using buffers and ranges to limit turnover when making constituent changes.

For example, MSCI ESG indexes specifically target sector weights to match MSCI's own underlying benchmarks. Calvert caps individual positions based on economic sector weights, but does not match a benchmark per se. The DJSI U.S. weights its holdings by market capitalization and relative industry weights, with individual constituents capped at 10%. The FTSE4Good US Index weights individual constituents based on their adjusted market cap, but does not adjust sector weights.



## Tracking error considerations for investors

Since long-term RI index performance is relatively indistinguishable from the broad market, tracking error may matter most to institutional investors subject to specific limits over shorter time periods. It's important to note, however, that RI indexes with lower tracking error are more likely to provide performance more consistent with a broad market benchmark.

Investors should carefully consider specific RI index construction, the ESG evaluation process, and the underlying market benchmark when selecting a strategy. Index methodology drives tracking error, resulting in meaningful differences among RI indexes.

To learn more,  
contact a  
relationship  
manager or visit  
[TIAA.org/  
assetmanagement](https://TIAA.org/assetmanagement)

## Conclusions

- RI indexes achieved long-term performance similar to broad market benchmarks while pursuing social goals.
- Incorporating ESG criteria did not result in higher risk levels, measured by Sharpe ratio and standard deviation. By constraining their investment universes, RI indexes introduce tracking error and greater short-term return variability, although the magnitude depended on how the index is constructed.
- RI indexes differed significantly in how closely they tracked broad market indexes. Understanding index methodology is critical to properly evaluating and selecting a specific index. In particular, investors should consider an index's breadth of holdings, market-cap and benchmark exposure, and steps taken to reduce tracking error and help improve consistency relative to its broad market benchmark.

## Endnotes

1. Total includes assets managed under ESG incorporation strategy alone or in combination with shareholder advocacy, but excludes assets only under shareholder advocacy strategy.
2. See Appendix 1.
3. See Appendix 2 for study methodology.
4. The Russell 3000 and S&P 500 indexes were selected as the most appropriate proxies against which most investors might measure RI performance. They have not necessarily been constructed for these comparisons and do not necessarily represent what would be an appropriate comparison as a parent index.
5. See Appendix 2.
6. Statistical comparison is based on the shorter time period common to both indexes, based on the 1/10/2003 inception of the FTSE4Good US Index. The average tracking error for the DJSI U.S. Index shown in Exhibit 5 is based on the index's 1/8/1999 inception date. See Appendix 3 for tracking error statistical analysis.

## Appendix 1

Many studies have analyzed the performance of RI-focused mutual funds versus their conventional counterparts and concluded the two are statistically similar. A May 2014 Empirical Research Partners publication reported there have been more than 60 separate academic studies on the subject and that 80% found no significant performance difference between RI funds and non-RI funds. Morningstar's 2015 review of 1,730 "socially conscious" funds found their star ratings for risk-adjusted performance "skew positively," supporting other research showing that performance generally meets or exceeds that of traditional funds. "The evidence is fairly clear that investors don't have to give up returns in exchange for having a sustainable portfolio," Morningstar concluded.

There has been much less research on the performance of RI *indexes*, the subject of our report. However, here too, research has shown the financial performance of RI indexes has been similar to conventional benchmarks.

Below is a sampling of academic papers and institutional research on the impact of ESG investing, including some that identify additional academic resources:

### General research

Morningstar Inc., October 2015, "Sustainable Investing Doesn't Have to Mean Sacrificing Returns."

Empirical Research Partners, May 2014, "Stock Selection: Research and Results, May 2014: Perspectives on Socially Responsible Investing"

Kidd D., May 2014, "Sustainable Investing: Reducing Risk to Create Alpha," *CFA Institute Investment Risk and Performance Feature Articles*, Vol. 2014, No. 1

Brière M., Peillex J., and Ureche-Rangauc L., January 2014, "Do Social Responsibility Screens Really Matter? A Comparison with Conventional Sources of Performance," *Social Science Research Network working papers series*

Gil-Bazo J., Ruiz-Verdu P, Portela A.A.P, 2010, "The performance of socially responsible mutual funds: the role of fees and management companies," *Journal of Business Ethics*, 94(2), 243–263

Cortez M.C., Silva F., Areal N., 2009, "The performance of European socially responsible funds," *Journal of Business Ethics*, 87(4), 573–588.

Hoepner A., McMillana D., August 2009, "Research on 'Responsible Investment': An Influential Literature Analysis comprising a rating, characterisation, categorisation & investigation," *Social Science Research Network working papers series*

Mercer white paper, Nov. 2009, "Shedding Light on Responsible Investment: Approaches, returns, and impacts."

### Index-related studies

Schröder M., 2007, "Is there a difference? The performance characteristics of RI equity indexes," *Journal of Business Finance & Accounting*, 34(1–2), 331–348.

Statman M., 2006, "Socially responsible indexes: composition, performance and tracking error," *Journal of Portfolio Management*, 32(3), 100–109

Sauer D.A., 1997, "The impact of social-responsibility screens on investment performance: evidence from the Domini 400 Social Index and Domini equity mutual fund," *Review of Financial Economics*, 6(2), 137–149.

## Appendix 2

**Survey methodology:** To measure the efficacy of RI strategies versus the broad market, we compared risk and return measures of RI indexes to appropriate benchmarks. We selected indexes with track records of at least 10 years. We focused on RI equity strategies because indexes with longer-term track records are readily available and represent the bulk of RI assets.

Our data came from FactSet Research Systems Inc., Morningstar, Inc., and MSCI Inc. and consisted of the daily, weekly and monthly total returns (gross of fees) for five widely followed U.S. RI indexes and the S&P 500 and Russell 3000 indexes. Returns were used to calculate mean performance, volatility and tracking error for periods when necessary data first became available following index inception.

**Performance analysis:** We conducted hypothesis tests to determine whether RI index returns were statistically similar to broad market benchmarks. The analysis showed that differences were more likely the result of random variation, rather than systematic causes. We calculated t-statistics to determine if there was a difference in mean returns, pairing an RI index with either S&P 500 or Russell 3000 index as appropriate. We assumed the two samples were independent, approximately normally distributed, and drawn from a population with the same underlying variance. We also calculated an F-distribution to test if return variances were the same. Analysis was for the period beginning when weekly performance data was first available for each RI index through 12/31/15. Exhibit 2A below shows the t- and F-test results: F-test allows us to conclude the sample pair variances were similar; t-test shows return pairs were statistically the same, indicating performance was comparable.

### Exhibit 2A: Hypothesis tests of returns for RI indexes and broad market benchmarks

#### F-Test Two-Sample for Variances

	MSCI		DJSI U.S.		FTSE		Russell		MSCI USA	Russell
	KLD 400	SP 500	DJSI U.S.	SP500	4Good US	SP500	Calvert	3000	IMI ESG	3000
Mean	0.21	0.20	0.11	0.12	0.18	0.18	0.10	0.12	0.15	0.16
Variance	6.00	5.85	6.35	6.33	5.91	5.67	6.96	6.43	6.54	6.36
Observations	1103.00	1103.00	886.00	886.00	677.00	677.00	817.00	817.00	769.00	769.00
Degrees of freedom	1102.00	1102.00	885.00	885.00	676.00	676.00	816.00	816.00	768.00	768.00
F stat	1.03		1.00		1.04		1.08		1.03	
P(F<=f) one-tail	0.33		0.48		0.29		0.13		0.34	
F Critical value one-tail	1.10		1.12		1.13		1.12		1.13	

If value of F stat is less than F Critical value, it can be assumed sample pairs have the same variances (i.e., are statistically similar).

#### t-Test: Two-Sample Assuming Equal Variances

	MSCI		DJSI U.S.		FTSE		Russell		MSCI USA	Russell
	KLD 400	SP 500	DJSI U.S.	SP500	4Good US	SP500	Calvert	3000	IMI ESG	3000
Mean	0.21	0.20	0.11	0.12	0.18	0.18	0.10	0.12	0.15	0.16
Variance	6.00	5.85	6.35	6.33	5.91	5.67	6.96	6.43	6.54	6.36
Observations	1103.00	1103.00	886.00	886.00	677.00	677.00	817.00	817.00	769.00	769.00
Pooled Variance	5.93		6.34		5.79		6.69		6.45	
Hypothesized mean diff.	0.00		0.00		0.00		0.00		0.00	
Degrees of freedom	2204.00		1770.00		1352.00		1632.00		1536.00	
t Stat	0.05		-0.10		-0.06		-0.13		-0.04	
P(T<=t) one-tail	0.48		0.46		0.48		0.45		0.48	
t Critical one-tail	1.65		1.65		1.65		1.65		1.65	
P(T<=t) two-tail	0.96		0.92		0.95		0.90		0.96	
t Critical value two-tail	1.96		1.96		1.96		1.96		1.96	

If value of t stat is less than t Critical value, it can be assumed sample pairs of returns are statistically the same.

Index weekly total returns through 12/31/15. Beginning dates: MSCI KLD 400 Social, 11/11/1994; DJSI U.S., 1/8/1999; Calvert U.S. Large Cap Core Responsible, 5/5/2000; MSCI USA IMI ESG, 4/1/2001; FTSE4Good U.S., 1/10/2003. Dates reflect first availability of weekly returns after index inception date.

Sources: FactSet, Morningstar, MSCI, and TIAA

### Appendix 3

**Tracking error analysis:** We performed a similar analysis of tracking error rates to determine if variances were random or statistically significant. We calculated t-statistics on pairs of RI indexes. We also calculated an F-distribution to test if variances for the tracking error rates were the same. Analysis was for the period beginning when weekly performance data was first available for both RI indexes in each pair tested through 12/31/15. Exhibit 3A below shows the t- and F-test results. F-test results allow us to conclude the sample pair variances were not similar; t-test results show that pairs of tracking error rates were statistically different and not the result of randomness.

#### Exhibit 3A: Hypothesis testing for RI index and broad market benchmark tracking error rates

##### F-Test Two-Sample for Variances

	Calvert	MSCI USA IMI ESG	FTSE 4Good US	DJSI U.S.	DJSI U.S.	MSCI KLD 400	FTSE 4Good US	MSCI KLD 400
Mean	2.57	1.76	2.89	2.79	3.43	2.53	2.89	2.16
Variance	0.91	0.36	0.96	0.41	1.87	0.86	0.96	0.30
Observations	718.00	718.00	626.00	626.00	835.00	835.00	626.00	626.00
Degrees of freedom	717.00	717.00	625.00	625.00	834.00	834.00	625.00	625.00
F stat	2.51		2.33		2.17		3.14	
P(F<=f) one-tail	0.00		0.00		0.00		0.00	
F Critical value one-tail	1.13		1.14		1.12		1.14	

If value of F stat is greater than F Critical value, it can be assumed the variances for the sample pairs are not statistically the same.

##### t-Test: Two-Sample Assuming Equal Variances

	Calvert	MSCI USA IMI ESG	FTSE 4Good US	DJSI U.S.	DJSI U.S.	MSCI KLD 400	FTSE 4Good US	MSCI KLD 400
Mean	2.57	1.76	2.89	2.79	3.43	2.53	2.89	2.16
Variance	0.91	0.36	0.96	0.41	1.87	0.86	0.96	0.30
Observations	718.00	718.00	626.00	626.00	835.00	835.00	626.00	626.00
Hypothesized mean diff.	0.00		0.00		0.00		0.00	
Degrees of freedom	1209.00		1079.00		1468.00		986.00	
t Stat	19.33		2.22		15.80		16.26	
P(T<=t) one-tail	0.00		0.01		0.00		0.00	
t Critical one-tail	1.65		1.65		1.65		1.65	
P(T<=t) two-tail	0.00		0.03		0.00		0.00	
t Critical value two-tail	1.96		1.96		1.96		1.96	

If value of t stat is greater than t Critical value, it can be assumed sample pairs of tracking error rates are not statistically the same or the result of randomness.

Weekly index total returns through 12/31/15. Beginning dates: DJSI U.S. vs. KLD 400 Social, 12/31/1999; Calvert vs. MSCI IMI ESG US, 3/29/2002; DJSI U.S. vs. FTSE4Good US, 1/2/2004; FTSE4Good US vs. MSCI KLD 400 Social, 1/2/2004. For each index pair, dates are 12 months after inception date of index with shortest track record. Initial TE calculations require a full year of returns data.

Sources: FactSet, Morningstar, MSCI, and TIAA

## Appendix 4

### Index methodology and ESG evaluation processes

The evaluation process used to assess ESG performance and identify securities for inclusion may be performed in-house or by a third party. It is important to note that there are differences in research approaches and methodologies for assessing company ESG performance. For example, sources of ESG data may come solely from publicly available, self-reported information, in-depth interviews with companies, proprietary research, or a combination of sources. The type and scope of ESG issues emphasized may vary. Companies may be assessed on the same set of criteria on an absolute basis, or on industry-specific issues and relative to peers. Familiarity with differences in ESG research approaches may be an additional dimension to help investors evaluate an RI index.

Below is a summary of index methodology and rating processes:

#### Exhibit 4A: RI index methodologies and rating processes

RI Index	Stock selection methodology and rating process
<b>Calvert U.S. Large Cap Core Responsible Index</b>	<ul style="list-style-type: none"> <li>Constituents from the S-Network 1000 index that meet Calvert's ESG criteria (The SN1000 includes the 1,000 largest capitalization stocks listed in the US and whose principal place of business is the US.) Sector weights mirror the SN1000.</li> <li>In-house ESG score selects approximately 700 stocks with the highest proprietary rankings.</li> </ul>
<b>Dow Jones Sustainability U.S. Index</b>	<ul style="list-style-type: none"> <li>Target percentage of companies ranked highest for sustainability (top 20% in each industry); constituents market-cap weighted; cap on individual positions</li> <li>Sustainability ranking generated through self-reporting by companies and third-party analysis.</li> </ul>
<b>FTSE4Good US Index</b>	<ul style="list-style-type: none"> <li>Companies from FTSE's broad market universe, meeting FTSE's ESG criteria; constituents market-cap weighted</li> <li>Excludes businesses involved with tobacco, weapons systems, and components for controversial weapons.</li> <li>Ratings based on publicly available data; process overseen by an independent committee.</li> </ul>
<b>MSCI USA IMI ESG</b>	<ul style="list-style-type: none"> <li>Companies at or above ESG rating threshold, constituents market-cap weighted; sector weight targets tied to benchmark to reduce tracking error</li> <li>Ratings from in-house research, looking at key ESG performance indicators and extensive data to create individual company ESG ratings</li> </ul>
<b>MSCI KLD 400 Social Index</b>	<ul style="list-style-type: none"> <li>Same starting MSCI USA IMI ESG universe, but limited to 400 constituents. (200 large-cap, 200 mid-cap); market-cap weighted</li> <li>Excludes companies involved with alcohol, gambling, tobacco, firearms, military weapons, nuclear energy, adult entertainment and genetically modified organisms.</li> </ul>

Source: Index providers

## Appendix 5

### Portfolio characteristics—a closer look

Despite efforts to reduce tracking error and more closely approximate a broad market profile, RI indexes look different than their benchmarks in other ways. A case in point is the MSCI USA IMI ESG index. Compared with the Russell 3000, this RI index had the lowest tracking error of the five we studied yet still had only 39% of the number of benchmark holdings, an average weighted market cap 66% of the benchmark, and P/E 10% lower than the broad market index.

**Exhibit 5A: Portfolio characteristics: MSCI USA IMI ESG vs. Russell 3000**  
(as of 12/31/15)

	MSCI USA IMI ESG	Russell 3000	% relative to Russell 3000
<b>Market Capitalization</b>			
Weighted Average	86,060.4	114,598.2	75%
Median	2,373.2	1,438.7	165%
Weighted Median	38,249.0	51,636.2	74%
<b># of Securities</b>	1,162	3,021	38%
<b>Dividend Yield</b>	2.04	2.01	101%
<b>Price/Earnings</b>			
Weighted Average	31.2	44.0	71%
<b>P/E using FY1 Est</b>			
Weighted Average	80.3	54.3	148%
<b>Price/Cash Flow</b>			
Weighted Average	21.1	19.4	109%
<b>Price/Book</b>			
Weighted Average	7.1	6.7	106%
<b>Price/Sales</b>			
Weighted Average	15.4	10.1	152%
<b>ROA</b>	7.8	7.2	108%
<b>ROE</b>	16.8	15.9	106%

Sources: FactSet and MSCI

## Appendix 6

### TIAA-CREF Social Choice Equity Fund – Institutional class Net of fees as of March 31, 2016\*

#### Average annual returns vs. Russell 3000 Index

Gross expense ratio: 0.18%/Net expense ratio: 0.18%

	3 Month	1 Year	3 Year	5 Year	10 Year	Since Inception
Social Choice Equity Fund (Institutional Class)	1.68%	-1.52%	9.44%	10.05%	6.50%	4.58%
Russell 3000 Index	0.97%	-0.34%	11.15%	11.01%	6.90%	4.87%

\* The Fund's inception date is July 01, 1999. Returns do not reflect the taxes that a shareholder would pay on fund distributions or on redemptions of fund shares. Investment return and principal value will fluctuate so that, when redeemed, shares may be worth more or less than the original cost.

The performance presented represents past performance and is not an indicator or guarantee of future results. Performance data for the Fund is presented net of management fees and expenses and includes the reinvestment of dividends through March 31, 2016. Current performance may be lower or higher than the performance presented herein. For performance current to the most recent month-end, please visit [TIAA.org](http://TIAA.org) or call 800-842-2252.

The expense ratio for this portfolio is 0.18% (gross)/0.18% (net). A contractual arrangement is in place that limits certain fees and/or expenses. Had fees/expenses not been limited ("capped"), currently or in the past, returns would have been lower. Expense Cap Expiration Date: February 28, 2017. Please see the prospectus for details.

The Russell 3000 Index measures the performance of the stocks of the 3,000 largest publicly traded U.S. companies, based on market capitalization. The index measures the performance of about 98% of the total market capitalization of the publicly traded U.S. equity market. You cannot invest directly in any index. Index returns do not reflect a deduction for fees or expenses.

To learn more,  
contact a  
relationship  
manager or visit  
[TIAA.org/  
assetmanagement](http://TIAA.org/assetmanagement)



Consider the investment objectives, risks, charges and expenses carefully before investing. Please call 877-518-9161 or go to [TIAA.org](http://TIAA.org) for product and fund prospectuses that contain this and other information. TIAA-CREF mutual funds are subject to market and other risk factors. Read the prospectuses carefully before investing.

Investment, insurance and annuity products are not FDIC insured, are not bank guaranteed, are not bank deposits, are not insured by any federal government agency, are not a condition to any banking service or activity, and may lose value.

Certain products and services may not be available to all entities or persons. Past performance does not guarantee future results.

Investments in Responsible Investment Funds are subject to the risk that because social criteria exclude securities of certain issuers for non-financial reasons, investors may forgo some market opportunities available to those that don't use these criteria. Investment products may be subject to market and other risk factors. See the applicable product literature, or visit [TIAA.org](http://TIAA.org) for details.

The material is for informational purposes only and should not be regarded as a recommendation or an offer to buy or sell any product or service to which this information may relate. Certain products and services may not be available to all entities or persons. Past performance does not guarantee future results.

TIAA Global Asset Management provides investment advice and portfolio management services through TIAA and over a dozen affiliated registered investment advisers.

TIAA-CREF Individual & Institutional Services, LLC, Teachers Personal Investors Services, Inc., and Nuveen Securities, LLC, Members FINRA and SIPC, distribute securities products.

©2016 Teachers Insurance and Annuity Association of America (TIAA), 730 Third Avenue, New York, NY 10017