"Tracking error minimization under varying sustainability criterion stringency: environmental ratings and U.S. stock portfolios"

AUTHORS	Rickard Olsson
ARTICLE INFO	Rickard Olsson (2010). Tracking error minimization under varying sustainability criterion stringency: environmental ratings and U.S. stock portfolios. *Insurance Markets and Companies*, *1*(3)
RELEASED ON	Wednesday, 29 December 2010
JOURNAL	"Insurance Markets and Companies"
FOUNDER	LLC "Consulting Publishing Company "Business Perspectives"

Rickard Olsson (Sweden)

Tracking error minimization under varying sustainability criterion stringency: environmental ratings and U.S. stock portfolios

Abstract

The study provides empirical evidence on how minimum tracking error varies, as the stringency of a sustainability criterion is varied. The sustainability criterion is based on environmental (EV) ratings for a universe of large capitalization U.S. firms. Increasingly sustainable portfolios are created from increasingly smaller subsets each containing stocks with increasingly higher EV ratings. Minimized tracking error standard deviation increases with sustainability stringency and varies from 0.4% per year for a portfolio, created from the 400 stocks with the highest EV ratings to 4.6% per year for a portfolio, created from the 20 stocks with the highest EV ratings. These sustainable portfolios' tracking errors appear to be equal or lower than those of existing sustainable funds from similar universes.

Keywords: tracking error optimization, sustainable investments, stocks.

Introduction

Tracking error measures how much a fund's performance deviates from that of its benchmark (index). A tracking error constraint is imposed on a fund, in order to limit the performance difference between the fund and its benchmark, and constitutes, along with the benchmark and the investment universe, a central part of the investment guidelines in most fund manager mandates. The focus on tracking error constraints is reported as a major impediment to a more widespread application of sustainable (also called socially responsible) investment strategies (Kurtz, 2005; Statman, 2006). Such investments are popular, for, as of 2007, 2.7 trillion USD or 11% of all assets under management in the U.S. were invested in sustainable portfolios. Institutional investors, including insurance companies, represent the largest and fastest growing segment of the sustainable investments category[1]. Tracking error constraints are problematic for sustainable investment strategies because such strategies often produce tracking errors that are large compared to those of typical institutional portfolios (e.g., Kurtz, 2005; Dimtcheva et al., 2002a; Jennings and Martin, 2007).

For a sustainability mandate, the investment guidelines also include the sustainability criteria, the fund must satisfy; the nature of the criteria may be quite variable, ranging from exclusion/inclusion of securities to sustainability score-based tilts. Since there is limited empirical evidence on what portfolio performance different sets of sustainable investment policy guidelines would allow, it is difficult for owners and trustees to set sensible investment guidelines for sustainable mandates.

The importance of the relation between minimum tracking error and the stringency of sustainability criteria is recognized in the literature (Kurtz, 2005). The empirical evidence on this relation is, however, limited, as the following brief review of recent studies indicate.

There are studies which examine the tracking error of sustainable indexes relative to ordinary indexes (e.g., Statman, 2006). Barnett and Salomon (2006) and Derwall et al. (2007) analyze the relation between mutual fund performance and the type and stringency of sustainability criteria.

Evidence on a single level of stringency and the corresponding minimum tracking error is given by studies that consider how (the set of stocks in) a sustainability index or a sustainable investable set could be reweighted to minimize tracking error relative to a benchmark index (diBartolomeo and Kurtz, 1999; Troutman, 2001). In other studies, the level of stringency is varied, and the associated (forecasted) minimum tracking error is observed (Dimtcheva et al., 2002a; Milevsky et al., 2006; Jennings and Martin, 2007). Except diBartelomeo and Kurz (1999), all said studies consider a single point in time and, thus, a forecasted or hypothetical tracking error estimate.

Several of the above studies are based on proprietary software and risk models (diBartolomeo and Kurtz, 1999; Troutman, 2001; Dimtcheva et al., 2002; Jennings and Martin, 2007). This could complicate the replication of results. Other studies, e.g., Barnett and Salomon (2006) and Derwall et al. (2007), are based on mutual funds and there is the risk that the observed relation is affected by confounding factors such as fund expenses and manager skills. Finally, there are studies that are case-based and their findings can be difficult to generalize (e.g., Milevsky et al., 2006; Troutman, 2001; Jennings and Martin, 2007).

The purpose of this study is to provide empirical evidence on how minimum tracking error varies as the sustainability criterion stringency is varied. The sustainability criterion is based on environmental (EV) ratings of U.S. firms. The ratings are produced by GES

[1] Source: www.socialinvest.org/resources/sriguide/srifacts.cfm.

Investments Services (GES), an analysis house specialized in responsible investments, founded in 1992 (www.ges-invest.com). The investment universe, used in this study, includes all U.S. stocks with GES EV ratings. A benchmark index for the universe, GIndex, is calculated and its performance is observed to be close to the performance of the S&P500 index (see Figure 2 and Table 3 below). Thus, the results of the study should be of interest to the many fund managers that have the S&P500 as benchmark/universe. Increasingly sustainable portfolios are created from increasingly smaller subsets of the universe, with each subset containing stocks with increasingly higher EV ratings. On a daily basis, the tracking error of each sustainable portfolio is minimized, relative to the benchmark. Various aspects of the performance of these portfolios are reported and analyzed.

The tracking error standard deviation of the sustainable portfolios increases with sustainability and is between 0.39%-4.63% per year, which is lower than the average of 5.9% for a set of existing sustainable funds with similar universes, as reported in Jennings and Martin (2007).

1. Data

Stock and S&P500 index data are collected from Thomson Datastream.

GES assigns each stock a specific and a general EV risk rating. Every six months, beginning in December 2003 and ending in June 2006, GES issued ratings for approximately the 1000 largest firms in the MSCI World Index. A firm's general EV risk rating (0 to 3, where 0 is the highest risk) is intended to reflect the EV risk of the firm's industry. The specific EV risk rating (0 to 3, where 0 is the highest risk) indicates the particular EV risk of a given firm. The specific EV risk rating is derived through analysis of the company along more than 60 dimensions, based on international standards for environmental management and industry-specific key indicators for environmental performance, among other things. Information sources, used in the analysis process include official company documents, dialogue with companies, non-governmental organizations, the media and GES partners in the SiRi Group (www.ges-invest.com and private communication with GES representatives).

The specific and general EV risk ratings are combined and converted to a summary EV score (*SumScEV*), where a higher score indicates less EV risk. Based on private communications with representatives of GES, *SumScEV* is calculated, using the theorem of Pythagoras, as the square root of the sum of the squared specific and the squared general EV measures. Descriptive statistics for the *SumScEV* per the dates ratings are published in Table 1.

Table 1. Descriptive statistics for the *SumScEV*

	SumScEV
Mean	2.14
Std	0.77
Number of observations	2636
Min	0
First quartile	1.53
Median	2.36
Third quartile	2.87
Max	3.79

2. Method

The number of firms in the investment universe of the study varies between 406 and 460 (Figure 1).

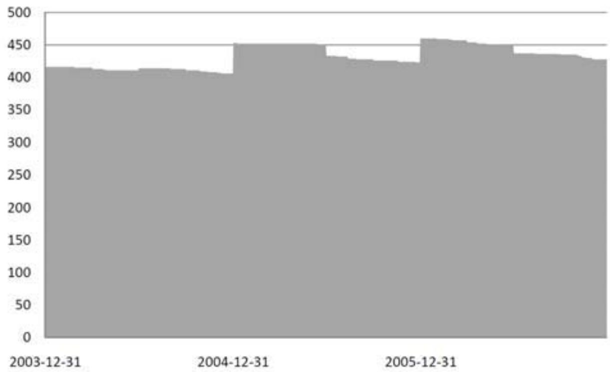

Fig. 1. Number of U.S. stocks with GES ratings

This universe is probably quite close to the S&P500 universe, which contains 500 of the largest U.S. listed companies. The reason is that universe contains 406 to 460 U.S. stocks and these are members of the MSCI World Index, which contains the largest stocks in terms of market capitalization in each country. Given this, it might seem obvious to use the S&P500 as the benchmark index in the study. Instead, based on the universe of U.S. stocks with GES ratings, an index is calculated following largely the capitalization-weighted construction methodology of the S&P500 index. Let this index be called GIndex. Calculating the GIndex serves there are two purposes. First, it ascertains that the sustainable portfolios are strict subsets of the universe and the benchmark. Second, exact data on index composition are important for tracking error minimization (Olsson, 2005). Such data are not freely available for the S&P500 on a daily basis. Calculating the GIndex gives exact daily data on the benchmark composition.

Five increasingly sustainable portfolios are created from the 400, 300, 200, 100 and 20 highest ranked stocks in terms of the most current *SumScEV* ratings at any given day. For each portfolio, tracking error relative to the benchmark is minimized. Optimizations are done daily from the end of 2003 up to the end of 2006.

Tracking error and tracking error constraints are often expressed in terms of tracking error standard deviation (*TESD*), which is a measure of the dispersion in tracking error. The greater a fund's *TESD* is, the greater the likelihood of large tracking errors, that is large return differences between the fund and its benchmark index. Below a quadratic program, equation (1), is formulated for minimizing *TESD* for each of the five increasingly sustainable funds.

Let x be the $n \times 1$ column vector of portfolio weights that minimizes *TESD*. m_i indicates whether a stock is eligible for inclusion based on its EV rating, that is $m_i = 1$ for stocks in the sustainable subset and $m_i = 0$ otherwise. b is the $n \times 1$ column vector of benchmark weights (for the GIndex effective the next day). R is the $T \times n$ matrix of historical returns over T prior periods for the n stocks; in the empirical study, $T = 750$ days (approx. 3 years) is used. $V = \text{Var}(R)$ represents the sample variance-covariance matrix of R. The portfolio's tracking error standard variance is minimized relative to the benchmark index under the constraint that weights sum to 1 and are non-negative, according to equation (1):

$$\min(x-b)'V(x-b), \qquad (1)$$

s.t. $1'x = 1$

$0 \le x_i \le m_i, \qquad i = 1, 2, ..., n.$

The solution to equation (1) gives the portfolio, defined by x with the minimum tracking error variance, based on historical returns R and the benchmark weights b effective the next day. *TESD* is the square root of the objective in the above program. The non-negativity constraint on the weights is intended to reflect the common restriction on institutional portfolios against short positions in stocks.

The performances of the implemented portfolios are evaluated, using several measures. Let r_{pt} and r_{bt} denote the realized returns of the tracking portfolio and the benchmark index, respectively, in day t. The tracking error for day t is then $TE_t = r_{pt} - r_{bt}$. An important gauge of tracking performance is the realized tracking error standard deviation $TESD = \sqrt{Var(TE)}$, where Var(*TE*) is the variance of TE_t from the beginning of 2004 to the end of 2006, a total of 754 days. Other measures calculated are the arithmetic time-average, minimum and maximum daily tracking error, that is *AvgTE*, *MinTE*, and *MaxTE*. Three performance measures, based on the ratio of the (net asset) value of the fund and the index level (scaled appropriately) are computed. *EndNAV* equals the ratio at the end of 2006, and provides a measure of the total performance of the fund relative to the index. *MinNAV* and *MaxNAV* are the minimum and maximum of the ratio, during the test period.

The so-called market model can be used to estimate the risk-adjusted performance of a portfolio relative to an index (see Frino and Gallagher, 2002). The market model is a linear regression of portfolio returns on index returns, $r_{pt} = \alpha_p + \beta_p r_{bt} + \varepsilon_{pt}$, where α_p and β_p are the parameters, and ε_{pt} ε_{pt} an error term. In this study, the parameters are estimated by the ordinary least squares method. For a portfolio p that perfectly tracks its benchmark index b, one would expect neutral performance relative to the index, that is $\alpha_p = 0$, and unit exposure to the returns of the index, that is $\beta_p = 1$ (Frino and Gallagher, 2002).

To give an indication of the level of diversification of the implemented portfolios, the average, minimum and maximum number of stocks in the portfolios are presented. The level of sustainability of the funds is described by the daily capitalization-weighted average *SumScEV* of the holdings of a given fund.

3. Results

According to Figure 2 and Table 2, column one, GIndex tracks S&P500 well, despite being based on fewer and not necessarily the same stocks as the S&P500 is. The annualized *TESD* is 0.53%. Relative to the S&P500 index, α is zero and β is close to one, 0.99, which is near what a perfect tracking portfolio would have, namely $\alpha = 0$ and $\beta = 1$. *EndNav* is one, which means that the value of the GIndex equals the value of the S&P500 index at the end of 2006. The similarity of the GIndex and the S&P500 index suggests that the findings of the study should be relevant to the many fund managers, which have the S&P500 as benchmark/universe.

The value-weighted *SumScEV* increases as intended with the sustainability of the funds (Figure 3 and Table 2).

Figure 4 portrays the behavior of funds and GIndex over time. The fund, restricted to the 100 stocks, having the highest *SumScEV*, gave the highest return of around 35% for the whole period.

According to Table 2, the tracking errors, *TESD*, of the sustainable funds relative to the GIndex are 0.39%-4.63% and increase with the level of sustainability. Tracking performance vs. S&P500 is very similar to that vs. GIndex, albeit larger. The levels of *TESD* relative to the S&P500 index for all the portfolios are comparable or lower than those of sustainable funds with similar universes, for which Jennings and Martin (2007) report an average *TESD* of 5.9% vs. S&P500.

Fig. 2. GIndex vs. S&P500

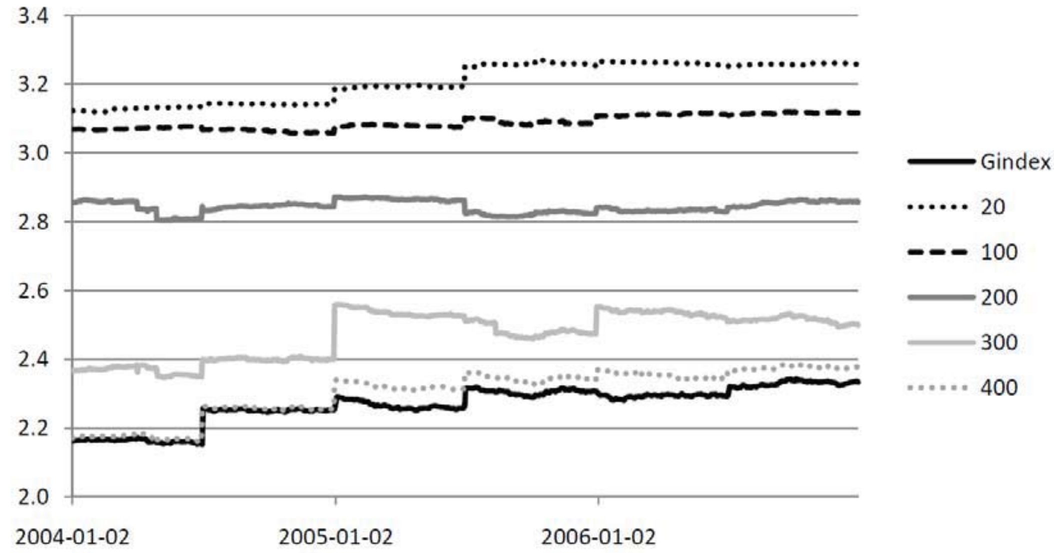

Fig. 3. *SumScEV* over time

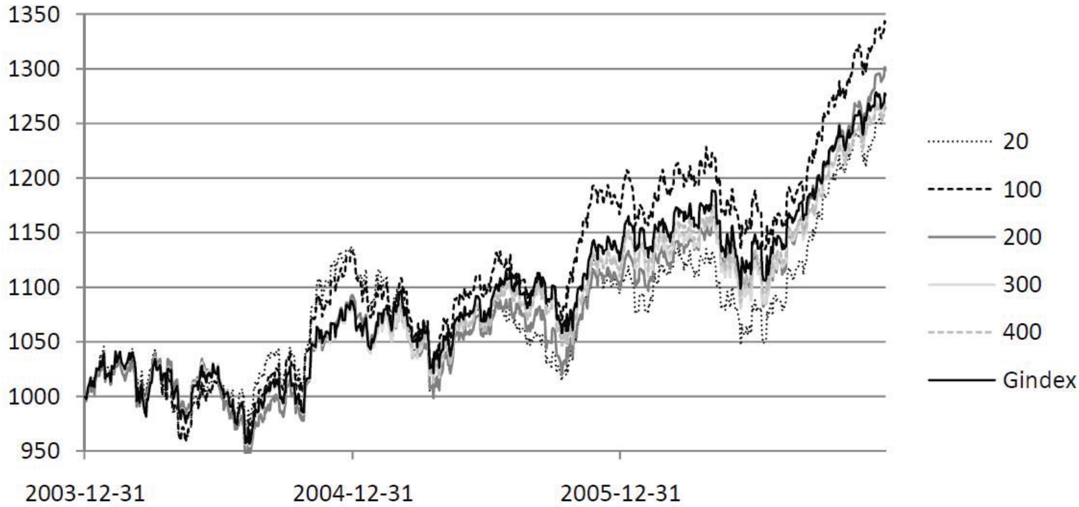

Fig. 4. Index and fund levels over time

Table 2. Results

Index	S&P500	G	G	G	G	G
Fund	G	20	100	200	300	400
SumScEV	2.30	3.21	3.09	2.84	2.47	2.30
AvgTE (%)	-0.001	-0.001	0.007	0.002	-0.001	-0.001
MinTE (%)	-0.29	-1.06	-0.92	-0.55	-0.27	-0.12
MaxTE (%)	0.19	0.89	1.02	0.54	0.30	0.12
TESD (%)	0.53	4.63	3.75	2.80	1.21	0.39
TESD vs S&P500 (%)		4.69	3.87	2.93	1.33	0.63
MinNAV	0.99	0.94	0.98	0.96	0.98	0.99
MaxNAV	1.00	1.05	1.05	1.02	1.00	1.00
EndNAV	1.00	0.99	1.05	1.02	0.99	0.99
α (%)	0.00	0.00	0.00	0.00	0.00	0.00
β	0.99	0.95	0.97	0.96	0.99	1.00
MinN	406	18	55	109	180	310
AvgN	431.9	19.6	64.6	121.1	216.5	356.4
MaxN	460	20	80	136	251	400

Notes: *SumScEV* – average of the average (value-weighted) daily *SumScEV* of the stocks in a given fund; *AvgTE* (%) – average daily tracking error (difference between fund and index return); *MinTE* (%) – minimum daily tracking error; *MaxTE* (%) – maximum daily tracking error; *TESD* (%) – annualized daily tracking error standard deviation (using 252 days); *MinNAV* – minimum ratio of fund net asset value and index level; *MaxNAV* – maximum ratio of fund net asset value and index level; *EndNAV* – ending ratio of fund net asset value and index level; α (%) – intercept of OLS regression of fund returns on index returns; β – slope of OLS regression of fund returns on index returns; *MinN* – minimum number of fund holdings; *AvgN* – average number of fund holdings; *MaxN* – maximum number of fund holdings.

Figure 5 depicts the behavior of the implemented funds in terms of their value as a fraction of the value of the index. The higher the *TESD* of given fund is, the higher the probability of outperforming or underperforming the index should be. This relation is reflected quite well in Figure 5 and by *MinNAV* and *MaxNAV* in Table 2. As well as by *MinTE* and *MaxTE*, which each varies monotonically with *TESD*, as one would expect, except *MaxTE* for the fund with 100 stocks which deviates.

Moreover, *EndNAV*, as a measure of total performance, is never below 0.99 for any of the sustainable portfolios. This means that none of the sustainable funds were beaten by the index by more than 1% over the studied period.

α and β are close to zero and one, respectively, for the majority of the sustainable funds. Minimization of *TESD*, thus, appears to lead to funds, whose risk-adjusted performance in terms of α is zero, and whose exposure β to the index is one.

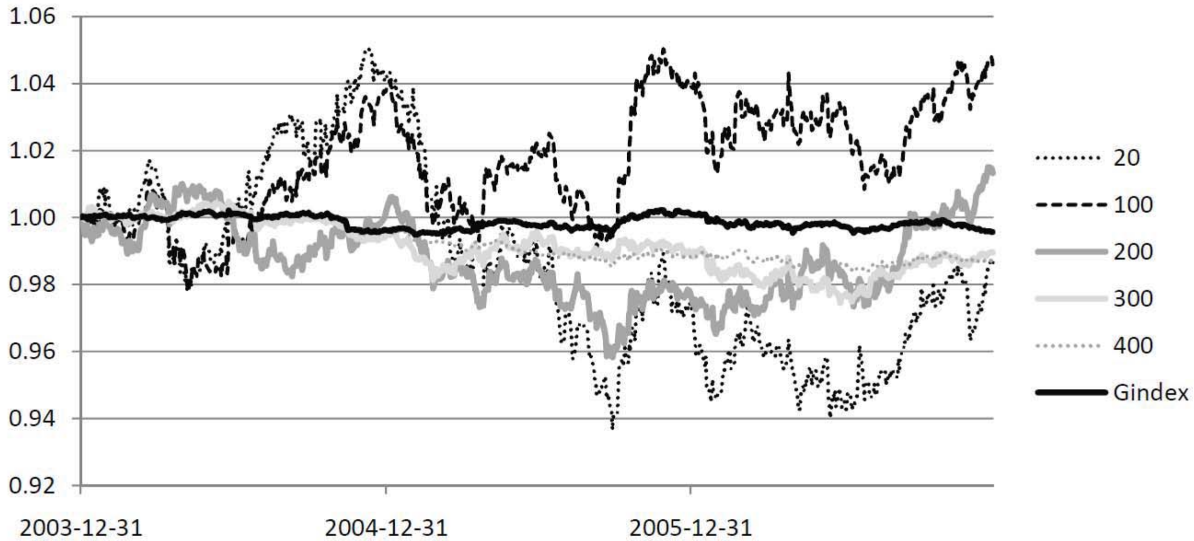

Fig. 5. GIndex and fund levels relative to S&P500

The funds seem to be well diversified in the sense that they contain relatively many stocks compared to the maximum number allowed (Table 2).

Tentative results indicate high turnover and many updates in some of the sustainable portfolios, particularly in the funds with the fewest number of holdings allowed. High turnover could produce high transaction costs which affect performance adversely. Future research will attempt to measure and control transaction costs for the kind of investment strategies analyzed here.

Conclusions

The study provides empirical evidence on how minimum tracking error standard deviation varies, as the stringency of a sustainability criterion is varied. The sustainability criterion is based on environmental (EV) ratings for a universe of U.S. firms. A benchmark index for the universe is calculated and its performance is observed to be close to the performance of the S&P500 index. The results of the study should, thus, be of interest to the many fund managers that have the S&P500 as benchmark/universe. Five increasingly sustainable portfolios are created from the 400, 300, 200, 100 and 20 highest ranked stocks in terms of the stocks EV ratings. The tracking error standard deviation of each sustainable portfolio is minimized relative to the benchmark index every day 2004-2006. Annualized tracking error standard deviation increases with sustainability stringency and varies from 0.4% for the portfolio, created from the 400 stocks with the highest EV ratings to 4.6% for a portfolio, created from the 20 stocks with the highest EV ratings. These sustainable portfolios' tracking errors appear to be equal or lower than those of existing sustainable funds from similar universes (Jennings and Martin, 2007).

Overall, the results suggest that tracking error minimization can produce EV sustainable portfolios, whose levels of tracking error in many cases could be accepted by typical institutional investors. Increased knowledge about the relation between minimal tracking error and sustainability should make it easier to set sensible investment guidelines for sustainable funds. Also performance evaluation should be facilitated.

In the current study, as in the majority of studies, several factors, which present potential challenges to fund managers, were not considered at all or treated in a simplified manner. Examples of such factors are investor and dividend cash flows, corporate actions and transaction costs, including price impact. Future research aims at analyzing the impact of such factors on sustainable investment strategies.

References

1. Barnett, M.L., R.M. Salomon (2006). Beyond dichotomy: the curvilinear relationship between social responsibility and financial performance, *Strategic Management Journal*, 27, pp. 1101-1122.
2. Derwall, J., R. Bauer, K. Koedjiik (2007). SRI mutual fund performance around the world, in J. Derwall, 2007, The Economic Virtues of SRI and CSR, Dissertation.
3. DiBartolomeo, D. and L. Kurtz (1999). Managing risk exposures of socially screened accounts, Northfield, Working Paper.
4. Dimtcheva, L., G. Morrison, and J. Marsland (2002). Green with envy, Commerzbank Securities, 19 July 2002a.
5. Dimtcheva, L., G. Morrison, and J. Marsland (2002). Boxing against green shadows, Commerzbank Securities, 19 July 2002b.
6. Frino, A., D. Gallagher (2002). Is index performance achievable? An analysis of Australian equity index funds, *Abacus* 38 (2), pp. 200-214.
7. Jennings, W.W., och G.W. Martin (2007). Socially enhanced indexing: applying enhanced indexing techniques to socially responsible investment, *The Journal of Investing*, 16 (2), pp. 18-31.
8. Kurtz L. (2005). Answers to four questions, *The Journal of Investing*, Fall, pp. 125-137.
9. Milevsky, M.A., A. Aziz, A. Goss, J. Thomson, and D. Wheeler (2006). Cleaning a passive index: how to use portfolio optimization to satisfy CSR constraints, *Journal of Portfolio Management*, Spring, pp. 110-118.
10. Olsson, R. (2005). Portfolio management under transaction costs: model development and Swedish evidence, Dissertation.
11. Statman, M. (2006). Socially responsible indexes, *Journal of Portfolio Management*, Spring, pp.100-109.
12. Troutman, M.L. (2001). Risk control techniques for social investing, *Journal of Investing*, 10, pp. 51-57.