Do Mutual Funds Dress their Greens?

Angie Andrikogiannopoulou*
King's College London

Filippos Papakonstantinou†
King's College London

Hao Song‡
King's College London

Chunyu Zu§
Renmin University of China

Abstract

We construct a novel measure to identify green window dressing in mutual funds by comparing the hypothetical returns-based ESG measure imputed from disclosed holdings of funds with the fund's actual returns-based ESG measure. We find that, since 2016, a substantial 15% of funds green-dress. We also find green dressing is associated with higher self-proclaimed ESG commitment, poor past performance, and weak oversight. Moreover, green-dressing funds attract comparable flows as genuinely green funds, indicating that investors are unable to differentiate between them. However, green dressers exhibit inferior performance relative to genuinely green funds. Our results provide valuable insights for investors as well as for regulators who seek to detect ESG-related misconduct.

Keywords: ESG, Greenwashing, Window Dressing, Green Dressing
Mutual Funds, Fund Flows, Fund Performance
JEL Classification: G11, G23

*King's Business School, Department of Banking & Finance, Bush House, 30 Aldwych, London WC2B 4BG, UK, e-mail: aandriko@kcl.ac.uk.

†Corresponding author. King's Business School, Department of Banking & Finance, Bush House, 30 Aldwych, London WC2B 4BG, UK, e-mail: fpapakon@kcl.ac.uk.

‡King's Business School, Department of Banking & Finance, Bush House, 30 Aldwych, London WC2B 4BG, UK, e-mail: hao.song@kcl.ac.uk.

§Renmin Business School, Department of Finance, Renmin University of China, Beijing 100872, PR China, e-mail: chunyuzu@ruc.edu.cn.

1 Introduction

Within the mutual fund industry, environmental, social, and governance (ESG) investing has gained growing attention over the last several years. In the beginning, this interest stemmed from socially responsible investors with a preference for making a positive impact on society, but investors are also increasingly being attracted by financial motives (Edmans and Kacperczyk, 2022). Driven by this increasing demand for sustainable investment, in recent years there has been a steady increase in the number of sustainable investment funds, reaching a total of about 600 by the end of 2022, according to Morningstar. However, the emergence of the phenomenon of greenwashing—the intentional divergence of proclaimed from actual ESG practices—undermines the credibility of ESG initiatives and raises skepticism among market participants about the authenticity of funds' ESG commitments. In this paper, we study a specific type of greenwashing behavior: funds' propensity to window dress their reported portfolio holdings to make them appear more heavily tilted toward ESG investments. That is, a fund may invest in high-ESG stocks and/or disinvest from low-ESG stocks right before its portfolio disclosure date and then reverse these trades right after. We refer to this cosmetic behavior of disclosing disproportionately higher holdings in high-ESG (often called green) stocks as *green dressing*.

Similar to other forms of window dressing, green dressing is not directly observable. This is because funds' holdings are not continuously observable by the public. According to the regulations of the U.S. Securities and Exchange Commission (SEC), mutual funds are only required to report to the SEC monthly portfolio holdings and, from these reports, only information for the end of each fund's fiscal quarter becomes publicly available.[1] The fact that information on funds' daily portfolio holdings is not public not only allows fund managers to engage in various forms of window dressing shortly before disclosure dates but also presents a challenge for investors to detect these manipulations.

In this paper, we propose a novel measure to address the challenge of identifying green dress-

[1]Since 2019, monthly portfolio holdings are reported on Form N-PORT; for detailed filing requirements, see https://www.sec.gov/files/formn-port.pdf. Before 2019, the SEC collected portfolio holdings on a quarterly basis on Forms N-Q and N-CSR, both of which are public filings.

ing. Our green-dressing measure relies on two ideas inspired by the work of Andrikogiannopoulou et al. (2022) on identifying on identifying funds' true ESG investing and the work of Agarwal, Gay and Ling (2014) on identifying performance-based window dressing. The first idea is that a portfolio's returns contain information about its tilt toward ESG investments. The second idea is that, over a period of time close to the disclosure date, a fund's returns—hence the ESG-related information they contain—should, in principle, be close to the returns of the portfolio disclosed by the fund, and so a systematic divergence between the two reveals green dressing. To implement these ideas, we first calculate a portfolio's returns-based ESG measure by conducting a Sharpe (1992) style analysis of its returns and calculating a weighted average of its exposure to stock portfolios sorted by their ESG scores. Then, we calculate a fund's green-dressing measure by comparing (i) the fund's *hypothetical* returns-based ESG measure computed from the returns of the portfolio disclosed by the fund and (ii) the fund's *actual* returns-based ESG measure computed from its actual returns. To be precise, our baseline measure of a fund's green-dressing is the difference between the hypothetical and actual returns-based ESG measures calculated using daily returns during the last calendar month of each disclosure period. High levels of discrepancy between the two measures would indicate that the portfolio held by the fund in the period leading *up to* the disclosure date is different from the portfolio held *on* the disclosure date, suggesting possible green-dressing behavior. In a simulation analysis, we find that, under a variety of reasonable assumptions, our measure is quite effective at capturing green-dressing activity.

Our analysis focuses on all U.S. domestic equity funds for the period January 2016 to June 2020. The beginning of this sample period coincides with the adoption of the Paris Agreement, a landmark event that largely raised both the public's awareness of climate change and investors' interest in ESG investing, hence it is during this period that fund managers are most likely to have an incentive to green-dress their holdings. Indeed, we start our analysis by verifying that, in our sample, investors respond to the ESG tilt of funds' holdings and hence that fund managers have an incentive to green-dress so they can attract more investment capital. Specifically, we calculate a fund's holdings-based ESG measure as the value-weighted ESG score of its portfolio holdings.

Using this measure, we find that, while investors do not respond linearly to increases in a fund's holdings-based ESG measure, they direct significantly more capital to funds when this measure ranks in the top decile within the fund's investment category. This suggests that investors focus on extreme levels of holdings-based ESG measures, which is consistent with the literature that highlights the importance of salience in investment decision making (Bordalo, Gennaioli and Shleifer, 2012; Bordalo, Gennaioli and Shleifer, 2013). It also consistent with the original finding by Hartzmark and Sussman (2019) that, when Morningstar introduced its sustainability ratings in early 2016, investors responded by investing significantly more in the top-rated funds.

Next, we use our green-dressing measure to estimate the prevalence of green-dressing behavior in the mutual fund industry. Specifically, for each fund, we conduct a paired t-test of the null hypothesis that mean difference between its hypothetical and actual returns-based ESG measure is zero. After controlling for false discoveries (see Storey, 2002), we find that the mean difference is positive—indicating green dressing—for about 15% of funds and negative—indicating the reverse behavior—for almost no funds. We note that our hypothetical returns-based ESG measure is backward-looking, as it utilizes the returns of the disclosed portfolio for the 30 days up until this portfolio is disclosed. This could raise the concern that our evidence that a substantial proportion of funds green-dress is driven by the increasing interest in ESG investments over time. To alleviate this concern, we also calculate a *forward-looking* version of our green-dressing measure, i.e., one that calculates the difference between hypothetical and actual returns-based ESG measures using daily returns during the calendar month *following* (rather than preceding) the disclosure date. Our results on the prevalence of green-dressing remain very similar.

We provide supplementary evidence of funds' green dressing by analyzing the return patterns of green and brown stocks around holdings disclosure dates, where green (brown) stocks are those with ESG score in the top (bottom) quintile. We hypothesize that the intensive purchasing of green stocks by green dressers before a disclosure date, followed by quick selling afterward, creates price pressure that drives green stocks' prices up before a disclosure date and down afterward; the reverse trend is expected for brown stocks. We empirically test this hypothesis using an event study

over a window of $(-3, +3)$ days around holdings disclosure dates. Indeed, we find that green stocks drift upward before disclosure dates, followed by a downward drift after disclosure. Specifically, green stocks, on average, earn a statistically significant cumulative average abnormal return (CAAR) of 0.156% during the three days before the disclosure date and a CAAR of −0.135% in the three days after. We find the opposite pattern for brown stocks' returns. Furthermore, in a panel regression analysis, we find a stock's ESG percentile rank is positively associated with its cumulative abnormal returns at quarter-end, but this relationship reverses at the beginning of the following quarter.

Our finding that green dressing behavior exists leads us to explore several additional research questions. First, which fund characteristics are associated with green dressing? Second, can investors distinguish between green-dressing funds and genuinely green funds? And, finally, is green dressing associated with superior or inferior performance?

We find that green-dressing behavior is more likely to be observed for funds with self-proclaimed ESG commitments and with poor past returns, which is consistent with the hypothesis that funds are more likely to engage in green dressing when they have a strong incentive to do so. In addition, green dressing is associated with funds that have higher expense ratios and those targeted to retail investors, consistent with the hypothesis that funds are more likely to green dress when the likelihood of detection is lower. These results indicate that green-dressing funds differ from non-green-dressing funds in meaningful and predictable ways, further validating our measure's ability to capture green-dressing behavior.

Next, we study how investors respond to green dressing. That is, do they recognize potential discrepancies between funds' disclosed and actual holdings, consequently penalizing green-dressing funds with lower flows, or not? We find that investors direct their flows to funds with exceptionally high holdings-based ESG scores regardless of whether these funds are green dressers or genuinely green. This means that, on average, investors are not sophisticated enough to detect green-dressing behavior. This result is consistent with extant literature in behavioral finance documenting investors' irrational reactions to cosmetic changes to investment products (e.g., Cooper,

Dimitrov and Rau, 2001; Cooper, Gulen and Rau, 2005; Andrikogiannopoulou et al., 2022).

Finally, we study whether green dressing is associated with higher or lower fund performance. The theoretical link between ESG risk and asset returns is ambiguous (see Giglio, Kelly and Stroebel, 2021; Pástor, Stambaugh and Taylor, 2021), and the empirical evidence and interpretations on this relationship are also inconclusive (see Bolton and Kacperczyk, 2021; Pástor, Stambaugh and Taylor, 2022; Atilgan et al., 2023). We find that the performance of funds extreme ESG holdings depends on whether they engage in green dressing or not. Specifically, we find that genuinely green funds—which continuously hold green stocks—outperform other funds. On the other hand, green-dressing funds—which only hold green stocks for a short period of time around disclosure dates—underperform the genuinely green funds. Our results hold for various performance measures, including market-adjusted returns, CAPM-adjusted returns, returns adjusted by Fama and French (1993) 3-factor model and by the Carhart (1997) 4-factor model.

Our paper contributes to several strands of the literature. First, it is related to the emerging literature on greenwashing in ESG investment within the asset management industry (e.g. Raghunandan and Rajgopal, 2022;Gibson Brandon et al., 2022;Kim and Yoon, 2023;Andrikogiannopoulou et al., 2022;Liang, Sun and Teo, 2022). Specifically, Kim and Yoon (2023) find that U.S. mutual funds that are the United Nations Principles for Responsible Investment (UNPRI) signatories do not have better holdings-based ESG scores than non-UNPRI funds. This lack of differentiation in holdings-based ESG scores, coupled with the absence of substantive ESG engagement or active ESG-oriented voting practices, is interpreted as evidence of greenwashing. Moreover, Andrikogiannopoulou et al., 2022 document widespread greenwashing in the mutual fund industry by revealing the discrepancies between ESG claims in fund prospectuses and their actual ESG-oriented holdings. Prior research relies on analyzing end-of-quarter portfolio holdings to evaluate funds' genuine level of ESG investment. However, as only end-of-quarter holdings are required to be disclosed, these disclosures may not reflect the unbiased composition of the investment portfolio over the reporting period. Our paper makes a key contribution by developing a method to detect this ESG-related holdings manipulation, through which we confirm the prevalence of green

window dressing in mutual fund industry.

In a contemporaneous study, Parise and Rubin (2023) examine a similar research question but use a different method. They measure green dressing by comparing fund's exposure to an ESG index shortly before and after mandatory portfolio disclosure date. This method, however, has two key limitations. First, it relies on a linear factor model to assess the ESG investment intensity, but multicollinearity among the factors may lead to imprecise statistical inferences (Jiang, Liang and Zhang, 2023). Second, the ESG exposures are estimated using a very short sample period—just 10 trading days in their main analysis—which undermines the reliability of the OLS estimates. Unlike Parise and Rubin (2023), our approach to determining a fund's actual ESG investment intensity leverages the quadratic regression framework embedded in Sharpe (1992) style analysis. Prior research has shown that this method is statistically robust and provides clear economic insights into fund portfolio allocation (Brown and Goetzmann, 1997; Fung and Hsieh, 1997; Annaert and Van Campenhout, 2007; Jiang, Liang and Zhang, 2022). We categorize U.S. stocks into quintiles according to their rank of MSCI ESG scores. Then, applying Sharpe (1992) style analysis to fund returns allows us to estimate the weights allocated by the fund among stock quintiles with varying ESG risk. In contrast, using a simple linear regression of fund returns on a long-only ESG index, as done by Parise and Rubin (2023), only captures exposure to a portfolio with the lowest ESG risk and thus misses the full spectrum of the fund's ESG positioning. In addition, in terms of research question, instead of merely concluding that green dressing exists in mutual fund industry due to systematically higher ESG exposure shortly before the mandatory disclosure date at the aggregate level as in Parise and Rubin (2023), we go further by providing evidence on the actual proportion of mutual funds engaging in green window dressing behavior.

Our paper also contributes to the long-standing literature on portfolio manipulations by fund managers. Prior studies document that asset managers window dress their holdings before quarter-end disclosures—by shifting into recently well-performing stocks to boost apparent performance (e.g. Lakonishok et al., 1991; Ng and Wang, 2004; Hu et al., 2014; Agarwal, Gay and Ling, 2014), reducing high-risk positions to appear safer (e.g. Musto, 1999; Agarwal, Daniel and Naik, 2011),

or purchasing more shares of already held stocks to artificially inflate portfolio value (e.g. Carhart et al., 2002; Ben-David et al., 2013; Wang, 2024). Our paper expands this strand of literature by evidencing a new form of window dressing: green dressing. Different from existing types of window dressing that aims at altering the portfolio's perceived risk or return profile, green dressing seeks to improve the portfolio's sustainability image. This practice exploits the growing investor demand for ESG investments, yet—like other window dressing strategies—presents a distorted representation of portfolio holdings.

Finally, our paper adds to the literature documenting that investor flows are irrationally influenced by the cosmetic actions of mutual funds. For instance, Cooper, Gulen and Rau (2005) find that funds experience significant cumulative abnormal flows after changing their names to reflect hot investment styles, even when they do not make corresponding changes to their holdings. Similarly, Solomon, Soltes and Sosyura (2014) show that investors direct their flows toward holdings with media-featured past-winning stocks, regardless of whether these stocks were included only shortly before the disclosure date.We extend this strand of literature by examining whether investors respond differently to high-ESG holdings resulting from green dressing, another cosmetic tactic by fund managers. Our findings show that investor flows do not distinguish between funds that engage in green dressing and those that do not. In other words, investors tend to chase funds with high ESG holdings, even if these holdings do not genuinely represent the fund's ESG investment intensity in the previous quarter. This also highlights the need for regulators to address green-dressing practices, which can mislead investors through the biased presentation of ESG holdings.

The rest of the paper is organized as follows. In Section 2, we introduce the data sources. In Section 3, we describe the construction of our green-dressing measures. In Section 4, we present evidence on green dressing. In Section 5, we present results on the determinants of green dressing and its impact on fund flows and fund performance. In Section 6, we conclude.

2 Data

To construct the variables we use, we rely on data from multiple sources that we outline below.

2.1 Fund information

Our sample begins in January 2016 and ends in June 2020. We constrain our sample to data from 2016 onward, as it is a period marked by heightened public awareness of climate change following the drafting of the Paris Agreement in December 2015. At this time, rising concerns over climate risk and expectations for regulatory actions prompted investors to pay closer attention to the ESG characteristics of investments. Thereby, we conjecture that, during this period, fund managers have greater incentives to manipulate the level of ESG investment in reported portfolio holdings with the aim of attracting socially responsible investors and of circumventing potential regulatory interventions.

We retrieve fund-level data from the Center for Research in Security Prices (CRSP) Survivorship-Bias-Free U.S. Mutual Fund Database. As our study focuses on U.S. mutual funds that predominantly invest in domestic equities, we exclude fund categories such as bond, money market, balanced, index, variable annuity funds and ETFs. We rely on a combination of CRSP's objective codes and flags along with keyword searches within the fund names to identify these fund categories and remove them from the sample.[2] The CRSP database provides an extensive range of share class-level attributes, including daily net returns, monthly total net assets (TNA), and annual expense ratios, portfolio turnover ratios, investment objectives, etc. It also updates information on fund portfolio holdings on a quarterly basis.[3] While the various share classes of a fund have identical holdings, they differ in TNA, fees, and net returns. Since our analysis is conducted at the

[2]The objective codes utilized in CRSP are aggregated from three distinct sources: Wiesenberger, Strategic Insight, and Lipper. ETFs and variable annuity funds are identified and removed based on the 'et_flag' and 'vau_fund' variables within CRSP. Index funds are filtered out using the 'index_fund_flag' variable, identifying specific terms like "INDEX", "S&P", "DOW 30", and "NASDAQ" in fund names, and by excluding funds from certain families (such as Dimensional Fund Advisors, Direxion, Potomac, ProFunds, and Rydex) known for index-like strategies.

[3]Notably, since 2008—encompassing our entire period of analysis—the CRSP mutual fund holdings database has demonstrated enhanced completeness and reliability compared to the Thomson Reuters s12 database.

fund level, we aggregate share-class data for TNA, fees, and net returns at the fund level. Next, we discuss in detail the definitions and calculations for various fund characteristics variables.

Fund size. We define the fund size as the log of fund-level TNA. We calculate a fund's TNA by summing the TNAs of its individual classes.

Fund return. We calculate a fund's daily net returns by computing the weighted average of returns for each class, using the beginning-of-month TNA values of each class as weights.

Expense ratio. Similarly, fund-level expense ratio is also a weighted average of expense ratios of its classes. As in Roussanov, Ruan and Wei (2021), we adjust a fund's expense ratio for amortized loans

Fund Flow. We calculate quarterly fund flows for fund i during fiscal quarter t as:

$$Flows_{i,t} = \frac{TNA_{i,t} - TNA_{i,t-1}\left(1 + r_{i,t}\right)}{TNA_{i,t-1}\left(1 + r_{i,t}\right)}, \tag{1}$$

where $TNA_{i,t}$ is total net assets and $r_{i,t}$ is the net return of fund i over quarter t.

Fund performance. Our main measure for fund performance is fund alpha estimated from the Carhart (1997) 4-factor model. Specifically, we first estimate daily alphas using daily fund and factor returns over the past 12 months. Subsequently, we sum up the daily alphas over a specified fiscal quarter to obtain the quarterly α. For robustness purposes, we also measure fund performance using the market-adjusted return, which is the fund return minus the return of the market portfolio. Additionally, we estimate fund alphas from alternative asset pricing models, i.e., the Capital Asset Pricing Model (CAPM) and the Fama and French (1993) 3-factor model. The market return is represented by the value-weighted return for all stocks listed on the NYSE, Amex and NASDAQ. The risk-free rate is represented by the one-month Treasury bill rate. Returns for factor-mimicking portfolios for size, value, and momentum factors are sourced from Kenneth French's online database.

Performance-based window dressing. We calculate the measure of performance-based window dressing, which is "backward holding return gap" defined in Agarwal, Gay and Ling (2014).

Other variables used in our analysis are *Fund age*; *Turnover*; *Fund for institutions*, defined as a dummy taking a value of 1 if a fund is targeted to institutional investors; *Fund style,* identified

by the CRSP investment objective code; and *Marketing expense*, defined as a combination of a fund's 12b-1 fee and front load. To reduce the effect of outliers, we winsorize fund characteristics variables at the 1% and 99% levels.

In addition to fund-level information, our analysis utilizes information on stock prices and returns, which is also available from CRSP.

Our final sample contains 2,144 unique funds, observed during the period January 2016 and June 2020. In Panel A of Table 1, we present summary statistics for the characteristics of all funds in our sample. The average fund is about 17.61 years old, has total net asset value of about 1.8 billion U.S. dollars, fund inflows of about 0.07% per quarter, and Carhart (1997) 4-factor alpha of about −0.72% per quarter, and is equally likely to be primarily targeted to institutional or to retail investors.

2.2 ESG ratings

We obtain firm-level ESG ratings data from MSCI, which we subsequently utilize to calculate fund-level ESG scores. MSCI's ESG ratings range from 0 to 10, with higher value indicating greater resilience to ESG-related risks. Given its extensive cross-sectional coverage and lowest noise levels for U.S. firms, as evidenced by Berg et al. (2022), we rely on the MSCI database as our primary source for ESG ratings. In addition, we obtain fund-level globe ratings data from Morningstar, which were first published in March 2016. The globe ratings are based on the percentile rank of the fund's sustainability score within its Morningstar category. Funds in the top 10% are rated Globe 5, whereas those in the bottom 10% are rated Globe 1. We also identify whether a fund is a UNPRI signatory by referencing the signatory directory provided on the UNPRI website, which provides information on each signatory's name and signature date.[4] Using this data, we calculate the following measures to capture a fund's ESG characteristics.

Holdings-based ESG measure. Our holdings-based ESG measure is calculated as the value-

[4]The UNPRI website is at https://www.unpri.org/signatories/signatory-resources/signatory-directory.

weighted average of the stock-level ESG scores. That is, we define

$$ESG^{H}_{i,t} := \frac{\sum_{s \in S_t} V_{i,s,t} ESG_{s,t}}{\sum_{s \in S_t} |V_{i,s,t}|}, \tag{2}$$

where $V_{i,s,t}$ is the value of fund i's quarterly reported holdings in stock s, $ESG_{s,t}$ is the ESG score for stock s in quarter t, and S_t is the universe of stocks for which we have an ESG score in quarter t.[5]

ESG commitment. This is an indicator variable that takes value 1 if, as of the end of a given fiscal quarter, a fund is a UNPRI signatory or its name contains ESG keywords.[6]

Border fund. As in Gantchev, Giannetti and Li (2024), this is an indicator variable that takes value of 1 if, at the end of a given fiscal quarter, the percentile rank of a fund's portfolio sustainability score is within ±2.5% of the cutoff between the Globe 4 and Globe 5 ratings.

3 Construction of the green-dressing measure

Our green-dressing measure builds on the returns-based ESG measure introduced by Andrikogiannopoulou et al. (2022). As we discuss below in detail, this returns-based ESG measure reflects the average level of ESG investment for a portfolio over a period of time. For each fund-quarter, we construct two sets of returns-based ESG measures. One is calculated using the daily returns of a hypothetical portfolio composed of a fund's quarter-end reported holdings as in Agarwal, Gay and Ling (2014), while the other is calculated using the fund's actual daily returns. Subsequently, we capture green-dressing behavior by calculating the difference between these two sets of measures.

3.1 Returns-based ESG measure

As in Andrikogiannopoulou et al. (2022), we calculate the returns-based ESG measure to capture a portfolio's average level of ESG investing throughout a period of time. This measure, in turn, builds on the returns-based style analysis of Sharpe (1992). In each fiscal quarter t, we first es-

[5]Note that $V_{i,s,t}$ is positive for long and negative for short positions. We divide by the sum of the absolute value of the holdings in each stock, because otherwise the measure is not well defined (e.g., if the total value of longs equals that of shorts).

[6]The list of ESG keywords we use is "ESG", "environmental", "social", "governance", "climate", "green", "sustainable", "ethic", "renewable", "alternative energy", "carbon", "human right", "SRI", "impact".

timate an asset-class factor model using the realized daily returns of the fund and of each asset class in the last month of that quarter.[7] Specifically, we estimate

$$r_{i,t,\tau} = \alpha_{i,t} + \mathbf{F}'_{t,\tau}\mathbf{w}_{i,t} + \varepsilon_{i,t,\tau}, \tag{3}$$

where $r_{i,t,\tau}$ is the net return of fund i in day τ of quarter t, $\mathbf{F}_{t,\tau}$ is the vector of returns of the various asset categories, $\mathbf{w}_{i,t}$ represents a collection of non-negative style weights on each asset category that sum up to 1, and $\varepsilon_{i,t,\tau}$ denotes the residual return which is due to selection rather than style. We categorize U.S. stocks into quintiles according to their rank of MSCI ESG scores, and include them in the model; the return of each ESG quintile is the value-weighted average return of all stocks in this quintile weighted by their market-capitalization. Consistent with Sharpe (1992), we also include other asset categories into the model: emerging markets stocks, developed markets stocks, government bonds, corporate bonds, and mortgage-backed securities.

We thus obtain an estimate, $\hat{w}^q_{i,t}$, of each fund's style weight on the q^{th} ESG quintile in each fiscal quarter t. Then, we assign scores q ranging from 1 to 5 to the five ESG quintiles, with 1 representing the lowest ESG quintile and 5 representing the highest ESG quintile. And, finally, we combine the estimated style weights and the scores to calculate the *actual* returns-based ESG measure for fund i in fiscal quarter t as

$$ESG^R_{i,t} = \sum_{q=1}^{5} q \cdot \hat{w}^q_{i,t}. \tag{4}$$

Then, we repeat the above calculations, replacing in Equation 3 the fund's actual returns with those of the portfolio that the fund reports in its quarter-end disclosure, to calculate the corresponding *hypothetical* returns-based ESG measure, denoted by $\widetilde{ESG}^R_{i,t}$.

An alternative approach to measuring ESG investment intensity is to estimate ESG factor loading using a linear factor model. However, this method often suffers from statistical inference issues arising from multicollinearity among different risk factors (Brown and Goetzmann, 1997; George Jiang, Bing Liang and Huacheng Zhang, 2023). In contrast, our method allows for a precise as-

[7] In Section 3.2 below, we explain why, in the context of constructing our green-dressing measure, it is preferable to use daily returns from the last month of a quarter rather than from the entire quarter.

sessment of portfolio weights allocated to stocks across different ESG-risk quintiles, rather than merely gauging exposure to stocks with extreme ESG risk levels. Moreover, it avoids the aforementioned statistical problems.

3.2 Measure of green dressing

Next, we construct the measure of green dressing for each fund in each fiscal quarter. In the spirit of Agarwal, Gay and Ling (2014), our green-dressing measure is defined as the difference between (i) the hypothetical returns-based ESG measure, $\widetilde{ESG}^R_{i,t}$, calculated from the returns of the portfolio consisting of a fund's disclosed end-of-quarter holdings, and (ii) the actual returns-based ESG measure, $ESG^R_{i,t}$, calculated from the fund's realized daily returns, i.e.,

$$GD_{i,t} = \widetilde{ESG}^R_{i,t} - ESG^R_{i,t}. \tag{5}$$

We note that, in our calculations, we use the daily returns in the last month of each fiscal quarter rather than of the entire fiscal quarter. This is because a fund's strategy, hence its holdings, may change throughout the fiscal quarter. Narrowing our focus to returns over a shorter period mitigates the impact of changes in fund strategy on the difference between the hypothetical and the returns-based ESG measures.

A contemporaneous study by Parise and Rubin (2023) measures green window dressing by comparing ESG loadings in the 10 days before and after the portfolio disclosure date. However, if fund managers do not immediately reverse the trades used to green their reported holdings after disclosure date (Wang, 2024), their method may fail to detect green dressers. Moreover, estimating factor loadings with so few observations introduces substantial noise into the OLS estimates. Our method, by comparison, overcomes these limitations. Indeed, in a simulation analysis, we find that, under a variety of reasonable assumptions for green-dressing behavior, our measure is quite effective at capturing green-dressing activity; we discuss more details on the analysis and its results in Section 3.3.

In addition to this continuous green-dressing measure, we define dummy versions—denoted

$GD_{i,t}^{\text{top-10\%}}$ and $GD_{i,t}^{\text{top-20\%}}$—that indicate whether a fund's continuous green-dressing measure, $GD_{i,t}$, is in the top 10% or the top 20%, respectively, within a fund's investment category as of the end of its fiscal quarter. These measures will be useful in our subsequent analyses, as they can help us uncover potential non-linear effects as suggested by the literature on salience.

3.3 Validity of the green-dressing measure

In this subsection, we conduct a simulation analysis to check the validity of our green-dressing measure. Let N be the number of funds in our fictitious data, where we set $N = 2{,}144$, which is the number of funds in our real data. Also let p be the proportion of funds that green-dress; we conduct three separate simulations with p set to 25%, 50% and 100%, respectively. Then, from our real data, we randomly pick $p \cdot N$ funds among those with high (i.e., above-median) holdings-based ESG measure, labeling them as *potential green dressers*. The remaining $(1 - p) \cdot N$ funds are classified as non–green dressers.

To generate the returns for the $p \cdot N$ green dressers in our fictitious data, we work as follows. The pre-disclosure window (i.e., the period during which funds are presumed to green-dress their holdings) is defined as the 10-trading-day period preceding the disclosure date for all funds. The returns over the pre-disclosure window, i.e., $(-10, 0)$, are those of the potential green dressers we picked from the real data. The remaining trading days in the month prior to the disclosure date are referred to as the intermediate window. For the returns over the intermediate window, i.e., $(-22, -10)$ [8]and post-disclosure window, i.e., $(0, +10)$, we use the returns of a randomly drawn matching fund that has relatively low (i.e., below-median) holdings-based ESG measure. By doing so, we ensure that the constructed returns of these funds mimic the behavior of green dressers, that is, holding high-ESG stocks only shortly before the disclosure date, while holding low-ESG stocks for the remainder of the disclosure period. The returns for the $(1 - p) \cdot N$ non-green-dressers in our fictitious data are the returns of the corresponding funds in the real data.

Next, we calculate our green-dressing measure defined in Equation 3 for funds in the fictitious

[8]Here, we quote the average number of trading days in month, though the exact number varies across months.

data. For each fund, we perform a paired t-test of the null hypothesis that the the mean difference between the hypothetical and actual returns-based ESG measure is zero. Then we calculate the proportion of funds with significantly positive and significantly negative mean difference and report the results in Table 3. In the leftmost columns of Panel A of Table 3, we present the results. For the fictitious data with 25% green-dressing funds, we observe that the proportion of funds with significantly positive mean difference between hypothetical and actual ESG measure—so the proportion of green dressers—is about 20% after accounting for false discoveries, using a test size of 5%. As expected, with a test size of 10% and 1%, respectively, we find higher and lower proportions, respectively, of 24% and 10%. We conclude that the estimated proportions are quite close to the true proportion of 25% green-dressing funds in this fictitious data. We also present the results for the fictitious data with 50% and 100% green-dressing funds in Panel A of Table 3. The performance of our measure here is a little worse, as respectively we estimate (after adjusting for false discoveries) that 32% and 60% of funds are green dressers using a test with size 5%, respectively. Despite being lower than the true proportion of 50% and 100%, these estimated proportions are much higher than those estimated in the fictitious data with 25% green-dressing funds. Furthermore, we note that the estimated proportions of funds with negative difference between hypothetical and actual holdings-based ESG measure (so exhibiting the opposite behavior to green dressing) is close to zero in all cases, as it should be.

For comparison, we also compute the green-dressing measure proposed by Parise and Rubin (2023), defined as the difference between pre- and post-disclosure ESG loadings, estimated using a linear factor model with a market factor and an ESG factor proxied by the Morningstar US Sustainability Index. Likewise, for each fund, we perform a paired t-test to assess whether the mean difference between pre- and post-disclosure ESG loadings is different from zero. The results are shown in the rightmost columns of Panel A of Table 3. We observe that, relative to our measure, this measure consistently identifies a substantially lower proportions of green dressers, across all test sizes and in every fictitious fund sample. Specifically, using a test size of 5%, it yields estimates of only about 4%, 10%, and 20% of green dressers for fictitious datasets in which the true

proportions of green dressers are 25%, 50%, and 100%, respectively. These results indicate that this measure substantially underestimates the proportion of green dressers in the sample.

Our previous data generation process assigns a uniform pre-disclosure window of 10 days to all green dressers. However, in practice, these windows may vary across funds. To better reflect this, we now randomly assign different pre-disclosure windows to different funds: 2 days, 5 days, 7 days, and 10 days, with each window applied to 25% of the sample. We then calculate both green-dressing measures for new fictitious datasets and rerun the aforementioned paired t-tests for each fund. The results are presented in Panel B of Table 3. We show that our green-dressing measure continues to identify a substantial proportion of green dressers, with estimates even slightly higher than those obtained when all funds were assigned the same pre-disclosure window. In contrast, the performance of Parise and Rubin (2023) deteriorates further when funds are assigned with varying pre-disclosure windows. These results highlight a key limitation of the Parise and Rubin (2023) measure: it requires the arbitrary selection of a short window before and after disclosure to estimate ESG loadings. If the chosen estimation window does not align with the actual period during which the manager engages in green dressing, the resulting estimates are subject to substantial noise. Taken together, these results indicate that our green-dressing measure can be quite effective at capturing green-dressing behavior.

3.4 Summary statistics

In Panel A of Table 2, we present summary statistics for the funds' ESG measures and characteristics. *Holdings ESG measure* is the value-weighted mean of the fund investments' ESG scores from MSCI; its mean/median is 4.76/4.84. *Returns ESG measure* is the fund's returns-based ESG score calculated using the Sharpe (1992) style analysis; its mean/median is 2.65/2.55. *ESG commitment* is an indicator variable taking value 1 if, at the end of the fiscal quarter, fund i is a UNPRI signatory or its name contains ESG keywords; in our sample, funds have made such a commitment to sustainability in about 20% of all fund-quarters. *Border Fund* is an indicator variable taking value 1 if, at the end of the fiscal quarter, the percentile rank of a fund's portfolio sustainability score

is within ±2.5% of the cutoff between the Globe 4 and Globe 5 ratings, as in Gantchev, Giannetti and Li (2024); 5% of all fund-quarter observations are classified as border funds. In Panel B, we present summary statistics for our green-dressing measure; its mean/median value is 0.05/0.02.

4 Evidence on green dressing

In this section, we show the empirical evidence on the motivation for and the existence of green dressing in the mutual fund industry. Specifically, we first study whether investors respond to the ESG tilt of funds' holdings and hence that fund managers have an incentive to green-dress so they can attract more investment capital. Then, for each fund, we perform a paired t-test to test whether there is significant difference between the means of their hypothetical and actual returns-based ESG measures. Finally, we provide further evidence of green dressing by investigating the patterns of cumulative abnormal returns of green and and brown stocks around fund disclosure dates.

4.1 Motivation for green dressing

For mutual fund investors, the ESG characteristics of the disclosed portfolio holdings serve, in principle, as a source of information for them to evaluate funds' ESG commitment and associated risks. If investors indeed pay attention to this information hence capital flows respond to it, fund managers might be motivated to manipulate their disclosed holdings' ESG characteristics. We examine the relation between fund flows and the holdings-based ESG characteristics, controlling for fund past performance and other fund characteristics. Our specification is

$$Flows_{i,t+1} = \alpha_{s,t} + \beta ESG^{H}_{i,t} + \gamma' \mathbf{x}_{i,t} + \varepsilon_{i,t}, \tag{6}$$

where $Flows_{i,t+1}$ is the quarterly net flow ratio of fund i in fiscal quarter $t+1$. As in Agarwal, Gay and Ling (2014), we focus on quarterly fund flows because fund managers can choose to delay the filing of Form N-PORT to the SEC (which contains the holdings information) by up to 60 days, and so quarterly flows may better capture investors' response to information contained in portfolio holdings; $ESG^{H}_{i,t}$ is the holdings-based ESG measure at the end of fiscal quarter t as

introduced in Section 2.1; and $\mathbf{x}_{i,t}$ is a vector of fund characteristics that might influence fund flows, including fund age, size, expense ratio, marketing expense, turnover ratio, funds for institutions, past performance and flows, the returns-based ESG measure and the performance-based window dressing measure introduced by Agarwal, Gay and Ling (2014). We also control for fund investment category-by-quarter fixed effects, with investment category identified by the CRSP investment objective code. Standard errors are clustered by fund and year-by-quarter.

The results from estimating this model are reported in column 1 of Table 4. We observe that the estimated coefficient on the holdings-based ESG measure is not statistically different from zero. This indicates either that investors do not care about the ESG characteristics of the disclosed portfolio, or possibly that investors' response is not linear. To test whether investors respond to *high* levels of ESG commitment, we define an indicator variable that takes value 1 if a fund's holdings-based ESG measure is in the top 10% within its investment category at the end of the fiscal quarter. We then re-estimate Equation 6, replacing the continuous holdings-based ESG measure with this indicator variable, and report the results in column 2 of the table. We see that investor flows respond positively and significantly (t-statistic of 2.39) to funds in the top decile of holdings-based ESG measure. Specifically, we estimate that, on average, top-10% ranked funds experience additional inflows of about 0.7% of their net asset value per quarter. This relationship serves as potential motivation for fund managers to green-dress their portfolio holdings.

We note that our results verify in our sample the original finding by Hartzmark and Sussman, 2019 that, when Morningstar introduced its sustainability ratings in 2016, investors responded by investing significantly more in the top-rated funds. More generally, our results are consistent with the extant literature in mutual funds showing that investors' investment decisions are not only influenced by funds' past performance, but also by information inferred from their portfolio holdings (see, e.g., Agarwal, Gay and Ling, 2014).

4.2 Mean difference between hypothetical and actual ESG measure

As introduced in Section 3.2, at the end of each fiscal quarter, we construct two sets of returns-based ESG measures for each fund: the hypothetical one (based on the returns of the fund's disclosed holdings at the fiscal quarter end) and the actual one (based on the fund's actual returns during the last month of the fiscal quarter). For each fund, we conduct a paired t-test to determine if there is a significant difference between the means of these two measures, with a positive difference corresponding to green-dressing behavior.[9] Thus, for each fund i, we test the null hypothesis

$$H_0^i : E\left[\widetilde{ESG}_{i,t}^R - ESG_{i,t}^R\right] = 0. \tag{7}$$

Next, we calculate the proportions of funds for which we reject this null hypothesis, and present the results in Panel A of Table 5. We note that, since we conduct a large number of hypotheses tests, we expect that some will be rejected even if all nulls are true. For example, with a test size of 5%, we would reject the null for about 5% of funds if all nulls were true; half of these rejections (so 2.5%) would correspond to a positive and half to a negative mean difference. Using a test size of 5%, we find that about 16% funds have significantly higher mean for the hypothetical than for the actual ESG measure so, after subtracting the proportion of "false discoveries", we calculate that about 14% of funds exhibit evidence of green dressing. As expected, we see in the table that using a test with larger (smaller) size hence higher (lower) power yields slightly higher (lower) estimated proportions of green-dressing funds. Specifically, using a test size of 10% and 1%, respectively, we estimate (after controlling for false discoveries) a proportion of about 16% and 9% green-dressing funds, respectively. Regardless, across test sizes, our finding that a substantial proportion of about 10% to 15% of funds green dress is quite robust. For comparison, the proportion of funds that have significantly *lower* mean for the hypothetical than for the actual ESG measure, hence that exhibit the opposite behavior to green dressing, is consistently close to zero after controlling for false discoveries, which is indeed reassuring.

We note that our hypothetical returns-based ESG measure is backward-looking, as it utilizes

[9]For this part of analysis, we keep funds with at least 10 quarterly observations over the sample period.

the returns of the disclosed portfolio for the 30 days up until this portfolio is disclosed. This could raise the concern that our evidence that a substantial proportion of funds green-dress is driven by the increasing interest in ESG investments over time. To alleviate this concern, we also calculate a *forward-looking* version of our green-dressing measure, i.e., one that calculates the difference between hypothetical and actual returns-based ESG measures using daily returns during the calendar month *following* (rather than preceding) the disclosure date. We repeat our analysis using this forward-looking measure and report the results in Panel B of Table 5. Our results here are very similar to those in Panel A of the table using the backward-looking measure: The proportion of funds that exhibit green-dressing behavior is slightly higher (e.g., 17% instead of 14%, for a test size of 5%), while the proportion of funds that exhibit the opposite behavior remains very close to zero.

To further check the validity of our results, we perform a placebo test. We compare the *actual* returns-based ESG measure for the last month of the previous quarter with that for the first month of the subsequent quarter. These should be similar for all funds, regardless of whether they green- (or brown-) dress or not. For each fund, we conduct a paired t-test of the null hypothesis that these two actual returns-based ESG measures have the same mean, and we present the results in Panel C of Table 5. We see that, after controlling for false discoveries, across all test sizes, the null is accepted for almost all funds. These results further validate our previous finding that a substantial proportion of mutual funds green-dress.

4.3 Evidence from the returns of green and brown stocks

In this subsection, we provide evidence of green dressing in the mutual fund industry by investigating the cumulative abnormal returns of green and brown stocks (i.e., high- and low-ESG stocks) surrounding fund disclosure dates. The idea is that, if fund managers practice green dressing by heavily purchasing (liquidating) green stocks shortly before (after) the disclosure dates, the resulting price pressure from intensive trading will drive the prices of green stocks upwards (downwards). Conversely, we should observe the opposite pattern in the price movements of brown stocks if fund managers substantially reduce their positions of brown stocks before the disclosure

date but reinvest in them immediately after. To examine this idea, we first employ an event-study methodology.

In this analysis, the events are the fund holdings disclosures, which occur at the end of each fund's fiscal quarter. Since the fiscal quarters for most funds in our sample align with the calendar quarters, the last trading day of each calendar quarter constitutes an event date. On each event date, we partition U.S. stocks into quintiles by sorting them in ascending order using their latest MSCI ESG score, and we classify as green stocks those in the top quintile and as brown stocks those in the bottom quintile. We then calculate CAARs from −3 to +3 days for both categories. Specifically, to calculate the CAAR for the green-stocks category, for each day τ in the event window $(-3, +3)$, we first estimate the abnormal daily return (AR) for green stock j using

$$AR_{j,\tau} = \left(r_{j,\tau} - r_{f,\tau}\right) - \left[\alpha_j + \beta_j^m \left(r_{m,\tau} - r_{f,\tau}\right) + \beta_j^{SMB} SMB_\tau + \beta_j^{HML} HML_\tau\right], \tag{8}$$

where r_j, r_f, and r_m are the daily returns of stock j, the risk-free asset, and the market portfolio, respectively, and *SMB* and *HML* are the daily returns of the small-minus-big and high-minus-low factor-mimicking portfolios, and where the parameters of the Fama and French (1993) 3-factor model are estimated using daily returns during the 150-day window ending 30 days before the event. Then, the average abnormal daily return (AAR) for N events in this category for each day τ is

$$AAR_\tau = \frac{1}{N} \sum_{j=1}^{N} AR_{j,\tau}. \tag{9}$$

Finally, we calculate the cumulative average abnormal return (CAAR) from the start to day τ of the event window as

$$CAAR = \sum_{\tau=-3}^{\tau} AAR_\tau. \tag{10}$$

We repeat the same procedure to calculate the CAARs for the brown-stocks category.

We report the the AARs for each day in the event window for both the green and brown stocks in Table 6, and we plot the CAARs for both categories in Figure 1. In Panel A of Table 6, we see that green stocks on average consistently earn positive abnormal returns from event day −2 to event day 0. As shown in Figure 1a, this implies an upward drift in CAARs leading up to the event

date, i.e., the disclosure date. In contrast, green stocks yield negative abnormal returns right after the event date, resulting in a downward drift in CAARs after the event date. Specifically, green stocks generate a positive CAAR of 0.156% leading up to the fund disclosure date, contrasted with a negative CAAR of −0.135% in the three days following it. These findings corroborate our earlier results on green-dressing behavior surrounding the disclosure dates. Furthermore, they illustrate that this behavior has asset pricing implications.

On the other hand, brown stocks exhibit the opposite pattern. As shown in Panel B of Table 6, The AAR is statistically significant and negative on event day −1 (−0.161%), but becomes statistically significant and positive on event day +1 (0.101%). Consequently, we observe from Figure 1b a decrease in CAARs leading up to the disclosure date, followed by a notable reversal thereafter. This suggests that funds divest from brown holdings prior to disclosure dates, but swiftly reinvest in these holdings after that.

Next, we investigate the patterns of stock returns at quarter-end and quarter-beginning periods under a panel regression setting. We begin by calculating the cumulative abnormal daily return (CAR) for each stock over 3 days preceding and following the last trading day of each calendar quarter, corresponding to event windows $(-2, 0)$ and $(+1, +3)$, respectively. If a substantial number of fund managers engage in green dressing around portfolio disclosure dates, we should observe a positive association between stock-level ESG performance and CAR prior to the disclosure date, and a negative association in the short period that follows. To test this hypothesis, we then estimate model 11 and 12

$$CAR(-2,0)_{j,t} = \alpha_{p,t} + \beta_1 ESGperformance_{j,t} + \gamma' \mathbf{x}_{j,t-1} + \varepsilon_{j,t}, \tag{11}$$

$$CAR(+1,+3)_{j,t} = \alpha_{p,t} + \beta_1 ESGperformance_{j,t} + \gamma' \mathbf{x}_{j,t-1} + \varepsilon_{j,t}, \tag{12}$$

where $CAR(-2,0)_{j,t}$ is the CAR earned by stock j over the 3 trading days prior to the end of quarter t, and $CAR(+1,+3)_{j,t}$ is the CAR earned by stock j over the 3 trading days following the end of quarter t. $ESGperformance_{j,t}$ is the percentile rank (in ascending order) of the most recent MSCI ESG score of stock j at the end of quarter t. Following Wang (2024), we control $\mathbf{x}_{j,t}$, which is

a vector of firm characteristics, including the logarithm of market capitalization, book-to-market ratio, return on assets, debt-to-equity ratio, net profit margin and past quarter returns. We also account for GICS-industry-by-quarter fixed effects.

Columns 1 and 2 in Panel A of Table 7 reports the results of estimating model 11 and 12, respectively. We observe that coefficient on $ESGperformance_{j,t}$ is positive and statistically significant (coefficient: 0.209, t-statistics: 2.795) in column 1, whereas it becomes significantly negative (coefficient: -0.132, t-statistics: -1.842) in column 2. These results are consistent with our hypothesis. That is, prior to the end-of-quarter disclosure, fund managers tend to purchase high-ranked ESG stocks aggressively while liquidating low-ranked ESG stocks. This demand-driven trading pressure results in higher CARs for high-ranked ESG stocks relative to their low-ranked counterparts. Following disclosure, however, funds unwind their positions in high-ranked ESG stocks and increase their holdings of low-ranked ESG stocks, producing a reversed relationship between stock's ESG percentile rank and CAR. We also construct an alternative measure to capture extreme stock-level ESG performance, defined as a dummy variable equal to 1 if a stock's MSCI ESG score ranks in the top 20% of the distribution for a given quarter. With this measure, we re-estimate model 11 and model 12 and report the results in Panel B of Table 7. We find that top-performing ESG stocks exhibit 0.122% higher CARs than other stocks in the three days preceding the quarter-end, whereas this outperformance disappears at the beginning of following quarter.

5 Analysis of green dressing

In this section, we start by studying which fund characteristics are associated with green-dressing behavior. This analysis is intended as an informal verification that our green-dressing measures truly capture green-dressing behavior. Next, we investigate how investors react to green dressing. That is, do they recognize the discrepancy between what funds claim to hold and what they actually hold, by responding differently to green dressers and genuinely green funds? If not, this would imply a clear benefit of green dressing, as higher flows lead to higher management fees. Finally, we check whether there is a difference in fund performance between green dressers and genuinely

green funds. On the one hand, as we have discussed previously, it is not clear ex ante whether green investments have higher or lower mean returns. On the other hand, the higher transaction costs incurred by green dressing could lead to lower returns.

5.1 Determinants of green dressing

In this subsection, we investigate the fund characteristics of green dressers. First, we hypothesize that funds are more likely to green-dress when they have strong incentives to do so. We anticipate that funds that have publicly stated their ESG commitments, either by signing the UNPRI or by including ESG keywords in their fund names, have more incentives to manipulate the perceived sustainability of their holdings in order to maintain their public image as responsible investors. Furthermore, motivated by the evidence by Hartzmark and Sussman (2019) that funds with the highest sustainability rating—Morningstar's Globe 5—attract more inflows, we conjecture that funds with rating near the cutoff between Globe 4 and Globe 5 are more likely to manipulate their holdings in an effort to be upgraded (or not to be downgraded). We also conjecture that overstating a fund's ESG commitments through green dressing could be an attention-grabbing strategy to divert investors' focus from poor past performance.

Additionally, we expect that funds are more likely to engage in green-dressing activities when there is a lower likelihood to be detected. This could be the case for funds with weaker governance, which could be proxied by higher expense ratios (see Gil-Bazo and Ruiz-Verdú, 2009). It could also be the case for funds targeted to retail investors, as they are less sophisticated in processing information (see Evans and Fahlenbrach, 2012).

Motivated by these conjectures, we estimate the model

$$\begin{aligned} GD^{\text{top}}_{i,t} =& \alpha_{s,t} + \beta_1 ESGcommitment_{i,t} + \beta_2 BorderFund_{i,t-1} + \beta_3 ExpenseRatio_{i,t} \\ &+ \beta_4 InstitutionalFund_{i,t} + \beta_5 PastReturns_{i,t} + \gamma' \mathbf{x}_{i,t} + \varepsilon_{i,t}, \end{aligned} \tag{13}$$

where $GD^{\text{top}}_{i,t}$ is an indicator variable that takes value 1 if our continuous green-dressing measure for fund i is ranked toward the top (10% or 20%) within its investment category at the end of fiscal

quarter t. $ESGcommitment_{i,t}$ is an indicator variable taking value 1 if fund i is a UNPRI signatory or its name contains ESG keywords in fiscal quarter t. $BorderFund_{i,t-1}$ is an indicator variable taking value 1 if fund i's portfolio sustainability rank at the end of fiscal quarter $t-1$ is within ±2.5% of the cutoff between the Globe 4 and Globe 5 ratings, as in Gantchev, Giannetti and Li (2024). $ExpenseRatio_{i,t}$ is the logarithm of the expense ratio charged by fund i in fiscal quarter t. $InstitutionalFund_{i,t}$ is an indicator variable equal to 1 if fund i is targeted to institutional investors. $PastReturns_{i,t}$ is the mean of fund i's monthly returns over the past 12-month period ending at the end of quarter t. $\mathbf{x}_{i,t}$ is a vector of other fund characteristics such as fund size, age, turnover ratio, and marketing and distribution fees. We also control for fund investment category-by-quarter fixed effects, and cluster standard errors by fund and year-by-quarter.

In Table 8, we report the results from OLS and logistic estimations of Equation 13, using two alternative definitions of our dummy green dressing measure—$GD_{i,t}^{\text{top-10\%}}$ and $GD_{i,t}^{\text{top-20\%}}$. The estimated coefficient on ESG commitment is significantly positive in all specifications. This is in line with recent literature showing that UNPRI signatories fail to deliver on their ESG commitments (see Gibson Brandon et al., 2022; Kim and Yoon, 2023). The coefficients on past returns are all significantly negative, indicating a negative relationship between green dressing and past returns. We also find evidence that funds with sustainability rating near the cutoff between Globe 4 and Globe 5 are more likely to green dress; estimated coefficient signs and magnitudes are consistent throughout, but the evidence is only statistically significant when using the $GD_{i,t}^{\text{top-20\%}}$ dummy. Moreover, we see that estimated coefficients on expense ratio and funds targeted on institutions are significantly positive and negative, respectively, confirming our hypothesis that funds are more likely to green dress when this activity is more likely to go undetected. Finally, as expected and consistent with the extant literature on window dressing in mutual funds (e.g., Agarwal, Gay and Ling, 2014), we find that portfolio turnover is strongly associated with green dressing.

Taken together, our results show that green dressing behavior is associated with stronger incentives to do so and a lower probability of being detected.

5.2 Do investor flows respond to green dressing?

In this subsection, we examine how investor flows respond to green dressing. If investors are sophisticated enough to detect funds' green-dressing activities, then the relationship between flows and the ESG characteristics of disclosed holdings should be weaker (or non-existent) for green-dressing than for genuinely green funds. On the other hand, if investors are deceived and misled by the cosmetic changes to fund holdings' ESG characteristics, then this relationship should be similar across both sets of funds. We test investors' reactions to green dressing by estimating the model

$$Flows_{i,t+1} = \alpha_{s,t} + \beta_1 ESG^{H,\text{top}}_{i,t} + \beta_2 GD^{\text{top}}_{i,t} + \beta_3 ESG^{H,\text{top}}_{i,t} \times GD^{\text{top}}_{i,t} + \gamma' \mathbf{x}_{i,t} + \varepsilon_{i,t}, \tag{14}$$

where $GD^{\text{top}}_{i,t}$ is our green-dressing dummy variable as above, and $ESG^{H,\text{top}}_{i,t}$ is a dummy variable indicating that fund i's holdings-based ESG measure in quarter t ranks in the top 10% within its investment objective in that quarter. If investors can correctly process the available information to identify the green dressers, then β_3—the coefficient on the interaction term $ESG^{H,\text{top}}_{i,t} \times GD^{\text{top}}_{i,t}$—should be negative. As in our other models, in this model we also control for various fund characteristics, including the returns-based ESG measure and the performance-based window dressing measure. We also control for investment category-by-quarter fixed effects and cluster standard errors by fund and year-by-quarter.

In columns 1 and 2 of Table 9, we present the results from estimating Equation 14 for alternative green dressing measures, i.e., the green dressing dummies based on the top 10% and 20% of the continuous green dressing measure. Across specifications, we see that the estimated coefficient on the dummy indicating extreme holdings-based ESG measure is significantly positive, while the coefficient on the interaction term $ESG^{H,\text{top}}_{i,t} \times GD^{\text{top}}_{i,t}$ is not statistically different from zero. This indicates that, while investors respond positively to funds with extremely high levels of holdings-based ESG measure, they cannot distinguish whether these funds are green dressers or not.

Overall, our findings show that investors do not uncover the inconsistency between what funds disclose about their holdings and what they actually invest in throughout the reporting period. This implies that investors have bounded rationality and cannot detect the cosmetic actions taken

by mutual funds, consistent with the findings of Cooper, Gulen and Rau (2005), and Andrikogiannopoulou et al. (2022).

5.3 Green dressing and fund performance

As discussed in the introduction, the link between ESG risk and asset returns remains ambiguous. Giglio, Kelly and Stroebel (2021) propose that the ESG risk premium is negative if government action to punish brown firms is more likely to take place in economic expansion where production and pollution are high. In contrast, Pástor, Stambaugh and Taylor (2021) suggest that the ESG risk premium is positive if investors dislike the harm caused by climate risk and require higher expected return to hold brown assets. However, they also suggest that green assets outperform brown assets if there are unexpectedly increasing climate change concerns. As a result, it is not clear whether we should expect green dressers to have lower returns (because they have higher transaction costs) or higher returns (because they avoid investing in green stocks) than genuinely green funds. Regardless, if we find that they have lower returns, this should weaken funds' incentive to green dress, and vice versa.

We estimate the model

$$Performance_{i,t+1} = \alpha_{s,t} + \beta_1 ESG^{H,\text{top}}_{i,t} + \beta_2 GD^{\text{top}}_{i,t} + \beta_3 ESG^{H,\text{top}}_{i,t} \times GD^{\text{top}}_{i,t} + \gamma' \mathbf{x}_{i,t} + \varepsilon_{i,t}, \tag{15}$$

where $Performance_{i,t+1}$ is fund i's performance in fiscal quarter $t+1$, and the other variables are as defined previously. We estimate this model using various measures of performance, specifically, market-adjusted returns as well as risk-adjusted returns based on the CAPM, the Fama and French (1993) 3-factor model and the Carhart (1997) 4-factor model. Across specifications, we control for past performance and various fund characteristics. We also include investment category-by-quarter fixed effects, and cluster standard errors by fund and year-by-quarter. If green-dressing funds outperform (underperform) genuinely green funds, then β_3—the coefficient on the interaction term $ESG^{H,\text{top}}_{i,t} \times GD^{\text{top}}_{i,t}$—should be positive (negative).

We report the results from estimating this model using our two dummy green-dressing mea-

sures ($GD_{i,t}^{\text{top-10\%}}$ and $GD_{i,t}^{\text{top-20\%}}$) in Panels A and B of Table 10, respectively. In both panels, and for all measures of fund performance, we see that the estimated coefficient on the top-10% holdings-based ESG indicator is significantly positive, while the estimated coefficient for the interaction term of top-10% holdings-based ESG indicator and the green-dressing indicator is significantly negative in all specifications. These results indicate that the performance of funds whose holdings-based ESG score is in the top decile depends on whether they are green dressers or genuinely green funds. Funds that consistently hold high proportions of green assets over time (i.e., they are genuinely green funds) outperform other funds. But funds that only invest in green assets heavily for a very short period of time around the disclosure dates (i.e., they are green dressers) do not outperform other funds. Indeed, we find that green dressers significant underperform green funds.

Notably, our finding on the outperformance of genuinely green funds further corroborates the recent findings of a positive relationship between ESG performance and asset returns in Pástor, Stambaugh and Taylor (2022) and Ardia et al. (2023).

6 Conclusion

In this paper, we investigate green window dressing within the mutual fund industry, i.e., the practice of tilting portfolio holdings toward green stocks and away from brown stocks right before holdings disclosure dates, to inflate investors' perception of the fund's ESG commitment. Since green dressing is not directly observable, we first develop a novel measure to identify funds' engagement in green-dressing activities. Our measure is the discrepancy between the hypothetical returns-based ESG measure imputed from a fund's disclosed holdings and the fund's actual returns-based ESG measure.

Subsequently, we use our measure to study green dressing, specifically, its prevalence, its determinants, and its implications for fund flows and fund performance. We find that a substantial 15% of all funds in our sample green-dress. Furthermore, we document that the cumulative ab-

normal returns of green (brown) stocks on average drift upwards (downwards) before holdings disclosure date, followed by a reversal afterwards. The return patterns of green assets and brown assets further corroborate the prevalence of green dressing. We also investigate the determinants of green dressing and its impact on fund flows and returns. Our results show that green dressing is associated with funds with self-proclaimed ESG commitment, poor past performance, and weak oversight. In terms of fund flows, we find that investors allocate similar amounts of capital to both green-dressing and genuinely green funds, suggesting that investors are unable to distinguish between them. On the other hand, when compared with genuinely green funds, green dressers have lower performance.

Overall, our findings uncover the prevalence of green window dressing in mutual funds, and provide valuable insights for investors, policy makers, and other market participants who seek to identify funds that overstate their ESG commitments through reported portfolio holdings.

References

Agarwal, V, G. D Gay, and L Ling. 2014. "Window dressing in mutual funds." *The Review of Financial Studies*, 27(11): 3133–3170.

Agarwal, V, N. D Daniel, and N. Y Naik. 2011. "Do hedge funds manage their reported returns?" *The Review of Financial Studies*, 24(10): 3281–3320.

Andrikogiannopoulou, A, P Krueger, S. F Mitali, and F Papakonstantinou. 2022. "Discretionary Information in ESG Investing: A Text Analysis of Mutual Fund Prospectuses." *Available at SSRN*.

Annaert, J, and G Van Campenhout. 2007. "Time Variation in Mutual Fund Style Exposures*." *Review of Finance*, 11(4): 633–661.

Ardia, D, K Bluteau, K Boudt, and K Inghelbrecht. 2023. "Climate change concerns and the performance of green vs. brown stocks." *Management Science*, 69(12): 7607–7632.

Atilgan, Y, K. O Demirtas, A Edmans, and A. D Gunaydin. 2023. "Does the carbon premium reflect risk or mispricing?" *Available at SSRN*.

Ben-David, I, F Franzoni, A Landier, and R Moussawi. 2013. "Do Hedge Funds Manipulate Stock Prices?" *The Journal of Finance*, 68(6): 2383–2434.

Berg, F, J. F Koelbel, A Pavlova, and R Rigobon. 2022. "ESG confusion and stock returns: Tackling the problem of noise." National Bureau of Economic Research.

Bolton, P, and M Kacperczyk. 2021. "Do investors care about carbon risk?" *Journal of Financial Economics*, 142(2): 517–549.

Bordalo, P, N Gennaioli, and A Shleifer. 2012. "Salience theory of choice under risk." *The Quarterly Journal of Economics*, 127(3): 1243–1285.

Bordalo, P, N Gennaioli, and A Shleifer. 2013. "Salience and asset prices." *American Economic Review*, 103(3): 623–628.

Brown, S. J, and W. N Goetzmann. 1997. "Mutual fund styles." *Journal of Financial Economics*, 43(3): 373–399.

Carhart, M. M. 1997. "On persistence in mutual fund performance." *The Journal of Finance*,

52(1): 57–82.

Carhart, M. M, R Kaniel, D. K Musto, and A. V Reed. 2002. "Leaning for the tape: Evidence of gaming behavior in equity mutual funds." *The Journal of Finance*, 57(2): 661–693.

Cooper, M. J, H Gulen, and P. R Rau. 2005. "Changing names with style: Mutual fund name changes and their effects on fund flows." *The Journal of Finance*, 60(6): 2825–2858.

Cooper, M. J, O Dimitrov, and P. R Rau. 2001. "A rose.com by any other name." *The Journal of Finance*, 56(6): 2371–2388.

Edmans, A, and M Kacperczyk. 2022. "Sustainable finance." *Review of Finance*, 26(6): 1309–1313.

Evans, R. B, and R Fahlenbrach. 2012. "Institutional investors and mutual fund governance: Evidence from retail–institutional fund twins." *The Review of Financial Studies*, 25(12): 3530–3571.

Fama, E. F, and K. R French. 1993. "Common risk factors in the returns on stocks and bonds." *Journal of Financial Economics*, 33(1): 3–56.

Fung, W, and D. A Hsieh. 1997. "Empirical Characteristics of Dynamic Trading Strategies: The Case of Hedge Funds." *The Review of Financial Studies*, 10(2): 275–302.

Gantchev, N, M Giannetti, and R Li. 2024. "Sustainability or performance? Ratings and fund managers' incentives." *Journal of Financial Economics*, 155: 103831.

Gibson Brandon, R, S Glossner, P Krueger, P Matos, and T Steffen. 2022. "Do responsible investors invest responsibly?" *Review of Finance*, 26(6): 1389–1432.

Giglio, S, B Kelly, and J Stroebel. 2021. "Climate finance." *Annual Review of Financial Economics*, 13: 15–36.

Gil-Bazo, J, and P Ruiz-Verdú. 2009. "The relation between price and performance in the mutual fund industry." *The Journal of Finance*, 64(5): 2153–2183.

Hartzmark, S. M, and A. B Sussman. 2019. "Do investors value sustainability? A natural experiment examining ranking and fund flows." *The Journal of Finance*, 74(6): 2789–2837.

Hu, G, R. D McLean, J Pontiff, and Q Wang. 2014. "The year-end trading activities of institutional investors: Evidence from daily trades." *The Review of Financial Studies*, 27(5): 1593–

1614.

Jiang, G, B Liang, and H Zhang. 2023. "Sharpening the Sharpe Analysis with Machine Learning: Evidence of Mutual Fund Style-Shifting Skill." *SSRN Electronic Journal*.

Jiang, G. J, B Liang, and H Zhang. 2022. "Hedge Fund Manager Skill and Style-Shifting." *Management Science*, 68(3): 2284–2307.

Kim, S, and A Yoon. 2023. "Analyzing active fund managers' commitment to ESG: Evidence from the United Nations Principles for Responsible Investment." *Management Science*, 69(2): 741–758.

Lakonishok, J, A Shleifer, R Thaler, and R Vishny. 1991. "Window Dressing by Pension Fund Managers." *The American Economic Review*, 81(2): 227–231.

Liang, H, L Sun, and M Teo. 2022. "Responsible hedge funds." *Review of Finance*, 26(6): 1585–1633.

Musto, D. K. 1999. "Investment decisions depend on portfolio disclosures." *The Journal of Finance*, 54(3): 935–952.

Ng, L, and Q Wang. 2004. "Institutional trading and the turn-of-the-year effect." *Journal of Financial Economics*, 74(2): 343–366.

Parise, G, and M Rubin. 2023. "Green Window Dressing." *Available at SSRN 4459352*.

Pástor, L, R. F Stambaugh, and L. A Taylor. 2021. "Sustainable investing in equilibrium." *Journal of Financial Economics*, 142(2): 550–571.

Pástor, L, R. F Stambaugh, and L. A Taylor. 2022. "Dissecting green returns." *Journal of Financial Economics*, 146(2): 403–424.

Raghunandan, A, and S Rajgopal. 2022. "Do ESG funds make stakeholder-friendly investments?" *Review of Accounting Studies*, 27(3): 822–863.

Roussanov, N, H Ruan, and Y Wei. 2021. "Marketing mutual funds." *The Review of Financial Studies*, 34(6): 3045–3094.

Sharpe, W. F. 1992. "Asset allocation: Management style and performance measurement." *Journal of Portfolio Management*, 18(2): 7–19.

Solomon, D. H, E Soltes, and D Sosyura. 2014. "Winners in the spotlight: Media coverage of

fund holdings as a driver of flows." *Journal of Financial Economics*, 113(1): 53–72.

Storey, J. D. 2002. "A direct approach to false discovery rates." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(3): 479–498.

Wang, P. 2024. "Portfolio pumping in mutual fund families." *Journal of Financial Economics*, 156: 103839.

Table 1: Summary statistics of fund characteristics

In Panel A, we present summary statistics for the entire sample of 26,810 fund-quarters (2,144 funds at an average of 12.50 quarters each from January 2016 to June 2020). In Panel B, we present summary statistics for the sample of 2,656 fund-quarters that have *GD* measure in the top 10% in the same category-quarter group. *Fund age* is the number of years since the fund's inception. *Total net asset value* is measured in millions of U.S. dollars. *Expense ratio* is total annual management, administrative, and 12b-1 fees and expenses, divided by year-end Total net asset value. *Turnover ratio* is the minimum of aggregate purchases and sales of securities, divided by the average Total net asset value over the calendar year. *Marketing expenses* are defined similar to Roussanov, Ruan and Wei (2021) as the combination of 12b-1 fees and front loads. *Fund inflows* are the ratio of quarterly fund flows to the beginning-of-quarter Total net asset value. The *Institutional-fund* dummy classifies each fund as targeted to institutional or retail investors using CRSP's institutional class indicator for its largest (by net asset value) share class. *Net returns* are the quarterly fund returns net of fees, expenses, and transaction costs. *Fund alpha* is the quarterly fund alpha estimated from the Carhart (1997) 4-factor model based on daily fund and factor returns in the past 12 months.

Panel A: Entire Sample

				Percentiles		
	# Obs	Mean	Std. Dev.	p10	p50	p90
Fund age	26,809	17.61	13.58	3.00	17.00	32.00
Total net asset value	26,676	7.60	6,213.00	18.30	313.40	3,769.00
Expense ratio (%)	26,614	1.06	0.42	0.66	1.02	1.48
Turnover ratio (%)	26,614	79.67	268.80	15.00	48.00	129.00
Marketing expenses (%)	26,633	0.46	0.46	0.00	0.25	1.00
Fund inflows (%)	26,603	0.07	4.46	-0.09	-0.02	0.07
Institutional-fund dummy	26,810	0.50	0.50	0.00	1.00	1.00
Fund returns (net) (%)	26,723	2.55	8.43	-6.93	3.39	10.35
Fund alpha (%)	26,248	-0.52	3.07	-3.51	-0.42	2.50

Panel B: Sample of top-10% green-dressers

				Percentiles		
	# Obs	Mean	Std. Dev.	p10	p50	p90
Fund age	2,656	15.60	14.44	2.00	13.00	31.00
Total net asset value	2,629	1,854.00	8,460.00	13.10	156.80	3,044.00
Expense ratio (%)	2,614	1.17	0.49	0.74	1.10	1.74
Turnover ratio (%)	2,614	154.80	566.10	11.00	54.00	219.00
Marketing expenses (%)	2,617	0.48	0.46	0.00	0.25	1.00
Fund inflows (%)	2,612	0.22	8.51	-0.10	-0.02	0.10
Institutional-fund dummy	2,656	0.48	0.50	0.00	0.00	1.00
Fund returns (net) (%)	2,634	2.20	7.42	-5.91	2.62	9.16
Fund alpha (%)	2,565	-0.56	3.24	-3.55	-0.43	2.35

Table 2: Summary statistics of fund ESG characteristics

In Panel A, we present summary statistics for the funds' ESG measures and ESG characteristics. *Holdings ESG score* is the value-weighted mean of the fund investments' ESG scores from MSCI, the average ESG score in our sample is 4.76. *Returns ESG score* is the fund's returns-based ESG score calculated using the Sharpe (1992) style analysis. *ESG commitment* is an indicator variable taking value 1 if fund i is a UNPRI signatory or its name contains ESG keywords in fiscal quarter t and 0 otherwise. *Border Fund* is an indicator variable taking value 1 if rank of fund i's portfolio sustainability score at the end of fiscal quarter $t-1$ is within ±2.5% from the cutoff of the Globe 5 and Globe 4 ratings (which is 10%), as in Gantchev, Giannetti and Li (2024). In Panel B, we present the summary statistics of our *GD* measure.

Panel A: ESG measures and characteristics

				Percentiles		
	# Obs	Mean	Std. Dev.	p10	p50	p90
Holdings ESG measure	26,041	4.76	0.66	4.06	4.84	5.43
Returns ESG measure	26,802	2.65	1.01	1.45	2.55	4.03
ESG Commitment	26,810	0.20	0.40	0.00	0.00	1.00
Border Fund	17,172	0.05	0.21	0.00	0.00	0.00

Panel B: Green-dressing measure

				Percentiles		
	# Obs	Mean	Std. Dev.	p10	p50	p90
GD	26,802	0.05	0.36	-0.18	0.02	0.27

Table 3: Validity of the green-dressing measure

This table reports the results of simulation tests on the validity of the green-dressing measure. We conduct three separate simulations with proportions of green dressers p set to 25%, 50% and 100% in the fictitious samples. Panel A reports results from the data generation processes in which a uniform pre-disclosure window of 10 trading days is applied to all green dressers, while Panel B reports results from the data generation processes in which pre-disclosure windows are randomly assigned across funds, with 25% of the sample allocated to each of four window lengths: 2, 5, 7, and 10 trading days. In each panel, we report the proportions of funds that are identified as green dressers using the green-dressing measure defined in Equation 5 (leftmost columns) and in Parise and Rubin (2023) (rightmost columns), at the 10%, 5%, and 1% significance levels.

Panel A: Uniform Pre-Disclosure Window (10 Days)

		Returns-based ESG Difference			Parise and Rubin (2023)		
	Signif. Level	10%	5%	1%	10%	5%	1%
25% Green Dressers	Signif. Positive (%)	27.51	20.27	10.63	7.70	4.78	1.59
	Signif. Negative (%)	3.59	1.79	0.86	2.85	1.06	0.13
50% Green Dressers	Signif. Positive (%)	39.87	34.02	22.99	15.26	10.68	4.31
	Signif. Negative (%)	3.06	1.66	0.93	2.39	0.86	0.07
100% Green Dressers	Signif. Positive (%)	68.54	61.18	45.17	29.29	20.49	8.37
	Signif. Negative (%)	0.87	0.72	0.43	2.02	0.87	0.14

Panel B: Randomly Assigned Pre-Disclosure Windows (2, 5, 7, and 10 Days)

		Returns-based ESG Difference			Parise and Rubin (2023)		
	Signif. Level	10%	5%	1%	10%	5%	1%
25% Green Dressers	Signif. Positive (%)	28.57	21.66	11.89	5.37	2.92	0.80
	Signif. Negative (%)	4.25	2.06	1.00	3.05	1.06	0.13
50% Green Dressers	Signif. Positive (%)	41.73	36.54	25.85	8.96	5.44	1.59
	Signif. Negative (%)	3.52	1.86	0.93	2.46	0.93	0.13
100% Green Dressers	Signif. Positive (%)	71.14	66.38	52.67	15.44	9.81	2.89
	Signif. Negative (%)	1.30	0.29	0.14	2.31	1.15	0.14

Table 4: Fund flows and ESG characteristics of portfolio holdings

This table shows how fund flows respond to holdings-based ESG score. The specifications differ in the measure of holdings-based ESG characteristics. Column 1 shows the results with continuous holdings-based ESG score, and column 2 shows the results with a dummy indicating if the holdings-based ESG score is in the top 10% within its investment category at the fiscal-quarter-end. All specifications include investment category-by-quarter fixed effects and controls for age, size, expense ratio, 12b-1 fees, turnover ratio, past 12-month mean return, past 12-month mean flow ratio and dummies indicating if prior 12-month α is in the bottom or top 10% for for the investment category at the fiscal-quarter-end, and a dummy indicating funds targeted on institutional investors. t-statistics from standard errors clustered two-ways at the fund and year-by-quarter levels are reported. $^{*}/^{**}/^{***}$ indicate significance at the 10%/5%/1% levels.

	(1)		(2)	
Holdings ESG score	0.001			
	0.219			
Holdings ESG top-10%			0.007	**
			2.392	
Returns ESG score	-0.002		-0.002	
	-1.342		-1.432	
Performance window dressing	-0.156		-0.123	
	-0.618		-0.501	
log (Fund age)	-0.020	***	-0.02	***
	-8.015		-7.894	
log (Fund size)	-0.002	***	-0.002	***
	-3.299		-3.337	
log (expense ratio)	-0.802		-0.888	*
	-1.689		-1.915	
log (effective 12-b1 fee)	0.737	**	0.755	**
	2.304		2.339	
Turnover ratio	-0.001		-0.000	
	-0.377		-0.160	
Funds for institutions	0.002		0.002	
	0.932		0.963	
Past 12-month mean alpha top 10%	0.040	***	0.039	***
	7.587		7.633	
Past 12-month mean alpha bottom 10%	-0.021	***	-0.022	***
	-8.785		-9.362	
Past 12-month mean return	2.264	***	2.247	***
	6.300		6.215	
Category-by-time Fixed Effects	Yes		Yes	
# of Observations	22,378		22,321	
Adjusted R^2	0.08		0.08	

Table 5: Hypothetical vs. actual ESG measure

This table reports the proportions of funds with significant positive (negative) mean difference between hypothetical returns-based ESG score and actual returns-based ESG scores at the 10%/5%/1% significance levels. In Panel A, the hypothetical returns-based ESG score are calculated from fund's daily return of a hypothetical portfolio comprising a fund's end-of-quarter holdings that are assumed to have been held throughout the last month of the preceding fiscal quarter, and the returns-based ESG score calculated from fund's actual daily return in the last month of the preceding fiscal quarter. In Panel B, the returns-based ESG scores are calculated from fund's daily return of a hypothetical portfolio comprising a fund's end-of-quarter holdings that are assumed to be held throughout the first month of the succeeding fiscal quarter, and the actual returns-based ESG score are calculated from fund's actual daily return in the first month of the succeeding fiscal quarter. In panel C, we conduct a placebo test, where we compare the actual returns-based ESG score for the previous fiscal quarter with that for the subsequent fiscal quarter. Again, we report the proportions of funds with significant positive (negative) mean difference at the 10%/5%/1% levels.

Panel A: Backward-looking green-dressing measure

Significance level	10%	5%	1%
Significant positive mean difference	20.47%	16.29%	9.55%
Significant negative mean difference	4.71%	2.88%	1.64%

Panel B: Forward-looking green-dressing measure

Significance level	10%	5%	1%
Significant positive mean difference	26.42%	19.36%	11.18%
Significant negative mean difference	3.60%	2.42%	1.50%

Panel C: Placebo test

Significance level	10%	5%	1%
Significant positive mean difference	5.62%	3.08%	0.80%
Significant negative mean difference	3.61%	1.74%	0.47%

Table 6: Abnormal returns of green and brown stocks surrounding fund disclosure date

This table presents the average abnormal stock returns over the event window from day −3 to day +3 relative to the fund disclosure date (event day 0). Panel A reports results for green stocks, defined as those in the top quintile of MSCI ESG scores, while Panel B reports results for brown stocks, defined as those in the bottom quintile. Abnormal stock returns are estimated using the Fama and French (1993) three-factor model with a 150-day estimation window that ends 30 days prior to the event day 0. */**/*** indicate significance at the 10%/5%/1% levels.

Panel A: Green stocks

	Event day						
	-3d	-2d	-1d	0d	1d	2d	3d
AAR	-0.021	0.112	0.016	0.049	-0.083	-0.001	-0.051
t-statistics	-0.752	5.105 ***	0.587	2.457 **	-3.411 ***	-0.040	-2.133 **

Panel B: Brown stocks

	Event day						
	-3d	-2d	-1d	0d	1d	2d	3d
AAR	-0.017	0.030	-0.161	-0.000	0.101	0.035	0.059
t-statistics	-0.548	1.233	-6.393 ***	-0.011	3.608 ***	1.362	2.260 **

Table 7: Stock-level ESG performance and cumulative abnormal returns at quarter start vs. end

This table shows the association between stock's ESG performance and its CARs around the turns of quarter. In Panel A, ESG performance is measured by the percentile rank of each stock's most recent MSCI ESG score, whereas in Panel B, it is represented by a dummy variable equal to 1 if the score falls within the top 20% of the distribution. For both panels, column 1 reports the effect of ESG performance on the cumulative abnormal return over the three trading days preceding the quarter-end, while column 2 reports the corresponding effect over the three trading days following the quarter-end. All specifications include industry-by-quarter fixed effects and firm controls for logarithm of market capitalization, book-to-market ratio, return on assets, debt-to-equity ratio, net profit margin and past quarter returns. *t*-statistics from standard errors clustered at the firm level are reported. */**/*** indicate significance at the 10%/5%/1% levels.

	Panel A: MSCI ESG percentile rank		Panel B: MSCI ESG top-20%	
	(1) CAR (-2, 0)	(2) CAR (+1, +3)	(1) CAR (-2, 0)	(2) CAR (+1, +3)
ESG performance	0.209 ***	-0.132 *	0.122 **	0.008
	2.795	-1.842	2.214	0.164
log (Market cap)	-0.040 ***	-0.012	-0.038 ***	-0.016
	-2.875	-0.808	-2.750	-1.094
Book-to-market	-0.108	0.394 ***	-0.108	0.395 ***
	-1.482	5.717	-1.491	5.710
Leverage	0.000 ***	-0.000 ***	0.000 ***	-0.000 ***
	3.600	-3.519	3.625	-3.550
ROA	1.031 ***	-1.357 ***	1.023 ***	-1.362 ***
	3.620	-6.005	3.584	-6.017
Net profit margin	-0.000	-0.000	-0.000	-0.000
	-0.082	-0.349	-0.094	-0.341
Past quarter return	-1.679 ***	-1.786 ***	-1.678 ***	-1.785 ***
	-8.700	-10.842	-8.696	-10.84
Constant	0.208	0.039	0.276 **	0.004
	1.585	0.279	2.123	0.031
# of Observations	30355	30355	30355	30355
Adjusted R^2	0.082	0.185	0.082	0.185
Industry-by-time Fixed Effect	Yes	Yes	Yes	Yes

Table 8: Fund characteristics related to green dressing

This table shows which fund characteristics are associated with green dressing. In column 1 and 2, we report the estimations from OLS regression. In column 3 and 4, we report the estimations from logistic regressions. *ESG commitment* indicates if a fund is a UNPRI signatory or its name contains ESG keywords in fiscal quarter t. *Border Fund* indicates if the rank of a fund's portfolio sustainability score at the end of fiscal quarter $t-1$ is within ±2.5% of the cutoff of the Globe 5 and Globe 4 ratings (which is 10%), as in Gantchev, Giannetti and Li (2024). All specifications include investment category-by-quarter fixed effects and fund controls for age, size, expense ratio, 12b-1 fees, turnover ratio, prior 12-month mean return, prior 12-month mean flow ratio, and a dummy indicating funds targeted to institutional investors. t-statistics from standard errors clustered two-ways at the fund and year-by-quarter levels are reported. $^{*}/^{**}/^{***}$ indicate significance at the 10%/5%/1% levels.

	OLS		Logistic	
	(1)	(2)	(3)	(4)
	GD 10% dummy	GD 20% dummy	GD 10% dummy	GD 20% dummy
ESG commitment	0.023 **	0.040 ***	0.270 **	0.251 ***
	2.453	2.997	2.520	3.077
Border Fund (4-5)	0.020	0.039 *	0.219	0.236 **
	1.434	2.086	1.528	2.037
log (Fund age)	-0.011 *	-0.014 *	-0.123 *	-0.091 *
	-1.823	-1.795	-1.799	-1.761
log (Fund size)	0.003	0.002	0.027	0.009
	1.060	0.542	0.842	0.445
log (expense ratio)	7.681 ***	7.856 ***	80.555 ***	47.841 ***
	4.495	3.296	4.709	3.368
log (effective 12b1 fee)	-3.173 ***	-3.325 **	-35.362 ***	-21.106 **
	-3.315	-2.284	-3.317	-2.319
Fund for institutions	-0.018 **	-0.032 **	-0.217 ***	-0.206 ***
	-2.586	-2.840	-2.604	-2.819
Turnover ratio	0.017 ***	0.012 **	0.140 ***	0.069 **
	3.096	2.173	3.158	2.058
Prior 12-month mean returns	-2.417 ***	-3.658 **	-25.882 ***	-22.475 ***
	-3.205	-2.810	-3.232	-2.788
Prior 12-month mean flows	0.017	-0.027	0.178	-0.183
	0.339	-0.457	0.316	-0.456
Category-by-time fixed effects	Yes	Yes	Yes	Yes
# of Observations	17,552	17,590	17,508	17,586
Adjusted R^2	0.01	0.01		
Pseudo R^2			0.02	0.01

Table 9: Fund flows and green dressing

This table shows how fund flows respond to top 10% holdings-based ESG score. The effect for green-dressing (GD) funds is shown in the row presenting the interaction of dummy of top 10% holdings-based ESG score with the dummy indicating GD funds. Column 1 shows the results with the GD dummy based on the top 10% values of the continuous GD measure, and column 2 shows the results with the GD dummy based on the top 20% values of the continuous GD measure. All specifications include investment category-by-quarter fixed effects and fund controls for age, size, expense ratio, 12b-1 fees, turnover ratio, past 12-month mean return, prior 12-month mean flows and dummies indicating if prior 12-month α is in the bottom or top 10% for the investment category at the fiscal-quarter-end, and a dummy indicating funds targeted on institutional investors. t-statistics from standard errors clustered two-ways at the fund and year-by-quarter levels are reported. $^{*}/^{**}/^{***}$ indicate significance at the 10%/5%/1% levels.

	(1)	(2)
Holdings ESG top-10%	0.007 **	0.006 **
	2.543	2.412
Holdings ESG top-10% * GD top-10%	0.002	
	0.271	
GD top-10%	0.000	
	0.010	
Holdings ESG top-10% * GD top-20%		0.004
		0.785
GD top-20%		-0.002
		-1.055
Category-by-time Fixed Effects	Yes	Yes
Controls	Yes	Yes
Observations	22,321	22,321
Adjusted R^2	0.079	0.079

Table 10: Fund performance and green dressing

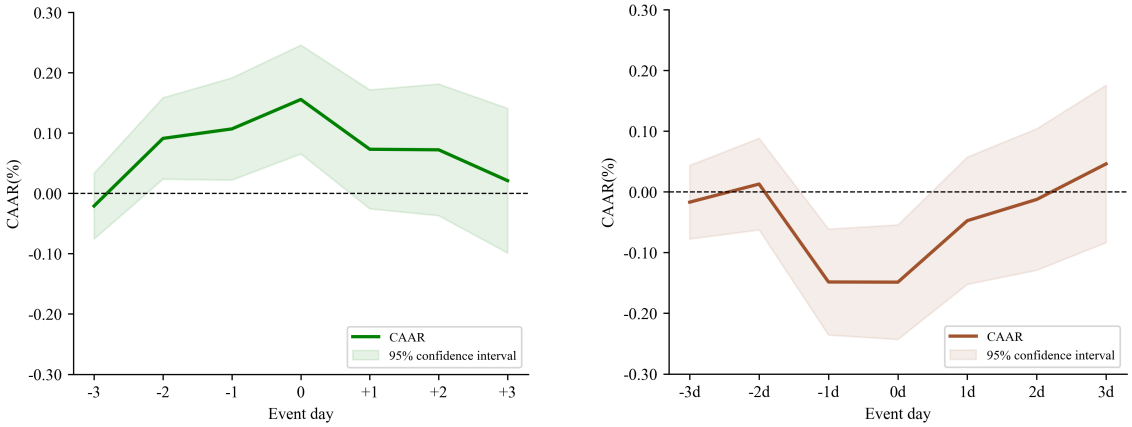
This table shows the effect on fund performance of funds holding stocks with high ESG scores, for non-green-dressers and for green dressers. The differential effect for green-dressing funds is shown in the row presenting the interaction of high ESG holdings with the dummy indicating green-dressing (GD) funds. The two panels differ in the definition of the green-dressing dummy: in Panel A (B), funds are classified as green dressers if their continuous green-dressing measure is in the top 10% (20%). The specifications within each panel differ in the way fund performance is estimated. In column 1, fund performance is the market-adjusted quarterly fund return. In columns 2–4, fund performance is α estimated from a factor model using daily fund and factor returns over the prior 12 months; in columns 2, 3, and 4, the factor model is the CAPM, the Fama and French (1993) 3-factor model, and the Carhart (1997) 4-factor model, respectively. All specifications include investment category-by-quarter fixed effects and fund controls for age, size, expense ratio, 12b-1 fees, past 12-month mean return and 12-month mean α, and a dummy indicating funds targeted to institutional investors. t-statistics from standard errors clustered two-ways at the fund and year-by-quarter levels are reported. $^{*}/^{**}/^{***}$ indicate significance at the 10%/5%/1% levels.

Panel A: With top-10% green-dressers

	(1) Market-Adjusted	(2) CAPM	(3) FF3	(4) Carhart4
Holdings ESG top-10%	0.003 **	0.003 ***	0.002 **	0.002 **
	2.328	3.693	2.143	2.803
Holdings ESG top-10% * GD top-10%	-0.004 **	-0.004 ***	-0.003 **	-0.004 ***
	-2.834	-3.373	-2.697	-3.883
GD top-10%	-0.000	0.002	0.001	0.001
	-0.162	1.248	1.056	1.233
Category-by-time Fixed Effects	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	23,055	22,382	22,382	22,382
Adjusted R^2	0.250	0.278	0.214	0.185

Panel B: With top-20% green-dressers

	(1) Market-Adjusted	(2) CAPM	(3) FF3	(4) Carhart4
Holdings ESG top-10%	0.004 **	0.004 ***	0.002 **	0.002 **
	2.619	3.997	2.223	2.806
Holdings ESG top-10% * GD top-20%	-0.005 ***	-0.004 ***	-0.003 **	-0.003 ***
	-3.394	-3.814	-2.866	-3.442
GD top-20%	0.000	0.002	0.000	0.000
	0.260	1.516	0.537	0.520
Category-by-time Fixed Effects	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	23,055	22,382	22,382	22,382
Adjusted R^2	0.250	0.279	0.214	0.185



(a) Green stocks.

(b) Brown stocks.

Figure 1: Event study of stock market reaction to quarterly fund-holdings disclosures for green stocks (in Panel a) and brown stocks (in Panel b). Event day 0 is the end of each calendar quarter, and the event window is the ±3 days around it. In each figure, the solid line plots the Cumulative Average Abnormal Return (CAAR), the shaded area indicates the corresponding 95% confidence intervals, and the horizontal dashed black line plots CAAR equal to zero. Green (brown) stocks are those whose MSCI ESG score is ranked in the top (bottom) 20% as of the event date.