

Can Efficiency of Returns Be Considered as a Pricing Factor?

J. Francisco Rubio¹ · Neal Maroney¹ ·
M. Kabir Hassan¹

Accepted: 13 January 2017 / Published online: 27 January 2017
© Springer Science+Business Media New York 2017

Abstract We add to the investments literature by employing new techniques to estimate asset performance. We estimate a data envelopment analysis based efficiency score that allows for direct comparison between ex-post efficiency rankings and test the ex-ante relevance of such scores by including them into asset pricing models. We find that knowing the fund efficiency score can help explain time-series returns. When efficiency is included in an asset pricing model, the absolute value of the average mispricing error is decreased, which we take as evidence of the explanatory power of efficiency scores. But more importantly, we show that efficacy scores can be used as next period predictors of stock returns. In addition, we further use the efficiency scores to differentiate between the performance of constrained and unconstrained investment assets, as in the case of socially responsible investments. Our findings give robustness to the literature on constrained investments showing significant underperformance of socially and responsible investments.

Keywords Mutual funds · Performance · Data envelopment analysis · DEA · Risk factors · SRI · Ethical investing

1 Introduction

According to Tobin (1958); Hanoch and Levy (1969); Arditti and Levy (1975); Leland (1999), and Joro and Na (2006) the mean-variance frontier is only consistent with traditional utility theory if either stock returns are normally distributed or the utility function is quadratic in nature. In fact, since both assumptions are violated if returns

✉ J. Francisco Rubio
j.rubio@ccsu.edu

¹ Central Connecticut State University, New Britain, CT, USA

are skewed, normality of returns is regarded as the only sufficient condition. Moreover, [Ang et al. \(2006\)](#) further this discussion by estimating the downside risk of equities; equities which are positively skewed should be regarded as more risky thereby providing higher returns. However, they estimate downside risk as market driven rather than asset specific.

Given that stock returns are not necessarily normally distributed ([Lau et al. 1990](#); [Turner and Weigel 1992](#); [Campbell and Hentschel 1992](#); [Arditti 1975](#); [Mandelbrot 1963](#); [Fama 1965](#); [Press 1967](#); [Praetz 1972](#); [Blattberg and Gonedes 1974](#); [Simkowitz and Beedles 1980](#); [Smith 1981](#); [Ball and Torous 1983](#); [Kon 1984](#); [So 1987](#); [Gray and French 1990](#)), investors need to be compensated for further moments ([Kraus and Litzenberger 1976](#); [Kane 1982](#); [Ho and Cheung 1991](#)). More specifically, the risk-return paradigm needs to be re-evaluated to account for skewness when pricing stocks and mutual funds.

We contribute to the literature by expanding the discussion on downside risk; that is, the extra risk associated with positively skewed returns. Unlike [Ang et al. \(2006\)](#), we develop a measurement for said downside risk as an asset specific measurement which is independent of overall market conditions. In this regard, we look for a measurement that can (1) depend on positive skewness which signals high downside risk, (2) be fund-specific in nature, and (3) be easily quantifiable; Data envelopment analysis meets all this requirements. Using DEA, we estimate an asset specific relative efficiency score which we expect can improve on current empirical asset pricing models. Our contribution works twofold: (1) we include efficiency scores in traditional empirical asset pricing model to test the hypothesis that investors do think about efficiency when investing, and (2) efficiency can help signal differences between investment assets, as it is the case of Socially and Responsible Investing assets.

The use of DEA allows for multiple inputs and outputs to be incorporated into a production framework while still offering a single ex-post performance index, thus allowing easy comparison between funds and even groups of funds. We use this to compare constrained and unconstrained investments. DEA estimates a production possibility frontier which envelops a production possibility set as tightly as possible. One can estimate a performance index based on the distance between a specific fund and the aforementioned frontier. Funds on said frontier are deemed 100% efficient.

Several studies have ranked the performance of mutual funds using DEA scores, but to our knowledge, such studies have been limited to only comparing efficiency scores among different assets [see for example [Murthi et al. \(1997\)](#), [McMullen and Strong \(1998\)](#), [Morey and Morey \(1999\)](#), [Choi and Murthi \(2001\)](#), [Basso and Funari \(2001, 2003, 2005\)](#), [Galagedera and Silvapulle \(2002\)](#), [Gregoriou \(2003, 2006\)](#), [Haslem and Scheraga \(2003\)](#), [Chang \(2004\)](#), [Darling et al. \(2004\)](#), [Gregoriou and McCarthy \(2005\)](#), [Gregoriou et al. \(2005\)](#), [Gregoriou and Chen \(2006\)](#), [Eling \(2006\)](#), among others].

Our study extends these findings by including efficiency scores in factor regressions. We establish an efficiency score as a subjective characteristic of a fund at a given month. The efficiency value is defined as $0 < 1 - P^{BCC} < 1$, where 1 represents a fund which is 100% inefficient. Overall, our results suggest that the use of our efficiency score can help decrease average mispricing errors. We find that including the efficiency score

in time-series regressions decreases average alphas from 18 basis points to 13 basis points in absolute values.

Furthermore, this study updates the literature by measuring the performance of smaller asset sub-samples that impose limiting rules on the available asset universe, as is the case of Socially Responsible Investing,¹ also known as Sustainable and Responsible Investing (SRI Investments are described in more depth in “Appendix”). This set of assets is self-restricted to investing only in specific types of companies, thus effectively influencing portfolio performance based on the mean-variance frontier. Hence, we expect that SRI holds more diversifiable risk. But the performance of these constrained assets has been debated in the literature without consensus (Markowitz 1952, 1959). Yet investment in such funds has increased during the last decades according to the Social Investing Forum.² We expand the literature by not only using the efficiency measurement to rank the funds, but also to measure the level of over- or under-performance.

SRI accounts for a small section of the overall American market, and as such, direct comparison with the asset universe will result in a size bias. To overcome this problem, we compare SRI with unconstrained investments as follows: first, given an efficiency score and a fund type, we rank funds in quintiles. Cross-tabulations³ of efficiency score quintiles for SRI and traditional funds show that there are more ethical funds within the low efficiency quintiles. There are 4.27% more socially responsible mutual funds in Q1 and Q2 compared to traditional investing, while there are 3.75% more orthodox funds in Q4 and Q5 compared to ethical funds.

Second, we include an ethical dummy in a cross-section regression, which includes the efficiency score, to test if efficiency has an effect on such asset sub-samples. We find that ethical funds receive a discount between 7 and 12 basis points in next period returns. Controlled for efficiency, this is taken as evidence that ethical mutual funds underperform traditional investment assets.

In addition, we provide evidence of a significant difference in performance pre- and post- crises. Based on a Chow test, we define 2 structural breaks in the data corresponding to the dot-com bubble and the real state bubble. Our findings show that alphas have increased from pre-crisis levels suggesting that mutual funds have enjoyed an extend period of abnormal returns following the structural breaks. Moreover, we provide evidence that although investors do seem consider efficiency when investing,

¹ SRI only invests in firms with high ethical principles and high social responsibility of business practices, where the common denominator is usually avoiding investing in alcohol, tobacco, gambling, pornography, and weapons. But the specific holdings of each fund depend on the managing firm; for example, Northwest and Ethical Funds Investing limit their assets to only those which meet their rigorous standards on environmental risk, social risk, and governance risk. *Environmental risk* Do the company's activities have a negative impact on the air, land or water? We look into issues, such as biodiversity, land-use planning, emissions, climate change and water use; *Social risk* Does the company respect human rights and workers' safety? *Governance risk* How is the company run? We look at factors such as board of director diversity, independence and executive compensation. www.ethicalfunds.com.

² <http://sriconference.com/index.jsp>.

³ Alternatively, we also estimated a size-adjusted comparison between the monthly efficiency scores of ethical and traditional funds. The results show that traditional investments systematically score higher than ethical investments. This is further evidence of ethical funds underperforming unconstrained investment.

this consideration was stronger for the pre-crisis period in comparison with the post-crisis periods.

Finally, it should be noted that our findings not only expand on the traditional empirical estimation of asset pricing models, they also provide the means of comparison between different asset classes, as exemplified by the use of SRI investing. The remainder of this paper is structured as follows. Section 3 develops the estimation of an efficiency score and its use under an investment strategy paradigm as well as describing the data. Finally, Sect. 4 provides concluding remarks.

2 Estimating Mutual Funds Efficiency: A DEA Approach

Portfolio performance is categorized by a mean-variance frontier, which results in a benchmark for comparison of the assets' risk adjusted returns. In this regard, risk is proxied by the portfolio's standard deviation [see [Treyner 1965](#); [Lintner 1965](#); [Sharpe 1966](#); [Jensen 1968](#); among others]. However, when the distribution is skewed, the risk-return relationship fails to account for the difference between the upside and downside risk. As positively skewed returns are inherently riskier due to the increased likelihood of a loss. Thus we look into an operational model which allows for multiple parameters of risk, allowing for the inclusion of higher moments.

Data envelopment analysis (DEA), as proposed by [Charnes et al. \(1978, 1979\)](#), focuses on an operational research methodology which allows the measurement of the relative efficiency of decision making units (assets) in the presence of multiple inputs and outputs. DEA ranks the efficiency of decision making assets and yields a single value which can be then easily assigned to the unit. In fact, with DEA, we can estimate a time specific score and assign that score as a characteristic of the asset, which we will call efficiency. In terms of investment assets, one can estimate monthly efficiency scores of each mutual fund and then compare them over time.

Several studies have analyzed the performance of mutual funds using DEA [see [Murthi et al. \(1997\)](#), [McMullen and Strong \(1998\)](#), [Morey and Morey \(1999\)](#), [Choi and Murthi \(2001\)](#), [Basso and Funari \(2001, 2003, 2005\)](#), [Galagedera and Silvapulle \(2002\)](#), [Gregoriou \(2003, 2006\)](#), [Haslem and Scheraga \(2003\)](#), [Chang \(2004\)](#), [Darling et al. \(2004\)](#), [Gregoriou and McCarthy \(2005\)](#), [Gregoriou et al. \(2005\)](#), [Gregoriou and Chen \(2006\)](#), [Eling \(2006\)](#), among others]. Our study extends their findings by incorporating the efficiency score into factor regressions.

We expect DEA to give a sound estimation of a performance index by incorporating multiple inputs and outputs simultaneously. DEA assumes a production framework in which one uses inputs to produce outputs—where inputs are defined as bearers of risk while outputs are bearers of wealth. Under normality of returns, only the variance and mean should suffice as measurements for risk and wealth. However, if distributions are non-normal, as it is often the case, then higher moments should be required ([Glawischnig and Sommersguter-Reichmann 2010](#)).

It is expected that different assets will have different statistical properties, such as skewed return distributions with fat tails, which makes relevant the inclusion of at least the third moment when characterizing the relationship between risk and return; it is thus paramount to check the normality of returns.

2.1 Data Specifications

We look at the performance of mutual funds listed in the Center of Research in Security Prices (CRSP). The data consists of monthly returns from January 1960 to September 2012. We also differentiate those funds that can be considered Socially Responsible Investments (SRI). We identify as SRI funds those that match the keywords “Ethical”, “Social”, “Socially”, and “SRI” within the fund name. This reveals that the first ethical mutual fund appeared in 1984, and thus we limit our sample to assets post-January 1984.

In order to use DEA, we require fund specific data based on the individual fund’s moments. This implies a restriction on the minimum number of observations per fund. We required funds to have a full trading history and a minimum of 3 years of existence. Finally, following Carhart (1997) we omit sector funds, international funds, balanced funds, and money market funds. The resulting dataset consists of an asset universe of 43,844 mutual funds, of which ethical funds account for only 142 funds.

Table 1 shows summary statistics regarding ethical funds and orthodox mutual funds from 1984 to 2012. Simple inspection reveals that the maximum yearly average returns of SRI funds are smaller than orthodox funds, while average minimum returns are usually never as low as that of traditional funds. Intuition will suggest that these types of investments are less risky than traditional investments and thus their returns are less volatile. Table 1 shows that the standard deviation of returns for ethical investment is usually lower than that of traditional investment except for brief periods of time. Finally, average returns seem to be higher for ethical funds than for orthodox funds –not adjusted for risk.

2.2 Normality Testing

In order to measure normality, we use the Jarque Bera (JB) test statistic (Jarque and Bera 1980, 1981, 1987), formally:

$$JB_j = \frac{T}{6} \left[S_j^2 + \frac{K_j^2}{4} \right] \sim \chi^2(k) \quad (1)$$

where

$$S_j = \frac{E_j (r_t - r_d)^3}{\sigma_j^3}, \text{ the skewness of monthly returns of fund } j$$

$$K_j = \frac{E_j (r_t - r_d)^4}{\sigma_j^4} - 3, \text{ the excess kurtosis of monthly returns of fund } j$$

r_t , monthly return

r_d , average returns

σ_j , standard deviation.

Normality is rejected when JB_j is larger than 5.99 for the 5% significance level and 9.21 for the 1% significance level. We find that at the 5% level, 20839 funds

Table 1 Monthly descriptive statistics of mutual funds

Year	Funds	TNA				Return					
		Average	Median	Min	Max	SD	Average (%)	Median (%)	Min (%)	Max (%)	SD (%)
<i>Panel A: Traditional mutual funds</i>											
1984	743	\$22.13	\$10.19	\$0.53	\$2319.28	\$112.53	0.30	0.46	-31.18	33.17	3.89
1985	919	\$23.08	\$10.95	\$0.54	\$2826.16	\$119.13	1.88	1.55	-18.27	59.77	3.44
1986	1202	\$23.11	\$11.64	\$0.53	\$3641.44	\$125.12	1.14	1.06	-23.24	45.60	3.80
1987	1619	\$20.25	\$10.86	\$0.62	\$3950.39	\$112.65	0.22	0.66	-49.62	45.20	6.31
1988	2005	\$16.71	\$10.28	\$0.58	\$3044.52	\$85.27	0.91	0.60	-20.73	89.67	2.60
1989	2202	\$17.70	\$10.68	\$0.34	\$3735.71	\$94.03	1.21	0.93	-17.96	28.30	2.58
1990	2402	\$15.39	\$10.24	\$0.29	\$3235.57	\$63.85	0.04	0.47	-29.79	52.46	4.00
1991	2712	\$15.03	\$10.64	\$0.57	\$1682.32	\$50.15	1.65	1.08	-49.71	101.62	3.43
1992	4355	\$12.73	\$10.82	\$0.62	\$1033.32	\$16.82	0.51	0.50	-40.59	95.92	2.41
1993	5664	\$12.87	\$11.08	\$0.80	\$442.94	\$11.39	0.88	0.66	-66.36	34.96	2.38
1994	7260	\$12.02	\$10.37	\$0.66	\$458.74	\$10.41	-0.33	-0.09	-51.08	102.43	2.60
1995	8656	\$12.54	\$10.74	\$0.59	\$539.32	\$11.27	1.46	1.24	-26.94	39.75	2.25
1996	9550	\$13.51	\$11.29	\$0.15	\$596.12	\$12.50	0.82	0.66	-33.44	99.28	2.89
1997	10,671	\$14.57	\$11.88	\$0.14	\$675.92	\$15.22	1.04	0.86	-78.27	532.64	4.15
1998	11,776	\$14.77	\$11.69	\$0.09	\$725.36	\$16.10	0.78	0.60	-60.64	72.04	5.71
1999	12,769	\$15.25	\$11.76	\$0.13	\$616.80	\$13.65	1.29	0.20	-65.88	169.89	4.76
2000	13,918	\$16.27	\$11.93	\$0.08	\$620.78	\$15.46	0.07	0.30	-86.97	97.11	6.34
2001	15,045	\$13.50	\$10.74	\$0.05	\$580.07	\$12.33	-0.40	0.36	-87.25	102.17	6.20
2002	15,537	\$12.24	\$10.35	\$0.05	\$524.54	\$10.79	-1.04	-0.13	-96.24	861.54	6.02

Table 1 continued

Year	Funds	TNA					Return				
		Average	Median	Min	Max	SD	Average (%)	Median (%)	Min (%)	Max (%)	SD (%)
2003	15,902	\$12.61	\$10.63	\$0.06	\$493.87	\$10.70	1.81	1.14	-33.33	33.33	3.41
2004	16,450	\$14.54	\$11.39	\$0.06	\$530.77	\$12.65	0.80	0.87	-34.11	30.00	2.84
2005	17,048	\$15.55	\$11.92	\$0.06	\$547.80	\$13.62	0.53	0.39	-46.56	27.74	2.72
2006	18,029	\$16.51	\$12.49	\$0.06	\$608.46	\$14.65	0.89	0.78	-50.00	29.64	2.57
2007	19,021	\$17.87	\$13.12	\$0.06	\$685.23	\$16.49	0.52	0.50	-33.43	32.17	2.92
2008	20,512	\$14.64	\$10.73	\$0.06	\$639.23	\$14.21	-2.96	-1.15	-64.42	378.94	6.96
2009	22,519	\$12.36	\$9.82	\$0.06	\$551.97	\$11.62	2.15	2.25	-67.84	58.86	6.03
2010	22,584	\$14.82	\$11.04	\$0.06	\$606.04	\$13.61	1.15	1.08	-42.42	43.84	4.88
2011	23,465	\$16.33	\$11.75	\$0.07	\$647.37	\$15.05	-0.04	0.09	-98.02	134.21	4.78
2012	18,288	\$16.49	\$11.84	\$0.72	\$673.47	\$15.06	1.14	1.33	-69.96	95.99	3.70
<i>Panel B: Ethical mutual funds</i>											
1984	1	\$16.59	\$16.74	\$15.74	\$17.21	\$0.46	0.56 ^a	0.79	-2.54	4.15	1.85
1985	1	\$19.17	\$19.27	\$18.17	\$20.61	\$0.75	2.04 ^a	2.15	-2.63	5.58	2.72
1986	1	\$23.58	\$23.56	\$20.92	\$25.19	\$1.16	1.47 ^a	2.07	-6.74	6.44	3.96
1987	3	\$18.65	\$15.25	\$11.90	\$27.61	\$5.87	-0.39	0.60	-17.29	8.11	5.64
1988	4	\$17.10	\$14.94	\$13.02	\$25.28	\$4.48	1.02 ^a	0.42	-2.22	7.07	2.26
1989	4	\$18.72	\$16.43	\$14.74	\$27.79	\$4.66	1.56 ^a	1.13	-1.71	6.06	1.84
1990	4	\$17.88	\$15.87	\$10.00	\$27.44	\$5.02	0.07	0.01	-7.81	7.36	3.26
1991	5	\$18.83	\$17.39	\$9.68	\$29.22	\$5.44	1.60 ^a	1.48	-4.22	11.40	3.21
1992	6	\$16.80	\$16.81	\$5.96	\$29.86	\$7.07	0.75	1.26	-2.97	5.54	1.87

Table 1 continued

Year	Funds	TNA				Return					
		Average	Median	Min	Max	SD	Average (%)	Median (%)	Min (%)	Max (%)	SD (%)
1993	8	\$17.75	\$17.57	\$6.69	\$31.17	\$6.56	0.74	0.86	-4.40	4.09	1.77
1994	13	\$15.50	\$15.92	\$4.41	\$30.84	\$7.00	-0.19	-0.03	-6.72	4.29	2.45
1995	15	\$16.27	\$16.27	\$4.47	\$32.91	\$7.18	2.05 ^a	1.98	-2.46	8.34	1.73
1996	15	\$17.76	\$17.52	\$6.20	\$33.12	\$6.63	1.07 ^a	1.14	-5.84	8.39	2.84
1997	19	\$17.99	\$17.61	\$6.84	\$35.79	\$7.29	1.92 ^a	2.43	-6.10	10.27	4.14
1998	29	\$21.29	\$20.26	\$7.56	\$40.21	\$8.33	1.55 ^a	2.45	-20.10	9.05	6.21
1999	42	\$22.54	\$24.04	\$8.98	\$41.89	\$8.02	0.97 ^a	0.69	-9.26	8.94	3.31
2000	51	\$20.85	\$19.77	\$8.34	\$43.03	\$8.85	-0.45	-0.49	-30.47	13.42	5.13
2001	64	\$16.11	\$13.40	\$3.86	\$36.09	\$8.10	-0.68	0.00	-34.54	30.61	7.06
2002	63	\$14.68	\$12.93	\$2.37	\$42.90	\$8.06	-1.55 ^a	-0.95	-19.94	19.08	6.06
2003	59	\$15.53	\$14.51	\$5.62	\$41.77	\$7.87	1.77	1.06	-4.77	12.45	3.03
2004	57	\$17.05	\$15.90	\$7.38	\$45.21	\$8.51	0.65	0.88	-5.44	5.69	2.03
2005	51	\$16.94	\$15.63	\$7.45	\$53.24	\$8.72	0.43 ^a	0.10	-6.42	6.63	2.25
2006	52	\$17.88	\$15.68	\$8.30	\$59.02	\$9.28	0.98	0.92	-6.74	8.32	1.88
2007	57	\$19.00	\$15.83	\$8.86	\$64.33	\$10.86	0.30 ^a	0.48	-8.32	10.44	2.56
2008	66	\$15.43	\$12.56	\$1.25	\$58.57	\$9.32	-3.53	-1.63	-29.06	11.01	6.40
2009	87	\$13.45	\$11.41	\$1.15	\$60.93	\$8.37	2.39 ^a	3.22	-13.07	19.70	5.80
2010	86	\$16.27	\$14.11	\$1.50	\$66.82	\$9.70	1.18 ^a	1.68	-10.21	12.02	5.01
2011	85	\$18.25	\$15.48	\$1.67	\$71.96	\$11.09	-0.02	-0.09	-17.12	17.52	4.41
2012	63	\$18.89	\$15.85	\$1.81	\$74.01	\$11.57	1.33 ^a	2.05	-15.24	11.69	3.39

Panel A shows traditional mutual funds while Panel B shows ethical mutual funds

^a Represents higher average return

Table 2 JB statistics for the fund universe accounting the sample period from January 1984 to September 2012

	Rejection rule for JB statistic			
	Panel A: Standard χ^2		Panel B: Bootstrapped χ^2	
	5% significance	1% significance	5% significance	1% significance
Traditional funds	21,563 69.57%	18,865 60.86%	18,388 59.32%	11,616 37.48%
Ethical funds	61 52.14%	52 44.44%	43 36.75%	31 26.50%

are non-normally distributed. That is, about 69.2% of ethical funds and only about 50.4% of traditional funds can be considered normally distributed. These results are summarized in Table 2 Panel A.

It is, however, expected that the results could be bias towards non-normality due to sample interdependence.⁴ Thus we bootstrap the returns to check how many funds will reject true normality under traditional χ^2 values. We estimate three-year-rolling-period betas, which we then use to create normal returns that mimic fund performance. Given the fund's specific $\beta_{i,t}$, we estimate what should be regarded as normally distributed returns given by $r_{i,t} = \alpha + \beta_i (r_m - r_f) + \varepsilon$ where $r_m \sim N(\mu, \sigma_m^2)$ and $\varepsilon \sim N(0, \sigma_r^2)$.

Using standard χ^2 values, we find that about 10% of the funds would reject normality with a 5% confidence level and 6% of funds would be rejected with a 1% confidence level. This proves the need of sample specific χ^2 values. These are 9.83 for the 5% level and 24.83 for the 1% level. Under the new χ^2 values, normality is rejected for a total of 18431 funds under a 5% confidence level, proving the need for moments beyond variance.

2.3 The DEA Model

We estimate fund-specific relative performance scores, which allow for ex-post ranking based on the data envelopment analysis. This value is estimated based on a 3-year monthly rolling window starting on January 1984. This results in the first efficiency score corresponding to December 1986 based on the time window January 1984 – December 1986.

Following Galagedera and Silvapulle (2002) and Glawischnig and Sommersguter-Reichmann (2010), we see the need for a DEA method that is input oriented with variable returns of scale (VRS). The use of this model can be justified because (1) constant returns of scale assume a constant ratio between inputs and outputs which is not consistent with the randomness of stock returns, and consequently mutual fund

⁴ This is expected because funds invest in a finite number of assets. In fact, the pool of ethical assets is even smaller, thus having more interdependence of funds, which increases the need to estimate sample specific χ^2 values.

returns; (2) the assumption of variable returns to scale is further justified by the fact that alternative investment funds might operate in regions of increasing or decreasing return to scale due to, for example, minimum investment requirements or fixed cost digression; (3) input oriented models are preferred because the output-oriented models work under the assumption of how much the outputs would have to increase for the fund to become efficient while maintaining inputs constant, which is not feasible from a mutual fund manager's stand point; and (4) Glawischnig and Sommersguter-Reichmann argue that the use of the BCC model is advisable whenever ratios are used as inputs or outputs. Evidence under input oriented with CRS can be seen in [Lim et al. \(2014\)](#) for the Korean market.

We, therefore, use the input oriented Banker, Charnes, and Cooper DEA model (BCC) under variable returns of scale. This DEA model estimates a production possibility frontier (PPF) to a production possibility set (PPS) that envelops said PPS as tightly as possible. When comparing the DEA frontier with standard portfolio theory (the Markowitz portfolio theory), the DEA frontier is essentially the same but with a different approach: the DEA frontier is the convex combination of best practices⁵ followed by industry, given the multiple hyperplanes resulting from the use of multiple inputs and outputs. Given that frontier, one can estimate a performance index based on the distance between a specific fund and the aforementioned frontier.

Formally, given $j = 1, \dots, N$ funds where each uses $x_m \forall m = 1, \dots, M$ inputs to produce $y_r \forall r = 1, \dots, R$ outputs, then the performance estimate for the k^{th} fund, P^{BCC} , is given by the solution to the linear programming (LP) problem:

$$\begin{aligned}
 P^{BCC} &= \min \theta \quad s.t. \\
 \sum_{j=1}^N x_{mj} \lambda_j &\leq \theta \cdot x_{mk} \quad \forall m = 1, \dots, M \\
 \sum_{j=1}^N y_{rj} \lambda_j &\geq y_{rk} \quad \forall r = 1, \dots, R \\
 \sum_{j=1}^N \lambda_j &= 1 \quad \text{with } \lambda_j \geq 0 \quad \forall j
 \end{aligned} \tag{2}$$

The LP given in 2 represents the percentage of efficiency of each particular fund k . Problem 2 minimizes the equiproportionate (radial) contraction θ of the inputs produced by unit k . P^{BCC} satisfies $0 \leq P^{BCC} \leq 1$ where $P^{BCC} = 1$ represents a fund which is 100% efficient. That is, given the input-output combination which characterizes the best practice frontier, $P^{BCC} = 1$ represents the k^{th} fund which is on the actual frontier and thus the fund is fully efficient. Any deviation from the PPF, given by $P^{BCC} < 1$, will result in inefficiencies. Furthermore, the linear program uses exogenous weights, λ_j , to fit the best linear combination of all funds given a specific input-output combination; $\lambda_j \neq 0$ represents the best practice units that delimit the

⁵ Best practices, in this context, refer to those funds with the highest level of outputs given the prevailing level of inputs.

frontier. That is, the process only uses those inputs and outputs delimiting the frontier and disregards everything else.

2.4 Inputs and Outputs Description

The specification of the input-output model is given by the risk-return relationship. In this case, all risk measures are considered as inputs while all return measurements will be considered outputs. Following the literature, initial input candidates are given by the fund’s standard deviation, the lower partial moments (*LPM*), and maximum drawdown period, while the output candidates are the expected returns, the upper partial moments (*UPM*) and the maximum consecutive gain.

The partial moments, both upper and lower, are estimated as the *m*th root of the *LPMs* and *UPMs* described by Eq. 3, in percentages. While the *LPMs* capture the downside or risk of holding a specific investment asset, the *UPMs* will capture the upside or benefit of the investment. We use *r_{min}* as the mean return to differentiate between the downside and the upside of the investment strategy.

$$\begin{aligned}
 LPM_{j,m} &= \frac{1}{\tilde{T}} \sum_{t=1}^{\tilde{T}} (r_{min} - \underline{r}_{t,j})^m \quad \forall m = 0, \dots, 4 \\
 UPM_{j,m} &= \frac{1}{\hat{T}} \sum_{t=1}^{\hat{T}} (\bar{r}_{t,j} - r_{min})^m \quad \forall m = 0, \dots, 4
 \end{aligned}
 \tag{3}$$

where

r_{min}, target rate

r_{t,j}, monthly return of fund *j* below target rate

T̃, number of returns of fund *j* below target rate

r̄_{t,j}, monthly return of fund *j* above target rate

T̂, number of returns of fund *j* above target rate.

It is worthwhile to note that *LPM0* accounts for the percentage of funds below the target rate and *LPM1* accounts for the percentage of funds above it, so using both *LPM0* and *UPM0* will be redundant.⁶ Thus using only one of them is sufficient. *LPM0* is preferred to *UPM0* because the risk measurement is regarded as more important.⁷

Finally, the maximum drawdown period (*MDP*) is estimated as the maximum number of months that fund *j* has been below historically high net asset value (*NAV*). And the maximum consecutive gain is the maximum number of months fund *j* has been above the minimum target rate.

Basso and Funari (2003, 2008) further include an output measurement that takes positive values from 0 ... 3 based on the degree of ethical exposure. However, as this

⁶ $LPM0_j = \frac{1}{\tilde{T}} \sum_{t=1}^{\tilde{T}} 1$; $LPM0_j = 1 - UPM0_j$.

⁷ Looking at the correlation between I/O also supports choosing *LPM0* instead of *UPM0*.

imposes a bonus for ethical investment which could bias the results in their favor, we avoid such a measurement.

Consistent with prior studies, a production framework that provides higher rewards for taking extra levels of risk is chosen, thus requiring inputs and outputs to be positively correlated. The correlation table (not reported) shows that only the inputs *LPM0-LPM4* and *MDP* as well as the outputs *UPM1-UPM4* are positively correlated. Further, the fourth moments, *LPM4* and *UPM4*, are disregarded due to the extreme high correlations; *LPM3* and *LPM4* as well as *UPM3* and *UPM4* show a correlation greater than 0.999.⁸ Still, using the third moment is enough to apply an extra penalty (bonus) to funds with high negative or positive skewness.

All inputs and outputs are estimated on a 3 year rolling window basis. We impose a restriction that each fund has to have a full history in NAV and returns to estimate its corresponding inputs and outputs. Since the first 36 observations are lost due to the estimation technique, future results are the average based on 309 months. Table 3 summarizes the descriptive statistics of the input-output specifications considered for the production framework used in further tests. Except for *MDP*, simple inspection shows that all inputs and outputs are usually higher for ethical funds. This is statistically significant only for *LPM0-LPM2* at the 10% level.⁹

It is worth mentioning that only orthodox funds achieve an average 100% on *LPM0*. This means that every month there are funds that have had only below average returns ($LPM0 = 1$) during the prior 36 months. Furthermore, the opposite is also true; only within traditional investment are there funds having all returns above average during the prior 36 months. Generally speaking, traditional investing seems perhaps more volatile than SRI.

The following subsection provides further analysis of each fund's monthly relative efficiency, as given by its subjective efficiency value.

2.5 Estimating Efficiency

Since the number of mutual funds increases overtime, finding a feasible solution for all funds might not be possible; the LP might not converge in a solution, and thus values of $P^{BCC} = 1$ might be misleading. This is corrected by setting the linear program with two passes, where each pass has a different starting value. P^{BCC} is then recorded only if the solution converges to the same numeric value regardless of its initial seed.

We solve problem 2 based on the inputs and outputs discussed in prior sections. We calculate monthly efficiency values only when a fund has no missing observation during the immediate past 3 years. This results in a panel with monthly efficiency values.

As an application of these efficiency scores, we compare the performance of two different samples: the unconstrained investment universe and the constrained SRI uni-

⁸ Just like Glawischnig and Sommersguter-Reichmann (2010), the results show high correlation among *LPM2* and *LPM3* (*UPM2* and *UPM3*), as well as *LPM2* and *LPM4* (*UPM2* and *UPM4*). Despite that high correlation, they are still incorporated into the model to apply an extra penalty/ benefit for higher moments due to non-normality of returns.

⁹ A test of differences in means was estimated but, for simplicity, it is not reported.

Table 3 Summary statistics of the proposed input–output specification used to estimate the DEA performance score under *P^{BCC}*

	Inputs				Outputs			
	MDP	LPM0 (%)	LPM1 (%)	LPM2 (%)	LPM3 (%)	UPM1 (%)	UPM2 (%)	UPM3 (%)
<i>Panel A: Traditional mutual funds</i>								
Mean	16	48.720	1.467	2.755	3.714	1.468	2.452	3.132
Median	14	47.222	1.207	2.354	3.173	1.367	2.263	2.865
Minimum	1	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Maximum	36	100.000	14.378	22.766	29.137	17.397	88.715	16.100
Standard deviation	13	14.484	1.020	1.875	2.551	1.073	1.834	2.419
<i>Panel B: Ethical mutual funds</i>								
Mean	15	47.291	1.569	2.927	3.922	1.523	2.559	3.260
Median	11	47.222	1.610	3.081	4.155	1.503	2.571	3.277
Minimum	1	2.778	0.005	0.029	0.052	0.001	0.003	0.006
Maximum	36	97.222	3.589	6.996	9.968	4.422	7.159	9.027
Standard deviation	12	10.058	0.750	1.385	1.880	0.746	1.206	1.528

All inputs and outputs are estimated based on the prior 3 years. Since the first 36 observations are lost due to estimations, results are the average based on 309 months

Table 4 Reports the cross-tabulation of average monthly P^{BCC} values for both ethical and unrestricted mutual funds

	Quintile Distribution of P^{BCC}					Total
	Low Q1	Q2	Q3	Q4	High Q5	
Traditional						
Funds	1626	1564	1580	1642	1286	
%Total	21.03	20.23	20.43	21.23	16.63	99.55
%Row	21.12	20.32	20.52	21.33	16.71	
%Column	99.57	99.43	99.56	99.58	99.61	
Ethical						
Funds	7	9	7	7	5	
%Total	0.09	0.12	0.09	0.09	0.06	0.45
%Row	20.00	25.71	20.00	20.00	14.29	
%Column	0.43	0.57	0.44	0.42	0.39	
Total	21.12	20.34	20.52	21.32	16.69	1

We report estimations based on the average number of monthly funds. Chi-square tests were conducted based on monthly cross-tabulations as well as the overall sample; the results are not reported but the P values are estimated as <0.001 , which rejects the null that the percentage among groups are equal

verse. We rank funds based on their P^{BCC} value and sort them into quintiles where Q1 is the lowest efficiency. Every month, we compute a cross-tabulation between the fund's monthly score with a differentiation based on whether a particular fund is catalogued as a socially responsible investment or not. Every time we run a cross-tabulation we also perform a chi-square test for difference between groups (not reported). We store the number of funds per category and estimate a time series average number of funds per category. Table 4 reports the results.

As expected, the number of traditional funds far outweighs that of ethical funds. Thus direct comparisons of the percentage number of funds per quintile are irrelevant to conclude difference in efficiency rankings. Instead, we look at the conditional probabilities of each quintile. It seems that traditional investment achieves the highest levels of efficiency. There is an average of 2.42% more traditional funds in the high efficiency quintile, Q5, than ethical funds.¹⁰ Though traditional investments have a higher amount of funds in the lowest quintile, Q1, when Q1 and Q2 are taken together, socially responsible investment has 4.27% more funds in the lower efficiency categories than orthodox investments (mainly resulting from an average 5.39% more ethical funds in Q2).

Unlike Hamilton et al. (1993), Diltz (1995), Guerard (1997), Sauer (1997), Bauer et al. (2005), Goldreyer and Diltz (1999), who have shown that ethical investing is not different than traditional investing, we find that SRI is more heavily distributed towards the low efficiency scores. These results suggest that traditional mutual funds

¹⁰ The difference is expected to be statistically significant given close inspection of monthly chi-square tests, but the test results are not reported for simplicity.

are more efficient than socially responsible investments.¹¹ However, a word of caution is advised given the presence of traditional funds in the very lowest quintile.

Having estimated an ex-post realization of efficiency, it remains unknown if investors take this into consideration when assessing their required rate of return. Though it is expected that highly efficient funds will have higher returns, it is uncertain if the efficiency score can be used as a predictor of future fund performance. The following sections looks at the explanatory power of the estimated efficiency scores and their plausible use as proxies for risk.

3 The Efficiency Score Explanatory Power in Time Series Regressions

Since ex-post estimations can still have an element of luck, we need to test whether knowing a fund’s efficiency score can produce an effective investment strategy. If investors care about fund efficiency, then it will become a valid component of a security’s pricing. Intuition suggests that investors should require higher rates of return for holding underperforming assets. We therefore argue that this will also be true for inefficient assets. We thus test the hypothesis that adding efficiency scores to would improve on the existing empirical asset pricing models; we expect that adding efficiency would decrease mispricing errors in CAPM and Carhart’s four-factor model.

Early attempts to fit the regressions resulted in a market beta well below 1. This would suggest a small correlation between mutual fund returns and the market excess returns. Further inspection revealed that a large number of small market beta funds rely on debt instruments.¹² Therefore, we include two bond factors to help explain the returns. These factors are constructed based on Fama and French’s (1993) suggested bond factors.

First, $TERM_t$ is expected to capture the term structure of bond yields, defined as the 10-year treasury security minus the 1-year treasury security. Second, $DEFAULT_t$ is set to capture the default premium of bond yields, measured as the spread between BAA and AAA corporate bonds.

Finally, because $0 < P^{BCC} < 1$, we could define $P^{BCC'} = 1 - P^{BCC}$ which measures total efficiency deviations rather than efficiency itself. This specification of $P^{BCC'}$ guarantees that no penalty will be given for full efficiency, but inefficient funds will receive a 100% penalty. In other words, $P^{BCC'}$ can be regarded as an efficiency discount measured as the percentage deviation from full efficiency. The equations are:

$$R_{i,t} - r_{f_t} = \hat{\alpha}_t + \hat{\beta}_{1,t} (R_{m_t} - r_{f_t}) + \hat{\beta}_{2,t} TERM_t + \hat{\beta}_{3,t} DEFAULT_t \tag{4}$$

$$R_{i,t} - r_{f_t} = \hat{\alpha}_t + \hat{\beta}_{1,t} (R_{m_t} - r_{f_t}) + \hat{\beta}_{2,t} TERM_t + \hat{\beta}_{3,t} DEFAULT_t + \hat{\beta}_{4,t} (P_t^{BCC'}) \tag{5}$$

¹¹ Earlier estimates looked at the monthly distribution of P^{BCC} based on a 1-1 match between ethical and traditional investment for a 3000-iteration bootstrap. Results were congruent with these findings: ethical funds appear less efficient than traditional investment assets, consistently after 1990. Results are available upon request.

¹² Even after deleting all money market mutual funds, the possibility of large debt holding is not removed. CRSP only reports the latest objective, and thus the fund could have historically had money market instruments. Furthermore, many funds rely in bonds.

$$R_{i,t} - r_{f,t} = \hat{\alpha}_t + \hat{\beta}_{1,t} (R_{m,t} - r_{f,t}) + \hat{\beta}_{2,t} TERM_t + \hat{\beta}_{3,t} DEFAULT_t + \hat{\beta}_{5,t} SMB_t + \hat{\beta}_{6,t} HML + \hat{\beta}_{7,t} MOM_t \quad (6)$$

$$R_{i,t} - r_{f,t} = \hat{\alpha}_t + \hat{\beta}_{1,t} (R_{m,t} - r_{f,t}) + \hat{\beta}_{2,t} TERM_t + \hat{\beta}_{3,t} DEFAULT_t + \hat{\beta}_{5,t} SMB_t + \hat{\beta}_{6,t} HML + \hat{\beta}_{7,t} MOM_t + \hat{\beta}_{4,t} (P_t^{BCC'}) \quad (7)$$

The betas are estimated based on a rolling window of 36 months, requiring full trading data. After estimating P^{BCC} and the betas, the starting date has moved to December 1989. The monthly market premium, $(R_{m,t} - r_{f,t})$, the monthly size factor, SMB_t , the monthly book to market factor, HML , the monthly momentum factor, MOM_t , as well as the monthly risk free rate all come from Kenneth French's website. We estimate monthly excess returns based on said risk free rate. Table 5, Panel A, summarizes the descriptive statistics.

As expected, there are funds which are completely efficient and those which are completely inefficient. The median efficiency score is 61.2% which is about 6% higher than the average score. This suggests a negative skewness of efficiency scores. As for the risk factors, TERM and DEFAULT are higher than all others, but also less volatile. All of Carhart's four factors seem small on average, but highly volatile. Surprisingly, SMB has a negative median which suggests that the returns of big firms are larger than those of small firms. A positive average momentum factor suggests that prior winners continue to outperform prior losers.

Table 5, Panel B, shows the cross-correlations of the proposed factors. Since P^{BCC} is fund specific, its correlation with all factors is estimated based on the pooled sample. One can immediately notice the moderate negative correlation between P^{BCC} and the two bond factors, $\rho_{P^{BCC}, TERM} = -0.5513$ and $\rho_{P^{BCC}, DEFAULT} = -0.4562$. This suggests that bond yields decrease as P^{BCC} scores increase. This could be due to investors requiring higher rates of return for holding inefficient bonds. This in turn should pass onto mutual funds in the same manner. This is more attenuated for long term debt of efficient assets which require lower compensation than long term debt of the less efficient assets. Moreover, both bond factors are positively correlated as expected.

Table 6, Panel A, shows the time-series monthly average mutual fund beta, as well as the time-series t-stat, for the period December 1989 to September 2012. The use of P^{BCC} seems to slightly increase the fit of the models, regardless of the original specification. R^2 increases in both cases by about 2 percentage points, and even Adjusted R^2 (not reported) increases by 1% in all cases. Further analysis between regressions shows that all factors are statistically significant, except for the maturity premium, which is only significant under the four factor model. Consistent with prior research, the use of the four factor model improves CAPM results. On one hand the value of R^2 increases from 54.6% to 65.61%, and on the other, the average mispricing error, α , decreases from -0.23% to -0.19% .

After including the efficiency penalty, we find that the beta on $P^{BCC'}$ is highly significant even at the 1% level. The results suggest that an average inefficient fund ($P^{BCC} = 0.5658$) will expect to receive about 72 basis points less than a fully efficient fund ($P^{BCC} = 1$) under the modified CAPM version (MCAPM hereafter), and about

Table 5 Summarizes statistical information on the proposed regressors

	Panel A: Descriptive statistics					Panel B: Correlations						
	Mean	Median	Standard deviation	Minimum	Maximum	p^{BCC}	$R_m - R_f$	TERM	DEFAULT	HML	SMB	Mom
p^{BCC}	0.5628	0.6121	0.3068	0	1	1.000						
$R_m - R_f$	0.43%	1.09%	4.72%	-23.24%	12.46%	0.099	1.0000					
TERM	1.57%	1.78%	1.15%	-0.41%	3.40%	-0.551	-0.0038	1.0000				
DEFAULT	1.09%	0.95%	0.49%	0.55%	3.38%	-0.456	-0.0559	0.2957	1.0000			
HML	0.22%	0.17%	3.17%	-12.60%	13.84%	0.013	-0.3027	0.0402	-0.0844	1.0000		
SMB	0.27%	-0.03%	3.26%	-16.39%	22.00%	-0.049	0.2284	0.1255	0.0296	-0.3279	1.0000	
Mom	0.32%	0.59%	5.59%	-34.74%	18.39%	0.149	-0.1874	-0.0826	-0.1979	-0.1303	0.0358	1

Panel A shows the descriptive statistics and Panel B shows the different correlations between them. The correlation on p^{BCC} is the pooled correlation estimation based on the fund specific scores and the time series of the proposed factors

Table 6 Comparison between time-series CAPM, Carhart's Four Factor Model, and subsequent models than includes a fund specific characteristic, P^{BCC} , which measures its relative efficiency from December 1989 to September 2012

$\hat{\alpha}$	$R_m - R_f$	TERM	Default	SMB	HML	MOM	$1 - P^{BCC}$	R^2
<i>Panel A: Average monthly betas for full sample</i>								
-0.227%	0.5824	0.0043	0.1900					0.546
(-4.1957)*	(98.0958)*	(0.3036)	(3.4329)*					
0.225%	0.5665	-0.0038	0.2713				-0.0164	0.573
(3.5355)*	(94.7863)*	(-0.246)	(4.881)*				(-20.4131)*	
-0.188%	0.5813	-0.0325	0.1599	0.0740	0.0276	0.0052		0.656
(-4.0443)*	(115.2118)*	(-2.2483)***	(3.625)*	(41.8206)*	(4.2559)*	(2.1761)***		
0.137%	0.5712	-0.0046	0.1959	0.0713	0.0262	0.0014	-0.0125	0.675
(2.5027)**	(113.1337)*	(-0.2891)	(4.0828)*	(37.334)*	(4.0126)*	(0.618)	(-21.6537)*	
<i>Panel B: Average monthly betas before December 2000</i>								
-0.235%	0.5401	-0.0163	0.2119					0.520
(-3.0482)*	(94.3984)*	(-0.8426)	(2.2865)**					
0.661%	0.5217	-0.0403	-0.0773				-0.0207	0.554
(7.5178)*	(89.9134)*	(-1.9055)	(-0.879)				(-17.5752)*	
-0.136%	0.5462	-0.0779	0.1517	0.0551	0.0455	0.0168		0.643
(-1.969)	(111.6577)*	(-4.4569)*	(1.9754)	(27.753)*	(7.0065)*	(6.493)*		
0.532%	0.5338	-0.0572	-0.0898	0.0521	0.0417	0.0123	-0.0157	0.668
(7.4334)*	(110.1443)*	(-2.775)**	(-1.1647)	(24.4314)*	(6.2005)*	(5.2437)*	(-19.4946)*	
<i>Panel C: Average monthly betas between January 2001 and March 2009</i>								
0.102%	0.5817	-0.0296	0.0441					0.517
(1.2738)	(55.703)*	(-1.2588)	(0.5074)					

Table 6 continued

$\hat{\alpha}$	$R_m - R_f$	TERM	Default	SMB	HML	MOM	$1 - p^{BCC}$	R^2
0.160%	0.5628	-0.0240	0.5892				-0.0153	0.540
(2.0572)***	(59.3737)*	(-0.946)	(7.5752)*				(-13.7196)*	
-0.076%	0.5733	-0.0873	0.2187	0.0956	0.0821	0.0150		0.628
(-1.207)	(72.6844)*	(-3.9377)*	(3.4567)*	(38.4591)*	(10.5355)*	(4.1972)*		
-0.029%	0.5629	-0.0688	0.5767	0.0931	0.0838	0.0116	-0.0113	0.644
(-0.428)	(78.049)*	(-2.985)*	(9.2446)*	(31.3999)*	(10.9179)*	(3.2727)*	(-12.0915)*	
<i>Panel D: Average monthly betas since April 2009</i>								
-0.960%	0.7153	0.1461	0.4584					0.693
(-9.1764)*	(189.127)*	(4.1324)*	(16.1136)*					
-0.971%	0.7138	0.1555	0.6175				-0.0054	0.705
(-8.2551)*	(184.4643)*	(4.0678)*	(9.834)*				(-3.5529)*	
-0.605%	0.7082	0.2340	0.0498	0.0829	-0.1535	-0.0535		0.760
(-4.6041)*	(168.0976)*	(7.0439)*	(1.1612)	(36.0194)*	(-22.6785)*	(-24.7804)*		
-0.707%	0.7057	0.3057	0.2028	0.0802	-0.1545	-0.0558	-0.0057	0.771
(-5.0333)*	(166.0191)*	(9.4664)*	(3.5013)*	(34.547)*	(-23.3045)*	(-28.5471)*	(-5.4354)*	

All models include two bond factors, TERM which is the 10 year minus the 1 year return in government securities and DEFAULT which is the AAA minus BAA corporate spread, to account for mutual funds with small correlation with equity markets. We estimate monthly betas based on a 3 year rolling window and report the average time series betas of said monthly betas. T-statistics are in parenthesis

*** Significant at the 10%; ** Significant at the 5% level; * Significant at the 1% level

55 basis points less under the modified four factor model. More importantly, a one standard deviation increase, $\Delta P^{BCC} = 0.3068$, will result in an increase of almost 50 basis points under MCAPM and 38 basis points under the modified four factor model.

Moreover, looking at Jensen's alphas, the average mispricing errors decrease when $P^{BCC'}$ is included to explain mutual fund returns. Under the MCAPM model, alpha decreases from -22.7 to 22.5 basis points. While this difference is statistically significant ($t = 12.04$), it is of no economic significance. However, under the modified four factor model, the decrease is of 5.1 basis points which is of both statistical ($t = 11.34$) and economic significance.

It is worth noting that, as expected given the correlation between $P^{BCC'}$ score and the bond factors, including the $P^{BCC'}$ score in the regression has a direct effect on the betas for both bond factors. The beta for the term structure decreases from 0.0043 to -0.0038 under CAPM, while increasing from -0.032 to -0.005 under the four factor model; the beta for the default premium increases from 0.19 to 0.27 and from 0.160 to 0.196 for both CAPM and Carhart's model respectively. These mixed effects are quite puzzling; it appears that including the efficiency score in the CAPM captures the combined effect of size, book to market, and momentum factors.

Having established the appropriate use of the efficiency score as a factor of mutual fund returns, we now check whether the effect is time-specific. We perform a Chow test for structural breaks (not reported) which not surprisingly matches the biggest recessions within our timeframe: the dot-com bubble and the global financial crisis. The Chow test shows that the biggest inflection points are in December 2000 and March 2009. We therefore split the sample in three to account for all different periods. Table 6, Panel B through D, shows the time-series average monthly betas for the subsamples. T-statistics are reported in parenthesis.

All in all, prior results continue to hold true. Regardless of the time period, P^{BCC} does provide statistical and even economic significance. Looking at MCAPM, the strongest effect is before the dot-com where an average efficient fund received a discount of 90 basis points. But during crisis periods, it seems like investors cared less about efficiency; after the dot-com the discount drops to 66 basis points, and it further drops to 23 basis points post 2009. The same holds true under the modified Carhart's model with discounts dropping from 68 basis points pre dot-com, to 25 basis points post 2009.

It seems, however, that the explanatory power of the efficiency values is time-subjective. Looking at the individual sub-periods, alphas increased rather than decrease regardless of the period or model, except for the Carhart's specification between 2000 and 2009. The results suggest that using the efficiency score shows a penalty applied to inefficient funds, but at the same time alphas increased in comparison. For example, pre dot-com, alphas increase by 90 basis points when adding $P^{BCC'}$ to the CAPM specification, but a penalty between $[0, 2\%]$ was added at the same time. This suggest that efficient funds would have enjoyed a 90-basis-point premium while inefficient funds would have received a 1.1% penalty. The same holds true under the Carhart specification. But after 2009, alphas actually decrease by 1 basis point with a penalty between $[0, 54 \text{ basis points}]$. This is consistent with investors not valuing efficiency in a post crisis period.

So far, we have shown evidence that efficiency seems to be a pricing feature of mutual funds. However, we have not provided any evidence that the efficiency scores have any forecasting power. We answer this question in the next section.

4 The Cross-Section of Mutual Funds Returns: Is P^{BCC} significant for pricing future returns?

An important aspect of asset pricing is whether a specific model can accurately predict future returns. In such a case, the model underlies a feasible investment strategy. We thus estimate a second pass regressions at $t + 1$ based on the factor loadings from the prior 3 years. Given an asset pricing model $R_{i,t} = r_{f_t} + \hat{\beta}_t \mathbf{X}_t$ where \mathbf{X}_t is a vector of risk factors, if the model has explanatory power, the returns at $t + 1$ will be given by $R_{i,t+1} = r_{f_{t+1}} + \delta_{t+1} \hat{\beta}_t$, where δ_{t+1} is the risk premium at $t + 1$. Also, the risk premium δ_{t+1} should be equal to \mathbf{X}_{t+1} . We thus estimate the empirical models:

$$R_{i,t+1} - r_{f_{t+1}} = \hat{\alpha}_{t+1} + \hat{\lambda}_{1,t+1} \hat{\beta}_{1,t} + \hat{\lambda}_{2,t+1} \hat{\beta}_{2,t} + \hat{\lambda}_{3,t+1} \hat{\beta}_{3,t} \tag{8}$$

$$R_{i,t+1} - r_{f_{t+1}} = \hat{\alpha}_{t+1} + \hat{\lambda}_{1,t+1} \hat{\beta}_{1,t} + \hat{\lambda}_{2,t+1} \hat{\beta}_{2,t} + \hat{\lambda}_{3,t+1} \hat{\beta}_{3,t} + \hat{\lambda}_{4,t+1} \hat{\beta}_{4,t} \tag{9}$$

$$R_{i,t+1} - r_{f_{t+1}} = \hat{\alpha}_{t+1} + \hat{\lambda}_{1,t+1} \hat{\beta}_{1,t} + \hat{\lambda}_{2,t+1} \hat{\beta}_{2,t} + \hat{\lambda}_{3,t+1} \hat{\beta}_{3,t} + \hat{\lambda}_{5,t+1} \hat{\beta}_{5,t} + \hat{\lambda}_{6,t+1} \hat{\beta}_{6,t} + \hat{\lambda}_{7,t+1} \hat{\beta}_{7,t} \tag{10}$$

$$R_{i,t+1} - r_{f_{t+1}} = \hat{\alpha}_{t+1} + \hat{\lambda}_{1,t+1} \hat{\beta}_{1,t} + \hat{\lambda}_{2,t+1} \hat{\beta}_{2,t} + \hat{\lambda}_{3,t+1} \hat{\beta}_{3,t} + \hat{\lambda}_{5,t+1} \hat{\beta}_{5,t} + \hat{\lambda}_{6,t+1} \hat{\beta}_{6,t} + \hat{\lambda}_{7,t+1} \hat{\beta}_{7,t} + \hat{\lambda}_{4,t+1} \hat{\beta}_{4,t} \tag{11}$$

Table 7 shows the time-series averages of the slopes of the month by month Fama-Macbeth type regression. Consistent with prior research, the average mispricing of future returns decreases under the four factor model. However, the use of an efficiency factor does not seem to improve results as R^2 values do not seem to improve.

Across models, it can be seen that both bond factors have no predictive power regarding excess returns. This is somewhat expected for the maturity premium since its beta is statistically insignificant. Yet the insignificance of the default premium is surprising. Moreover, average factors are 0.43%, 0.22%, 0.27%, and 0.32% for market premium, HML, SMB, and MOM. The estimation shows that only the factor loadings on the market and HML, with factor loading of 0.38% and 0.22%, are close to the averages.

Regarding efficiency, the estimation shows a magnitude problem, but the dimension seems to be accurate. Average $P^{BCC'} = 0.4372$ and average $\hat{\lambda}_4 = 0.0044$, significant at the 5% level, under the MCAPM specification and average $\hat{\lambda}_4 = 0.0045$, statistically significant at the 1% level, under the modified Carhart's specification. This suggest that the efficiency score can indeed be a predictor of next period performance. Given an efficiency loading at $T = t$, we can predict that $r_{T=t+1}$ will decrease by $0.0044 \hat{\beta}_{4,t}$ in percentage terms. Finally, it's worth noting that adding $P^{BCC'}$ actually increases average mispricing errors at $T = t + 1$.

Further analysis of different time periods provides interesting results. While alphas are small and insignificant for the pre-2000s period, they are large and highly sig-

Table 7 Comparison between the cross-section performance at time $t + 1$ of CAPM, Carhart Four Factor Model, and the funds relative efficiency values given by the individual factor loadings estimated at time t given 3 years of prior information

$\hat{\alpha}_{t+1}$	$\hat{\beta}_{1,t}$	$\hat{\beta}_{2,t}$	$\hat{\beta}_{3,t}$	$\hat{\beta}_{5,t}$	$\hat{\beta}_{6,t}$	$\hat{\beta}_{7,t}$	$\hat{\beta}_{4,t}$	SRI	R^2
<i>Panel A: Average monthly cross-section slopes for full sample</i>									
0.107%	0.0038	0.0000	0.0000					-0.0010	0.447
(1.8003)	(1.3969)	(0.0884)	(0.2495)					(-2.032)**	
0.184%	0.0041	0.0003	0.0003				-0.0044	-0.0008	0.455
(3.0051)*	(1.4486)	(0.6533)	(1.9767)***				(-2.6589)**	(-1.729)***	
0.071%	0.0035	0.0000	0.0000	0.0022	0.0009	-0.0003		-0.0006	0.566
(1.362)	(1.2836)	(0.0918)	(0.0372)	(1.2486)	(0.5377)	(-0.1354)		(-1.596)	
0.174%	0.0038	-0.0000	0.0001	0.0023	0.0005	-0.0012	-0.0045	-0.0006	0.563
(3.5121)*	(1.391)	(-0.0815)	(1.0461)	(1.2618)	(0.2583)	(-0.496)	(-4.4786)*	(-1.558)	
<i>Panel B: Average monthly betas before December 2000</i>									
-0.020%	0.0077	0.0002	0.0000					-0.0008	0.397
(-0.2312)	(2.0966)***	(0.3457)	(0.1176)					(-1.178)	
0.101%	0.0085	-0.0003	0.0001				-0.0056	-0.0007	0.405
(1.3126)	(2.2123)***	(-0.5482)	(0.7457)				(-2.0408)***	(-1.027)	
-0.014%	0.0074	0.0003	-0.0001	-0.0001	-0.0028	0.0012		-0.0012	0.514
(-0.21)	(2.1011)***	(0.669)	(-0.6268)	(-0.0188)	(-1.0224)	(0.3836)		(-1.861)***	
0.131%	0.0080	-0.0002	-0.0001	-0.0002	-0.0035	0.0003	-0.0059	-0.0012	0.516
(2.1173)***	(2.2293)***	(-0.6162)	(-0.515)	(-0.0573)	(-1.2097)	(0.0913)	(-3.7237)*	(-1.921)***	

Table 7 continued

$\hat{\alpha}_{t+1}$	$\hat{\beta}_{1,t}$	$\hat{\beta}_{2,t}$	$\hat{\beta}_{3,t}$	$\hat{\beta}_{5,t}$	$\hat{\beta}_{6,t}$	$\hat{\beta}_{7,t}$	$\hat{\beta}_{4,t}$	SRI	R^2
<i>Panel C: Average monthly betas between January 2001 and March 2009</i>									
0.123% (1.2177)	-0.0047 (-1.0394)	0.0002 (0.1726)	0.0002 (0.9499)					-0.0008 (-1.178)	0.468
0.248% (2.024)***	-0.0047 (-1.0065)	0.0006 (0.9157)	0.0006 (2.6782)**				-0.0053 (-2.0117)***	-0.0007 (-1.027)	0.476
0.011% (0.1087)	-0.0051 (-1.0884)	0.0005 (0.5419)	0.0003 (1.4346)	0.0043 (1.747)	0.0049 (1.7882)	0.0005 (0.1026)		-0.0012 (-1.861)***	0.604
0.161% (1.599)	-0.0049 (-1.0474)	0.0002 (0.3161)	0.0003 (1.8863)	0.0044 (1.8053)	0.0046 (1.668)	-0.0002 (-0.0409)	-0.0051 (-2.9781)*	-0.0012 (-1.921)***	0.594
<i>Panel D: Average monthly betas since April 2009</i>									
0.474% (4.2248)*	0.0113 (1.4668)	-0.0008 (-0.4447)	-0.0003 (-0.3347)					-0.0012 (-1.656)***	0.555
0.297% (2.2697)***	0.0108 (1.4019)	0.0012 (0.8823)	0.0004 (0.4767)				0.0013 (0.6996)	-0.0009 (-1.408)	0.563
0.485% (5.2141)*	0.0111 (1.4392)	-0.0018 (-1.6169)	-0.0005 (-0.6302)	0.0044 (1.2631)	0.0033 (1.0585)	-0.0073 (-1.0147)		-0.0001 (-0.231)	0.639
0.342% (3.432)*	0.0111 (1.4425)	0.0000 (0.0455)	0.0002 (0.3337)	0.0048 (1.364)	0.0032 (1.062)	-0.0082 (-1.1162)	0.0011 (0.8517)	0.0000 (-0.099)	0.639

All models include the factor loadings of two bond factors, TERM which is the 10 year minus the 1 year return in government securities and DEFAULT which is the AAA minus BAA corporate spread, to account for mutual funds with small correlation with equity markets. Table 7 reports the time-series averages of the slopes of monthly cross-section regression. Finally, we differentiate between Socially Responsible Investing and Traditional Investing by using a dummy variable for the slopes and intercepts *** Significant at the 10%; ** Significant at the 5% level; * Significant at the 1% level

nificant after January 2001. This suggests difficulty pricing future returns under the rapid expansionary period following the dot com bubble and even the subsequent 2009 financial crisis. Clearly mutual funds have taken advantage of this long bull market.

Looking at efficiency, once again there was a heavy discount before the dot-com. Average $\hat{\lambda}_4 = -0.0056$ under MCAPM, significant at the 10% level, and $\hat{\lambda}_4 = -0.0059$ under the modified Carhart version, significant at the 1% level. But post 2009, this effect drops to average premium of $\hat{\lambda}_4 = 0.0013$ under MCAPM, insignificant at the traditional levels, and $\hat{\lambda}_4 = 0.0011$ under the modified Carhart version, also insignificant at traditional levels.

4.1 DEA as an Ex-Post Efficiency Measurement: The Case Of Socially Responsible Investing

In prior sections, we stated that ethical funds do not seem to be as efficient as traditional unconstrained investing. However, it remains unknown if investors use this information to value ethical funds future returns. If they do, it is expected that the returns of ethical funds will be lower, on average, than those of orthodox mutual funds. In other words, there will be a relative underperformance of ethical investing.

To capture the effect of Socially Responsible Investing in future returns, we include a dummy variable to the estimations proposed in Eqs. 8–11. The dummy takes the value of 1 if a fund can be regarded as ethical and 0 otherwise. We include dummies not only for the intercept but for all factors. This allows to estimate two independent models while at the same time increasing sample size and improving standard errors.

Table 6 reports the intercept dummy for all periods. Looking at the whole sample, SRI funds would face a discount of 8 basis points under the MCAPM model, which decreases to 6 basis points under the modified Carhart version. This is consistent with our previous findings; Socially and Responsible Investments underperform unconstrained mutual funds.

Deeper analysis of the individual sub periods shows that the effect is somehow persistent through time. According to the MCAPM specification, average excess returns for ethical investments are 7 basis points lower for the pre-2000s era (not statistically significant), and its effect attenuates to an average of 9 basis points less after 2009 crisis. According to the four factor model, however, the returns of ethical funds are about 12 basis points before the year 2000, and zero after 2009. But it should be noted that the ethical dummy is not statistically significant across time and across models. It is only significant under CAPM post 2009, and under Carhart pre 2000s.

5 Concluding Remarks

The traditional Markowitz efficiency frontier describes the relationship between equities' risk and return under the assumption that assets are normally distributed. That is, we only need the variance to know the appropriate return of a given security. In this paper we have extended this view to account for higher moments, including skewness. In this case, we estimate an efficiency score based on the best practices of decision making units as they adjust the level of risk (inputs) to achieve maximum efficiency.

This follows the idea that investors will prefer highly efficient funds which minimize the risk that takes to provide a return. We measure efficiency as a non-parametrical measurement based on data envelopment analysis.

Though the use of an efficiency score might be enough for classification purposes, from an investment perspective, it does not bring any new opportunity to the field. Thus, we expand the literature on empirical asset pricing methods by testing the predictability of future mutual fund returns using efficiency scores. We show that the fund's efficiency score alone can improve on known asset pricing models given that the absolute value of the average mispricing error is decreased throughout different model specifications. But more importantly, estimating a 3-year factor loading on efficiency scores can help predict future mutual fund's performance.

Ultimately, following the specification of the efficiency score, we use the technique to differentiate between the performance of an unconstrained universe and that of a constrained subsample. Although intuition will suggest the constrained investment, by its own nature, would underperform unconstrained investments, consistent evidence has not been provided in the literature. This paper sheds light into this debate showing that, by using efficiency, the constrained investment does underperform.

Appendix

A Description of Ethical Investing

According to the Social Investment Forum in the U.S. (SIF), over the past 20 years the total dollars invested in SRIs has grown exponentially, as has the number of institutional, professional, and individual investors involved in the field. Between 1995 and 2010, total dollars under professional management in SRI grew from \$639 billion to \$3.07 trillion, outpacing the overall market. As of 2015, it has further increased to 6.57 trillion dollars.¹³ SRI investing has become a mainstream asset, and as a result, a number of conventional companies now offer SRI products to their clients. An increasing number of investors focus on ethical funds not only because of the ethical aspect, but because returns are comparable to those of more conventional investments.¹⁴

Managers of ethical funds have to follow socially responsible constraints on environmental risk, social risk, and governance risk. For example, they would not buy shares from the weapons, gambling, alcohol, or tobacco industries; nor would they buy shares from companies known for polluting the environment, using "sweatshops," discriminating against their employees, etc. On the other hand, these managers would increase their number of shares of ethical companies, philanthropic institutions, and firms well known for their social activities.¹⁵ This is the social screening process. Bar-

¹³ The Social Investment Forum defines Sustainable and Responsible Investing as an investment discipline that considers environmental, social and corporate governance (ESG) criteria to generate long-term competitive financial returns and positive societal impact. (<http://www.ussif.org/>).

¹⁴ Abstract taken from SIF's Performance and SRI investments page: <http://ussif.org/resources/performance.cfm>; the full report on Sustainable and Responsible Investment trends can be downloaded from: <http://ussif.org/resources/research/documents/2010TrendsES.pdf>.

¹⁵ See Basso and Funari (2008), page 26 for further exemplification of social responsible screening.

nett and Salomon (2006) find that when the number of social screens (both positive and negative) increases, the fund's annual return initially declines, but then rebounds as the number of screens reaches a maximum. Furthermore, SRI's managers also have to analyze the companies' financial statements. An example can be seen in Goldreyer and Diltz (1999) who find that the performance of ethical investment also depends on the level of financial screening.

The question still remains whether or not ethical investment reduces profitability, thereby reducing efficiency. Asked another way, are the risk-adjusted returns of SRI portfolios higher than traditional investing? In response to this matter, Hamilton et al. (1993) proposed three alternative hypotheses: (1) the risk-adjusted expected returns of socially responsible portfolios are equal to the risk-adjusted expected returns of conventional portfolios, as the social responsibility feature of stocks is not priced; (2) the risk-adjusted expected returns of socially responsible portfolios are lower than the expected returns of conventional portfolios, as the market prices the social responsibility characteristic by increasing the value of socially responsible companies relative to the value of conventional companies by driving down the expected returns and the cost of capital of socially responsible companies; and (3) (also suggested by Moskowitz 1972), the risk-adjusted expected returns of socially responsible portfolios are higher than the expected returns of conventional portfolios, as the market prices social responsibility (incorrectly) in the case of "doing well while doing good."

Overall, the literature is divided on the exact consequences of investing in ethical mutual funds. For example, it has been shown that volatility in said market is lower than the volatility of conventional funds (Bollen 2007; Renneboog et al. 2005); but this does not include a discussion on the effects of reducing diversification. Investing in a closed universe is expected to limit the power of diversification and thus it can result in lower risk-adjusted returns. Yet consistent evidence of underperformance of SRIs has been missing.

Looking at the ethical aspect, traditional numerical indexes used to measure the performance of mutual funds do not take such features into consideration. Particularly, the Sharpe ratio (Sharpe 1966), the reward-to-half-variance index (Ang and Chua 1979), and the Treynor index (Treynor 1965) are computed as ratios between the expected excess return and a risk indicator without considering additional features (Basso and Funari 2003). Still, Goldreyer and Diltz (1999), compute Sharpe and Treynor ratios to show that social screening does not affect the investment performance of ethical mutual funds in any systematic way. Furthermore, Statman (2000) shows that the DSI index¹⁶ (which is one of the most well-known SRI indexes) has a higher Sharpe ratio than the S&P 500, which implies that investors seeking to optimize the mean-variance would prefer the ethical investment.

Believers in the efficient market hypothesis argue that it is impossible that SRI funds outperform their conventional peers (Renneboog et al. 2008). At most, prior research

¹⁶ Domini 400 Social Index: A market cap weighted stock index of 400 publicly traded companies that have met certain standards of social and environmental excellence. Potential candidates for this index will have positive records on issues such as employee and human relations, product safety, environmental safety, and corporate governance. Companies engaged in the business of alcohol, tobacco, firearms, gambling, nuclear power and military weapons are automatically excluded. http://www.investopedia.com/terms/d/domini_400.asp#ixzz26NyXf2Ft.

has found that socially responsible investing performs at least as good as traditional investment portfolios.

Hamilton and Statman (1993) and Statman (2000) use a sample of SRI funds and non-SRI funds for the periods 1981–1990 and 1990–1998, respectively, to estimate Jensen's alphas based on CAPM, using the DSI 400 index and the S&P500 index as the benchmark returns for SRI and non-SRI. In the end, the authors find no statistical difference between the Jensen's alphas, thus favoring the first hypothesis. These results are consistent with Diltz (1995), Guerard (1997) and Sauer (1997).

Later attempts included the use of Fama and French's three factor model as well as the inclusion of Carhart's momentum factor. Bauer et al. (2005) find no statistical difference in the performance of SRIs for the overall period 1990–2001, but they do find a "catching up" period in their early life. Their results are consistent with Renneboog, et al. (2008a) who reiterates the study for the period 1991–2003. Renneboog, et al. (2008a) also find said catching up period in the sub-sample 1991–1995. Surprisingly, though, the authors also find that ethical funds outperform their counterparts for the sub-period 2001–2003 (though it is not statistically significant).

Still using Carhart's four factor model, Gil-Bazo et al. (2010) compare the pre and after fees performance of ethical funds against non-ethical funds for the period 1997–2005. They find that ethical funds have better before and after fees performance when compared to non-ethical funds of the same characteristics.

Most of the literature, however, focuses on the social factor of ethical funds and how SRIs outperform conventional investing due to that factor. Yet, several authors (Kurtz and DiBartolomeo 1996; Guerard 1997; Kurtz 1997) report that the reason for the over performance of the DSI 400 is due to the large investing on growth stocks and not the social factor.

Geczy et al. (2005) show that investing only in ethical funds carried a significant financial cost for the period 1963–2001. Yet on the other hand, Bollen (2007) justifies the profitability of SRI given that investors see beyond the traditional risk-reward optimization problem, as they may possess a multi-attribute utility function which incorporates a set of personal and societal values.

References

- Ang, A., Chen, J., & Xing, Y. (2006). Downside risk. *Review of Financial Studies*, 19(4), 1191–1239.
- Ang, J. S., & Chua, J. H. (1979). Composite measures for the evaluation of investment performance. *Journal of Financial and Quantitative Analysis*, 14(02), 361–384.
- Arditti, F. D. (1975). Skewness and investors' decisions: a reply. *Journal of Financial and Quantitative Analysis*, 10(01), 173–176.
- Arditti, F. D., & Levy, H. (1975). Portfolio efficiency analysis in three moments: the multiperiod case. *The Journal of Finance*, 30(3), 797–809.
- Ball, C. A., & Torous, W. N. (1983). A simplified jump process for common stock returns. *Journal of Financial and Quantitative Analysis*, 18(1), 53–65.
- Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, 30(9), 1078–1092.
- Barnett, M. L., & Salomon, R. M. (2006). Beyond dichotomy: The curvilinear relationship between social responsibility and financial performance. *Strategic Management Journal*, 27(11), 1101–1122.
- Basso, A., & Funari, S. (2001). A data envelopment analysis approach to measure the mutual fund performance. *European Journal of Operational Research*, 135(3), 477–492.

- Basso, A., & Funari, S. (2003). Measuring the performance of ethical mutual funds: A DEA approach. *Journal of the Operational Research Society*, 54(5), 521–531.
- Basso, A., & Funari, S. (2005). A generalized performance attribution technique for mutual funds. *Central European Journal of Operations Research*, 13(1), 65.
- Basso, A., & Funari, S. (2008). DEA models for ethical and non ethical mutual funds. *Mathematical Methods in Economics and Finance*, 2(1), 21–40.
- Bauer, R., Koedijk, K., & Otten, R. (2005). International evidence on ethical mutual fund performance and investment style. *Journal of Banking & Finance*, 29(7), 1751–1767.
- Bera, A. K., & Jarque, C. M. (1981). Efficient tests for normality, homoscedasticity and serial independence of regression residuals: Monte Carlo evidence. *Economics Letters*, 7(4), 313–318.
- Blattberg, R. C., & Gonedes, N. J. (1974). A comparison of the stable and student distributions as statistical models for stock prices. *The Journal of Business*, 47(2), 244–280.
- Bollen, N. P. (2007). Mutual fund attributes and investor behavior. *Journal of Financial and Quantitative Analysis*, 42(3), 683. (Renneboog et al. 2005).
- Campbell, J. Y., & Hentschel, L. (1992). No news is good news: An asymmetric model of changing volatility in stock returns. *Journal of financial Economics*, 31(3), 281–318.
- Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, 52(1), 57–82.
- Chang, K. P. (2004). Evaluating mutual fund performance: An application of minimum convex input requirement set approach. *Computers & Operations Research*, 31(6), 929–940.
- Charnes, A., Cooper, W. W., & Rhodes, E., (1978, 1979). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2(6), 429–444.
- Choi, Y. K., & Murthi, B. P. S. (2001). Relative performance evaluation of mutual funds: A non-parametric approach. *Journal of Business Finance & Accounting*, 28(7–8), 853–876.
- Darling, G., Mukherjee, K., & Wilkens, K. (2004). CTA performance evaluation with data envelopment analysis. *Commodity trading advisors: Risk, performance analysis and selection*. Hoboken, NJ: Wiley.
- Diltz, D. J. (1995). The private cost of socially responsible investing. *Applied Financial Economics*, 5(2), 69–77.
- Eling, M. (2006). Performance measurement of hedge funds using data envelopment analysis. *Financial Markets and Portfolio Management*, 20(4), 442–471.
- Fama, E. F. (1965). The behavior of stock-market prices. *The Journal of Business*, 38(1), 34–105.
- Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *The Journal of Finance*, 47(2), 427–465.
- Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1), 3–56.
- Fama, E. F., & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *The Journal of Political Economy*, 81, 607–636.
- Galagedera, D. U., & Silvapulle, P. (2002). Australian mutual fund performance appraisal using data envelopment analysis. *Managerial Finance*, 28(9), 60–73.
- Geczy, C., Stambaugh, R., & Levin, D. (2005). *Investing in socially responsible mutual funds*. Available at SSRN 416380.
- Gil-Bazo, J., Ruiz-Verdú, P., & Santos, A. A. (2010). The performance of socially responsible mutual funds: The role of fees and management companies. *Journal of Business Ethics*, 94(2), 243–263.
- Glawischnig, M., & Sommersguter-Reichmann, M. (2010). Assessing the performance of alternative investments using non-parametric efficiency measurement approaches: Is it convincing? *Journal of Banking & Finance*, 34(2), 295–303.
- Goldreyer, E. F., & Diltz, J. D. (1999). The performance of socially responsible mutual funds: Incorporating sociopolitical information in portfolio selection. *Managerial Finance*, 25(1), 23–36.
- Gray, J. B., & French, D. W. (1990). Empirical comparisons of distributional models for stock index returns. *Journal of Business Finance & Accounting*, 17(3), 451–459.
- Gregoriou, G. N. (2003). Performance appraisal of funds of hedge funds using data envelopment analysis. *Journal of Wealth Management*, 5(4), 88–95.
- Gregoriou, G. N. (2006). Trading efficiency of commodity trading advisors using data envelopment analysis. *Derivatives Use, Trading Regulation*, 12(1), 102–114.
- Gregoriou, G. N., & Chen, Y. (2006). Evaluation of commodity trading advisors using fixed and variable and benchmark models. *Annals of Operations Research*, 145(1), 183–200.

- Gregoriou, G. N., & McCarthy, K. (2005). *Efficiency of funds of hedge funds: A data envelopment analysis approach. Hedge Funds: Insights in Performance Measurement, Risk Analysis, and Portfolio Allocation*. New Jersey: Wiley.
- Gregoriou, G. N., Rouah, F., Satchell, S., & Diz, F. (2005). Simple and cross efficiency of CTAs using data envelopment analysis. *The European Journal of Finance*, 11(5), 393–409.
- Guerard, J. B., Jr. (1997). Additional evidence on the cost of being socially responsible in investing. *The Journal of Investing*, 6(4), 31–36.
- Hamilton, S., Jo, H., & Statman, M. (1993). Doing well while doing good? The investment performance of socially responsible mutual funds. *Financial Analysts Journal*, 49, 62–66.
- Hanoch, G., & Levy, H. (1969). The efficiency analysis of choices involving risk. *The Review of Economic Studies*, 36(3), 335–346.
- Haslem, J., & Scheraga, C. (2003). Data envelopment analysis of Morningstar's large-cap mutual funds. *Journal of Investing*, 12(4), 41–48.
- Ho, Y. K., & Cheung, Y. L. (1991). Behaviour of intra-daily stock return on an Asian emerging market-Hong Kong 1. *Applied Economics*, 23(5), 957–966.
- Jarque, C. M., & Bera, A. K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*, 6(3), 255–259.
- Jarque, C. M., & Bera, A. K. (1987). A test for normality of observations and regression residuals. *International Statistical Review/Revue Internationale de Statistique*, 55, 163–172.
- Jensen, M. C. (1968). The performance of mutual funds in the period 1945–1964. *The Journal of Finance*, 23(2), 389–416.
- Joro, T., & Na, P. (2006). Portfolio performance evaluation in a mean-variance-skewness framework. *European Journal of Operational Research*, 175(1), 446–461.
- Kane, A. (1982). Skewness preference and portfolio choice. *Journal of Financial and Quantitative Analysis*, 17(01), 15–25.
- Kraus, A., & Litzenberger, R. H. (1976). Skewness preference and the valuation of risk assets. *The Journal of Finance*, 31(4), 1085–1100.
- Kon, S. J. (1984). Models of stock returns—A comparison. *The Journal of Finance*, 39(1), 147–165.
- Kurtz, L. (1997). No effect, or no net effect? Studies on socially responsible investing. *The Journal of Investing*, 6(4), 37–49.
- Kurtz, L., & DiBartolomeo, D. (1996). Socially screened portfolios: An attribution analysis of relative performance. *The Journal of Investing*, 5(3), 35–41.
- Lau, A. H. L., Lau, H. S., & Wingender, J. R. (1990). The distribution of stock returns: New evidence against the stable model. *Journal of Business & Economic Statistics*, 8(2), 217–223.
- Leland, H. E. (1999). Beyond mean-variance: Performance measurement in a nonsymmetrical world. *Financial analysts journal*, 27–36.
- Lim, S., Oh, K. W., & Zhu, J. (2014). Use of DEA cross-efficiency evaluation in portfolio selection: An application to Korean stock market. *European Journal of Operational Research*, 236(1), 361–368.
- Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics*, 47(1), 13–37.
- Mandelbrot, B. (1963). The stable Paretian income distribution when the apparent exponent is near two. *International Economic Review*, 4(1), 111–115.
- Markowitz, H. (1952). Portfolio selection*. *The Journal of Finance*, 7(1), 77–91.
- Markowitz, H. (1959). *Portfolio selection: Efficient diversification of investments (No. 16)*. New Haven: Yale University Press.
- McMullen, P. R., & Strong, R. A. (1998). Selection of mutual funds using data envelopment analysis. *Journal of Business and Economic Studies*, 4(1), 1–12.
- Morey, M. R., & Morey, R. C. (1999). Mutual fund performance appraisals: A multi-horizon perspective with endogenous benchmarking. *Omega*, 27(2), 241–258.
- Moskowitz, M. (1972). Choosing socially responsible stocks. *Business and Society Review*, 1(1), 71–75.
- Murthi, B. P. S., Choi, Y. K., & Desai, P. (1997). Efficiency of mutual funds and portfolio performance measurement: A non-parametric approach. *European Journal of Operational Research*, 98(2), 408–418.
- Praetz, P. D. (1972). The distribution of share price changes. *Journal of Business*, 45, 49–55.
- Press, S. J. (1967). A compound events model for security prices. *Journal of Business*, 40, 317–335.
- Renneboog, L., Ter Horst, J., & Zhang, C. (2008). The price of ethics and stakeholder governance: The performance of socially responsible mutual funds. *Journal of Corporate Finance*, 14(3), 302–322.

- 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.
- Sharpe, W. F. (1966). Mutual fund performance. *The Journal of Business*, 39(1), 119–138.
- Simkowitz, M. A., & Beedles, W. L. (1980). Asymmetric stable distributed security returns. *Journal of the American Statistical Association*, 75(370), 306–312.
- Smith, J. B. (1981). *The probability distribution of market returns: A logistic hypothesis*. Doctoral dissertation, Graduate School of Business, University of Utah.
- So, J. C. (1987). The distribution of foreign exchange price changes: Trading day effects and risk measurement-A comment. *The Journal of Finance*, 42(1), 181–188.
- Statman, M. (2000). Socially responsible mutual funds. *Financial Analysts Journal*, 56, 30–39.
- Tobin, J. (1958). Liquidity preference as behavior towards risk. *The review of economic studies*, 25(2), 65–86.
- Treynor, J. L. (1965). How to rate management of investment funds. *Harvard Business Review*, 43, 63–75.
- Turner, A. L., & Weigel, E. J. (1992). Daily stock market volatility: 1928–1989. *Management Science*, 38(11), 1586–1609.