International Journal of Business and Economics, 2019, Vol. 18, No. 2, 141-163

Are there Really Long-Run Diversification Benefits from Sustainable Investments?[#]

Nicholas Apergis
Department of Law and Social Sciences,
University of Derby, UK

Vassilios Babalos
Department of Accounting and Finance,
University of Peloponnese, Greece

Christina Christou
Department of Economics and Management,
Open University of Cyprus, Cyprus

Rangan Gupta[*]
Department of Economics,
University of Pretoria, South Africa

Abstract

Socially responsible investments have turned into popular investment vehicles over the last decade. By employing a standard cointegration methodology along with a novel time-varying quantile cointegration approach, this paper estimates whether the U.S. Dow Jones Sustainability Index (DJSI) and its conventional counterpart are integrated. The results confirm the presence of an asymmetric long-run relationship between the two indices that is not picked up by the standard methodology of cointegration, rendering the cointegrating relationship to be quantile-dependent. Similar results appear for the world and European sustainability indices relative to their conventional counterparts, implying the robustness of our approach. These findings place any long-run diversification benefits under scrutiny and contain significant implications for international market participants.

Key words: Socially responsible investments, quantile cointegration, diversification benefits

JEL Classification: C5, G1, Q5

Received June 11, 2018, revised October 14, 2018, accepted October 17, 2018.

[#] We would like to thank two anonymous referees for many helpful comments. However, any remaining errors are solely ours.

[*] Correspondence to: Address: Department of Economics, University of Pretoria, Pretoria, 0002, South Africa. Email: rangan.gupta@up.ac.za. Phone # +27 12 420 3460.

1. Introduction

In the wake of the 2008 global financial crisis, a need has been felt for exploring alternatives to conventional financial practices in order to reduce investment risks, increase returns, enhance financial stability, and reassure investors and financial markets. In this regard, academic research on socially responsible investing (SRI) has intensified. One reason for the increased interest in SRI is that it combines the pursuit of financial returns with non-financial considerations relating to the environment, social issues, and governance (ESG) and hence can be less risky compared to conventional alternatives. There are various reasons that have led to the global SRI (sustainable investment) market to grow steadily both in absolute and relative terms in all regions except Europe, because it tightened its definition of sustainable investment. According to the Global Sustainable Investment Review of 2016 released by Global Sustainable Investment Association (GSIA), there is now \$22.89 trillion in assets being professionally managed globally under responsible investment strategies, which is a 25% increase since 2014.

The literature has investigated SRI from the following aspects, primarily through the lens of mutual funds and through regional SRI indices in the U.S., Europe, and other major developed economies: (a) performance (i.e., risk-return characteristics relative to conventional indices) using mutual fund portfolios and indices (Luther *et al*., 1992; Hamilton *et al*., 1993; Luther and Matatko, 1994; Mallin *et al*., 1995; White, 1995; Kurtz and DiBartolomeo, 1996; Gregory *et al*., 1997; Russo and Fouts, 1997; Sauer, 1997; DiBartolomeo and Kurtz, 1999; Goldreyer and Diltz, 1999; Statman, 2000; Stone *et al*., 2001; Garz *et al*., 2002; Kreander *et al*., 2002, 2005; Orlitzky *et al*., 2003; Bauer *et al*., 2005; Shank *et al*., 2005; Bauer *et al*., 2007; Girard *et al*., 2007; Schröder, 2007; Galema *et al*., 2008; Renneboog *et al*., 2008a, b; Edmans, 2011; Leite and Cortez, 2015) and at the firm level (Derwall *et al*. 2005, 2011; Kempf and Osthoff, 2007; Hong and Kacperczyk, 2009; Statman and Glushkov, 2009; Edmans, 2011; Kim and Venkatachalam, 2011; Guenster, 2012; Borgers *et al*. 2013; Nofsinger and Varma 2014); (b) ratings (Angel and Rivoli, 1997; Guenster *et al*., 2011); (c) screenings

(Guerard, 1997); (d) predictability and determinants of returns and volatility (Lean and Nguyen, 2014; Antonakakis *et al*., 2016); and (e) co-movements within SRI indices across regions (Roca *et al*., 2010). From these studies, research on SRI has primarily focused on the risk-return characteristics of these securities in relation to conventional investments, with no clear-cut empirical evidence on whether SRI does yield higher returns after adjusting for risks. One missing area of research in this regard is whether these securities offer diversification opportunities for conventional investments.

Against this backdrop, the novelty of this study is to explore, for the first time, within the context of a time–varying cointegrating model the presence of a long-run relationship between the Dow Jones Sustainability Indices for the U.S., Europe, and the world and their conventional counterparts. The only related study to ours is by Balcilar *et al.* (2017). Our paper, however, analyzes whether there are short-run diversification opportunities based on a Markov-switching DCC-GARCH model.

It is widely known that most financial time series display non-linear dynamics and have non-elliptic distributions. In view of these properties, we implement the quantile cointegration methodology, proposed by Xiao (2009), that allows time-varying cointegrating parameters. In other words, this is the first paper to address the issue of whether there exist any long-run diversification benefits from SRI relative to its conventional counterparts, based on cointegration models.

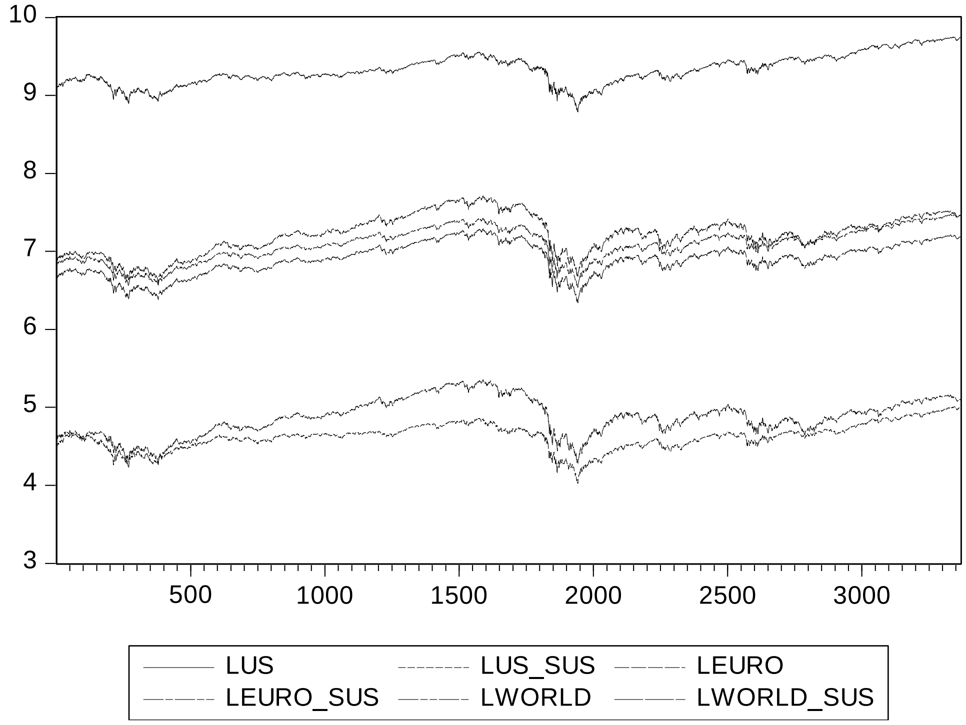
The remainder of the paper is organized as follows. Section 2 presents the data and the quantile cointegration methodology. Section 3 discusses the results. Section 4 provides a robustness analysis based on the European and world indices. Finally, Section 5 concludes.

2. Data Description and Methodology

2.1. Quantile Cointegration Analysis

Our time series span fourteen years, from September 28, 2001 to August 26, 2014 on a daily basis (3368 observations), allowing us to investigate whether the cointegrating vector remains constant over time. The data come from Datastream. Figure 1 plots the natural logarithmic values of two series.

Figure 1. Plot of Conventional (LUS, LEURO, LWORLD) and Sustainability Indices (LUS_SUS, LEURO_SUS, LWORLD_SUS) in Natural Logarithmic Values for the U.S., Europe, and Aggregate World.



Quantile methodological approaches are capable of capturing asymmetric/non-linear types of behavior, implying different responses at different points of the conditional distribution of the DJ Sustainability Index. Therefore, a novelty of this paper is that it investigates the whole conditional distribution of stock prices by estimating quantile cointegrating regressions for a sequence of quantiles. To this direction, the analysis implements the quantile cointegration methodology proposed by Xiao (2009), which allows us to explore the whole distribution of returns and allows for time-varying cointegration coefficients, which are a key issue in this analysis. The quantile cointegration model of Xiao (2009) captures systematic influences of conditioning variables on the location, scale, and shape of the conditional distribution of emissions.

We thus consider the following contegrating regression:

$$P_t = \alpha + \beta_t S_t + \varepsilon_t, \tag{1}$$

where S_t is the Dow Jones Sustainability Index, P_t is the Dow Jones Stock Price Index, while the cointegrating coefficient is allowed to be time-varying and thus quantile dependent.[2] Following Saikkonen (1991), Xiao (2009) suggests adding leads and lags of the dependent variables to deal with the endogeneity of the traditional cointegration model:

$$P_t = \alpha + \beta_t S_t + \sum_{i=-K}^{K} \pi_{it} \Delta S_t + \varepsilon_t. \tag{2}$$

In the above model the values of cointegrating coefficients are affected by the shocks received in each period and thus are quantile dependent. The τ^{th} quantile representation yields:

$$Q_{S_t}(\tau/\mathcal{J}_t) = \alpha(\tau) + \beta(\tau) S_t + \sum_{i=-K}^{K} \pi_i(\tau) \Delta S. \tag{3}$$

Estimation of the parameters in Eq. (3) results in:
$\Theta = \big(\alpha(\tau), \beta(\tau), \pi_{-K}(\tau), \dots, \pi_K(\tau)\big)$, which involves the solution to the problem:

$$\hat{\Theta}(\tau) = \arg\min_{\Theta} \sum_{t=1}^{\mathrm{T}} \rho_\tau \left(S_t - Q_{S_t}(\tau/\mathcal{J}_t)\right), \tag{3}$$

where $\rho_\tau(u) = u(\tau - I(< 0)$.

If we consider testing the null hypothesis $H_0 \colon \beta(\tau) = 1$, then we may construct the following Wald statistic:

$$W_T(\tau) = \frac{f_\varepsilon\left(\widehat{F_\varepsilon^{-1}(\tau)}\right)^2}{\hat{\omega}_\psi^{*2}} \left(\hat{\beta}(\tau) - 1\right)^2 \sum_t (S_t - \bar{S})^2, \tag{4}$$

where $\hat{\beta}(\tau)$ is the estimator of $\beta(\tau)$ given by Eq. (3), $f(\cdot)$ and $F(\cdot)$ are the respective p.d.f. and c.d.f. of $\{\varepsilon_t\}$, $f\left(\widehat{F_\varepsilon^{-1}(\tau)}\right)$ is a consistent non-parametric estimator of

$f_\varepsilon\big(F_\varepsilon^{-1}(\tau)\big)$ (Bofinger, 1975; Chamberlain, 1994), and $\widehat{\omega}_\psi^{*2}$ is a consistent estimator of the long-run variance of $\psi_\tau(\varepsilon_{t\tau}) = \tau - I(\varepsilon_{t\tau} < 0)$ with $\varepsilon_{t\tau} = \varepsilon_t - F_\varepsilon^{-1}(\tau)$.

Xiao (2009) shows that $W_T(\tau)$ asymptotically follows the chi-square distribution and also suggests a formal test for the constancy of the cointegrating coefficients. Specifically, he highlights that the varying-coefficient behavior can be tested using the Kolmogoroff-Smirnoff statistic $sup_\tau\big|\hat{V}_T(\tau)\big|$, where $\hat{V}_T(\tau) = T\big(\hat{\beta}(\tau) - \tilde{\beta}\big)$, $\hat{\beta}(\tau)$ is the quantile estimator from (3), and $\tilde{\beta}$ is a T-consistent estimator of β. The $sup_\tau\big|\hat{V}_T(\tau)\big|$ statistic has a non-standard asymptotic distribution, while critical values are calculated by bootstrap methodologies.

Following Xiao (2009), we specifically calculate the critical values by a five-step re-sampling procedure. (1) From Equation (2) we obtain the estimates $\hat{\beta}(\tau)$ and $\hat{\beta}$ through quantile regression and OLS regression, respectively, calculate the residuals $\hat{u}_t = P_t - \hat{\alpha} - \hat{\beta}S_t$, and then construct $\hat{V}_T(\tau) = T\big(\hat{\beta}(\tau) - \hat{\beta}\big)$. (2) We define $\widehat{w}_t = (\upsilon_t - \hat{u}_t)$, $\upsilon_t = \Delta S_t$, and get the fitted residuals $\hat{\varepsilon}_t$ from $\widehat{w}_t = \sum_{j=1}^{q} \hat{B}_j\, \widehat{w}_{t-j} + \hat{\varepsilon}_t$, $t = q+1, \dots, T$. (3) We draw i.i.d. variables $\{e_t^*\}_{t=q+1}^{T}$ from the centered residuals $\hat{\varepsilon}_t - \frac{1}{T-q}\sum_{j=q+1}^{T} \hat{\varepsilon}_t$ and generate w_t^* as follows:

$$w_t^* = \begin{cases} \widehat{w}_t = \displaystyle\sum_{j=1}^{q} \hat{B}_j\, \widehat{w}_{t-j} + e_t^*, & t = q+1, \dots, T \\ \qquad\qquad w_j^* = \widehat{w}_j, & j = 1, \dots\dots, q \end{cases}.$$

(4) We generate the bootstrap samples (y_t^*, x_t^*) as $S_t^* = S_{t-1}^* + \upsilon_t^*$ with $S_1^* = S_1$ and $P_t^* = \hat{\alpha} + \hat{\beta}S_t^* + u_t^*$, where $w_t^* = (\upsilon_t^*, u_t^*)$. (5) We use the bootstrapped samples (y_t^*, x_t^*) and the procedure described in step (1) to calculate the bootstrapped versions $\hat{\beta}^*(\tau)$, $\hat{\beta}^*$ and $\hat{V}_T^*(\tau) = T\big(\hat{\beta}^*(\tau), -\hat{\beta}^*\big)$ of $\hat{\beta}(\tau)$, $\hat{\beta}$ and $\hat{V}_T(\tau) = T\big(\hat{\beta}(\tau) - \hat{\beta}\big)$, respectively.

Furthermore, Xiao (2009) suggests the $sup_\tau\big|\hat{Y}_T(\tau)\big|$ statistic as a robust test for the quantile cointegration null, where $\hat{Y}_T(\tau) = \frac{1}{\sqrt{T\widehat{\omega}_\psi^{*2}}}\sum_{t=1}^{[T\tau]} \psi_\tau(\hat{\varepsilon}_{t\tau})$.

3. Empirical Analysis

Before resorting to the cointegration results, we take unit root tests to ensure that both the DJ Sustainability Index and the Dow Jones Stock Price Index follow an integrated process. These unit root tests are the Augmented Dickey-Fuller test (ADF, 1979), the Philips-Perron (PP, 1988) test, the Elliot-Rothemberg-Stock (ERS, 1996) test, and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS, 1992) test. To achieve good size and power properties, the lag length is selected through the MAIC proposed by Ng and Perron (2001).

Table 1 reports the results, which indicate that the unit-root null cannot be rejected by the ADF, the PP, and the ERS tests for both variables under investigation, while the null of stationarity in the case of the KPSS test is accepted for both variables at the 1% significance level. Realizing that our sample includes the 2008 global financial crisis, we also implement the Zivot and Andrews (ZA, 1992) unit root test with a structural break in both constant and the trend. The two stock price indices are found again to be unit root processes at conventional levels of significance, with the identified break being at and around the 2008 global financial crisis event. The results are also in Table 1.

Table 1. Unit Root Test Results

Test	US Sustainability Index		U.S. Dow Jones Stock Price Index	
	Level	First Differences	Level	First Differences
		With an Intercept		
ADF	-2.292	-66.197***	-1.849	-65.966***
PP	-1.294	-66.471***	-0.881	-66.065***
ERS	-1.396	-61.335***	-0.647	-2.556***
KPSS	1.756***	0.184	3.901***	0.146
		With an Intercept and a Trend		
ADF	-1.980	-66.209***	-2.216	-65.976***
PP	-2.015	-66.584***	-2.460	-66.086***
ERS	-1.485	-2.990	-1.600	-5.028***
KPSS	0.574***	0.056	0.580***	0.044
ZA	-5.012*		-4.900*	

Notes: *, **, and *** denote rejection of the null hypothesis at the 10%, 5%, and 1% levels, respectively. ADF, PP, ERS, KPSS, and ZA stand for Augmented Dickey-Fuller (1979), Phillips-Perron (1988), Elliot-Rothemberg-Stock (1996), Kwiatkowski-Phillips-Schmidt-Shin (1992), and Zivot-Andrews (1992) unit root tests. The null hypothesis of the ADF, PP, ERS, and ZA tests is “unit root”, whereas the null hypothesis of the KPSS test is “no unit root”.

We start off with the standard cointegration methodology of Engle and Granger (1987). The obtained *p*-value of the τ test is 0.61, given the test statistic of -1.83, while that of the *z*-statistic is -4.71 with a *p*-value of 0.76, indicating clear evidence against cointegration. This is possibly due to the presence of structural breaks and non-linearity in the relationship between the two stock indices, which in turn requires a time-varying approach.

The Brock *et al.* (1996, BDS) test applied to the residuals of equation (1) rejects the null of *i.i.d.* residuals across all dimensions at the 1% significance level. In addition, the Bai and Perron (2003) test of multiple structural breaks, applied to the constant and slope of equation (1), reveals the presence of five breaks. These two observations statistically explain the failure of the constant parameter cointegration test to detect a long-run relationship and hence motivate the need to look at a time-varying approach (i.e., quantile cointegration). We also conduct the Gregory and Hansen (1996) test of non-linear cointegration, with breaks in constant, constant and trend, and constant and slope. However, the null of no cointegration cannot be rejected even at 10% significance level. Complete details of these test outcomes are available upon request from the authors.

The results motivate us to look into quantile cointegration next. Table 2 reports the findings of quantile cointegration for a range of quantiles, including the estimated values of constants, cointegrating coefficients, and the Wald, $\sup_{\tau}\left|\widehat{V}_{T}(\tau)\right|$, and $\sup_{\tau}\left|\widehat{Y}_{T}(\tau)\right|$ tests.

The *p*-value for the constant terms helps us investigate the null of zero with student-*t* tests, while the counterpart for the Wald test looks for the null that the coefficient equals one. The $\sup_{\tau}\left|\widehat{Y}_{T}(\tau)\right|$ test provides an overall viewpoint of the long-run relationship between the two variables under investigation. The results provide supportive evidence that the two variables display a long-run equilibrium relationship across all selected quantiles, since the null hypothesis of quantile cointegration is not rejected. Moreover, the quantile-varying cointegrating coefficients are further confirmed strongly by the $\sup_{\tau}\left|\widehat{V}_{T}(\tau)\right|$ test, implying that the cointegration model with constant coefficients is subject to misspecifications. These findings validate Xiao (2009), whereby the presence of time-varying cointegrating coefficients is the major factor causing conventional cointegration methodologies to lack the ability to uncover any long-run relationship across variables as suggested by economic theory.

We next investigate the long-run relationship between the two variables in each specific quantile. The estimated values for intercepts and cointegrating coefficients differ across various quantiles. The results in Table 2 indicate the estimated values of the coefficients and their corresponding Wald tests, which determine the impact of the Sustainability Index on the DJ price index in each quantile. The findings highlight that the estimates are less than one across all quantiles and statistically significant, while the corresponding Wald tests generally reject the unit-coefficient null at the conventional 5% level, except for the quantile range of 0.30-0.55. The positive β coefficients imply across all quantiles that the inclusion of a Dow Jones Stock Price Index firm in a sustainability index provides a bonus to its stock prices, probably due to higher reputational gains associated with the reputation it shares as a reliable indicator for sustainability performance. Furthermore, the evidence through the Wald tests, i.e. the β coefficients are less than one across all quantiles, highlights that although the impact of the sustainability index participation is positive, the rewards are proportionately less, indicating that market participants do not get sufficient reward, since they already demonstrate exceptionally high financial performance. In other words, firm performance moderates the effects of status gains out of the participation in the Sustainability Index. Although investors often have uncertainty about how to assess such effects, the indicators of current and expected firm performance help them evaluate the value of status signals. These results contribute to the manner in which investors' perceptions are built, especially aligned with how Wall Street traders and market analysts form their expectations differently from that of the more general society (Lamin and Zaheer, 2012).

We also conduct the time-varying cointegration test of Bierens and Martins (2010). Not surprisingly, consistent with the quantile cointegration approach, we detect time-varying cointegration. These results are reported in Table A1 of the Appendix. However, we decide to focus on quantile cointegration as it allows us to study the entire conditional distribution and also the testing of a hypothesis is relatively easier with this approach.

Table 2. Quantiles' Cointegration Results

Quantile	Constant [p-value]	Beta [p-value]	Wald test [p-value] $H_0: \beta(\tau) = 1$
0.05	-4.0592	0.9221	19.3102
	[0.000]	[0.000]	[0.000]
0.10	-3.9619	0.9124	5.8726
	[0.000]	[0.000]	[0.015]
0.15	-3.6974	0.8853	4.7813
	[0.000]	[0.000]	[0.029]
0.20	-3.5641	0.8719	5.2968
	[0.000]	[0.000]	[0.021]
0.25	-3.4913	0.8648	4.4512
	[0.000]	[0.000]	[0.035]
0.30	-3.3277	0.8485	3.2799
	[0.000]	[0.000]	[0.070]
0.35	-3.2391	0.8399	2.4597
	[0.000]	[0.000]	[0.117]
0.40	-3.2172	0.8388	2.1366
	[0.000]	[0.000]	[0.144]
0.45	-3.1595	0.8341	1.9862
	[0.000]	[0.000]	[0.159]
0.50	-3.1173	0.8310	2.4177
	[0.000]	[0.000]	[0.120]
0.55	-3.0295	0.8231	3.4109
	[0.000]	[0.000]	[0.065]
0.60	-2.8854	0.8088	6.1370
	[0.000]	[0.000]	[0.013]
0.65	-2.7800	0.8000	10.7534
	[0.000]	[0.000]	[0.001]
0.70	-2.6621	0.7861	20.6476
	[0.000]	[0.000]	[0.000]
0.75	-2.4765	0.7672	33.5711
	[0.000]	[0.000]	[0.000]
0.80	-2.3571	0.7550	54.7840
	[0.000]	[0.000]	[0.000]
0.85	-2.2209	0.7411	97.1942
	[0.000]	[0.000]	[0.000]
0.90	-2.0847	0.7272	113.3236
	[0.000]	[0.000]	[0.000]
0.95	-1.7954	0.6977	105.2272
	[0.000]	[0.000]	[0.000]

$\boldsymbol{sup_\tau\left|\hat{V}_T(\tau)\right|} = 1377.7[p - value = 0.00]$

(CV1, CV5, CV10) = (919.992, 665.991, 557.318)

$sup_\tau\left|\hat{Y}_T(\tau)\right| = 0.682[p - value = 0.79]$

Notes: Figures in square brackets are p-values. Here, $\sup_\tau\left|\hat{V}_T(\tau)\right|$ is the bootstrapped-based statistic for testing the null of constant cointegrating coefficients. CV1, CV5, and CV10 are the bootstrapped critical values of statistical significance at the 1%, 5%, and 10% levels, respectively; $\sup_\tau\left|\hat{V}_T(\tau)\right|$ rejects the null of constant cointegrating coefficients when its value is greater than the critical value, and the number of repetitions in bootstrapping is 3000; $\sup_\tau\left|\hat{Y}_T(\tau)\right|$ tests the null of the existence of quantile cointegration. The quantile cointegration analysis is performed from Xiao (2009).

4. Robustness Analysis

For robustness test, we repeat the analysis above for the European market and the aggregate world market, with data again sourced from Datastream and with the period covering daily observations over September 28, 2001 to August 26, 2014. The natural logarithmic values of the four series are plotted in Figure 1. This also allows us to ensure whether our results are only specific to the U.S. economy or not.

We start off with the unit root tests in Tables 3A and 3B and find that the sustainability and the conventional indices of the European and world economy are I(1). Again, the ZA unit root test with a structural break also confirms that all the four series under consideration are in fact non-stationary. The results are reported in Tables 3A and 3B.

As with the U.S. case, we start off with the standard cointegration approach by Engle and Granger (1987). The obtained *p*-values of the τ test are 0.39 and 0.88 for the European and world economy cases, respectively, with test statistics of -2.26 and -1.09, while the corresponding *z*-test statistics are -8.08 (*p*-value=0.49) and -3.67 (*p*-value=0.83), respectively. All these results provide evidence against cointegration. As with the U.S., the Brock *et al.* (1996, BDS) test applied to the residuals of equation (1) for Europe and the world rejects the null of *i.i.d.* residuals across all dimensions at the 1% significance level. In addition, the Bai and Perron (2003) test of multiple structural breaks, applied to equation (1), reveals the presence of five breaks. Just like for the U.S. economy, we also conduct the Gregory and Hansen (1996) test of non-linear cointegration, however, the null of no cointegration is not rejected even at the 10% significance level. Complete details of these test results are available upon request from the authors.

We nextN look into the quantile cointegration results in Tables 4 and 5. The $\sup_{\tau}\left|\hat{Y}_T(\tau)\right|$test, which provides an overall viewpoint of the long-run relationship between the two variables under investigation, displays supportive evidence that the two variables share a long-run equilibrium relationship across all selected quantiles for both the European and world cases, since the null hypothesis of quantile cointegration is not rejected. Moreover, for both the European and world cases, the quantile-varying cointegrating coefficients are strongly confirmed by the $\sup_{\tau}|\hat{V}_T(\tau)|$ test, implying that the cointegration model with constant coefficients is subject to misspecifications, as under the U.S. case.

We now investigate the long-run relationship between the two variables in each specific quantile. The estimated values for intercepts and cointegrating coefficients differ across various quantiles as indicated in Tables 5 and 6. The findings highlight for the European case that the estimates of β are less than one only from the quantile 0.60 and above; below it, the estimate are greater than one and statistically significant. The corresponding Wald tests reject the unit-coefficient null at quantiles 0.45-0.50 and 0.80-0.95. For the world case, the estimates of β are greater than one from 0.50 and above with the unit-null bein rejected at quantiles 0.65 and above. Thus, while there are differences across the three cases (U.S., Europe, and the world) in terms of the estimated relationship, there is strong evidence that these two indices share a long-run time-varying or quantile-specific relationship irrespective of the geographic region under consideration.

Table 3. Unit Root Test Results

Panel A: European Case

Test	European Dow Jones Sustainability Index		European MSCI	
	Level	First Differences	Level	First Differences
		With an Intercept		
ADF	-1.876	-28.120***	-1.581	-28.278***
PP	-1.876	-59.362***	-1.809	-58.367***
ERS	-0.820	-2.412**	-0.357	-2.510**
KPSS	1.607***	0.072	2.129***	0.071
		With an Intercept and a Trend		
ADF	-2.022	-28.195***	-1.736	-66.209***
PP	-2.001	-59.353***	-1.958	-66.584***
ERS	-1.955	-4.796***	-1.649	-4.943***
KPSS	0.789***	0.071	0.835***	0.066
ZA	-4.792		-4.055	

Panel B: World Case

Test	World Dow Jones Sustainability Index		World MSCI	
	Level	First Differences	Level	First Differences
		With an Intercept		
ADF	-1.488	-27.576***	-1.218	-41.081***
PP	-1.640	-52.062***	-1.224	-51.706***
ERS	-0.403	-1.774**	-0.056	-6.428***
KPSS	2.139***	0.071	2.917***	0.085
		With an Intercept and a Trend		
ADF	-1.854	-27.573***	-1.837	-41.077***
PP	-1.991	-52.053***	-1.832	-51.699***
ERS	-1.861	-5.596***	-1.873	-50.098***
KPSS	0.622***	0.072	0.575***	0.080
ZA	-4.850*			-4.755

Notes: *, **, and *** denote rejection of the null hypothesis at the 10%, 5%, and 1% levels, respectively. ADF, PP, ERS, KPSS, and ZA stand for Augmented Dickey-Fuller (1979), Phillips-Perron (1988), Elliot-Rothemberg-Stock (1996), Kwiatkowski-Phillips-Schmidt-Shin (1992), and Zivot-Andrews (1992) unit root tests. The null hypothesis of ADF, PP, ERS, and ZA tests is "unit root", whereas the null hypothesis of the KPSS test is "no unit root".

Table 4. Quantiles' Cointegration Results – European Case

Quantile	Constant [p-value]	Beta [p-value]	Wald test [p-value] $H_0: \beta(\tau) = 1$
0.05	2.0297	1.0584	2.0937
	[0.000]	[0.000]	[0.148]
0.10	2.0542	1.0544	3.0818
	[0.000]	[0.000]	[0.079]
0.15	2.0721	1.0513	3.1452
	[0.000]	[0.000]	[0.076]
0.20	2.0831	1.0495	3.0118
	[0.000]	[0.000]	[0.0826]
0.25	2.0827	1.0503	3.1114
	[0.000]	[0.000]	[0.077]
0.30	2.0916	1.0491	2.9601
	[0.000]	[0.000]	[0.085]
0.35	2.0973	1.0483	2.9093
	[0.000]	[0.000]	[0.088]
0.40	2.0958	1.0492	3.1714
	[0.000]	[0.000]	[0.075]
0.45	2.0788	1.0533	4.1455
	[0.000]	[0.000]	[0.042]
0.50	2.0653	1.0574	5.6983
	[0.000]	[0.000]	[0.017]
0.55	2.3519	1.0026	0.0146
	[0.000]	[0.000]	[0.904]
0.60	2.3834	0.9968	0.0293
	[0.000]	[0.000]	[0.864]
0.65	2.3966	0.9945	0.1270
	[0.000]	[0.000]	[0.721]
0.70	2.4169	0.9908	0.556
	[0.000]	[0.000]	[0.455]
0.75	2.4415	0.9862	2.0626
	[0.000]	[0.000]	[0.151]
0.80	2.4682	0.9813	6.7534
	[0.000]	[0.000]	[0.009]
0.85	2.4986	0.9755	22.6575
	[0.000]	[0.000]	[0.000]
0.90	2.5161	0.9723	67.7754
	[0.000]	[0.000]	[0.000]
0.95	2.5374	0.9686	315.7557
	[0.000]	[0.000]	[0.000]

$\boldsymbol{sup_\tau \left|\hat{V}_T(\tau)\right|} = 209.9^{**}$

(CV1, CV5, CV10) = (267.650, 194.392, 165.060)

$sup_\tau \left|\hat{Y}_T(\tau)\right| = 0.462[p - value = 0.81]$

Notes: Figures in square brackets are p-values. Here, $\sup_\tau\left|\hat{V}_T(\tau)\right|$ is the bootstrapped-based statistic for testing the null of constant cointegrating coefficients. CV1, CV5, and CV10 are the bootstrapped critical values of statistical significance at 1%, 5%, and 10% levels, respectively. $\sup_\tau\left|\hat{V}_T(\tau)\right|$ rejects the null of constant cointegrating coefficients when its value is greater than the critical value, and the number of repetitions in bootstrapping is 3000; $\sup_\tau\left|\hat{Y}_T(\tau)\right|$ tests the null of the existence of quantile cointegration. The quantile cointegration analysis is performed from Xiao (2009).

Table 5. Quantiles' Cointegration Results – World Case

Quantile	Constant [p-value]	Beta [p-value]	Wald test [p-value] $H_0: \beta(\tau) = 1$
0.05	0.3547	0.9695	0.5286
	[0.000]	[0.000]	[0.467]
0.10	0.3250	0.9743	0.5068
	[0.000]	[0.000]	[0.476]
0.15	0.3199	0.9756	0.4765
	[0.000]	[0.000]	[0.490]
0.20	0.3077	0.9778	0.3982
	[0.000]	[0.000]	[0.528]
0.25	0.2890	0.9811	0.2930
	[0.000]	[0.000]	[0.588]
0.30	0.2844	0.9821	0.2644
	[0.000]	[0.000]	[0.607]
0.35	0.2763	0.9836	0.2275
	[0.000]	[0.000]	[0.633]
0.40	0.2543	0.9872	0.1484
	[0.000]	[0.000]	[0.700]
0.45	0.2159	0.9931	0.0474
	[0.000]	[0.000]	[0.827]
0.50	0.1571	1.0022	0.0057
	[0.000]	[0.000]	[0.939]
0.55	0.0675	1.0158	0.3834
	[0.000]	[0.000]	[0.536]
0.60	-0.0027	1.0265	1.4760
	[0.000]	[0.000]	[0.224]
0.65	-0.1385	1.0473	6.8008
	[0.000]	[0.000]	[0.009]
0.70	-0.3561	1.0809	29.8756
	[0.000]	[0.000]	[0.000]
0.75	-0.6512	1.1265	105.2164
	[0.000]	[0.000]	[0.000]
0.80	-0.6519	1.1275	169.2368
	[0.000]	[0.000]	[0.000]
0.85	-0.7026	1.1361	299.9622
	[0.000]	[0.000]	[0.000]
0.90	-0.6749	1.1328	463.8700
	[0.000]	[0.000]	[0.000]
0.95	-0.5371	1.1141	548.5422
	[0.000]	[0.000]	[0.000]

$\boldsymbol{sup_{\tau}|\hat{V}_T(\tau)|} = 368.8.7^{**}]$

(CV1, CV5, CV10) = (505.082, 363.389, 299.718)

$sup_{\tau}|\hat{Y}_T(\tau)| = 0.511[p - value = 0.74]$

Notes: Figures in square brackets are p-values. Here, $\sup_{\tau}|\hat{V}_T(\tau)|$ is the bootstrapped-based statistic for testing the null of constant cointegrating coefficients. CV1, CV5, and CV10 are the bootstrapped critical values of statistical significance at the 1%, 5%, and 10% levels, respectively. $\sup_{\tau}|\hat{V}_T(\tau)|$ rejects the null of constant cointegrating coefficients when its value is greater than the critical value, and the number of repetitions in bootstrapping is 3000; $\sup_{\tau}|\hat{Y}_T(\tau)|$ tests the null of the existence of quantile cointegration. The quantile cointegration analysis is performed from Xiao (2009).

5. Conclusion

This paper investigates the presence of a long-run asymmetric equilibrium relationship between the U.S. Dow Jones Sustainability Index and its conventional counterpart. By employing the time-varying cointegrating model of Xiao (2009), the empirical findings document a long-run relationship between the selected indices that varies across various quantiles of the returns' distribution, while standard cointegration methodologies provide evidence against cointegration. Our results also carry over for the world and European cases, indicating the robustness of our approach. The implications for investors are that any long-run diversification benefits are doubtful, but formal testing in the future is still needed using a portfolio allocation exercise, along with time-varying short-run analysis (for example, using the DCC-GARCH approach).

Endnotes

1. Based on the suggestion of any anonymous referee, we switch the dependent and independent variables in equation (1) and re-conduct the empirical analyses. Our results remain robust to such a reverse regression, as pointed out by Christou and Pittis (1999). Complete details of these results are available upon request from the authors.

Appendix:

Table A1. Bierens and Martins (2010) Time-varying Cointegration Test Results

	m=1	m=2	m=3	m=4	m=5	m=6
	11.98	15.60	18.97	30.41	30.53	40.12
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
	m=7	m=8	m=9	m=10	m=11	m=12
	46.74	**52.90**	**62.43**	**75.41**	**75.85**	**79.33**
	[0.00]	**[0.00]**	**[0.00]**	**[0.00]**	**[0.00]**	**[0.00]**

Notes: Figures in square brackets denote p-values, while m is the number of the Chebychev time polynomials. The null hypothesis tested is that the cointegration vector is time invariant.

References

Angel, J. J. and P. Rivoli, (1997), "Does ethical investing impose a cost upon the firm? A theoretical perspective," *The Journal of Investing*, 6(4), 57-61.

Antonakakis, N., V. Babalos, and C. Kyei, (2016), "Predictability of sustainable investments and the role of uncertainty: Evidence from a non-parametric causality-in-quantiles test," *Applied Economics*, 48(48), 4655-4665.

Bai, J. and P. Perron, (2003), "Computation and analysis of multiple structural change models," *Journal of Applied Econometrics*, 6, 72–78.

Balcilar, M., R. Demirer, and R. Gupta, (2017), "Do Sustainable Stocks Offer Diversification Benefits for Conventional Portfolios? An Empirical Analysis of Risk Spillovers and Dynamic Correlations," *Sustainatility*, 9, 1799.

Bauer, R., J. Derwall, and R. Otten, (2007), "The ethical mutual fund performance debate: New evidence from Canada," *Journal of Business Ethics*, 70(2), 111-124.

Bauer, R., K. Koedijk, and R. Otten, (2005), "International evidence on ethical mutual fund performance and investment style," *Journal of Banking & Finance*, 29(7), 1751-1767.

Bierens, H. J. and L. F. Martins, (2010), "Time Varying Cointegration," *Econometric Theory*, 26, 1453–1490.

Bofinger, E., (1975), "Estimation of a density function using order statistics," *Australian Journal of Statistics*, 17, 1–17.

Borgers, A., J. Derwall, K. Koedijk, and J. T. Horst, (2013), "Stakeholder relations and stock returns: On errors in investors' expectations and learning," *Journal of Empirical Finance*, 22, 159-175.

Brock, W., D. Dechert, J. Scheinkman, and B. LeBaron, (1996), "A test for independence based on the correlation dimension," *Econometric Reviews*, 15, 197–235.

Chamberlain, G., (1994), "Quantile regression, censoring, and the structure of wages," *In Advances in Econometrics*, Vol. 1, Sixth World Congress, ed. C. A. Sims, Cambridge: Cambridge University Press, 171–209.

Christou, C. and N. Pittis, (1999), "Forward versus Reverse Regression and Cointegration," *Economic Letters*, 65, 157-163.

Derwall, J., N. Guenster, R. Bauer, and K. Koedijk, (2005), "The eco-efficiency premium puzzle," *Financial Analysts Journal*, 61(2), 51-63.

Derwall, J., K. Koedijk, and J. T. Horst, (2011), "A Tale of Value-seeking versus Profit-driven Investors," *Journal of Banking and Finance*, 35(8), 2137-2147.

DiBartolomeo, D. and L. Kurtz, (1999), "Managing risk exposures of socially screened portfolios," *Northfield Information Services*, 1-17.

Dickey, D. A. and W. A. Fuller, (1979), "Distribution of the estimators for autoregressive time series with a unit root," *Journal of the American Statistical Association*, 74, 427–431.

Edmans, A., (2011), "Does the stock market fully value intangibles? Employee satisfaction and equity prices," *Journal of Financial Economics*, 101(3), 621-640.

Engle, R. F. and C. W. J. Granger, (1987), "Co-integration and error correction: Representation, estimation and testing," *Econometrica*, 55(2), 251–276.

Galema, R., A. Plantinga, and B. Scholtens, (2008), "The stocks at stake: Return and risk in socially responsible investment," *Journal of Banking & Finance*, 32(12), 2646–2654.

Garz, H., C. Volk, and M. Gilles, (2002), "More Gain than Pain, SRI: Sustainability Pays Off," WestLB Panmure. Available at http://www.westlbpanmure.com/sri/pdf/sri_nov2002.pdf.

Girard, E. C., H. Rahman, and B. A. Stone, (2007), "Socially responsible investments: Goody-two-shoes or bad to the bone?" *The Journal of Investing*, 16(1), 96–110.

Goldreyer, E. F. and D. J. Diltz, (1999), "The performance of socially responsible mutual funds: incorporating sociopolitical information in portfolio selection," *Managerial Finance*, 25(1), 23–36.

Gregory, A., J. Matatko, and R. Luther, (1997), "Ethical unit trust financial performance: small company effects and fund size effects," *Journal of Business Finance & Accounting*, 24(5), 705–725.

Gregory, A.W. and B. E. Hansen, (1996), "Residual-based tests for cointegration in models with regime shifts," *Journal of Econometrics*, 70, 99–126.

Guenster, N., (2012), "Performance implications of SR investing: past versus future," *Socially Responsible Finance and Investing: Financial Institutions, Corporations, Investors, and Activists*, 443–454.

Guenster, N., R. Bauer, J. Derwall, and K. Koedijk, (2011), "The Economic Value of Corporate Eco-Efficiency," *European Financial Management*, 17(4), 679–704.

Guerard, Jr., J. B., (1997), "Is there a cost to being socially responsible in investing?" *The Journal of Investing*, 6(2), 11–18.

Hamilton, S., H. Jo, and M. Statman, (1993), "Doing well while doing good? The investment performance of socially responsible mutual funds," *Financial Analysts Journal*, 49(6), 62–66.

Hong, H. and M. Kacperczyk, (2009), "The price of sin: The effects of social norms on markets," *Journal of Financial Economics*, 93(1), 15–36.

Kempf, A. and P. Osthoff, (2007), "The effect of socially responsible investing on portfolio performance," *European Financial Management*, 13(5), 908–922.

Kim, I. and M. Venkatachalam, (2011), "Are sin stocks paying the price for accounting sins?" *Journal of Accounting, Auditing & Finance*, 26(2), 415–442.

Kreander, N., R. H. Gray, D. M. Power, and C. D. Sinclair, (2005), "Evaluating the performance of ethical and non-ethical funds: a matched pair analysis," *Journal of Business Finance & Accounting*, 32(7-8), 1465–1493.

Kreander, N., R. Gray, D. Power, and C. D. Sinclair, (2002), "The Financial Performance of European Ethical Funds 1996-1998," *Journal of Accounting and Finance*, 1, 3–22.

Kuriyama, N., (2016), "Testing cointegration in quantile regressions with an application to the term structure of interest rates," *Studies in Nonlinear Dynamics & Econometrics*, 20(2), 107–121.

Kurtz, L. and D. DiBartolomeo, (1996), "Socially screened portfolios: an attribution analysis of relative performance," *The Journal of Investing*, 5(3), 35–41.

Kwiatkowski, D., P. C. B. Phillips, P. Schmidt, and Y. Shin, (1992), "Testing the null hypothesis of stationary against the alternative of a nit root," *Journal of Econometrics*, 54, 159–178.

Lamin, A., and S. Zaheer, (2012), "Wall Street vs. Main Street: Firm strategies for defending legitimacy and their impact on different stakeholders," *Organization Science*, 23, 47–66.

Lean, H. H., and D. K. Nguyen, (2014), "Policy uncertainty and performance characteristics of sustainable investments across regions around the global financial crisis," *Applied Financial Economics*, 24(21), 1367–1373.

Leite, P. and M. C. Cortez, (2015), "Performance of European socially responsible funds during market crises: Evidence from France," *International Review of Financial Analysis*, 40, 132–141.

Luther, R. G. and J. Matatko, (1994), "The performance of ethical unit trusts: choosing an appropriate benchmark," *The British Accounting Review*, 26(1), 77–89.

Luther, R. G., J. Matatko, and D. C. Corner, (1992), "The Investment Performance of UK "Ethical" Unit Trusts," *Accounting, Auditing & Accountability Journal*, 5(4), 57–70.

Mallin, C. A., B. Saadouni, and R. J. Briston, (1995), "The financial performance of ethical investment funds," *Journal of Business Finance & Accounting*, 22(4), 483–496.

Ng, S. and P. Perron, (2001), "Lag length selection and the construction of unit root tests with good size and power," *Econometrica*, 69, 1519–1554.

Nofsinger, J. and A. Varma, (2014), "Socially responsible funds and market crises," *Journal of Banking & Finance*, 48, 180–193.

Orlitzky, M., F. L. Schmidt, and S. L. Rynes, (2003), "Corporate social and financial performance: A meta-analysis," *Organization Studies*, 24(3), 403–441.

Phillips, P. C. B. and P. Perron, (1988), "Testing for a Unit Root in Time Series Regression," *Biometrika*, 75(2): 335–346.

Renneboog, L., J. T. Horst, and C. Zhang, (2008a), "Socially responsible investments: Institutional aspects, performance, and investor behaviour," *Journal of Banking & Finance*, 32(9), 1723–1742.

Renneboog, L., J. T. Horst, and C. Zhang, (2008b), "The price of ethics and stakeholder governance: The performance of socially responsible mutual funds," *Journal of Corporate Finance*, 14(3), 302–322.

Roca, E., V. S. Wong, and G. A. Tularam, (2010), "Are socially responsible investment markets worldwide integrated?" *Accounting Research Journal*, 23(3), 281–301.

Russo, M. V. and P. A. Fouts, (1997), "A resource-based perspective on corporate environmental performance and profitability," *Academy of Management Journal*, 40(3), 534-559.

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.

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

Shank, T., D. Manullang, and R. Hill, (2005), "Doing well while doing good revisited: a study of socially responsible firms' short-term versus long-term performance," *Managerial Finance*, 31(8), 33–46.

Statman, M., (2000), "Socially responsible mutual funds," *Financial Analysts Journal*, 56(3), 30–39.

Statman, M., and D. Glushkov, (2009), "The wages of social responsibility," *Financial Analysts Journal*, 65(4), 33–46.

Stone, B. K., J. B. Guerard, M. N. Gultekin, and G. Adams, (2001), "Socially responsible investment screening: strong evidence of no significant cost for actively managed portfolios," In *Social Investment Forum & Co-op America, www. socialinvest. org (June)*.

White, M. A., (1995), "The performance of environmental mutual funds in the United States and Germany: is there economic hope for green investors," *Research in Corporate Social Performance and Policy*, 1, 323–344.

Xiao, Z., (2009), "Quantile cointegrating regression," *Journal of Econometrics*, 150, 248–260.

Zivot, E. and K. Andrews, (1992), "Further Evidence on the Great Crash, the Oil Price Shock, and the Unit Root Hypothesis," *Journal of Business and Economic Statistics*, 10(10), 251–70.