

CIRPÉE

Centre interuniversitaire sur le risque, les politiques économiques et l'emploi

Cahier de recherche/Working Paper **05-01**

The Impact of Ethical Ratings on Canadian Security Performance: Portfolio Management and Corporate Governance Implications

Klaus Fischer
Nabil Khoury

Février/February 2005

Fischer: Professor of Finance Laval University

Khoury: Professor Emeritus, Desjardins Chair in the management of derivative products, University of Quebec in Montreal. *Corresponding author*: Nabil Khoury, 1637 Des Rocs Street, Sainte-Foy (Quebec), Canada G1W 3J7
Tel. (418) 653-0177, (514) 987-3000 (ext. 6717)

nabilkhoury@sympatico.ca

Financial support from the NATCAN and from FCAR are gratefully acknowledged. Security screening data were graciously provided by the Canadian Social Investment Database. The authors wish to thank Pierre-Marie Rousseau, Marie-Diane Deslauriers and Martin Desrochers for their able assistance. They also thank G. Charest, B. M'Zali, J.-M. Gagnon and participants at the doctoral seminar at Laval University for their insightful comments.
Type of software used: Microsoft Word

Abstract: One approach that is gaining in popularity among portfolio managers uses ethical ratings, published by specialized research organizations, to screen securities for portfolio selection. Portfolio managers can thus gain a better understanding of the phenomenon and adopt a better and more consistent approach to ethical investment. By the same token, board of directors can measure the impact of their ethical policies on the market performance of the stock of their company. This paper provides new evidence about the impact of ethical ratings published in Canada on the risk-adjusted returns of the securities concerned, within the framework of a multi-factor Capital Asset Pricing Model, and gives an interpretation of the results from the perspective of portfolio composition and of corporate governance.

Keywords: Ethical Ratings and Security Performance

Résumé: La philosophie de l'investissement socialement responsable en valeurs mobilières s'est répandue à un rythme accéléré en Amérique du Nord au cours des dernières années. La sélection des titres tient compte alors du comportement socialement responsable des entreprises, ce qui se fait de plus en plus à l'aide de critères de notation éthique. En se basant sur la version multi-facteur du modèle d'évaluation des actifs financiers à l'équilibre, cette étude examine l'impact du nombre d'infractions éthiques sur le rendement des titres ainsi que l'impact du ciblage éthique sur la performance des portefeuilles. Les résultats des tests empiriques sont interprétés dans une perspective de gestion de portefeuille et de gouvernance d'entreprise.

JEL Classification: G11

**THE IMPACT OF ETHICAL RATINGS ON CANADIAN SECURITY
PERFORMANCE: PORTFOLIO MANAGEMENT AND CORPORATE GOVERNANCE
IMPLICATIONS**

Ethical screening of securities for portfolio composition has received considerable attention in recent years, particularly from institutional investors and corporate managers. The major preoccupation of investors, in this regard, is their ability to achieve the same or even a better risk/return tradeoff from portfolios restricted to securities of socially responsible companies as from portfolios without such constraints. For corporate managers, the question is whether ethical screening will affect the market evaluation of the securities of their companies. Either way, this new attention to the ethical behavior of companies can have an impact on business policies and practices particularly as regards the choice of operating procedures, products mix, sector diversification, and the relationship with stakeholders. It can also have an impact on the extent to which investors may make ethical concerns a corporate governance issue in an effort to protect the value of their investment.

Various approaches can be followed by investors in order to construct a universe of investment opportunities on an ethical basis. One approach that is gaining in popularity among portfolio managers uses ethical ratings, published by specialized research organizations, to screen securities for portfolio selection. These ratings are usually built around a number of criteria pertaining to ethical behaviour and lines of business. As more investors adopt this framework to define their universe of eligible securities, it becomes important to clarify the range of criteria included in these ratings, and to measure their impact on security risk-adjusted returns. By analyzing the influence of these criteria, portfolio

managers can gain a better understanding of the phenomenon and adopt a better and more consistent approach to ethical investment. By the same token, board of directors can measure the impact of their ethical policies on the market performance of the stock of their company. As seen in this way, investors' perceptions of ethical business behaviour becomes a corporate governance issue. Ignoring ethical concerns may result in discount on the price of stock, as investors defect to firms that display higher ethical standards. While adopting strict ethical standards may increase operating costs, if they are priced by the market ignoring them could result in a loss of wealth to shareholders that may exceed those costs. This could result in ethical issues becoming a concern for corporate governance that may be impossible to ignore by firm managers and directors without penalty, since investors concerned about their wealth are likely to increase the pressure on directors to adopt sound ethical standards to forego discount by the market.

So far, no study has addressed this issue in a market equilibrium framework. To be sure, a number of event studies have examined the impact of specific ethical news or specific ethical policies on security returns. For example, Gunthorpe (1997) examined whether the financial markets penalize public corporations at the announcement of unethical business behavior. His study uses a sample of 69 US corporations that were involved in some form of alleged illegal activity over the period 1988-1992. He uses the Market Model to estimate for each firm the average and cumulative average abnormal return five days before and five days after the announcement of the alleged unethical event. His results show that public announcements of unethical behavior carry a statistically significant penalty for stock prices of the firms concerned of about 1.3% over a day and 2.3% over a week. Klassen and McLaughlin (1996) also used the Market Model

approach and found a significant relationship between positive and negative environmental event announcements affecting companies and abnormal returns for their stocks. Their study covers a sample of 96 publicly traded US firms for which a positive public announcement was made in regard to strong environmental performance, and 16 firms for which a negative public announcement was made thus signaling weak environmental performance over the period 1985-1991. They report a positive cumulative average abnormal return of 0.82% following positive environmental events and a negative cumulative average abnormal return of 1.5% following negative environmental events. In the same vein, Feldman, Soyka and Ameer (1997) using the CAPM methodology found that improving both the environmental management system and environmental performance can reduce firms' perceived risk in the market place and increase their stock price by as much as 5%.

Other studies have also examined the impact of ethical constraints from the perspective of portfolio performance. Thus, Guerard (1997) compared the average monthly return of 950 stocks of companies designated as socially responsible with that of 1300 stocks of companies not having that designation for the period 1987-1994. He also conducted a multifactor regression analysis of the total return of each stock on a number of accounting variables, on a quarterly basis, in order to rank stocks in terms of their expected returns, and form portfolios on this basis. His results show that by and large no significant difference exists in average monthly returns between portfolios of socially screened and unscreened stocks. He also shows that a composite model using both value and growth components produces no significant differences in stock selection modeling in screened and unscreened universes. In the wake of the divestment movement of South African equities in the 1970's, Rudd (1979)

looked at the impact of such exclusions on portfolio risk. He began by excluding from the S&P 500 Index the 116 companies listed by the Investor Responsibility Research Center as having links with South Africa. He then optimized the remaining securities to form a portfolio that matched the S&P 500 Index as closely as possible for September 1978. His results show that this optimal portfolio although very well diversified with an R^2 of .989 has an annual residual standard deviation of 2.21% in comparison to a residual standard deviation of zero for a pure index fund and 1.5% for a typical marketed index fund. In the same vein, Wagner, Emkin and Dixon (1984) examined the possibility of replacing 152 companies of the S&P 500 index operating in South Africa with the largest “unrestricted” companies of their respective industries and compared this modified portfolio with the original index from the first quarter of 1979 to the first quarter of 1984. They concluded that the modified portfolio, although very well diversified (R^2 of 0.968) is riskier than the index (Beta of 1.08) and entails additional costs of research, trading and administration. Kahn, Lekander and Leimkuhler (1997) examined the investment implications of a forced divestiture of the tobacco holdings of pension funds for passive as well as for active managers. By using the S&P 500 as benchmark for passive fund managers, the study shows that removing tobacco stocks from the index reduces its performance by 21 basis points per year, over the period 1987-1996, with a negligible effect on risk. As for active managers, the study indicates that the tobacco divestiture decision reduces their potential for outperformance by reducing the number of stocks in the opportunity set, and creates measurement problems associated with imperfect benchmarks. The study also shows that a strategy of optimally weighting the tobacco-free portfolio to match the beta and other risk characteristics of the S&P 500 as closely as possible, would result in higher transaction costs due to higher turnover. On the other hand, Diltz (1995)

looked at the impact of various ethical screens on portfolio performance for the period 1989-1991 and concluded that eleven ethical screens and combinations of them had no significant effect on performance. He also found that the good environmental behavior screen and the nuclear and military exclusion screen had positive impact on performance, whereas the provision of family-related benefits screen had a negative impact.

Attempts were also made to study the relationship between ethical behavior and financial performance at the company level. Thus, Waddock and Graves (2000) find that companies that successfully pass a social screen and those that don't perform about the same financially. Verschoor and Murphy (2002), on the other hand, report that the 100 companies included in the S&P 500 index which qualified for the "Business Ethics Best Citizen" title awarded by Business Week in 2001 ranked higher by ten percentile points on average than the mean ranking of the remaining companies in this index in terms of overall financial performance as measured by Business Week on the basis of eight financial ratios. A range of studies, such as those by Hart and Ahuja (1996) and Russo and Fouts (1997) draw on the resource-based theory of the firm to evaluate the impact of environmental policies and show that environmental performance and economic performance are positively linked¹. In all cases, the results suggest a positive association between ethical behavior and profitability although it should be noted that causality has not been demonstrated by these tests.

In this paper we provide new evidence about the impact of ethical ratings published in Canada on the risk-adjusted returns of the securities concerned, within the framework of a multi-factor Capital Asset Pricing Model, and give an interpretation of the results from the perspective of portfolio composition and that

of corporate governance. The next section describes the data set and presents the methodology used. The third section discusses the statistical estimation procedures used in the tests and gives the empirical results and their interpretation. The fourth section concludes the paper.

Data and methodology

The initial sample in this study includes 448 Canadian companies over the period 1997-2000. Ethical scoring of these companies come from the Canadian Social Investment Database (CSID) and reflects the ratings given to each company in regard to qualitative as well as exclusionary screens. The first category of screens encompasses seven criteria pertaining to community, diversity, employee relations, environment, international operations, product and practices considerations, and corporate governance. Each of these criteria is sub-divided into a number of sub-criteria² for which a standard of ethical conduct is defined. The sub-criteria that are associated with positive ethical business conduct result in strength scores, while the sub-criteria associated with unethical behavior result in concern scores. Each individual score takes the value of 0 or 1 depending upon whether the standard of the relevant sub-criterion is met or not. The second category of screens includes five criteria pertaining to activities related to alcohol, gambling, tobacco, military contracting and nuclear power. Each of these five criteria is sub-divided into two sub-criteria only, reflecting either a minor involvement (with a score of 1) or a major involvement (with a score of 2) in each one of these lines of businesses.

Ratings are assigned to companies with respect to the qualitative screens, in terms of strengths and concerns scores assigned on the basis of whether or not the

companies meet the standard of each strength and concern sub-criterion defined for each criterion. As for exclusionary screens, their ratings are only in terms of major or minor concerns, since they are intended to screen out companies engaged in specific lines of businesses. Company ratings are updated by groups of companies every quarter, in order to have the whole sample reviewed over the course of each year.

Monthly stock returns come from the TSE-Western tape (CFMRC) and accounting information is collected from the Stockguide tape.

The first part of our analysis seeks to estimate stock performance relative to ethical scores of strengths, concerns and exclusions in order to establish the relative impact of the criteria responsible for these scores on returns. Following Fama and French (1993), we estimate stock performance relative to the criteria in ethical ratings using the following model:

$$R_{it} - R_{Ft} = \alpha + \beta_{1it} (R_{Mt} - R_{Ft}) + \beta_{2it} (S_t - B_t) + \beta_{3it} (H_t - L_t) + \sum_{j=1}^n \omega_{jit} (S_{jit}) + \sum_{h=1}^n \chi_{hit} (C_{hit}) + \sum_{k=1}^n \nu_{kit} (E_{kit}) + \varepsilon_{it} \quad (1-a)$$

where:

$(R_{it} - R_{Ft})$ is the difference between the return on stock *i* and the return on a 90 day Canadian Treasury Bill for each year *t*;

$(R_{Mt} - R_{Ft})$ is the difference between the return on the Toronto Stock Exchange 300 Index (capped)³ and that on a 90 day Canadian Treasury Bill for year *t*;

$(S_t - B_t)$ is the difference between the return on a portfolio of small capitalization stocks and that on a portfolio of large capitalization stocks for year t ;

$(H_t - L_t)$ is the difference between the return on a portfolio of stocks with a high ratio of Book Value / Market Value and that of a portfolio of stocks with a low such ratio for year t ;

S_{jit} are the strength scores for the various sub-criteria (j) reported for stock i in year t with respect to the criteria of the qualitative screens;

C_{hit} are the concern scores for the various sub-criteria (h) reported for stock i in year t with respect to the criteria of the qualitative screens;

E_{kit} are the scores for the various sub-criteria (k) reported for stock i in year t with respect to the exclusionary screens;

Within the limits of the data available, the risk factors $(S_t - B_t)$ and $(H_t - L_t)$ were estimated from six different portfolios comprising 380 stocks in 1997, 377 stocks in 1998, 374 stocks in 1999 and 387 stocks in 2000. The six portfolios were organized as in Fama and French (1993) with three portfolios of small capitalization stocks having a high (SH) an average (SM) and a low (SL) ratio of Book Value/Market Value, and three other portfolios of large capitalization stocks with a high (BH) an average (BM) and a low (BL) ratio of Book Value / Market Value. The factor $(S_t - B_t)$ represents the difference between the aggregate return of SH, SM and SL portfolios and that of BH, BM and BL portfolios for each month t , whereas the factor $(H_t - L_t)$ represents the difference between the aggregate return of BH and SH portfolios and that of BL and SL portfolios. The correlation coefficient between these two factors for the period under study is -0.0225 which indicates their orthogonality.

It is also interesting, in this regard, to examine the impact of the total number of strengths, concerns and/or exclusions pertaining to each stock on its performance. To perform this test a modified version of model (1-a) is used as follows:

$$R_{it} - R_{Ft} = \alpha + \beta_{1it} (R_{Mt} - R_{Ft}) + \beta_{2it} (S_t - B_t) + \beta_{3it} (H_t - L_t) + \omega_{it} (SS_{it}) + \chi_{it} (CC_{it}) + \nu_{it} (EE_{it}) + \varepsilon_{it} \quad (1-b)$$

where the SS_{it} , CC_{it} and EE_{it} are the total number of strengths, concerns and exclusions scores affecting stock i in year t .

The second part of our analysis seeks to determine the total number of concern scores beyond which portfolio performance is adversely affected. The analysis is limited to the number of concerns as they represent the decision variable in portfolio composition that determines the degree to which a portfolio of non-excluded stocks can be ethical without sacrificing performance⁴. In this regard, portfolio performance is estimated using the following model:

$$R_{it} - R_{Ft} = \alpha + \beta_{1it} (R_{Mt} - R_{Ft}) + \beta_{2it} (S_t - B_t) + \beta_{3it} (H_t - L_t) + \omega_{1it} (D_1) + \omega_{2it} (D_2) + \omega_{3it} (D_3) + \varepsilon_{it} \quad (2)$$

where:

$R_{it} - R_{ft}$ is the difference between the return of portfolio i and that of a 90 Day Canadian Treasury Bill for each month t . The other risk factors are as defined before, except that they are calculated on a monthly basis.

Four different portfolios were composed for this test with stocks exhibiting either zero, one, two or three concerns. The dummy variables D_1 , D_2 and D_3 take on the value of 0 or 1 with respect to these portfolios as follows:

	D_1	D_2	D_3
Portfolio of stocks with a score of zero concerns	0	0	0
Portfolio of stocks with a score of one concern	1	0	0
Portfolio of stocks with a score of two concerns	0	1	0
Portfolio of stocks with a score of three and more concerns	0	0	1

The four portfolios in this test are equally weighted, and are rebalanced on a yearly basis in order to keep their basic features⁵ constant.

Before testing, the data set was rearranged to accommodate each of the models to be estimated. For Model (1-a) the data set includes all the 448 companies for which complete ratings are available. To estimate this model, annual observations of excess returns, the three risk factors and the annual scores with respect to strengths (S_{jit}), concerns (C_{hit}) and exclusions (E_{kit}) are attributed to each company. For model (1-b) a second very similar data set was created with returns and the risk factors still calculated on a yearly basis, but with strengths, concerns and exclusions scores aggregated for each company in each year. To estimate model (2), a data base was constituted with monthly observations of returns and risk factors, and with a binary variable D_1 to D_3 attached to each portfolio. The portfolio with zero concerns (D_0) represents the intercept. As a result of this data organization, we end up with all data sets in panel format. For Models (1-a) and (1-b), the data set has $i=1, \dots, 448$ “individuals” and $t=1, \dots, 4$ giving a total of 1792 observations. However, due to the fact that the rating

agency has expanded its sample gradually over the period under study, we end up with a total of 1420 usable observations. For Model (2), the data set consists of $i=1, \dots, 4$ and $t=1, \dots, 48$, giving a total of 192 observations.

It should be mentioned that the relatively small number of monthly observations and the fact that the data takes a panel format precludes the use of GARCH procedures. Instead, we will control for heteroskedasticity by other means suitable for panel data.

Empirical Results

Summary statistics of the data used in models (1-a) and (1-b) are presented in Table 1. The table reveals, among other things, the high variability of returns of individual stocks (expressed in decimals), compared to that of the market. As regards ratings, it is interesting to note first that the maximum observed strength scores *vis-à-vis* a criterion is 4, implying that for this particular criterion (diversity) a company has exhibited strength on four related sub-criteria. The maximum cumulative strength scores obtained by a company with respect to all criteria in this category of screens is 10. As regards concerns, the maximum observed scores obtained for a criterion is 2 and the maximum total obtained by a company with respect to all criteria combined in this category of screens is 6. In the case of exclusions, the maximum number of concern scores obtained for each sub-criterion is of course 2, signifying a major concern, and the maximum total of all scores in regard of this category of screens is 3.

Given the panel nature of the data, a number of precautions must be taken in the estimation procedure. More specifically, OLS can be a suitable estimation

approach only under very specific conditions that are rarely found with this type of data. Thus, with respect to models (1-a and 1-b), we proceeded first by estimating the models by means of OLS to recover the residuals. These residuals were then subjected to an ANOVA that evaluates the potential presence of individual (firm) effects and time effects. The presence of one or both of these effects could bias the covariance matrix of parameter estimates. This would be the case if unobservable individual effects, or if time effects are correlated to the explicitly stated variables (risk or rating), that is if either $E[x|\mu_i] \neq 0$ or $E[x|\tau_t] \neq 0$, where μ_i and τ_t are individual or time unobservable effects respectively. The presence of individual effects would mean that specific firms returns for all periods may be explained by factors other than the risk factors and the scores. On the other hand, the presence of time effects would mean that there are time-varying macro factors not accounted for explicitly by the risk factors and the rating scores that explain variations of returns of all firms. The ANOVA on the residuals would provide information on whether such effects are present in the data or not.

The presence of either or both of the above mentioned effects requires an adaptation of the estimation methodology. Assuming the absence of endogeneity (e.g. that past return influences future rating) we can choose between two relatively straightforward procedures: estimation via “fixed effects” (FE) or estimation via “random effects” (RE). The choice of the appropriate procedure can be decided upon by a Hausman specification test, in which the vectors of coefficients estimated by both procedures are compared.⁶ If the equality of vectors coefficients is rejected, it would be a clear indication that the RE estimation procedure, while more efficient, yields biased parameters. In this instance, the use of an FE based estimation procedure becomes mandatory.

However, if the equality of vector coefficients is not rejected, it would suggest that either procedure will generate unbiased estimates, but RE would provide in this case, more efficient estimates of the covariance matrix. Once we have established the right estimation procedure, we go on to present the results of the tests, using either the FE or the RE procedure, and to perform a Wald test of exclusion on the variables of interest.

Given the somewhat different nature of the dataset used to estimate Model 2, we will be forced to introduce some variations to the above procedure that will be explained later in the paper.

Table 2 shows the results of the ANOVA for the residuals of Model (1-a) estimated by OLS.

The F-statistics obtained in the ANOVA suggest that we need to take into consideration the existence of both individual and time effects. In a more economic sense, this test implies that the variations of returns are not fully accounted for by the variables explicitly included in the model.

When we perform the Hausman test that compares vectors coefficients obtained with the FE and RE procedures, the Chi-Square stands at 228.127 with a significance level of 0.0000, suggesting the inequality of coefficients obtained with both procedures. Estimation by FE thus becomes mandatory. The result of estimating Model (1-a) by FE are presented in Table 3.

Table 3 reveals the near absence of any significant coefficient. The only exception is the coefficient for S7 (the corporate governance strength criterion)

with a significance level of 2.5% and the right sign. While corporate governance strength is no doubt an important factor that may cause this result, its high significance may also be due to a type II error. The coefficients of the risk factors are also non-significant, and all Chi Square values for the Wald joint exclusion test for groups of coefficients (all Ss, Cs, and Es respectively) are quite small and non-significant. This suggests that individually taken, almost no single rating score, be it a *strength*, a *concern* or an *exclusion*, appears to affect stock returns. In other words, investors appear to be reasonably tolerant in regard to individual scores of ethical issues, and do not seem to penalize or reward stocks for one or the other of these scores. The fact that the Wald test of joint exclusion is non-significant does not contradict this conclusion since this test does not measure a cumulative effect.

With respect to Model (1-b), the ANOVA for the residuals estimated by OLS yield the results shown in Table 4.

This ANOVA suggests that in the case of this series there is no time nor individual effects that may invalidate the OLS estimates. In other words, estimation by FE or RE should yield consistent results with those of OLS. The results of estimating Model (1-b) by OLS are reported in Table 5.

In this regression, as in that for Model (1-a), no particular risk factor or aggregate rating score stands out as significant in explaining stock returns. The only significant exception is the coefficient related to aggregate concerns, which has a level of significance of 4% and the expected sign. This means that when strength scores, concern scores and exclusion scores are aggregated separately for each company and regressed on stock returns, the only significant variable that shows

in the results turns out to be total concerns, and its impact on returns is, as expected, negative. Thus, even though investors appear to be indifferent to the nature of individual concerns, strength or exclusions, they seem to pay attention to the level of aggregate ethical concerns. We will have the opportunity to validate these results with those obtained from Model (2) below. It is also interesting to mention, in this regard, that estimations by RE yield exactly the same results as those obtained by OLS.

Turning now to Model (2), we present in Table 6 some summary statistics about the data. The table is organized in terms of portfolios with the number of concerns increasing from one to three or more.

In estimating Model (2) we are implicitly using an FE approach where individual effects are controlled thru the dummies (see Hsiao (1986)). However, this estimation procedure does not control for potential unobservable time effects. While risk factors should account for most time variations that affect all individuals, there is no guarantee that this, in fact, is the case.⁷ Ignoring other non-observable time effects could thus bias coefficients, since we would be assuming that $E[\mathbf{x}|\tau_t] = 0$ when this is not the case.

To test whether residual time effects are present despite the use of the risk factors we estimate Model (2) using an OLS regression and then perform an ANOVA on the residuals in order to test whether unaccounted time effects are still present. The results of the ANOVA are shown in Table 7.

We have, of course, ignored individual effects in estimating the ANOVA because these have already been eliminated thru the dummies. The F-statistic suggests

unambiguously the presence of time effects not accounted for by the risk factors. The t-statistics obtained from the OLS estimates may thus be biased⁸. To correct for this bias we must use an adapted estimation procedure. Assuming the absence of endogeneity, the two procedures used for Models (1-a) and (1-b) can also be applied here, namely the estimation by FE with correction for time effects and the estimation by RE also with correction for time effects. However, as noted before when discussing the estimation procedure for Models (1-a) and (1-b), the choice of the appropriate procedure depends upon the outcome of the Hausman test. When this test is performed for Model (2) we obtain a trivially small value of 3.44×10^{-28} , which is obviously non-significant. This result is not surprising considering the fact that in this dataset, in contrast to the one used to estimate Models (1-a) and (1-b), the value of $T=48$ is much larger than that of $N=4$. This means that the RE estimation procedure is the correct one. Table 8 presents the results of the estimation of Model (2) using RE.

Several important observations can be made from this table. In the first place, as the relatively high value of the adjusted R^2 shows, over 70% of the total variation in returns is explained by the model. Secondly, despite the fact that the risk factors and the concerns scores do not account for all the variations in returns, they are all highly significant. Most importantly, the number of concerns seems to affect returns in an increasing fashion. To start with, the portfolio with zero concerns (portfolio D_0) earns no abnormal return; its Jensen's alpha is not different from zero. However, the three other portfolios which have increasing number of concerns are penalized quite heavily, roughly in proportion to the level of concerns. Thus, the portfolio with stocks exhibiting one concern yields a return that is lower than that of the portfolio with zero concerns by 1.3% per month, at a confidence level of 8%. The portfolio of stocks with two concerns

yields a monthly return lower by 1.2% at a confidence level of 12%, than that of the portfolio with zero concerns. Finally, the portfolio with three and more concerns yields a full 2.0% per month less than the portfolio with zero concerns, at the 1% confidence level⁹. Furthermore, when we perform a Wald test of exclusion on the three dummies, we obtain a Chi-Square of 7.466 with a significance level of 0.0584. It should be remembered, in this regard, that the average monthly return of the market during the period under study was 1.33% with a standard deviation of 5.39%. It should also be noted that the correct interpretation of the coefficients for the dummies implies that the losses are not cumulative. It is also interesting to note that when model (2) is estimated with an FE procedure, the coefficients obtained are numerically similar to those in table 8 but at much lower significance levels.

The above results are consistent with those obtained from Model (1-b) where we observed that investors appear to pay attention to aggregate concerns rather than to the presence of a particular concern. In other words, investors do not seem to follow a simplistic yes/no exclusion screen in defining their universe of eligible securities. Rather, their decision process reflects personal choices regarding the level of ethical behaviour they wish to express in their portfolios. Their basic concern, therefore, does not lie with strength or exclusion scores, but with the overall level of concerns afflicting each company. More importantly, for the period under study, the portfolio of stocks with zero concerns outperformed the portfolios of stocks with either one, two or three or more concerns. This last observation can be of significance in the debate regarding the impact of ethical screens on portfolio performance.

The results obtained through the estimation of both models are thus consistent. On one hand, investors that ignore ethical concerns are likely to obtain statistically significant and sizable lower returns on their portfolios. On the other hand, company directors who ignore ethical issues in their decision making process are likely to inflict losses on their shareholders' wealth. The financial viability of ethical screening of securities, illustrated in the findings of this paper, also suggest that the issue of ethical rating is not likely to go away soon. In a market with these characteristics, corporations should pay attention to the level of aggregate ethical concerns that is signalled to the market concerning their activities and to continually clarify their ethical policies and practices to investors.

Conclusion

Increasingly, investors are stressing the importance of the ethical posture of corporations in choosing their universe of investment opportunities. The strategy that is gaining momentum in this regard is to screen stocks on the basis of ethical ratings published by specialized research organisations. These ratings relate to various aspects of the business behaviour of corporations as observed and evaluated by the research organisation with respect to an accepted standard.

The purpose of this study is to investigate the market response to the ethical business behaviour of Canadian publicly traded corporations, as signalled to the public thru ratings of ethical strengths concerns and exclusions. By examining whether or not stock returns are rewarded for positive ethical practices and policies and/or penalized for the negative ones, insight might be gained into two related questions, namely:

- (1) should ethical screening scores be considered in portfolio composition, and
- (2) is there a financial incentive to incorporate ethics into business culture and decision making.

The findings of this research clearly show that investors react only to the level of concern scores signalled to them by the scoring agency, without regard to the particular aspect of unethical behaviour covered by each concern. The analysis also shows that a portfolio of stocks with zero concerns outperforms portfolios comprising securities with one, two and three or more concerns. Furthermore, there is a significant decline in portfolio risk-adjusted returns as exposure to the number of concerns increases.

From the investor point of view, our research indicates that there is good reason for relying on the number of concern scores in screening securities for portfolio composition. The viability of this strategy in terms of portfolio performance indicates that such ethical screening is not about to go away soon.

From the point of view of corporate governance, the findings of this paper suggest that ignoring ethical concerns triggers a negative reaction from investors, as they defect to firms that display higher ethical standards. Thus, although the implementation of strict ethical standards can undoubtedly increase operating costs for corporations, ignoring them could result in considerable loss of wealth to shareholders that may well exceed those costs. These findings also suggest that ethical issues may become in the future a major consideration in corporate governance that cannot simply be overlooked by corporate directors without considerable penalty to the market value of their companies. Indeed, if the trend identified in this research persists, investors are likely to increase the pressure on directors for the implementation of sound ethical standards and practices in order to avoid the market penalties. Alternatively, corporate raiders may be able to extract rent by acquiring undervalued firms and devising ethical standards and policies for them. Corporations should therefore be proactive as regards the level of ethical concerns they wish to signal to the market, and to conduct their business policy accordingly.

REFERENCES

- Diltz J. (1995): "The Private Cost of Socially Responsible Investing", *Applied Financial Economics*, 5.
- Fama E.F. and K.R. French (1993): "Common Risk Factors in the Returns on Stocks and Bonds", *Journal of Financial Economics*, 33.
- Feldman S.J., P.A. Soyka and P.G. Ameer (1997): "Does Improving a Firm's Environmental Management System and Environmental Performance Result in a Higher Stock Price?", *Journal of Investing*, 6 (4).
- Guerard J.B. (1997): "Is There a Cost to Being Socially Responsible in Investing", *Journal of Investing*, 6 (2).
- Gunthorpe D.L. (1997): "A Quantitative Analysis of the Impact of Unethical Behavior by Publicly Traded Corporations", *Journal of Business Ethics*, 16 (5).
- Hart S.L. and G. Ahuja (1996): "Does it Pay to Be Green? An Empirical Examination of the Relationship Between Emission Reduction And Firm Performance", *Business Strategy & the Environment*, 5 (1).
- Hausman, J.A. (1978): "Specification tests in econometrics" *Econometrica*, 46.
- Hsiao, C. (1986): *Analysis of Panel Data* Cambridge: Cambridge University Press.
- Kahn R.N., C. Lekander and T. Leimkuhler (1997): "Just Say No? The Investment Implications of Tobacco Divestiture", *Journal of Investing*, 6 (4).
- Klassen R.D. and C.P. McLaughlin (1996): "The Impact of Environmental Management on Firm Performance", *Management Science*, 42 (8).
- Rudd A. (1979): "Divestment of South African Equities: How risky?", *The Journal of Portfolio Management*, 5 (3).
- Russo M.V. and P.A. Fouts (1997): "A Resource-Based Perspective on Corporate Environmental Performance and Profitability", *Academy of Management Journal*, 40 (3).

Vreschoor C.C. and E.A. Murphy (2002): “Best Corporate Citizens Have Better Financial Performance”, *Strategic Finance Magazine*.

Waddock S. and S.B. Graves (2000): “Performance Characteristics of Social and Traditional Investments”, *Journal of Investing*, 9 (2).

Wagner W.H., A. Emkin and R.L. Dixon (1984): “South African Divestment: The Investment Issues”, *Financial Analysts Journal*, 40 (6).

FOOTNOTES

¹ Russo and Fouts also show that returns to environmental performance are higher in high-growth industries. Their tests cover a sample of 243 firms over a period of two years.

² For the community criterion five strength and four concern sub-criteria are defined, while for the diversity criterion, six strength and three concern sub-criteria are enumerated. Employee relations criterion comprises six strength and five concern sub-criteria, whereas the environment criterion comprises six strength and eight concern sub-criteria. The international criterion encompasses five strength and seven concern sub-criteria, while product and practices criterion has three strength and five concern sub-criteria. Finally, the last criterion, identified as corporate governance, is made up of four strength and four concern sub-criteria

³ This version of the Index imposes a limit of 10% on the weight of any security's market value.

⁴ By definition, exclusion criteria are of a binary nature and affect a relatively small number of stocks.

⁵ The portfolio of stocks with zero concerns is made up of 29 securities in 1997, 31 securities in 1998, 38 securities in 1999 and 46 securities in 2000. The portfolio of stocks with one concern is made up of 55 securities in 1997, 55 securities in 1998, 56 securities in 1999 and 60 securities in 2000. The portfolio of stocks with two concerns includes 49 securities in 1997, 50 securities in 1998, 49 securities in 1999 and 41 securities in 2000. Finally, the portfolio of stocks with three concerns is made up of 24 securities in 1997, 23 securities in 1998, 19 securities in 1999 and 17 securities in 2000.

Although the focus of this test is on the number of concerns scores, it is worth mentioning that the distribution of strength scores within each of the four portfolios over the period under study is almost similar.

⁶ The Hausman specification test is typically applied when two estimation procedures may be used, one efficient but potentially biased and the other unbiased but inefficient. In our case RE is efficient but potentially biased, while FE is unbiased but inefficient. The Hausman test consists of computing the difference of the vector of coefficients and comparing it against the covariance matrix of the coefficients. The statistic is distributed Chi-squared. See Hausman (1978).

⁷ Indeed, the Fama and French factors may explain a considerable portion of this time variation. However other time related sources of variation may exist that are not proxied by these factors. One simple example would be the technology stock crash that could have an across-the-board impact on the market around certain dates. One simple way to control for unaccounted time variation would be to introduce time dummy variables. These dummies would capture the unexplained time variation. But this is an inefficient approach.

⁸ We can only speculate about the source of this additional variation in the data.

⁹ It is interesting to note that using an index of all securities listed on the Toronto Stock Exchange instead of the TSE 300 Index (capped) in this regression does not yield significantly different results.

Table 1
Summary Statistics for Stocks and Ratings (annual data)

Series	Mean	Std Error	Minimum	Maximum
RI - RF	0.2656	2.4500	-1.0302	62.3798
RM - RF	0.0769	0.2403	-0.2148	0.4246
SMALL - BIG	-0.2397	0.2045	-0.3788	0.1132
HIGH - LOW	-0.1921	0.4297	-0.8876	0.2303
S1: Community	0.2736	0.5117	0	2
S2: Diversity	0.3325	0.6194	0	4
S3: Employee relation	0.6109	0.7392	0	3
S4: Environment	0.1961	0.4664	0	2
S5: International operations	0.0185	0.1350	0	1
S6: Product & business practices	0.1856	0.4150	0	2
S7: Corporate governance	0.1937	0.4453	0	2
SS Total	1.2536	1.7180	0	10
C1: Community	0.0153	0.1229	0	1
C2: Diversity	0.3179	0.4676	0	2
C3: Employee relation	0.0742	0.2683	0	2
C4: Environment	0.2510	0.5216	0	2
C5: International operations	0.0451	0.2154	0	2
C6: Product & business practices	0.0532	0.2386	0	2
C7: Corporate governance	0.5504	0.6497	0	3
CC Total	0.9050	1.1199	0	6
E1: Alcohol	0.0161	0.1790	0	2
E2: Gambling	0.0080	0.1060	0	2
E3: Tobacco	0.0129	0.1551	0	2
E4: Military contracting	0.0548	0.2902	0	2
E5: Nuclear power	0.0347	0.2039	0	2
EE Total	0.8771	0.3894	0	3

Table 2
ANOVA for OLS residuals of Model (1-a)

Source	Sum of Squares	DF	Mean Square	F-Statistic	Signif Level
INDIV	2690.161	405	6.6423	3.598	0.0000
TIME	21.9439	3	7.3146	3.9622	0.0081
JOINT	2690.161	408	6.5935	3.5716	0.0000
ERROR	1336.584	724	1.8461		
TOTAL	4026.745	1132			

Table 3***Estimation of Model (1-a) Using a Fixed Effect Approach***

Variable	Coeff	Std Error	T-Stat	Signif
Constant	1.8140	16207.3442	0.0001	0.9999
β_1	-0.9442	16919.5780	-0.0001	1.0000
β_2	4.3851	48817.2302	0.0001	0.9999
β_3	1.5662	19208.5380	0.0001	0.9999
S1	-0.1354	0.1976	-0.6853	0.4934
S2	-0.0096	0.1724	-0.0559	0.9555
S3	0.0157	0.1347	0.1167	0.9071
S4	-0.2175	0.2330	-0.9335	0.3509
S5	-0.0258	0.8552	-0.0302	0.9759
S6	0.0728	0.2729	0.2668	0.7897
S7	0.6267	0.2798	2.2401	0.0254
C1	-0.1634	0.5090	-0.3211	0.7482
C2	-0.1429	0.2183	-0.6545	0.5131
C3	-0.1068	0.3615	-0.2954	0.7678
C4	0.1208	0.2426	0.4979	0.6187
C5	0.2725	0.4882	0.5581	0.5769
C6	0.1073	0.3818	0.2810	0.7788
C7	-0.2214	0.1545	-1.4334	0.1522
E1	-0.1092	1584.1356	-0.0001	0.9999
E2	-0.0781	0.7762	-0.1007	0.9199
E3	-0.2265	1494.9771	-0.0002	0.9999
E4	-0.1078	0.7165	-0.1505	0.8805
E5	0.1563	0.8235	0.1898	0.8495

*Adjusted R*² = 0.449 *F* = 2.9505 0.0000 *N* = 448; *T* = 4

Table 4
ANOVA for OLS residuals of Model (1-b)

Source	Sum of Squares	DF	Mean Square	F-Statistic	Signif Level
INDIV	2009.711	406	4.95002928	0.7829	0.998008
TIME	0.57949	3	0.1931663	0.0306	0.9928151
JOINT	2009.711	409	4.913721	0.7772	0.9985254
ERROR	6392.277	1011	6.32272794		
TOTAL	8401.98983	1420			

Table 5
Estimation of Model (1-b) Using OLS

Variable	Coeff	Std Error	T-Stat	Signif
Constant	1.0139	0.4860	2.0860	0.0372
β_1	0.4996	0.4966	1.0059	0.3146
β_2	1.9758	1.4268	1.3847	0.1664
β_3	0.7627	0.5598	1.3623	0.1733
ω	-0.0243	0.0388	-0.6265	0.5311
χ	-0.1249	0.0608	-2.0543	0.0401
ν	-0.0297	0.1570	-0.1889	0.8502

Adjusted $R^2 = 0.01$ *DW*=2.056 $F=3.4230$ $N=448; T=4$

Table 6

***Summary Statistics for Portfolios exhibiting one, two or three
and more concerns
(monthly data)***

Series	Mean	Std Error	Minimum	Maximum
$R_i - R_F$	0.0072	0.0676	-0.2349	0.2080
$R_M - R_F$	0.0065	0.0714	-0.2139	0.2363
$S-B$	-0.0214	0.0395	-0.1113	0.0904
$H-L$	-0.0016	0.0571	-0.2079	0.1184

Table 7

Analysis of Variance for the residuals of OLS estimate of Model (2)

Source	Sum of Squares	DF	Mean Square	F-Statistic	Signif Level
TIME	47.0	47 0	.00333	2.3277	0.0000699
ERROR	0.2061678	144 0	.00143		
TOTAL	0.3628001	191.0			

Table 8

Estimation of Model (2) using RE and controlling for time effects

Variable	Coeff	Std Error	T-Stat	Signif
Constant	0.00002	0.00689	0.0028	0.9977
β_1	0.93838	0.07665	12.241	0.0000
β_2	-0.63430	0.11987	-5.2913	0.0000
β_3	0.62211	0.08704	-7.1473	0.0000
ϖ_1	-0.01338	0.00772	-1.7327	0.0831
ϖ_2	-0.01200	0.00772	-1.5537	0.1202
ϖ_3	-0.02080	0.00772	-2.6941	0.0070

Adjusted R2=0.722

N=4; T=48