What Drives the Dynamics of CSR Returns?*

Renxuan Wang† Russ Wermers‡ Ruoke Yang§

Tianchen (Hugo) Zhao¶

March 16, 2025

Abstract

We investigate the relationship between corporate social responsibility (CSR) and stock returns during economic downturns and the underlying economic mechanisms. First, we successfully replicate two influential studies that reach opposite conclusions on this topic and propose several explanations that may reconcile their findings, including differences in CSR scoring, empirical specifications, and adjustments for systematic risk exposures. We find that high-CSR stocks exhibit time-varying factor loadings, particularly a stronger tilt toward the value factor during bad economic times. From the standpoint of cash flow (CF) vs. discount rate (DR) news, we show that CSR returns are largely driven by aggregate DR shocks, highlighting the role of shifts in investor sentiment and preferences. CSR spread portfolio returns are lower in bad times, but increase in response to positive DR shocks. Finally, we analyze flows to socially responsible (SR) mutual funds and find evidence that is fully consistent with the documented return patterns: while SR funds typically face outflows during downturns, these outflows lessen when aggregate DR shocks increase. These findings suggest that investors in general view CSR as a luxury good, but some investors retain trust in high-CSR firms despite challenging market conditions.

*For helpful comments, we thank Brandon Yueyang Han, Pablo Slutzky, Ania Zalewska, and participants at the 2025 Junior Academics Research Seminars (JARS) in Finance.

†China Europe International Business School (CEIBS). *Email:* rxwang@ceibs.edu

‡Robert H. Smith School of Business, The University of Maryland. *Email:* wermers@umd.edu

§U.S. Securities and Exchange Commission. *Email:* ruoke.yang@gmail.com The Securities and Exchange Commission disclaims responsibility for any private publication or statement of any SEC employee or Commissioner. This article expresses the author's views and does not necessarily reflect those of the Commission, the Commissioners, or other members of the staff.

¶Robert H. Smith School of Business, The University of Maryland. *Email:* tzhao97@umd.edu

1. Introduction

Despite the tremendous expansion and popularity of socially responsible investing, there is considerable debate over the stock return performance of socially responsible firms. In particular, whether firms with good CSR performances earn higher returns during bad economic times (e.g. financial crisis) has surprisingly remained an open question.

The existing finance literature offers seemingly conflicting evidence. Lins, Servaes, and Tamayo (2017) (LST) study the Global Financial Crisis (GFC) from late 2008 to early 2009 and show that stocks with high CSR scores before the GFC achieved superior abnormal returns during the GFC, compared to firms with low CSR scores. In contrast, Bansal, Wu, and Yaron (2022) (BWY) find that a long-short CSR portfolio - buying high-CSR stocks and shorting low-CSR stocks - generates negative abnormal returns during periods with low stock market valuations, including the GFC.

In this paper, we reexamine the question proposed by LST and BWY and reconcile the differences in their findings. More importantly, we adopt the classic asset pricing framework that distinguishes between cash flow (CF) and discount rate (DR) news to assess the underlying mechanisms that drive the returns to CSR investment strategy over good and bad economic times. Lastly, we analyze flows to socially responsible (SR) mutual funds to provide insights into how shifts in investor demand influence the pricing of CSR.[1]

As a starting point, we follow the methodologies of LST and BWY and confirm that the baseline findings in both studies are reproducible. Despite their apparent disagreements, each paper employs reasonable empirical approaches and arrives at generally replicable results within its specific framework.

We next explore several potential reasons why LST and BWY reach different conclusions. While each paper's methodology has merit, LST and BWY differ in certain assumptions and empirical choices. First, they adopt distinct definitions of CSR scores. LST focus

[1]For prior work that studies demand associated with socially responsible mutual funds, see Riedl and Smeets (2017) and Hartzmark and Sussman (2019).

on a subset of categories they consider central to measuring CSR performance, whereas BWY incorporate all available categories to offer a more comprehensive CSR profile. This divergence reflects each study's viewpoint on which CSR components are relevant to firms and investors. Nevertheless, our analysis shows that the variation in CSR scoring alone is unlikely to resolve the discrepancy: when we apply BWY's approach using LST's CSR scores, the long-short CSR portfolio earns larger but still negative abnormal returns during the GFC.

The second distinction between LST and BWY lies in their baseline model specifications. LST estimates a cross-sectional regression, using CSR scores from 2006 to predict returns during the GFC. With this specification, their emphasis is on the predictive power of pre-crisis CSR attributes. In contrast, BWY employ a calendar-time portfolio approach that updates CSR scores continuously, mirroring the behavior of a real-time investor who regularly rebalances a long-short portfolio based on the latest CSR data. While both designs are valid for understanding how CSR affects crisis-period performance, they have potential implications for different audiences.

The third distinction involves how systematic risk exposures are accounted for in two studies. Specifically, when studying abnormal returns, LST estimate CAPM and Fama-French-Carhart (FFC) factor loadings using pre-GFC data, effectively treating stocks' betas as fixed and not affected by the crisis. On the other hand, BWY adopt rolling-window regressions to measure monthly FFC alphas of the long-short CSR portfolio, implicitly allowing the betas to be updated and vary across different economic episodes. While neither assumption is inherently correct or incorrect, the discrepancy here could potentially explain the differences in results. In additional analyses, we find that firms with higher CSR scores experience notable changes in factor loadings during bad economic times, which is consistent with the broader literature on conditional and time-varying betas (Jagannathan and Wang, 1996; Fama and French, 1997; Ferson and Harvey, 1999; Lettau and Ludvigson, 2001; Lustig and Van Nieuwerburgh, 2005; Zhang, 2005; Santos and Veronesi, 2006). In

particular, consistent across methodologies, we find that stocks with high CSR scores exhibit a significant increase in exposure to the HML (value) factor when market conditions deteriorate, implying that their distress risk rises in bad times (Fama and French, 1998). Given that the value factor did poorly during the GFC,[2] accounting for the increase in exposure to such factor (as in BWY) would put negative pressures on abnormal returns. Overall, we find the differences in measuring exposure to systematic factors likely drive most of the differences in the results of the two studies.

What drives returns to CSR investing over time, and what accounts for their dynamic exposures to systematic risks? According to the classic Campbell and Shiller (1988) decomposition, stock price fluctuations stem from two sources: expected cash flow (CF) changes and discount rate (DR) changes. CF shocks generally capture shifts in firm-level fundamentals, whereas DR shocks reflect changes in investor sentiment, preferences, and required risk premiums. Building on an asset pricing literature, we apply the CF vs. DR framework to determine how aggregate news in these two dimensions affects CSR returns over time.[3]

Understanding how CSR returns loads on CF vs. DR news helps distinguish theories proposed in the literature. LST and BWY present direct but distinct predictions regarding how CSR returns are related to CF news and DR news. LST contend that high-CSR firms benefit from stakeholder trust, which improves fundamentals (e.g., higher profit margins, sales growth, etc) during recessions. This trust may also induce shareholders to retain or even expand their investments in high-CSR firms during downturns, suggesting that both CF and DR channels play important roles. In contrast, BWY view high-CSR stocks as "luxury goods" that investors favor when wealth is abundant but discard during bad times, implying that the return dynamics of CSR are driven primarily by shifts in DR.[4]

[2]The value factor fell by 15.7% during the GFC period defined in LST.

[3]In asset pricing, a host of studies aim to explain conditional betas and anomaly returns using CF and DR perspectives (Bansal and Yaron, 2004; Campbell and Vuolteenaho, 2004; Barberis et al., 2005; Santos and Veronesi, 2005; Da and Warachka, 2009; Campbell et al., 2010; Chen et al., 2013; Lochstoer and Tetlock, 2020)

[4]Hong et al. (2012) make an analogous argument about firm managers pursuing corporate social responsibility when times are well.

Following a recent study by Lan and Wermers (2024), we decompose aggregate market unexpected returns into CF and DR components using sell-side analyst earnings forecasts. We then link the returns of the long-short CSR portfolio directly to these two sources of news. Our analysis shows several interesting results. First, unconditionally, CSR returns load positively and strongly on DR shocks, rather than CF shocks, indicating that CSR returns are significantly influenced by shifts in overall investor sentiment and preferences. Second, during bad times, CSR returns decrease, in line with BWY's "luxury good" hypothesis. However, we also observe that the CSR returns' exposure to DR shocks becomes even more pronounced in bad times, implying that high-CSR stocks outperform low-CSR stocks specifically when the market declines for reasons unrelated to fundamentals. This pattern is consistent with LST's trust-based hypothesis, in which some investors continue to support trustworthy high-CSR firms during downturns.

Next, we directly test the role of DR shocks in explaining CSR returns by examining investor demand for high-CSR stocks, proxied by flows to socially responsible (SR) mutual funds. Following Hwang et al. (2022) and Cao et al. (2023), we compute quarterly CSR scores for each fund based on its most recent stock holdings, and we define SR funds as those in the top CSR tercile each quarter. Our findings reveal patterns that are fully consistent with the above results on CSR returns. Specifically, SR funds, compared to non-SR counterparts, generally face more outflows in bad times, consistent with BWY's prediction of lower investor demand for high-CSR stocks during economic stress ("luxury good"). However, when discount-rate spikes occur — i.e., when non-fundamental shocks drive down the market — the outflows from SR funds become comparatively smaller. This suggests that at least some investors uphold or even reinforce their commitment to high-CSR investments in difficult periods, consistent with LST's perspective on investor trust.

Overall, we document a series of empirical findings that reconcile the seemingly contradictory evidence on CSR and crisis-period stock returns. By further dissecting the

underlying drivers of CSR returns and evaluating existing theories, we demonstrate that investor preferences and trust indeed plays a crucial role when market conditions worsen.

Our paper contributes to the ongoing debate among researchers regarding the relationship between corporate social responsibility and returns. As first posited by Heinkel et al. (2001) and later discussed in Pástor et al. (2021), nonpecuniary investor preferences associated with corporate social responsibility can play an important role in shaping returns: socially preferred firms in equilibrium exhibit lower expected returns (and therefore enjoy cheaper cost of capital) as socially responsible investors sacrifice financial performance from the portfolio standpoint in order to invest according to their tastes.[5] Evidence drawn from various settings and different asset classes documents this impact on returns from investor preferences (Hong and Kacperczyk, 2009; Chava, 2014; Riedl and Smeets, 2017; Barber et al., 2021). On the other hand, studies stemming from primarily corporate finance and industry have argued for corporate social responsibility as a quality intrinsically valuable to a firm's business and can lead to a positive effect on returns if the benefits of corporate social responsibility are not fully valued by the market (Edmans, 2011; Deng et al., 2013; Dimson et al., 2015).[6] Our paper takes a different approach from these earlier works.[7] Instead of providing another example (e.g. evidence of an effect of a specific type of corporate social responsibility on the returns of a new asset class) that offers empirical support for one of these two strands of prior literature, our paper applies a unified framework and sheds light on the empirical significance of these two strands, one of which reflects DR shocks and the other reflecting primarily CF shocks.

The remainder of the paper is structured as follows: Section 2 describes data sources and variable construction. Section 3 presents our replication results of LST and BWY and potential explanations for the discrepancies in their findings. Section 4 explores the

[5]Under special circumstances, like an unexpected increase in demand from investors with social preferences, socially preferred firms may enjoy higher returns, see Pástor et al. (2022).

[6]The effect may not necessarily be positive if the particular form of socially responsible activity is a manifestation of managerial agency problems as highlighted by Masulis and Reza (2015) in their study of corporate philanthropy.

[7]For a review of additional studies related to CSR, see Hong and Shore (2023).

relationship between CSR portfolio returns and aggregate CF and DR news, as well as flows to SR mutual funds. Section 5 concludes.

2. Data

This section provides an overview of the datasets we use to investigate the time-varying relationship between CSR and stock returns and the underlying mechanisms. We describe our sample selection criteria, data cleaning process, and the construction of variables.

2.1. CSR Ratings

We obtain social responsibility ratings from MSCI KLD, which provides ratings on the CSR practices of large public firms. As indicated in both LST and BWY, KLD has been widely used by practitioners and academics, including both LST and BWY. Unlike ratings that typically score on a letter or numerical scale, KLD ratings are formed, at the end of each year, as a set of indicator variables for each firm. These indicators span the following ESG categories: environment, community, diversity, employee relations, human rights, product, governance, and sin (i.e., alcohol, tobacco, gambling, firearms, nuclear power, and military). Within each category, KLD identifies both strengths and concerns, offering a granular view of a firm's CSR profile.

We construct two sets of CSR ratings, CSR_LST and CSR_BWY, by following the methodologies of LST and BWY, respectively. LST focus on five KLD categories: environment, community, diversity, employee relations, and human rights. Based on their approach, to measure a firm's involvement in each CSR category in each year, we sum up its annual strengths (concerns) for each of these five categories, divide the sums by the maximum number of strengths (concerns) possible for each category in each year, and then subtract the sum of concerns from the sum of strengths to arrive at the rating for each category. Finally, the aggregate rating, CSR_LST, which ranges from -5 to +5, is calculated by adding

together the five net sum scores from the five selected categories. BWY, on the other hand, consider all eight categories relevant to CSR performances. Based on their approach, we sum up a firm's annual strengths (concerns) for each of these eight categories and then subtract the sum of concerns from the sum of strengths to arrive at the rating for each category. The aggregate rating, CSR_BWY, is calculated by adding together the eight net sum scores from the eight selected categories.[8]

2.2. *Financial Information*

We collect stock return data from CRSP and firm-level accounting data from Compustat. To link these two databases, we use the CRSP-Compustat Merged file, which maps CRSP PERMNO identifiers to Compustat GVKEY identifiers.[9] For firms with more than one class of shares, we use the primary share as indicated by CRSP-Compustat Merged. We then merge the CRSP-Compustat data with the KLD ratings using ticker symbols, CUSIPs, and firm names, and manually verify all matches to ensure accurate linkage.

2.3. *Aggregate Cashflow News and Discount Rate News*

We construct the time-series of aggregate cashflow (CF) news and discount rate (DR) news, following the methodology of Lan and Wermers (2024). Specifically, we start with the Campbell and Shiller (1988) decomposition identity, which states that stock price fluctuations arise from two fundamental dynamics: expected CF changes and DR changes:

$$r_{t+1} - E_t(r_{t+1}) = (E_{t+1} - E_t) \sum_{k=0}^{\infty} \rho^k \, \Delta d_{t+1+k} \; - \; (E_{t+1} - E_t) \sum_{k=1}^{\infty} \rho^k \, r_{t+1+k} = N_{\mathrm{CF},t+1} \; - \; N_{\mathrm{DR},t+1}, \tag{1}$$

where r_t is the log stock market return at time t, Δd_t is the aggregate dividend growth at time-t, and ρ is a log-linearization constant.

[8] Unlike LST, BWY do not scale the sums of strengths or concerns.
[9] The CRSP-Compustat Merged file is available via the Wharton Research Data Services (WRDS).

We then estimate the market CF news component ($N_{\text{CF},t+1}$) by using the IBES sell-side analyst earnings forecasts at a monthly frequency, in a similar fashion as Da and Warachka (2009) and Chen et al. (2013). These analyst earnings forecasts dynamics (e.g. revisions) are able to explain stock return variability and are correlated with changes in trading activities and liquidity, providing incremental information to investors (Lys and Sohn, 1990; Lang and Lundholm, 1996; Asquith et al., 2005; Frankel et al., 2006; Chen et al., 2013).[10] We estimate the unexpected market returns in IBES ($r_{t+1} - E_t(r_{t+1})$) following Campbell and Vuolteenaho (2004), and finally, we back out the market DR news component ($N_{\text{DR},t+1}$) as the CF news component ($N_{\text{CF},t+1}$) minus the unexpected market returns ($r_{t+1} - E_t(r_{t+1})$).[11]

2.4. *Socially Responsible (SR) Mutual Funds*

We obtain returns, total net assets (TNA), and other characteristics of U.S. diversified, actively managed equity mutual funds from the CRSP Survivor Bias-Free Mutual Fund Database. Quarterly fund holdings (s12) data comes from Thomson Reuters, which we link to the CRSP data using MFLINKS files from WRDS. For each stock holding reported in the portfolio, we retrieve its market capitalization from the CRSP equity files by matching the relevant stock identifiers.

For funds with multiple share classes, we treat them as a single fund by summing the total net assets (TNA) of all share classes. For other fund-level characteristics, we take the TNA-weighted average across share classes. Following Elton et al. (2001), we exclude portfolios with TNAs below $15 million to mitigate potential small-fund return biases. Moreover, to address the incubation bias highlighted by Evans (2010), we remove each fund's first eighteen months of data.

Following recent literature (Hwang et al., 2022; Cao et al., 2023), we measure the

[10] Moreover, estimating cash flow expectations using earnings forecasts is a model-free approach, which avoids predictive regressions and sensitivity issues with state variables in VAR models (Chen and Zhao, 2009).

[11] For details on the construction of the series, see Lan and Wermers (2024) and the associated Internet Appendix.

social preferences of mutual funds by evaluating the CSR performances of their holdings. Specifically, for each fund in each quarter, we compute the portfolio-level CSR score by taking the value-weighted average of the CSR scores of all stocks in its portfolio. We exclude portfolios with fewer than 30 stock holdings that have nonmissing CSR scores. Next, at the end of each quarter, we rank all funds and sort them into terciles based on their portfolio-level CSR scores. Funds in the top tercile are classified as SR mutual funds.[12]

The main variable of interest in our test is the fund flows for fund i in month t, defined as

$$Flows = \frac{TNA_{i,t} - TNA_{i,t-1}}{TNA_{i,t-1}} - r_{i,t} \tag{2}$$

, where $r_{i,t}$ is the net fund returns.

3. CSR and Stock Returns Over Time

In this section, we begin by outlining the main empirical methodologies employed by LST and BWY to investigate the relationship between CSR and stock returns during bad economic times. Next, we demonstrate that the key results of both studies can be qualitatively replicated. Finally, we discuss the different considerations in these studies and offer our attempt to reconcile the seemingly contradictory results.

3.1. Empirical Methodologies

3.1.1. LST's Approach

LST study how firm-level CSR performances predict stock returns during the Great Financial Crisis (GFC, hereafter). In their baseline empirical test, they estimate the following

[12]This classification status lasts for a quarter, until the portfolio-level CSR scores are recalculated and sorted. For example, if a fund is classified as an SR fund at the end of Q4 2004 (based on its holdings at the end of Q4 2004), it is treated as an SR fund in January 2005, February 2005, and March 2005.

cross-sectional regression:

$$R_{i,GFC} = \beta * CSR_LST_{i,2006} + \gamma * Factor_Loadings_{i,July2008} + \rho * Controls_i + \epsilon_{i,GFC} \quad (3)$$

LST define the GFC as the period of August 2008 to March 2009. For firm i, $R_{i,GFC}$ is either the buy-and-hold stock return during the GFC or the abnormal stock return, calculated from the market model estimated over the past 60 months ending July 2008, one month before the start of the GFC. $CSR_LST_{i,2006}$ is firm i's CSR rating in 2006. $Factor_Loadings_{i,July2008}$ is a 4-by-1 vector of loadings for the Fama-French-Carhart (FFC) risk factors (market, size, value and momentum), estimated from the 60-month window ending in July 2008. $Controls_i$ include the following firm characteristics: log market cap, cash over total assets, short-term debt over total assets, long-term debt over total assets, operating income over total assets, book-to-market ratio, momentum, idiosyncratic risk, a dummy that equals to one for negative book-to-market ratios and 0 otherwise. Idiosyncratic risk is estimated as the residual variance from the market model estimated from the 60-month window ending in July 2008. These variables, with the exception of momentum and idiosyncratic risk, are measured as of December 2007 or as close to it as possible for firms whose fiscal year does not end in December. Within these cross-sectional tests, industry fixed effects are included and they are defined at the two-digit SIC level. LST also remove financial firms and micro-cap firms with a market capitalization below $\$250$ million as of the end of 2007.

3.1.2. BWY's Approach

BWY study the abnormal returns from investing in socially responsible stocks during good vs. bad economic times, defined as periods of high vs. low valuations between 1993 and 2013. Specifically, BWY employ a calendar-time portfolio approach, which involves sorting their sample into decile portfolios based on CSR scores and comparing the average Fama-French-Carhart (FFC) four-factor alphas of the top- and bottom equal-weighted

decile portfolios, computed using 36-month rolling-window regressions. One valuation ratio that BWY use to define bad economic times is the cyclicality-adjusted real P/E (CAPE) ratios from Shiller (2005). Each month is classified as being in bad economic times if the CAPE ratios fall in the lower half of their respective 10-year rolling distributions. Based on this classification, the bad economic times during BWY's sample period are May 2002 to April 2013 and June 2013, which includes the period of the GFC (August 2008-March 2009), as defined in LST.

3.2. Replication Results

We begin by replicating the core results of LST and BWY, adhering to their respective methodologies. In our attempt to replicate BWY, we start our sample period in 1996 and end it in 2019, consistent with the availability and coverage of CSR scores in KLD.[13] To address the lag in KLD's CSR score releases and ensure that these scores are observable to investors, we form decile portfolios based on the previous-year CSR scores.[14]

Table 1 presents our estimation results, which confirm that the main findings of both LST and BWY are indeed reproducible. In Panel A of Table 1, exactly as LST argue, we find that higher firm-level CSR scores in 2006 predict higher buy-and-hold returns and abnormal returns during the GFC period, regardless of whether firm fundamentals are included as controls. In Panel B, consistent with BWY, we find that the High − Low (top decile − bottom decile) CSR portfolio has strong, positive average FFC four-factor alphas during favorable economic times, but almost zero alphas during challenging economic downturns (which includes the GFC period). The economic magnitude of our estimated coefficients closely aligns with those reported in both studies.

To further illustrate the contrast between LST and BWY's primary conclusions, Figure 1 plots the rolling-window FFC four-factor alphas of the CSR spread portfolio, following

[13] In untabulated results, we confirm that our results are robust if we use the original BWY's sample period.
[14] Our results are robust if we use the current-year CSR scores to sort stocks into decile portfolios.

BWY's approach. The risk-adjusted abnormal returns associated with the CSR strategy vary over time, and specifically, the long-short portfolio based on CSR scores generates negative FFC four-factor alphas during the GFC period.

Taken together, these replication results confirm both LST and BWY's key findings but also highlight the apparent tension in their conclusions, which we investigate next.

3.3. Potential Explanations

Whether superior CSR performance translates into higher stock returns during periods of economic distress and financial crisis remains an intriguing yet unresolved question. Both LST and BWY investigate this issue through rigorous empirical analyses, each presenting distinct but compelling evidence, and our findings confirm the validity of both. Despite the robustness of these outcomes, one might ask what drives their seemingly contradictory conclusions. In this section, we propose a range of potential explanations—some of which may reconcile the differences between LST and BWY's results, while others may be less conclusive.

3.3.1. The Construction of CSR Scores

A key difference between LST and BWY's methodologies lies in the construction of CSR scores. As discussed in Section 2.1, LST focus on five categories to measure firms' CSR performance, while BWY expand the scope to eight by incorporating corporate governance, product, and sin. Both approaches have merit. For example, LST argue that the product category may include elements—such as product quality and innovation—that extend beyond the scope of traditional CSR. Furthermore, whether a firm operates in a controversial industry may not be an ideal measure of CSR performance because improving such a firm's score would require the firm to exit the industry, which may be impractical. In contrast, BWY's broader framework captures additional attributes that (some) investors could perceive as relevant to CSR. By incorporating all attributes provided by KLD, BWY's

CSR scores represent the most comprehensive information set available to investors who evaluate the CSR performance of stocks using KLD data. In our view, both LST and BWY provide valid, albeit distinct, measures of CSR performance.

Despite these differences in CSR scoring, our evidence suggests that they do not fully account for the divergent findings between LST and BWY. In Table OA1, we replicate BWY's main results but sort stocks into decile portfolios using LST's CSR scores instead. Even under LST's scoring methodology, the rolling-window risk-adjusted abnormal returns (FFC four-factor alphas) of the long-short CSR investment strategy remain positive on average during favorable economic periods and are close to zero during downturns. Figure OA1 further illustrates that, although the long-short portfolio formed using LST's CSR scores generates larger alphas during the GFC than the portfolio formed using BWY's CSR scores, the alphas are still negative in this period. Overall, these results imply that differences in CSR scoring alone are unlikely to explain the contrasting conclusions in LST and BWY.

3.3.2. Empirical Model

Another key distinction between LST and BWY lies in their empirical models. LST employ a cross-sectional predictive regression, using each firm's 2006 CSR score to forecast its abnormal returns during the GFC period — approximately two years later. By focusing on a single cross-section, LST's approach assesses the return predictability of pre-crisis CSR measures. In contrast, BWY adopt a real calendar-time portfolio strategy, continuously updating CSR information across multiple years when constructing portfolios. In doing so, BWY's framework simulates an investor who regularly rebalances her portfolios based on the most recent CSR data. Both methods have merit for examining whether CSR performances predict stock returns, but they emphasize different aspects of the investment decision-making process.

3.3.3. *Exposures to Systematic Risks*

LST and BWY take notably different approaches to adjust for risk exposures in their empirical analyses. In LST's cross-sectional regressions, the authors control for FFC four-factor betas that are estimated over the five years preceding the GFC, effectively imposing the exposures to systematic risk factors to be constant and not affected by the crisis.[15] BWY, on the other hand, evaluates the performance of a long-short CSR portfolio by using rolling-window FFC four-factor alphas. In any month, the FFC alphas are derived from betas that are estimated using the most recent three-year data. Consequently, BWY's approach implicitly allows stocks' exposures to systematic risk factors to change and update across different economic episodes.

In our view, neither assumption is inherently right or wrong; each has merits and reflects a plausible viewpoint on how researchers/investors should account for risk exposures. Moreover, as discussed above, the existing literature does not prescribe a single, universally accepted method of controlling for time-varying risk exposures, so both modeling choices represent plausible approaches.

Nevertheless, differences in these assumptions may partly explain why LST and BWY reach divergent conclusions. For instance, if high-CSR stocks experience significantly stronger exposures to systematic risk factors that did poorly in a period, BWY's method would capture this dynamics by estimating and updating betas with the most recent data, and the estimated FFC alphas would decrease.

To test this hypothesis, we conuduct several analyses to investigate whether CSR performance itself is indeed associated with time-varying risk exposures. In Table 2, following LST's cross-sectional model, we regress changes in FFC four-factor betas — calculated as the difference between betas measured at the end of March 2009 and those measured at the end of July 2008 — on firms' 2006 CSR scores. In Panel A, we find that higher CSR scores

[15] Consistent with this assumption, one of the main dependent variables in LST is the GFC-period abnormal returns implied by the CAPM, and the CAPM betas are estimated using pre-crisis data and held fixed during the GFC.

predict a statistically significant and economically meaningful decrease in the market betas and an increase in the value betas during the GFC. In Panel B, once we control for firm-level characteristics, higher CSR scores still predict lower market betas, but the magnitude is much smaller (-0.0235 vs. -0.0797) and the coefficient is no longer statistically significant. However, higher CSR scores again predict a significant increase in the value betas.

Next, we use BWY's calendar-time portfolio approach and explore the risk exposures of the long-short CSR investment strategy. In Table 3, we regress the monthly returns of the High − Low (top decile − bottom decile) CSR portfolio on the FFC four factors (MktRF, SMB, HML, UMD), indicators for the GFC and bad economic times, and interactions between these indicators and risk factors. Consistent with the findings in Table 2, the results show that the High − Low CSR portfolio exhibits a positive and significant exposure to the value factor during the GFC and other challenging economic periods.[16] The magnitude of the increase in value betas during bad times and the GFC is also sizable.[17]

Finally, Figure 2 plots the rolling-window FFC four-factor betas of the same long-short CSR portfolio. Again, we observe that during the GFC and the subsequent downturns, the spread portfolio exhibits a pronounced increase in exposure to the value factor.

Taken together, this evidence suggests that CSR returns have dynamic exposures to systematic risks, in line with the implications from a large asset pricing literature on conditional and time-varying betas (Jagannathan and Wang, 1996; Fama and French, 1997; Ferson and Harvey, 1999; Lettau and Ludvigson, 2001; Lustig and Van Nieuwerburgh, 2005; Zhang, 2005; Santos and Veronesi, 2006). We also document that firms with higher CSR scores become more exposed to systematic risk along the value dimension during crisis, which, according to Fama and French (1998), is linked to distress, an important feature during the GFC. During this period in our sample, the cumulative return of the value factor is -15.7%. Therefore, accounting for CSR returns' surge in loadings on the value factor (as

[16]Note that the estimated coefficient of the indicator for bad economic times is negative (-0.6751) and statistically significant, which aligns with BWY's findings.

[17]During good economic times, as defined by the CAPE ratios in BWY, the High − Low CSR portfolio has a value beta of -0.2557, whereas it has a value beta of 0.3093 (-0.2557 + 0.5650) during bad economic times.

in BWY) would likely drag down FFC alphas, which partly explains why they are negative during the GFC.

4. Mechanisms Behind CSR Returns Dynamics

In this section, we examine how aggregate cash flow (CF) news and discount rate (DR) news — two fundamental drivers of stock market return movements — affect the returns of the CSR investment strategy. We conduct this exercise for two reasons.

First, our findings in Section 3.3.3 suggest that the returns of the CSR investment strategy exhibit time-varying exposures to systematic risks. This naturally raises the question of what might account for these changing exposures. A large literature in asset pricing adopts the CF vs. DR framework to explain equity premiums and anomaly returns (e.g., the value premium)(Bansal and Yaron, 2004; Campbell and Vuolteenaho, 2004; Barberis et al., 2005; Santos and Veronesi, 2005; Da and Warachka, 2009; Campbell et al., 2010; Chen et al., 2013; Lochstoer and Tetlock, 2020). By distinguishing between these two channels, we aim to better understand the economic drivers of CSR returns and their relationship to systematic risk exposures.

Second, interpreting CSR returns through the lens of CF and DR shocks offers valuable insights into the mechanisms behind the dynamic link between CSR and stock returns (why CSR matters for stock performance). LST posit that good CSR performances improve firm fundamentals during the GFC and thus explain stronger returns.[18] In their view, firms' socially responsible practices create social capital, fostering trust among stakeholders and shareholders, who in turn may be more willing to support the firm during economic downturns. For instance, employees may reciprocate fair treatment and a positive work environment by working harder to create value for the firm in challenging times, and

[18]We successfully replicate LST's results showing that higher CSR scores predict improved firm fundamentals (e.g., ROA, gross margins, sales growth) during the GFC. These replication results are available upon request.

investors might hold or even increase their positions in high-CSR firms, viewing them as more reliable under uncertainty or low-trust periods. Consequently, LST suggest that both CF and DR channels could contribute to the link between CSR and stock returns during crisis.

In contrast, BWY characterize high-CSR stocks as a sort of "luxury good," appealing to investors primarily when disposable resources are in great supply, such as in good economic times. Under economic stress, however, investors become more constrained and therefore are likely to reduce their positions in these stocks. BWY also discover no substantial difference in fundamentals between high- and low-CSR firms, implying that the return dynamics of CSR investing may be driven predominantly by DR shocks.

Against this backdrop, we directly test how CSR returns relate to aggregate CF and DR shocks. We also investigate the flows into socially responsible (SR) mutual funds during different economic periods to shed further light on investor demand for CSR.

4.1. Aggregate CF News, DR News, and CSR Returns

We rely on Lan and Wermers (2024) to construct monthly time series of aggregate (market) CF news and DR news from IBES sell-side analyst forecasts. As described in Section 2.3, we first obtain unexpected market returns and aggregate CF news from IBES, then use the Campbell and Shiller (1988) decomposition identity to back out the aggregate DR news component.

Table 4 Panel A presents summary statistics of the key variables. In line with Campbell et al. (2013) and Lan and Wermers (2024), aggregate DR news exhibits greater volatility (i.e., a higher standard deviation) than aggregate CF news. Panel B shows that aggregate DR news is positively correlated with aggregate CF news and negatively correlated with both unexpected market returns and excess market returns, consistent with the implications of the decomposition identity. Additionally, the returns of the High − Low CSR portfolio (constructed using BWY's methodology) are positively correlated with both CF and DR

news in this univariate setting. The correlation with DR news is stronger (0.1404), compared with CF news (0.0671).

To analyze how CSR returns respond to market cash flow and discount rate components, we conduct monthly time-series regressions, regressing the long-short CSR portfolio returns on aggregate CF news, DR news, or both. The results, reported in Table 5, indicate that the CSR returns are driven primarily by shocks in aggregate DR news rather than CF news. In Columns (1) and (2), we use CF and DR news separately as predictors; both have a positive relationship with CSR returns, but DR news shows a larger and statistically significant effect. When we include both CF and DR news in Column (3), the coefficient on CF news becomes negligible, whereas the coefficient on DR news remains strongly positive and statistically significant. A one standard deviation increase in aggregate monthly DR news (6.9061%) is associated with 0.3667% increase in monthly CSR returns ($6.9061\% \times 0.0531$). In Column (4), we first show that CSR returns decrease in bad times, as defined by BWY. Moreover, we further find that the CSR spread portfolio's exposure (beta) to DR news is concentrated entirely in this period. The magnitude of the coefficient of aggregate DR news in bad times (0.1103) is around twice as large as the baseline estimate (0.0531) in Column (3).

Overall, our findings suggest that aggregate DR news drives the returns of the CSR investment strategy. In other words, shocks to investor sentiment and preferences — rather than stock market-level cash flows — appear to be the primary mechanism through which CSR influences stock returns. This interpretation aligns with recent research that incorporates investors' non-pecuniary preferences in ESG investing (Pástor et al., 2021, 2022; Pedersen et al., 2021). Moreover, we find evidence that is consistent with both BWY and LST's hypothesis. CSR returns decrease during bad economic times, which is consistent with CSR being a "luxury good" suggested by BWY. At the same time, when aggregate DR news increases, required returns tend to rise while investor sentiment declines, CSR spread portfolio returns move upward. This implies that high-CSR stocks outperform low-CSR

stocks specifically when the market declines for reasons unrelated to fundamentals. This corroborates LST's hypothesis on shareholder trust: there could exist investors that support trustworthy high-CSR firms during downturns.

4.2. Aggregate CF News, DR News, and Flows to SR Funds

So far, our findings point toward time-varying investor demand for high-CSR stocks. Both LST and BWY posit that investor demand can drive the returns of CSR-oriented firms across different economic episodes. However, they offer contrasting views: LST predict that some investors continue to trust high-CSR firms during bad economic times, while BWY suggest that investors reduce their exposure to high-CSR firms when market valuations deteriorate.

In this section, we test these hypotheses directly by examining mutual fund flows within the cash flow (CF) and discount rate (DR) news framework described earlier. We measure investor demand for high-CSR stocks using flows to socially responsible (SR) mutual funds. As detailed in Section 2.4, we calculate each fund's portfolio-level CSR score based on the CSR scores of its previous quarter's holdings. We then rank and classify funds whose portfolio-level CSR scores fall in the top decile as SR mutual funds.

We begin by assessing what happens to flows into SR mutual funds during bad economic times. To this end, we estimate panel regressions of monthly fund flows on (i) an indicator for SR mutual funds (SR_MF), (ii) an indicator for bad economic times as defined by CAPE ratios in BWY (Bad_Times), (iii) the interaction SR_MF $\times$ Bad_Times, and (iv) various fund-level control variables (previous-month fund flows, the log of total net assets, net returns, and the expense ratio). We also include fund and year fixed effects, and cluster standard errors at the fund and year level.

Column (1) of Table 6 shows that, during bad economic times, SR mutual funds experience net outflows relative to non-SR mutual funds. The coefficient on SR_MF $\times$ Bad_Times is negative and statistically significant at the 10% level. This implies that investor demand

for high-CSR stocks wanes when valuations are low, consistent with BWY's view that high-CSR stocks function as "luxury goods" that investors are less willing to hold under adverse market conditions.

However, once we account for aggregate CF news, DR news, and their interactions with SR_MF and Bad_Times, we uncover a more nuanced pattern (Column (2) Table 6). While SR funds do experience outflows during bad times, these outflows are mitigated in response to non-fundamental shocks (i.e., changes in investor sentiment rather than cash-flow fundamentals) during market downturns. Specifically, the coefficient on the triple interaction term DR_News $\times$ SR_MF $\times$ Bad_Times is positive and significant, indicating that at least some investors support or trust high-CSR stocks during downturns by refraining from divestment. In contrast, the coefficient on CF_News $\times$ SR_MF $\times$ Bad_Times is not statistically significant, suggesting that flows into SR mutual funds do not respond as strongly to fundamentals-based shocks during bad economic times.

Overall, our results on fund flows appear to be fully consistent with returns patterns documented in Table 5. These results lend support to both LST and BWY's arguments. By linking SR fund flows to CF and DR news, we demonstrate that investors' demand for high-CSR stocks depends not only on whether the market is in a "bad time", but also on changes in market fundamentals (CF news) and sentiment (DR news) during a "bad time".

5. Conclusion

Whether high CSR firms have better stock returns performance during economic downturns remains surprisingly an open question. In the finance literature, both Lins, Servaes, and Tamayo (2017) (LST) and Bansal, Wu, and Yaron (2022) (BWY) study this question but arrive at opposite findings. In this paper, we examine the relationship between CSR and stock returns over time by revisiting evidence in LST and BWY and investigating the underlying mechanisms.

We begin by demonstrating that the results in LST and BWY can indeed be replicated if we follow their respective methodologies, indicating that both studies are reproducible and each provides valid insights. We then investigate several factors that may explain their differing conclusions, including variations in how CSR scores are constructed, empirical specification choices (cross-sectional predictive regressions vs. calendar-time portfolio approach), and assumptions on systematic risk exposures. We find that long-short CSR portfolios display dynamic risk exposures across economic regimes, with a pronounced increase in their tilt toward the value (HML) factor during downturns.

Next, to understand the drivers of CSR returns over time, especially in crises, we employ the classic framework on cash flow (CF) and discount rate (DR) news, two sources of stock return variations. In this setting, while LST and BWY put forward distinct economic narratives tied to CF and DR dynamics, our direct analysis shows that CSR returns predominantly load on aggregate DR shocks, rather than CF shocks. This effect is concentrated in bad times, suggesting that changes in investor sentiment and preferences are central to explaining the returns of CSR strategies. Specifically, we observe that as DR shocks intensify and investor sentiment deteriorates, long-short CSR portfolio returns increase, consistent with some shareholders demonstrating trust in high-CSR firms under adverse conditions.

Finally, to further assess the role of investor sentiment and demand, we examine flows to socially responsible (SR) mutual funds (versus non-SR funds). We find that SR funds generally experience more outflows during bad economic times, in line with BWY's "luxury good" perspective. However, when we incorporate aggregate CF and DR news into our analysis, a more nuanced picture emerges: outflows from SR funds decrease in response to heightened DR shocks in bad times. This pattern is consistent with LST's "trust" perspective and indicates that at least some investors retain or even strengthen their positions in high-CSR holdings during crises, particularly when the market is driven down by sentiment-based rather than fundamental forces.

Overall, our paper adds to the literature on sustainable investing by offering additional evidence on the time-varying CSR and stock return relationship. Moreover, we use classic asset pricing framework (CF news vs. DR news) and show that investor sentiment and preferences are important in explaining CSR returns, and investor trust is indeed present to support CSR returns during downturns.

References

Asquith, P., Mikhail, M. B., Au, A. S., 2005. Information content of equity analyst reports. Journal of financial economics 75, 245–282.

Bansal, R., Wu, D., Yaron, A., 2022. Socially responsible investing in good and bad times. The Review of Financial Studies 35, 2067–2099.

Bansal, R., Yaron, A., 2004. Risks for the long run: A potential resolution of asset pricing puzzles. The journal of Finance 59, 1481–1509.

Barber, B. M., Morse, A., Yasuda, A., 2021. Impact investing. Journal of Financial Economics 139, 162–185.

Barberis, N., Shleifer, A., Wurgler, J., 2005. Comovement. Journal of financial economics 75, 283–317.

Campbell, J. Y., Giglio, S., Polk, C., 2013. Hard times. The Review of Asset Pricing Studies 3, 95–132.

Campbell, J. Y., Polk, C., Vuolteenaho, T., 2010. Growth or glamour? fundamentals and systematic risk in stock returns. The Review of Financial Studies 23, 305–344.

Campbell, J. Y., Shiller, R. J., 1988. The dividend-price ratio and expectations of future dividends and discount factors. The review of financial studies 1, 195–228.

Campbell, J. Y., Vuolteenaho, T., 2004. Bad beta, good beta. American Economic Review 94, 1249–1275.

Cao, J., Titman, S., Zhan, X., Zhang, W., 2023. Esg preference, institutional trading, and stock return patterns. Journal of financial and quantitative analysis 58, 1843–1877.

Chava, S., 2014. Environmental externalities and cost of capital. Management science 60, 2223–2247.

Chen, L., Da, Z., Zhao, X., 2013. What drives stock price movements? The Review of Financial Studies 26, 841–876.

Chen, L., Zhao, X., 2009. Return decomposition. The Review of Financial Studies 22, 5213–5249.

Da, Z., Warachka, M. C., 2009. Cashflow risk, systematic earnings revisions, and the cross-section of stock returns. Journal of Financial Economics 94, 448–468.

Deng, X., Kang, J.-k., Low, B. S., 2013. Corporate social responsibility and stakeholder value maximization: Evidence from mergers. Journal of financial Economics 110, 87–109.

Dimson, E., Karakaş, O., Li, X., 2015. Active ownership. The Review of Financial Studies 28, 3225–3268.

Edmans, A., 2011. Does the stock market fully value intangibles? employee satisfaction and equity prices. Journal of Financial economics 101, 621–640.

Elton, E. J., Gruber, M. J., Blake, C. R., 2001. A first look at the accuracy of the crsp mutual fund database and a comparison of the crsp and morningstar mutual fund databases. The Journal of Finance 56, 2415–2430.

Evans, R. B., 2010. Mutual fund incubation. The Journal of Finance 65, 1581–1611.

Fama, E. F., French, K. R., 1997. Industry costs of equity. Journal of financial economics 43, 153–193.

Fama, E. F., French, K. R., 1998. Value versus growth: The international evidence. The journal of finance 53, 1975–1999.

Ferson, W. E., Harvey, C. R., 1999. Conditioning variables and the cross section of stock returns. The Journal of Finance 54, 1325–1360.

Frankel, R., Kothari, S., Weber, J., 2006. Determinants of the informativeness of analyst research. Journal of Accounting and Economics 41, 29–54.

Hartzmark, S. M., Sussman, A. B., 2019. Do investors value sustainability? a natural experiment examining ranking and fund flows. The Journal of Finance 74, 2789–2837.

Heinkel, R., Kraus, A., Zechner, J., 2001. The effect of green investment on corporate behavior. Journal of financial and quantitative analysis 36, 431–449.

Hong, H., Kacperczyk, M., 2009. The price of sin: The effects of social norms on markets. Journal of financial economics 93, 15–36.

Hong, H., Kubik, J. D., Scheinkman, J. A., 2012. Financial constraints on corporate goodness. Tech. rep., National Bureau of Economic Research.

Hong, H., Shore, E., 2023. Corporate social responsibility. Annual Review of Financial Economics 15, 327–350.

Hwang, C. Y., Titman, S., Wang, Y., 2022. Investor tastes, corporate behavior, and stock returns: An analysis of corporate social responsibility. Management Science 68, 7131–7152.

Jagannathan, R., Wang, Z., 1996. The conditional capm and the cross-section of expected returns. The Journal of finance 51, 3–53.

Lan, C., Wermers, R., 2024. Cashflow timing vs. discount-rate timing: An examination of mutual fund market-timing skills. Management Science 70, 694–713.

Lang, M. H., Lundholm, R. J., 1996. Corporate disclosure policy and analyst behavior. Accounting review pp. 467–492.

Lettau, M., Ludvigson, S., 2001. Consumption, aggregate wealth, and expected stock returns. the Journal of Finance 56, 815–849.

Lins, K. V., Servaes, H., Tamayo, A., 2017. Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis. the Journal of Finance 72, 1785–1824.

Lochstoer, L. A., Tetlock, P. C., 2020. What drives anomaly returns? The Journal of Finance 75, 1417–1455.

Lustig, H. N., Van Nieuwerburgh, S. G., 2005. Housing collateral, consumption insurance, and risk premia: An empirical perspective. The Journal of Finance 60, 1167–1219.

Lys, T., Sohn, S., 1990. The association between revisions of financial analysts' earnings forecasts and security-price changes. Journal of accounting and economics 13, 341–363.

Masulis, R. W., Reza, S. W., 2015. Agency problems of corporate philanthropy. The Review of Financial Studies 28, 592–636.

Pástor, L., Stambaugh, R. F., Taylor, L. A., 2021. Sustainable investing in equilibrium. Journal of financial economics 142, 550–571.

Pástor, L., Stambaugh, R. F., Taylor, L. A., 2022. Dissecting green returns. Journal of financial economics 146, 403–424.

Pedersen, L. H., Fitzgibbons, S., Pomorski, L., 2021. Responsible investing: The esg-efficient frontier. Journal of financial economics 142, 572–597.

Riedl, A., Smeets, P., 2017. Why do investors hold socially responsible mutual funds? The Journal of Finance 72, 2505–2550.

Santos, T., Veronesi, P., 2005. Cash-flow risk, discount risk, and the value premium.

Santos, T., Veronesi, P., 2006. Labor income and predictable stock returns. The Review of Financial Studies 19, 1–44.

Shiller, R. J., 2005. Irrational exuberance, princeton, nj stiglitz, joseph e., 1990. In: *Symposium on bubbles, in: The Journal of Economic Perspectives*, vol. 4, pp. 13–18.

Zhang, L., 2005. The value premium. The Journal of Finance 60, 67–103.

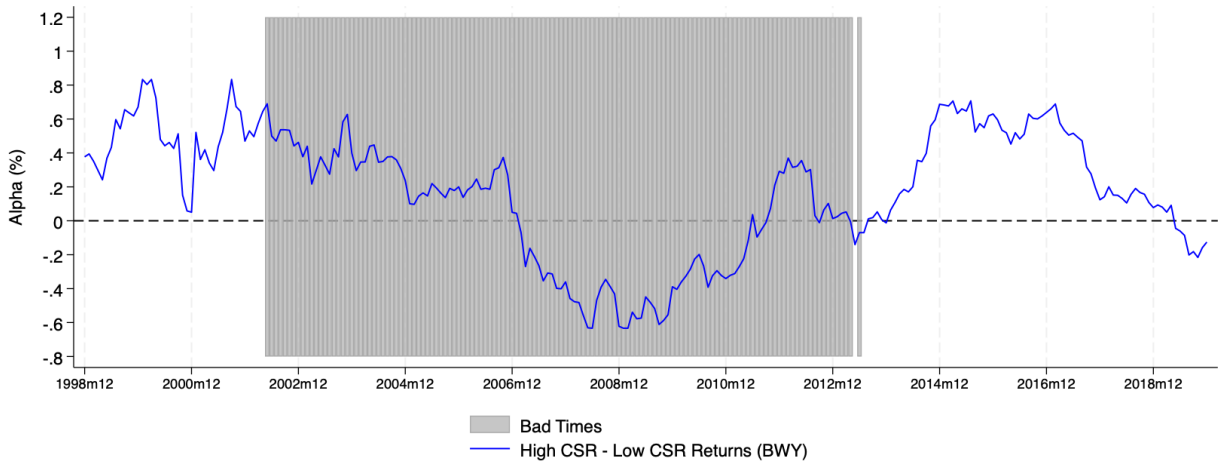


Fig. 1 Rolling-Window FFC Four-Factor Alphas of The CSR Investment Strategy. This figure reports the rolling-window FFC four-factor alphas of the High − Low (top decile − bottom decile) CSR portfolio. We sort stocks based on their previous-year CSR scores and create equal-weighted portfolio returns. The FFC four-factor alphas are estimated using 36-month rolling window regressions, and each stock is required to have at least 6 months of returns data available in CRSP. CSR scores are constructed following BWY's methodology. The FFC factor returns come from Kenneth French's website. The alphas are in percentage and winsorized at the 1st and 99th percentile. The grey-shaded area corresponds to the bad economic times as defined by the CAPE ratios in BWY.

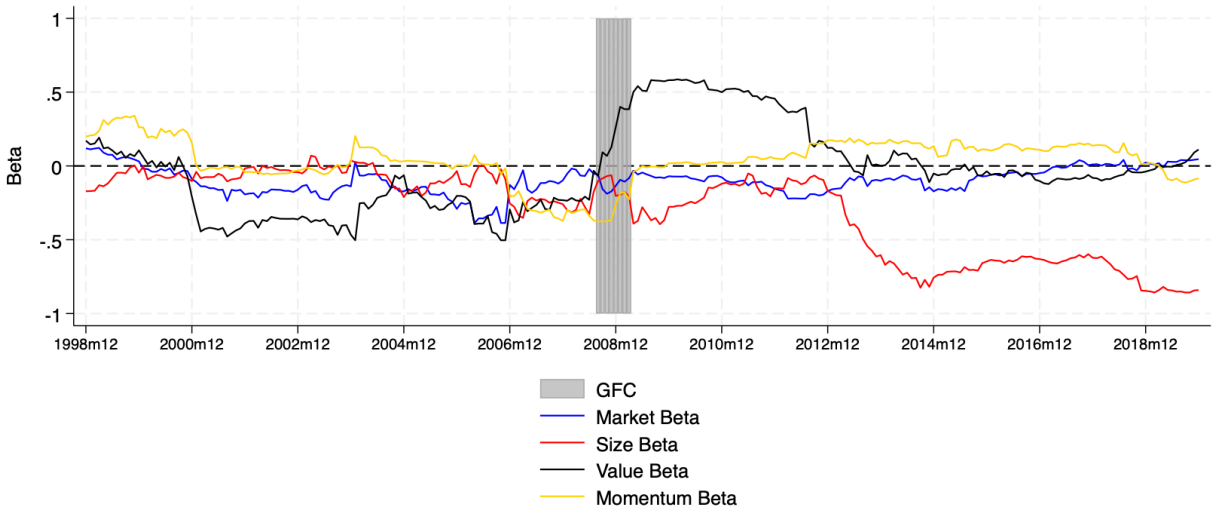


Fig. 2 Rolling-Window FFC Four-Factor Betas of The CSR Investment Strategy. This figure reports the rolling-window FFC four-factor betas of the High − Low (top decile − bottom decile) CSR portfolio. We sort stocks based on their previous-year CSR scores and create equal-weighted portfolio returns. The FFC four-factor betas are estimated using 36-month rolling window regressions, and each stock is required to have at least 6 months of returns data available in CRSP. CSR scores are constructed following BWY's methodology. The FFC factor returns come from Kenneth French's website. The betas are winsorized at the 1st and 99th percentile.The grey-shaded area corresponds to the GFC period (August 2008 - March 2009) as defined in LST.

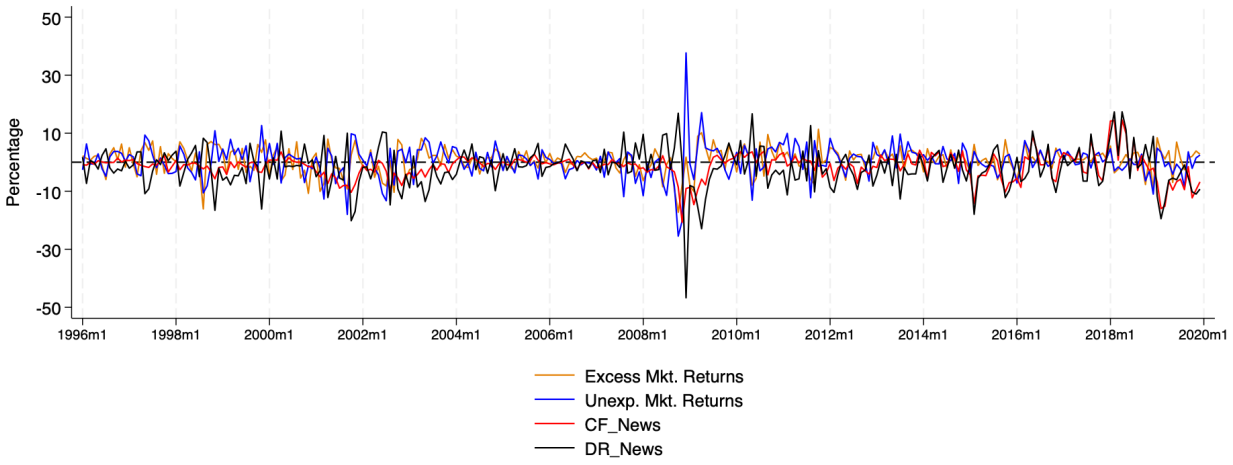


Fig. 3 Aggregate CF News, DR News, and Market Returns. This figure plots IBES-monthly aggregate cash flow (CF) news, aggregate discount rate (DR) news, aggregate stock market unexpected returns, and excess market returns (MktRf). CF news, DR news, and aggregate stock market unexpected returns are constructed following Lan and Wermers (2024). Excess market returns (MktRf) come from Kenneth French's website.

Table 1
Replication of LST and BWY's Main Results

This table presents the replication of the main results in LST and BWY. In Panel A, the dependent variable is either the raw return or the abnormal return over the GFC period of August 2008 - March 2009, as defined in LST. In Columns (2) and (4), the abnormal return is calculated based on the market model estimated over the fixed window covering August 2003 - July 2008. CSR is the key independent variable of interest that measures a firm's degree of corporate social responsibility as of the end of 2006 (constructed using LST's methodology). Firm characteristics, measured prior to the crisis period, include log market cap, cash over total assets, short-term debt over total assets, long-term debt over total assets, operating income over total assets, book-to-market ratio, momentum, and idiosyncratic risk. Additional control variables include factor loadings from the Fama-French-Cahart (FFC) four-factor model, which are estimated over August 2003 - July 2008. We also include a dummy that equals to 1 for negative book-to-market ratios and 0 otherwise. Control variables are winsorized at the 1st and 99th percentiles. Heteroskedasticity-consistent standard errors are reported in parentheses. Panel B reports the mean of monthly FFC four-factor alphas for the High CSR (top decile) and Low CSR (bottom decile) portfolios, as well as the High − Low CSR portfolio, during and outside of good economic states defined by the long-term P/E ratios (CAPE) in BWY. We sort stocks based on their previous-year CSR scores and create equal-weighted decile portfolio returns. The FFC four-factor alphas are estimated using 36-month rolling window regressions, and each stock is required to have at least 6 months of returns data available in CRSP. CSR scores are constructed following BWY's methodology. The FFC factor returns come from Kenneth French's website. The sample period is from December 1998 to December 2019. The Good Times include December 1998 - April 2002, May 2013, July 2013 - December 2019. Bad Times include May 2002 - April 2013 and June 2013. Standard errors are calculated using Newey-West procedures with 36 lags. The alphas are in percentage and winsorized at the 1st and 99th percentile. Significance at the 10% level is given by *; at the 5% level **, at the 1% level ***.

Panel A: Replication of LST

Dep. Var.:	Raw Ret. (1)	Abn. Ret. (2)	Raw Ret. (3)	Abn. Ret. (4)
CSR_LST_{2006}	0.0605***	0.102***	0.0448**	0.0747**
	[0.0166]	[0.0308]	[0.0174]	[0.0323]
Factor Loadings	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Firm Characteristics	No	No	Yes	Yes
Observations	1679	1679	1679	1679
Adj. R^2	0.159	0.274	0.202	0.309

Panel B: Replication of BWY

FFC4 α (%)	Good Times			Bad Times		
	Low	High	High − Low	Low	High	High − Low
CSR_BWY	-0.1364**	0.2369***	0.3728***	0.2246***	0.2258**	0.0012
	[0.0688]	[0.0575]	[0.0738]	[0.0696]	[0.1004]	[0.1468]
Observations	120	120	120	133	133	133

Table 2
CSR and Beta Change During the GFC: Cross-Sectional Regressions

This table reports the estimation results of cross-sectional regressions where the dependent variables are the changes in betas with respect to market, size, value and momentum. The change in beta is calculated as beta measured at the end of March 2009 minus the beta measured at the end of July 2008. CSR is the key independent variable of interest that measures a firm's degree of corporate social responsibility as of the end of 2006 (constructed using LST's methodology). In Panel B, we include firm characteristics, measured prior to the crisis period, as controls. These characteristics consist of log market cap, cash over total assets, short-term debt over total assets, long-term debt over total assets, operating income over total assets, book-to-market ratio, momentum, idiosyncratic risk. We also include a dummy that equals to 1 for negative book-to-market ratios and 0 otherwise. Control variables are winsorized at the 1st and 99th percentiles. Heteroskedasticity-consistent standard errors are reported in parentheses. Significance at the 10% level is given by *; at the 5% level **, at the 1% level ***.

Panel A: No Controls

Dep. Var.:	Mkt. Beta Δ	Size Beta Δ	Value Beta Δ	Mom. Beta Δ
	(1)	(2)	(3)	(4)
CSR_LST_{2006}	-0.0797**	-0.0420	0.1453***	-0.0355
	[0.0333]	[0.0341]	[0.0451]	[0.0251]
Observations	1679	1679	1679	1679
Adj. R^2	0.002	0.000	0.004	0.000

Panel B: With Controls

Dep. Var.:	Mkt. Beta Δ	Size Beta Δ	Value Beta Δ	Mom. Beta Δ
	(1)	(2)	(3)	(4)
CSR_LST_{2006}	-0.0235	-0.0286	0.0943**	-0.0402
	[0.0348]	[0.0388]	[0.0487]	[0.0274]
Observations	1679	1679	1679	1679
Adj. R^2	0.160	0.055	0.060	0.102

Table 3
CSR and Beta Change During the Bad Economic Times and GFC: Time-Series Regressions

This table reports the average FFC four-factor betas of the High − Low CSR portfolio. The sample period is from January 1996 to December 2019. Monthly time-series regressions are estimated, where the dependent variable is the monthly returns of the High − Low (top decile − bottom decile) CSR portfolio. CSR scores are constructed using BWY's methodology. We sort stocks based on their previous-year CSR scores and create equal-weighted portfolio returns. The factor returns (MktRf, SMB, HML, UMD) come from Kenneth French's website. All returns are in percentage. GFC is an indicator variable that equals to 1 for the period of August 2008 to March 2009, 0 for other periods. Bad_Times is an indicator variable, defined in BWY (based on the CAPE ratios), that equals to 1 for the period of May 2002 to April 2013 and June 2013, 0 for other periods. Standard errors are reported in the brackets. Significance at the 10% level is given by *; at the 5% level **, at the 1% level ***.

Dep. Var: CSR spread port. returns	(1)	(2)	(3)
GFC		0.0834	
		[1.5662]	
Bad_Times			-0.6751**
			[0.2671]
MktRf	-0.0210	-0.0642*	-0.0739
	[0.0344]	[0.0357]	[0.0472]
MktRf × GFC		-0.0608	
		[0.2463]	
MktRf × Bad_Times			-0.0251
			[0.0697]
SMB	-0.1714***	-0.1855***	-0.2084***
	[0.0450]	[0.0433]	[0.0497]
SMB × GFC		-0.5315	
		[0.5016]	
SMB × Bad_Times			-0.0259
			[0.1022]
HML	-0.0785*	-0.1681***	-0.2557***
	[0.0464]	[0.0478]	[0.0577]
HML × GFC		0.8756***	
		[0.1874]	
HML × Bad_Times			0.5650***
			[0.1022]
UMD	0.0182	0.0118	0.0338
	[0.0295]	[0.0287]	[0.0375]
UMD × GFC		-0.2990	
		[0.3050]	
UMD × Bad_Times			-0.0569
			[0.0561]
Constant	0.1127	0.1851	0.4729**
	[0.1407]	[0.1387]	[0.1857]
Observations	288	288	288
Adjusted_R^2	.0524	.134	.162

Table 4
Summary Statistics and Correlations

This table reports the summary statistics and the piece-wise correlations among High − Low CSR Returns (CSR Spread Portfolio Returns), Excess Market Returns (MktRf), Unexpected Market Returns, IBES-monthly aggregate cash flow (CF) news and discount rate (DR) news. The sample period is January 1996 to December 2019. CSR scores are constructed using BWY's methodology. We sort stocks based on their previous-year CSR scores and create equal-weighted portfolio returns. Excess market returns come from Kenneth French's website. Unexepcted market returns, CF News, and DR News are constructed following Lan and Wermers (2024). All returns are in percentage.

Panel A: Summary Statistics

Variables	Obs	Mean	Std. Dev.	P25	P50	P75
CSR Spread Port. Returns	288	.0687	2.396	-1.4622	.0928	1.6299
Excess Mkt. Returns	288	.6607	4.4021	-1.96	1.24	3.445
Unexp. Mkt. Returns	288	.3606	5.6193	-2.0297	.6901	3.505
CF_News	288	-1.4574	4.2266	-3.2977	-.8124	.7063
DR_News	288	-1.818	6.9061	-5.4666	-1.8086	2.0703

Panel B: Correlation Matrix

Variables	CSR Spread Port. Returns	Excess Mkt. Returns	Unexp. Mkt. Returns	CF_ News	DR_ News
CSR Spread Port. Returns	1.0000				
Excess Mkt. Returns	-0.0863	1.0000			
Unexp. Mkt. Returns	-0.1220	0.5604	1.0000		
CF_News	0.0671	0.0554	0.0368	1.0000	
DR_News	0.1404	-0.4221	-0.7912	0.5821	1.0000

Table 5
Aggregate CF News, DR News, and CSR Returns

This table reports the average market cashflow (CF) and discount rate (DR) betas of the High − Low CSR portfolio. The sample period is January 1996 to December 2019. Monthly time-series regressions are estimated. The dependent variable is the monthly returns of the High − Low (top decile − bottom decile) CSR score portfolios, constructed using BWY's methodology. We sort stocks based on their previous-year CSR scores and create equal-weighted portfolio returns. Standard errors are reported in brackets. CF_News and DR_News are aggregate cash flow news and discount rate news constructed from IBES, following Lan and Wermers (2024). All returns are in percentage. Bad_Times is an indicator variable, defined in BWY (based on the CAPE ratios), that equals to 1 for the period of May 2002 to April 2013 and June 2013, 0 for other periods. Standard errors are reported in brackets. Significance at the 10% level is given by *; at the 5% level **, at the 1% level ***.

Dep. Var.:	CSR Spread Portfolio Returns			
	(1)	(2)	(3)	(4)
Bad_Times				-0.6271**
				[0.2965]
CF_News	0.0381		-0.0125	0.0538
	[0.0334]		[0.0409]	[0.0563]
CF_News × Bad_Times				-0.0975
				[0.0860]
DR_News		0.0487**	0.0531**	-0.0166
		[0.0203]	[0.0250]	[0.0414]
DR_News × Bad_Times				0.1103**
				[0.0516]
Constant	0.1242	0.1572	0.1471	0.4275**
	[0.1493]	[0.1448]	[0.1488]	[0.1989]
Observations	288	288	288	288
Adjusted_R^2	.00103	.0163	.0131	.0392

Table 6

Aggregate CF News, DR News, and flows to Socially Responsible (SR) Mutual Funds

This table reports the impact of aggregate CF and DR news on flows into socially responsible (SR) mutual funds over time. The sample period is January 1996 to December 2019. The dependent variable is the current-month fund flows (in percentage). SR_MF is an indicator variable that equals to 1 if the fund has a portfolio-level CSR score that is in the top tercile at the end of the previous quarter. The portfolio-level CSR score is calculated as the value-weighted average of the CSR scores of the stock holdings. See Section 2.4 for more details on the construction of the mutual fund sample and variables. CF_News and DR_News are aggregate cash flow news and discount rate news (in percentage) constructed from IBES, following Lan and Wermers (2024). Both CF_News and DR_News are lagged one month. Bad_Times is an indicator variable, defined in BWY (based on the CAPE ratios), that equals to 1 for the period of May 2002 to April 2013 and June 2013, 0 for other periods. We control for the following fund-level characteristics: previous-month fund flows, previous-month log fund TNA, previous-month fund net returns, and the expense ratio. Fund and year fixed effects are included. Standard errors are clustered at the fund and year level. Significance at the 10% level is given by *; at the 5% level **, at the 1% level ***.

Dep. Var.:	Monthly Fund Flows	
	(1)	(2)
SR_MF	-0.0038	-0.0033
	[0.0695]	[0.0680]
Bad_Times	-0.1395	-0.2315
	[0.1441]	[0.1585]
SR_MF × Bad_Times	-0.1605*	-0.1549*
	[0.0813]	[0.0809]
CF_News		-0.0331*
		[0.0181]
CF_News × SR_MF		0.0273**
		[0.0126]
CF_News × Bad_Times		0.0180
		[0.0207]
CF_News × SR_MF × Bad_Times		-0.0259
		[0.0198]
DR_News		0.0480***
		[0.0133]
DR_News × SR_MF		-0.0208**
		[0.0085]
DR_News × Bad_Times		-0.0363**
		[0.0152]
DR_News × SR_MF × Bad_Times		0.0220**
		[0.0096]
Controls	Yes	Yes
Fund FE	Yes	Yes
Year FE	Yes	Yes
Observations	223776	223776
Adjusted R^2	.2	.201

Online Appendix to

What Drives the Dynamics of CSR Returns?

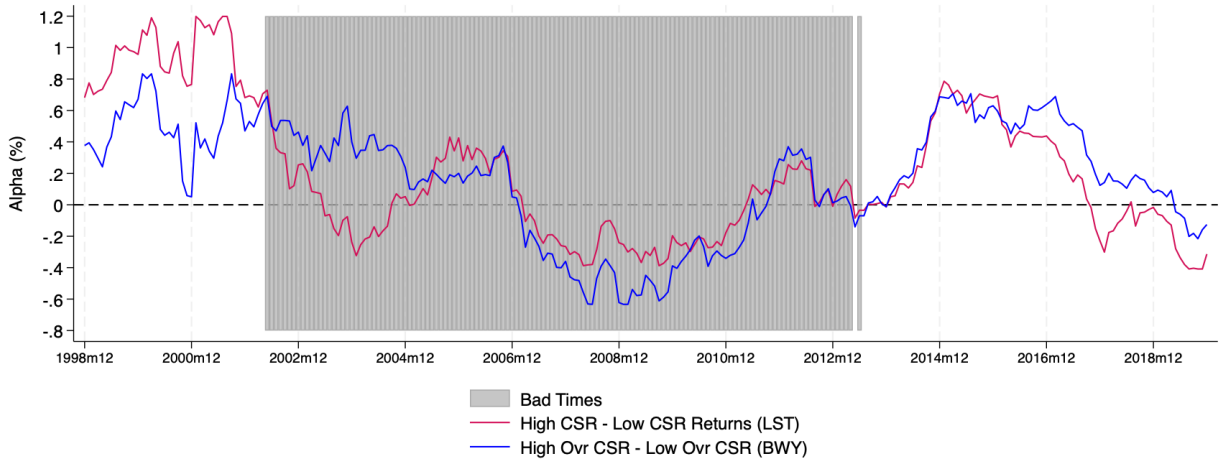


Fig. OA1 Rolling-Window FFC Four-Factor Alphas of The CSR Investment Strategy (Portfolios Created Using LST's CSR scores) This figure reports the rolling-window FFC four-factor alphas of the High − Low (top decile − bottom decile) CSR portfolio. We sort stocks based on their previous-year CSR scores and create equal-weighted portfolio returns. The FFC four-factor alphas are estimated using 36-month rolling window regressions, and each stock is required to have at least 6 months of returns data available in CRSP. CSR scores are constructed following LST's methodology. The FFC factor returns come from Kenneth French's website. The alphas are in percentage and winsorized at the 1st and 99th percentile. The grey-shaded area corresponds to the bad economic times as defined by the CAPE ratios in BWY.

Table OA1
Replication of BWY's Main Results Using LST's CSR Scores

This table presents the replication of the main results in BWY, using CSR scores from LST. This table is constructed exactly the same as in Table 1 Panel B, except that we use LST's CSR scores when sorting stocks into decile portfolios. The sample period is from December 1998 to December 2019. The Good Times include December 1998 - April 2002, May 2013, July 2013 - December 2019. Bad Times include May 2002 - April 2013 and June 2013. Standard errors are calculated using Newey-West procedures with 36 lags. The alphas are winsorized at the 1st and 99th percentile. Significance at the 10% level is given by *; at the 5% level **, at the 1% level ***.

FFC4 α (%)	Good Times			Bad Times		
	Low	High	High − Low	Low	High	High − Low
CSR_LST	-0.1278**	0.3069*	0.4357**	0.2333***	0.2200***	-0.0136
	[0.0549]	[0.1570]	[0.1787]	[0.0749]	[0.0651]	[0.0693]
Observations	120	120	120	133	133	133