Stock market reaction to *changes* in environmental, social and governance performance (or "Doing well while doing good" revisited: doing better while doing more good?)

Cristiana Manescu, 15th March 2010 [1]

Very preliminary! Please do not cite without permission of the author

Abstract

Several empirical studies have shown that positive abnormal returns can be earned on investment strategies based on individual or aggregated environmental, social and governance (ESG) ratings, such as buying the high performers and selling the low performers (e.g. Kempf and Osthoff, 2007; Statman and Glushkov, 2008). These abnormal returns cannot be attributed to differences in exposures to the size, value or momentum risk factors. This paper aims at investigating in greater detail the relation between stock returns and ESG performance, by measuring the impact that changes in the ESG performance have on stock returns. The hypothesis is that if high ESG rating leads to high returns then higher ESG rating should be rewarded with a higher return, but not necessarily a one-for-one relation. A correction for potentially reversed feedback, i.e. from returns to ESG performance is also carefully addressed. The analysis is performed on the members of the SP500 and DS400 indexes for which ratings on seven ESG dimensions are provided by Kinder, Lydenberg and Domini Research & Analytics (KLD). The findings of this study will deepen the understanding on the intricate dynamics between ESG performance and financial performance, which is all the more needed given the recent increased demand for corporate ESG engagement.

[1] Center for Finance, School of Business, Economics and Law, University of Gothenburg, Box 600, Gothenburg, SE-40 530. Email: cristiana.manescu@cff.gu.se.

1. Introduction

The concept of responsible or sustainable investing has attracted a lot of attention both from the investor and the academic communities. While responsible investment might mean different things to different people, it generally refers to incorporating into the investment strategy some from a number of environmental, social and governance (ESG) issues.

In an attempt to answer this decade's hot question of whether responsible investing is profitable or financially harmful, the academic literature has put forward various theoretical hypotheses as well as empirical findings based on a multitude of empirical designs and ESG measurement data. This effort seems to indicate that responsible investing is most of the times profitable rather than harmful but also that the direction of causation might go both ways.

Many studies using Kinder, Lydenberg and Domini Research & Analytics (KLD) ratings of U.S. firms show that portfolios built on various KLD dimensions (as well as an aggregate KLD measure) earn abnormal returns over long periods of time even when accounting for the Carhart four risk factors (e.g. Kempf and Osthoff, 2007; Statman and Glushkov, 2008, etc.). However, when the analysis is performed at firm level, by analyzing the direct impact of KLD scores on either ex-ante cost of equity capital or the ex-post stock returns, mixed results are found with respect to certain ESG issues. While Derwall and Verwijmeren (2007) find a positive impact of a social index and a negative impact of product, corporate governance and environment indicators over 1992-2006, Manescu (2010) identifies a positive impact of community and employee relations indicator over June 1992-July 2003 and a negative impact of employee relations, environment and human rights over July 2003- June 2008. Hence the implication that portfolio analyses and cross-sectional analyses might reveal different aspects of the relationship between ESG performance and stock performance.

The empirical analyses based on other data providers provide similar findings. A portfolio of US stocks with high rating on an aggregate environment indicator evaluated by Innovest Strategic Value Advisors is found to earn substantially higher average returns than its low counterpart over the 1995-2003 period (Derwall et al., 2005). Using UK microdata on environmental, employment and community activities in 2002 provided by the Ethical Investment Research Service (EIRIS), Brammer et al. (2006) find that environmental and employment indicators are negatively correlated to returns while the community indicator is weakly positively related. von Arx and Ziegler

(2008) analyze US and European stock returns reaction to environment performance and a social index reflecting relations to several stakeholders as provided by Sarasin&Cie in 2002. Both dimensions are found to have a positive impact on stock returns but weaker in the European sample.

The reverse causality issue of ESG performance and financial performance is investigated to a lesser extent which is not surprising given the difficulty to properly correct for the endogenous nature of these two.(To be elaborated.)

This study aims at investigating in greater detail the direct impact of the KLD assessed ESG performance on stock returns by pursuing three main directions. It is worth exploring the KLD dataset into greater detail also due to the unique way in which it measures ESG performance with a clear distinction between ESG strengths and concerns.

First, we analyze the effect of changes in the KLD scores - rather than that of absolute scores - on abnormal stock returns; in so doing one can conclude that it is the dynamics within ESG performance which determine a reaction in returns and not the absolute standing in ESG performance which might be achieved once and for good. This way we intend to provide a complementary analysis to the prevailing positive documented relation between most KLD scores and stock returns. If high absolute KLD scores lead to high stock returns, then it is to be expected that an increase in KLD score should lead to an increase in stock returns as well; if not, the relation found between KLD scores and stock returns could be due to underlying characteristics of the firm which overlap the KLD score and which have not been properly taken into account in the analysis.

Second, we intend to evaluate the separate impact that strengths and concerns (in changes) might have on stock returns. This way we can investigate whether it is the good ESG performance that is rewarded in the market or the bad ESG performance that is penalized, an important question which affects the incentives of corporations. There are a handful of studies pointing out that the market reaction is stronger to concerns than to strengths (e.g. Galema et al 2008) which could partly be due to the documented fact that concerns offer better indication of future risks (Chatterji et al 2009).

Third, and potentially most importantly, we also want to address the reverse causality issue by analyzing the impact of abnormal returns on future changes in KLD scores. Studies documenting higher financial performance for firms with higher ESG engagement are often criticized for not properly accounting for the endogeneity between high financial and ESG performance.

By working on changes in ESG performance and carefully addressing the reverse causality issue, this is the first study - to the best of our knowledge - to be taking a closer look at the complex mechanism between ESG performance and stock returns.

For preliminary findings please see the Discussion and Conclusions section.

2. Theoretical Hypotheses and Literature Review

To be elaborated.

3. Empirical Design

I) **Research Question:** Do changes in KLD strengths and concerns affect abnormal returns?

$$AR_{i,t} = \Sigma_{j=1}^{14} \Delta KLD_{j,i,t} + \Sigma_{k=1}^{15} Year_{k,i} + \epsilon_{i,t} \tag{1}$$

where $AR_{i,t}$ is the abnormal return measured according to equations 11 and 12. $\Sigma_{j=1}^{14} \Delta KLD_{i,t}$ are the changes in the number of strengths and the number of concerns separately in each of the seven KLD dimensions. The change in the KLD score is measured as a discrete variable, reflecting the yearly change in the number of strengths and concerns, respectively, for each of the seven dimensions: community, diversity, corporate governance, employee relations, environment, human rights and product safety. All variables in this model are contemporaneous.

An alternative is to estimate the impact of current level of changes in KLD scores on future abnormal returns which is justified by either the information availability or the economic argument. By the information availability argument, the KLD scores pertaining to year t are made public late January early February year $t+1$ and therefore could be reflected, if at all, in the abnormal returns of year $t+1$. By the economic argument, the benefits of ESG performance as assessed by the KLD scores could materialize one period later.

No other control variables than year dummies would be justified in the model for abnormal returns (Model 1) because in a four-factor world, as we assumed when computing the abnormal returns, i.e. Model 11 and 12, it is these four factors only that could impact returns and any other factor should have zero systematic impact. These potentially additional factors are captured by the $\epsilon_{i,t}$ which is normally distributed with mean zero and diagonal covariance matrix.

Model 1 can be estimated with Fixed Effects or Random Effects - the Hausman test will indicate which one is preferable. Alternatively, pooled OLS with clustered errors can also be used.

A probit/logit model (alternative to Model 1) of the impact of changes in KLD scores (strengths and concerns separately) on the likelihood of recording an increase in abnormal returns above a certain threshold can also be estimated. In fact, a (non-linear) probit/logit model might be preferred as the impact of changes in KLD score may not be linearly translated into stock return changes.

Interpretation: A statistically significant coefficient on $\Delta KLD_{i,t}$ in Model 1 would reconfirm the value-relevance of ESG performance which can be explained either as mispricing or compensation for risk. In line with prevailing empirical evidence that high KLD scores lead to high returns, we would expect a positive significant coefficient on the variable measuring the strengths and a negative coefficient on the variable measuring concerns.

Extension to these models: include the KLD scores in absolute values among the explanatory variables in order to control for current ESG performance.

II) **Research Question:** Does past financial performance (i.e. previous yearly abnormal returns) affect the change in the number of ESG strengths and concerns?

Basically, at this stage I worry that the change in KLD scores in previous model, i.e. model 1, might be endogenous and influenced by past abnormal returns. There are claims (cite reference) that ESG consideration could be initiated most often in firms that are financially successful (which can afford to invest in ESG, cite reference). Therefore, the purpose of this model is to uncover more subtle feedback from financial performance to changes in ESG performance.

Such a model could be estimated simultaneously, using a system estimation technique such as Two or Three Stage Least Squares in which case the estimated coefficients are usually difficult to interpret. Alternatively, we could estimate the model sequentially by the instrumental variables-two stage least squares approach, replacing the endogenous variables in the main equation with their estimates based on selected instrumental variables.

Either way, as we have 14 endogenous variables, i.e. 14 changes in KLD scores we will need fourteen instruments in order to identify the main equation for abnormal returns, Equation 1. The characteristic of a good instrument is that it is correlated with the endogenous variable (i.e., change in KLD scores) but not correlated with the dependent variable (i.e., abnormal returns).

Size, leverage and R&D expenses and industry dummies would make good instruments as they have usually been claimed to be correlated to the changes in the KLD scores, but should not have a systematic relation to abnormal returns as long as the four-factor model holds; and we believe it does. And finally, past abnormal return is another instrument given our initial suspicion that changes in KLD returns might be driven by past financial performance. At the same time, past abnormal return is not expected to be a driver of next period abnormal return. With a 10 industry classification, we have only 13 instruments (less than the 14 needed), given that we include an intercept in the equation for each KLD change (i.e., the 9 industry dummies plus the four other instruments). We could possibly disaggregate even further the industrial sectors in order to obtain the sufficient number of instruments. However, for now, I will just leave out the human strengths KLD score given its questionable history: it started being recorded only in the recent years and some of its constituents suffered renaming from other categories.

The new model would thus be:

$$AR_{i,t} = \Sigma_{j=1}^{13}\Delta KLD_{j,i,t} + \Sigma_{k=1}^{15}Year_k + \epsilon_{i,t} \tag{2}$$

$$\Delta KLD_{1,i,t} = AR_{i,t-1} + Size_{i,t-1} + Leverage_{i,t-1} + R\&D_{i,t-1} + \Sigma_{k=1}^{9}Industry_{k,i} + \eta_{1,i,t} \tag{3}$$

$$\Delta KLD_{2,i,t} = AR_{i,t-1} + Size_{i,t-1} + Leverage_{i,t-1} + R\&D_{i,t-1} + \Sigma_{k=1}^{9}Industry_{k,i} + \eta_{2,i,t} \tag{4}$$

$$... \tag{5}$$

$$\Delta KLD_{13,i,t} = AR_{i,t-1} + Size_{i,t-1} + Leverage_{i,t-1} + R\&D_{i,t-1} + \Sigma_{k=1}^{9}Industry_{k,i} + \eta_{13,i,t} \tag{6}$$

where $\epsilon_{i,t}$ is uncorrelated (potentially only within clusters of firms)

$\eta_{j,i,t}$ could potentially be contemporaneously correlated across j (engagement with one KLD dimension might trigger engagement with another dimension), but it is independent of $\epsilon_{i,t}$

If I cannot find enough instruments to identify the abnormal returns equation, i.e., Equation 1, or if the instruments I came up with are not good enough, a reasonable alternative approach would be to reduce the number of endogenous changes in KLD scores. It would also make sense to try subsets of strengths and concerns and see their effect. However, as we assume there might be some correlation between the strengths and concerns across different dimensions at a given time, excluding certain aspects might adversely impact the estimates for the remaining others.

First, I run the regression based Hausman test of endogeneity for $\Sigma_{j=1}^{13}\Delta KLD_{j,i,t}$ in equation 2, i.e. test if the joint effect of the residuals obtained from estimating equations 7 to 10 is zero when these residuals are included as explanatory variable in Equation 2. Equations 7 to 10 represent the endogenous variables regressed on all exogenous variables in Model 2 to 6 (both included and excluded from the equation for $\Delta KLD_{j,i,t}$). The estimation was done with pooled OLS and standard errors clustered at firm level to correct for the within-firm group correlation. The null hypothesis was rejected (p-value=0.00) => the changes in KLD scores seem to be endogenous. However, this conclusion holds conditional on the validity of the instruments.

$$\Delta KLD_{1,i,t} = AR_{i,t-1} + Size_{i,t-1} + Leverage_{i,t-1} + R\&D_{i,t-1} + \Sigma_{k=1}^{9} Industry_{i,k} + \Sigma_{k=1}^{15} Year_k + \nu_{i,t} \quad (7)$$

$$\Delta KLD_{2,i,t} = AR_{i,t-1} + Size_{i,t-1} + Leverage_{i,t-1} + R\&D_{i,t-1} + \Sigma_{k=1}^{9} Industry_{i,k} + \Sigma_{k=1}^{15} Year_k + \nu_{i,t} \quad (8)$$

$$\ldots \quad (9)$$

$$\Delta KLD_{14,i,t} = AR_{i,t-1} + Size_{i,t-1} + Leverage_{i,t-1} + R\&D_{i,t-1} + \Sigma_{k=1}^{9} Industry_{i,k} + \Sigma_{k=1}^{15} Year_k + \nu_{i,t} \quad (10)$$

Thus, our prior expectations of endogeneity of the change in KLD scores are confirmed by the econometric test.

The next step is to estimate the entire system of Equations 2 to 6 either with an Instrumental Variable (IV) (IV-2SLS) approach or the System 3 Stage Least Squares (3SLS) or 2 Stage Least Squares (2SLS). System estimators, either 2SLS or 3SLS, are more efficient that the IV-2SLS estimators of equation by equation as long as the equations are correctly specified. However, in case of equation misspecification, the system estimators will be inconsistent (Wooldridge, 2002, Chapter 9). Moreover, it is usually difficult to interpret the estimators resulting from a system estimation.

As there is some suspicion on the validity of the instruments, it would make sense to discuss the results of the test for the overidentifying restrictions, where possible.

4. Model Estimation

4.1. **General ideas.** It makes sense to focus on the DS400 and SP400 list of firms from the point of view of public availability of ESG data but also to minimize the risk of 0 rating meaning no rating. Statman and Glushkov (2008) adopt a similar strategy for the same reason. Moreover, by looking at summary statistics (unreported) when going from the larger to the smaller sample, the number of firms with changing number of strengths or concerns is only marginally diminished compared to the total number of firms which almost halves, thus justifying the use of the smaller sample. One notable exception are the samples' dynamics with respect to the corporate governance dimension.

4.2. **Data details.** The abnormal return for each security (i), each year(t), ($AbnRet_{i,t}$) is determined as the difference between its actual return ($ActRet_{i,t}$) and its expected return ($ExpRet_{i,t}$). The security expected return for year t is the risk free rate ($RiskFreeR_t$) plus the product of the security sensitivities to the four risk factors, i.e. the market portfolio and the size, value and momentum risk-factor mimicking portfolios, and the expected risk premiums on these factors in year t (see Equation 12). The expected risk premiums on these factors are their realized returns in year t. The actual return is the simple return for year t.

$$AbnRet_{i,t} = ActRet_{i,t} - ExpRet_{i,t} \tag{11}$$

$$\begin{aligned} ExpRet_{i,t} &= RiskFreeR_t + \beta^{Market}_{i,t-1} ExpRet_{(Market_t - Rf_t)} + \beta^{Size}_{i,t-1} ExpRet_{(Small_t - Big_t)} \\ &+ \beta^{Value}_{i,t-1} ExpRet_{(High_t - Low_t)} + \beta^{Mom}_{i,t-1} ExpRet_{(Winners_t - Loosers_t)} \end{aligned} \tag{12}$$

where $t = 1991...2007$ and $i = 1...N$, N is the number of securities in each year.

The sensitivities $\beta^k_{i,t-1}, k = 1...4$ are estimated on a sample of 60 (or at least 24) monthly returns ending December year $t-1$ and the associated monthly returns on the four factors, using a regression on all factors simultaneously, see Model 13. This way each sensitivity β^k captures the security's sensitivity to that particular risk factor k net of the impact of the other factors.

$$\begin{aligned} ActRet_{i,s} - Rf_s &= \alpha_i + \beta^{Market}_{i,t-1} ActRet_{(Market_s - Rf_s)} + \beta^{Size}_{i,t-1} ActRet_{(Small_s - Big_s)} \\ &+ \beta^{Value}_{i,t-1} ActRet_{(High_s - Low_s)} + \beta^{Mom}_{i,t-1} ActRet_{(Winners_s - Loosers_s)} + \gamma_{i,s} \end{aligned} \tag{13}$$

where $s = 1...60$ months up to December year $t-1$.

Moreover, the $\beta^k_{i,t-1}, k = 1-4$, parameters are estimated by Model 13 such that they satisfy the constraint that the expected return on security i, obtained by Model 12, cannot be lower than -100%, meaning that you cannot expect to lose more than your entire investment. Specifically, when estimating regression 13, the constraint on $\beta^k_{i,t-1}, k = 1-4$, parameters is that their product with the expected risk premiums on the risk factors next period t, i.e. the expected return component in excess of the risk free rate used in Model 12, cannot go below -100%.

The summary statistics of the beta estimates in Table 1, indicate that both Market and Size betas are tightly distributed around their mean, while the other two are more dispersed. Also, the mean estimate market beta of 1 is in line with expectations.

	(1) Mean	(2) SD	(3) p25	(4) p75	(5) $\Sigma_{j=1}^N \beta_j * w_j$
Market beta	1.01	0.03	0.99	1.03	1.01
Size beta	0.65	0.07	0.60	0.69	-0.05
Value beta	0.21	0.13	0.12	0.35	0.04
Momentum beta	-0.13	0.05	-0.18	-0.09	-0.05

TABLE 1. Mean, standard deviation (SD), 25th (p25) and 75th(p75) percentiles of the yearly average beta sensitivities to the for risk factors (Columns 1 to 4). The yearly average of the market capitalization weighted sum of each year's securities betas is in Column 5.

Moreover, there is a post-estimation validity check with respect to the $\beta^k_i, k = 1...4$ risk factor sensitivity estimates that needs to hold, which is that if by assumption one owns all stocks in the market, then one should earn no more and no less than the market portfolio return. That is equivalent to taking the sum over all securities on the left hand side of equation 13, weighted by their market capitalization ratio, and obtaining the market portfolio excess return on the right hand side of this equation. More specifically, the following conditions should hold every year: $\Sigma_{j=1}^N \beta_j^{Market} * w_j = 1$, $\Sigma_{j=1}^N \beta_j^{Size} * w_j = 0$, $\Sigma_{j=1}^N \beta_j^{Value} * w_j = 0$, $\Sigma_{j=1}^N \beta_j^{Momentum} * w_j = 0$, where w_j is the market weight of security j . Column 5 in Table 1 indicates that these conditions hold on average and therefore equation 12 can be used to estimate expected excess returns.

The exogenous variables used in Model 2 to 6, are the following:

- Firm Size: log of assets
- Firm Leverage: long-term liabilities%Total capital (First tried with long-term liabilities to common equity, but often firms have very small value for equity leading to unusually high ratios)
- Firm Research&Development expenses: measured as R&D%Total Sales.
- Year dummies
- Industry dummies

After the panel data construction for period 1991-2007, several outlier elimination was performed sequentially: abnormal returns are required to be larger than -150% and lower than 200%

(0.006% observations lost), leverage is imposed to be positive and lower than 150% (another 0.005% observations lost) and R&D expenses ratio has to be lower than 150% (another 0.011% observations lost).

In Table 2, summary statistics are presented for the main variables of interest in the sample consisting of SP500 and DS400 members.

4.3. **Preliminary estimation results.** Table 3 displays estimation results of Model 1 with pooled OLS (&clustered standard errors at firm level), random effects (RE), fixed effects (FE) and random effects with the KLD scores in levels among the explanatory variables. Table 4 displays the marginal effects of various probit estimations of Model 1. In Table 5, the estimation results of a lead-lag variant to Model 1 are presented. Finally, the instrumental variables-two stage least squares (IV-2SLS) estimation as well as a system three stage least squares (3SLS) and a system two stage least squares (2SLS) estimation of Model 2 to 6 are shown in Table 6.

Some preliminary implications are as follows.

In table 3, the pooled OLS, RE and FE deliver similar results of model 1 estimation and the Hausman test indicates that the RE estimation is consistent. Therefore we use the RE estimates as main estimates. Moreover, this verified consistency of the RE estimates confirms the strict exogeneity of firm specific effects, i.e., they are uncorrelated with the error term. This means that we might have succeeded in not having any omitted variables in the error term of the model for abnormal returns, model 1; or, that the four factor model used to estimate abnormal returns succeeds in capturing all relevant information for stock returns.

An estimation (random effects) of a contemporaneous relation between abnormal returns and changes in the KLD scores (Table 3, Column 2) indicates that a higher number of concerns with respect to community, governance as well as product dimensions leads to lower abnormal returns. However, when controlling for the level of KLD scores, it is only the governance concern that shows a statistically significant relation to abnormal returns (still negative) (Table 3, Column 4).

A change in strengths in any dimension seems to be having no impact on abnormal returns.

The negative impact of an increase in governance concerns is consistent when estimated on the likelihood of higher abnormal returns, although somewhat weaker. The threshold for which the abnormal return dummy is equal to one is 1%, which is true for 4579 out of 10066 stocks.

With probit estimation, changes in more of the strength dimensions have negative effects than with levels estimation, but it is only the community ones that are also statistically significant in one of the specifications (Table 4, Column 3).

The estimation of the impact of lagged changes in KLD scores on abnormal returns (Table 3) provides somewhat opposite results to the contemporaneous estimation (Table 2). By contrast, the impact of lagged change in governance concerns has a *positive* impact on lead abnormal returns and statistically significant across the various estimations.

Overall, the discrepancy between contemporaneous and lagged impact of changes in KLD scores on abnormal returns might be an indication of reversed causality between these two. Or, it could also have to do with the dynamics between ESG performance and abnormal returns. It certainly requires more thinking about the implications of these effects. Using a lag-lead model could be an alternative way to come around the endogeneity problem.

The Hausman test for endogeneity has a p-value of 0.00 indicating that, given the instruments used are valid, the change in the KLD scores is endogenous in Model 1, and therefore the coefficient estimates in Tables 3 and 4 could be biased.

The various estimation techniques aimed at correcting for endogeneity, i.e. Table 6, deliver somewhat strange results as the coefficient estimates have unusual large values.

A second remark on Table 6 is that the IV-2SLS and system 2SLS deliver identical coefficient estimates, though with different standard errors- Columns 1 and 3 (therefore the different p-values in parentheses), which is related to the consistency and efficiency properties of the two models (to be elaborated).

Moreover, another interesting observation is that these three models seem to agree on a *positive* estimated impact for changes in KLD governance score, similar to the effect estimated with a lead-lag model for abnormal returns (Table 5).

More concerning, however, is the fact that the F-test of the lack of joint explanatory power of the variables in Model 2 to 6 is not rejected, thus raising doubts about its well-specification. This fact corroborated with the odd coefficient estimates require a reconsideration of the way in which the instruments have been used, which will be the next step in the development of this paper.

5. Discussion and Conclusions

Based on the preliminary estimation results, several conclusions can be drawn. First: it seems that it is the KLD concerns and much less the KLD strengths that have an impact on abnormal returns.

Second: among the various KLD concerns, it is primarily the change in governance concerns that negatively affects contemporaneous abnormal returns and positively affects next period abnormal returns. This finding can be explained by an overreaction to an increase in the number of governance concerns in the period, followed by a catching up phase in the next period, given the estimated coefficients are trustworthy.

Third: there is some evidence that the change in KLD scores and the abnormal returns might be determined simultaneously. However, more effort needs to be put into properly accounting for this issue.

(To be elaborated further)

References

Brammer, S., Brooks, C., and Pavelin, S. (2006). Corporate social performance and stock returns: Uk evidence from disaggregate measures. *Financial Management*, 35(3):75–116.

Derwall, J., Gunster, N., Bauer, R., and Koedijk, K. (2005). The ecoefficiency premium puzzle. *Financial Analysts Journal*, 61:61–63.

Derwall, J. and Verwijmeren, P. (2007). Corporate social responsibility and the implied cost of equity capital. *Mimeo*, Available at http://publishing.eur.nl/ir/repub/asset/8986/ESP2007101F26A9058921328DERWALL.pdf.

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

Manescu, C. (2010). Stock returns reaction to environmental social and governance performance: mispricing of compensation for risk? *Mimeo*.

Statman, M. and Glushkov, D. (2008). The wages of social responsibility. *Mimeo*, Available at $http://papers.ssrn.com/sol3/papers.cfm?abstract\_id = 1372848$.

von Arx, U. and Ziegler, A. (2008). The effect of csr on stock performance: New evidence for the usa and europe. *Mimeo*, Available at http://econpapers.repec.org/paper/ethwpswif/08-85.htm.

Wooldridge, J. (2002). *Econometric Analysis of Cross Section and Panel Data*. The MIT Press.

Table 2 Summary Statistics for firms in SP500 or DS400

Variable	Obs	Mean	Std. Dev.	Min	Max
AbnormalRet.	10066	.8032097	36.82052	-145.2536	196.5411
Size	10066	15.4149	1.714955	9.732165	21.50608
Leverage	10066	32.87071	22.92327	0	149.81
R&D	10066	-35.7141	50.87787	-100	85.83
deltacomstr	9082	.011231	.3630395	-3	3
deltacomcon	9082	.0151949	.2447338	-2	2
deltacgovstr	9082	.0144241	.2585634	-2	2
deltacgovcon	9082	.0546135	.4613592	-2	3
deltadivstr	9082	.0883065	.5343789	-3	4
deltadivcon	9082	.0137635	.3086821	-2	2
deltaempstr	9082	.0279674	.3888231	-2	2
deltaempcon	9082	.052081	.478173	-3	3
deltaenvstr	9082	.0198194	.3038556	-2	4
deltaenvcon	9082	.0221317	.3699051	-3	3
deltahumstr	9082	.0013213	.0726912	-1	1
deltahumcon	9082	.0016516	.2248163	-2	2
deltaprostr	9082	.0014314	.1891744	-1	2
deltaprocon	9082	.0341335	.3532876	-4	2
Financials	10066	.129843	.3361474	0	1
Basic Mater.	10066	.0769919	.2665918	0	1
Industrials	10066	.1883568	.3910163	0	1
Technology	10066	.0926883	.2900095	0	1
Consum.Goods	10066	.1390821	.3460493	0	1
Oil&Gas	10066	.0496722	.2172775	0	1
Healthcare	10066	.0747069	.2629309	0	1
Consum. Serv	10066	.1535863	.3605697	0	1
Telecom	10066	.0215577	.1452415	0	1
Utilities	10066	.0735148	.2609926	0	1

Table 3. Estimation results of Model (1) with pooled OLS and clustered standard errors (Column 1), panel fixed effects (2)and random effects (3), and panel random effects with KLD scores in levels among explanatories (4). P-value of the Hausman test on the consistency of the Random effects at the bottom.

	(1)	(2)	(3)	(4)
	(P-OLS)	(RE)	(FE)	(RE)
VARIABLES	**AR-level**	**AR-level**	**AR-level**	**AR-level**
Includes KLD Levels?	**NO**	**NO**	**NO**	**YES**
deltacomstr	-0.82	-1.09	-1.45	-1.46
	(0.36)	(0.28)	(0.16)	(0.17)
deltacomcon	-2.65**	-2.68*	-2.38	-2.68
	(0.04)	(0.07)	(0.12)	(0.11)
deltacgovstr	-1.33	-1.68	-2.03	-1.69
	(0.35)	(0.23)	(0.16)	(0.30)
deltacgovcon	-2.29***	-2.49***	-2.67***	-2.54***
	(0.00)	(0.00)	(0.00)	(0.01)
deltadivstr	0.28	-0.12	-0.61	0.27
	(0.68)	(0.86)	(0.39)	(0.71)
deltadivcon	-0.27	-0.12	0.09	0.77
	(0.82)	(0.92)	(0.94)	(0.56)
deltaempstr	0.88	0.56	0.01	-0.01
	(0.38)	(0.55)	(0.99)	(0.99)
deltaempcon	-0.13	0.04	0.15	-0.36
	(0.87)	(0.95)	(0.85)	(0.68)
deltaenvstr	-0.59	-1.00	-1.42	0.03
	(0.54)	(0.41)	(0.26)	(0.98)
deltaenvcon	-0.43	-0.39	-0.40	-0.31
	(0.64)	(0.69)	(0.69)	(0.76)
deltahumstr	-0.67	-0.69	-0.40	-1.22
	(0.91)	(0.89)	(0.94)	(0.83)
deltahumcon	0.82	0.51	0.02	-0.19
	(0.58)	(0.76)	(0.99)	(0.91)
deltaprostr	0.49	0.56	0.76	0.29
	(0.80)	(0.77)	(0.70)	(0.89)
deltaprocon	-1.64*	-1.75*	-1.81*	-1.74
	(0.06)	(0.09)	(0.09)	(0.12)
Year Dummy	YES	YES	YES	YES
Constant	-2.91**	-0.65	-0.36	16.44
	(0.02)	(0.68)	(0.81)	(0.72)
Observations	9082	9082	9082	9082
Number of dscdID		1034	1034	1034
R-squared	0.03	.	0.04	.

p-value Hausman Test =0.17

Robust p values in parentheses
***** p<0.01, ** p<0.05, * p<0.1**

Table 4. (Probit) Marginal effects estimation of a modified version of Model (1) such that the dependent variable, Abnormal returns, is a dummy variable =1 if AR>1%. The estimation is done with pooled probit (Column 1), probit Random Effects (2) and probit Random effects with KLD scores in levels among explanatories (3).

VARIABLES	(1) (Pooled) AR-Dummy (>1%)	(2) (RE) AR-Dummy (>1%)	(3) (RE) AR-Dummy (>1%)
Includes KLD Levels?	NO	NO	YES
deltacomstr	-0.05	-0.05	-0.09**
	(0.17)	(0.14)	(0.03)
deltacomcon	-0.07	-0.07	-0.06
	(0.20)	(0.22)	(0.32)
deltacgovstr	-0.05	-0.05	-0.07
	(0.38)	(0.34)	(0.22)
deltacgovcon	-0.05*	-0.05*	-0.08**
	(0.08)	(0.08)	(0.03)
deltadivstr	-0.02	-0.02	-0.02
	(0.42)	(0.36)	(0.42)
deltadivcon	-0.01	-0.00	0.03
	(0.90)	(0.91)	(0.55)
deltaempstr	0.03	0.03	0.02
	(0.41)	(0.41)	(0.50)
deltaempcon	-0.03	-0.03	-0.04
	(0.28)	(0.30)	(0.22)
Deltaenvstr	-0.02	-0.02	-0.01
	(0.65)	(0.63)	(0.88)
deltaenvcon	-0.04	-0.04	-0.03
	(0.25)	(0.27)	(0.46)
deltahumstr	-0.06	-0.06	-0.10
	(0.76)	(0.75)	(0.63)
deltahumcon	0.06	0.06	0.02
	(0.31)	(0.32)	(0.80)
deltaprostr	-0.02	-0.03	-0.06
	(0.73)	(0.72)	(0.42)
deltaprocon	-0.03	-0.03	-0.03
	(0.46)	(0.46)	(0.42)
Year Dummy	YES	YES	YES
Constant	-0.12**	-0.12**	-0.37
	(0.03)	(0.03)	(0.49)
Observations	9082	9082	9082
Number of dscdID		1034	1034

Robust p values in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5. Estimation results of a modified version of Model (1) such that the KLD explanatory variables are lagged one period. The estimation is done with pooled OLS and clustered standard errors (Column 1), panel random effects (2) and fixed effects (3).

	(1)	(2)	(3)
	(P-OLS)	(RE)	(FE)
VARIABLES	AR-level	AR-level	AR-level
L.deltacomstr	0.52	0.22	-0.21
	(0.61)	(0.83)	(0.84)
L.deltacomcon	-0.90	-0.97	-0.88
	(0.52)	(0.53)	(0.58)
L.deltacgovstr	-1.34	-1.55	-1.66
	(0.35)	(0.30)	(0.28)
L.deltacgovcon	1.84**	1.70**	1.43*
	(0.03)	(0.04)	(0.09)
L.deltadivstr	-0.06	-0.36	-0.70
	(0.93)	(0.61)	(0.34)
L.deltadivcon	-0.52	-0.45	-0.46
	(0.68)	(0.71)	(0.72)
L.deltaempstr	1.47	1.22	0.86
	(0.19)	(0.22)	(0.41)
L.deltaempcon	0.62	0.80	0.90
	(0.47)	(0.32)	(0.28)
L.deltaenvstr	-0.93	-1.25	-1.55
	(0.42)	(0.32)	(0.23)
L.deltaenvcon	-1.59*	-1.62	-1.63
	(0.07)	(0.11)	(0.12)
L.deltahumstr	3.37	3.55	3.79
	(0.43)	(0.49)	(0.47)
L.deltahumcon	-0.13	-0.52	-1.03
	(0.93)	(0.76)	(0.55)
L.deltaprostr	-1.78	-1.58	-1.12
	(0.43)	(0.43)	(0.59)
L.deltaprocon	0.02	-0.08	-0.13
	(0.98)	(0.94)	(0.91)
Year Dummy	YES	YES	YES
Constant	-0.63	16.66***	3.48**
	(0.69)	(0.00)	(0.03)
Observations	8131	8131	8131
Number of dscdID		996	996
R-squared	0.04	.	0.04

*** p<0.01, ** p<0.05, * p<0.1
Robust p values in parentheses

Table 6. Estimation results of Equations (2) with Instrumental Variables-IV-2SLS (Column 1), three Stage Least Squares- 3SLS (2) and two Stage Least Squares (3). Several post estimation test statistics are reported at the bottom of the table.

	(1) (IV-2SLS)	(2) (3SLS)	(3) (2SLS)
VARIABLES	AR-level	AR-level	AR-level
deltacomstr	548.57	-1,185.63	548.57
	(0.84)	(0.78)	(0.90)
deltacomcon	541.45	-1,084.54	541.45
	(0.82)	(0.77)	(0.88)
deltacgovstr	-412.38	-648.97	-412.38
	(0.54)	(0.55)	(0.71)
deltacgovcon	68.54	350.87	68.54
	(0.91)	(0.72)	(0.94)
deltadivstr	-67.10	139.73	-67.10
	(0.85)	(0.80)	(0.90)
deltadivcon	400.73	-1,011.26	400.73
	(0.85)	(0.75)	(0.90)
deltaempstr	-68.04	346.97	-68.04
	(0.93)	(0.78)	(0.96)
deltaempcon	-91.99	88.56	-91.99
	(0.74)	(0.83)	(0.82)
deltaenvstr	-24.70	644.97	-24.70
	(0.98)	(0.68)	(0.99)
Deltaenvcon	-216.29	201.67	-216.29
	(0.73)	(0.83)	(0.82)
deltahumcon	202.91	514.44	202.91
	(0.79)	(0.66)	(0.86)
deltaprostr	157.22	-149.53	157.22
	(0.88)	(0.93)	(0.93)
deltaprocon	52.99	74.93	52.99
	(0.76)	(0.80)	(0.86)
Year Dummy	YES	YES	YES
Constant	-9.41	(.)	-9.41
	(0.79)	(.)	(0.86)
Observations	5596	5596	5596
Chi2- p-value	(.)	0.00	(.)
F-test p-value	0.99	(.)	1.00

Anderson canon. corr. LR statistic (identification/IV relevance test):
Chi-sq(1) P-val = 0.8826 (for IV estimation)
Hansen J statistic (overidentification test of all instruments):
p-value 0.00 (equation exactly identified)

Robust p values in parentheses
***** p<0.01, ** p<0.05, * p<0.1**