NBER WORKING PAPER SERIES

MUTUAL FUND PERFORMANCE AND FLOWS DURING THE COVID-19 CRISIS

Lubos Pastor
M. Blair Vorsatz

Working Paper 27551
http://www.nber.org/papers/w27551

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2020

We are grateful to the University of Chicago Booth School of Business for research support. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed a financial relationship of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w27551.ack

Mutual Fund Performance and Flows During the COVID-19 Crisis
Lubos Pastor and M. Blair Vorsatz
NBER Working Paper No. 27551
July 2020
JEL No. G01,G11,G12,G14,G23

ABSTRACT

We present a comprehensive analysis of the performance and flows of U.S. actively-managed equity mutual funds during the COVID-19 crisis of 2020. We find that most active funds underperform passive benchmarks during the crisis, contradicting a popular hypothesis. Funds with high sustainability ratings perform well, as do funds with high star ratings. Fund outflows largely extend pre-crisis trends. Investors favor funds that apply exclusion criteria and funds with high sustainability ratings, especially environmental ones. Our finding that investors remain focused on sustainability during this major crisis suggests they view sustainability as a necessity rather than a luxury good.

Lubos Pastor
University of Chicago
Booth School of Business
5807 South Woodlawn Ave
Chicago, IL 60637
and NBER
lubos.pastor@chicagobooth.edu

M. Blair Vorsatz
University of Chicago
Booth School of Business
5807 South Woodlawn Ave
Chicago, IL 60637
mvorsatz@chicagobooth.edu

1. Introduction

Active equity mutual funds are well known to have underperformed passive benchmarks, net of fees.[1] Despite its long-lasting underperformance, the active management industry remains large, managing tens of trillions of dollars. The existence of a large underperforming industry appears puzzling because an alternative—passive funds—is easily available to investors.

One popular hypothesis is that investors are willing to tolerate this underperformance because active funds outperform in periods that are particularly important to investors. This hypothesis is first formulated by Moskowitz (2000) who asks whether mutual funds provide a hedge against recessions. Glode (2011) formalizes this hypothesis by building a model in which a fund manager generates active returns that depend on the state of the economy. In equilibrium, the manager chooses to work harder in periods when investors' marginal utility of consumption is higher because investors are willing to pay for this insurance. If active funds deliver high returns in periods when investors need them the most then these funds' unconditional performance understates the funds' true abilities.

We test this hypothesis by analyzing the performance of active mutual funds during the COVID-19 crisis of 2020. This crisis is particularly suitable for the task at hand for two reasons. First, it has led to an unprecedented output contraction and the fastest increase in unemployment on record. Investors surely want to hedge against such a severe crisis. Second, active managers have an opportunity to perform well during this crisis because the crisis has created unusually large price dislocations in financial markets. In the equity market, the S&P 500 index experienced its steepest descent in living memory, losing 34% of its value in the five-week period between February 19 and March 23, 2020 before bouncing back by over 30% by the end of April. The sharp response of equity markets to COVID-19 is analyzed in a growing number of studies.[2] In the bond market, liquidity evaporated in March 2020, not only for corporate bonds (e.g., Kargar et al., 2020, and O'Hara and Zhou, 2020) but also for the usually-liquid Treasuries (e.g., Schrimpf, Shin, and Sushko, 2020). Until liquidity improved following the interventions from the Federal Reserve, its temporary shortage created massive market disruptions. For example, in the corporate bond market, bonds traded at large discounts to credit default swaps, and ETFs traded at large discounts to net asset values (Haddad, Moreira, and Muir, 2020). In addition, the Treasury market witnessed significant

[1] See Jensen (1968), Elton, Gruber, Das, and Hlavka (1993), Malkiel (1995), Gruber (1996), Carhart (1997), Wermers (2000), Pástor and Stambaugh (2002), and Fama and French (2010), and others.

[2] For evidence at the aggregate stock market level, see, for example, Alfaro et al. (2020) and Gormsen and Koijen (2020). For cross-sectional evidence, see Bretscher, Hsu, and Tamoni (2020), Ding et al. (2020), Fahlenbrach, Rageth, and Stulz (2020), Gerding, Martin, and Nagler (2020), Pagano, Wagner, and Zechner (2020), Ramelli and Wagner (2020), and others.

mispricing between bonds and bond futures (Schrimpf, Shin, and Sushko, 2020). These price dislocations are due to a combination of factors including record-high volatility and traders working from home. Under the hypothesis that active funds outperform during recessions, they should find it particularly easy to outperform when markets are rife with mispricing.

Contrary to this hypothesis, we find that active funds underperform their passive benchmarks during the COVID-19 crisis. We define the crisis period as the ten-week period between February 20 and April 30, 2020. We choose February 20 as the starting date because the stock market peaked on February 19 before its rapid descent. We choose April 30 as the ending date because it is a month-end by which the market largely rebounded, and also because it puts the market bottom on March 23 roughly in the middle of the crisis period. The ten-week crisis period is thus roughly evenly split between the crash and the recovery. Our evidence is based on daily returns of all U.S. active equity mutual funds.

The underperformance of active funds is particularly strong when measured relative to the S&P 500 benchmark. We find that 74.2% of active funds—about three quarters!—underperform the S&P 500 during the COVID-19 crisis. The average fund underperformance is −5.6% (t = −5.37) during the ten-week period, or −29.1% on an annualized basis.

While the S&P 500 is the most popular benchmark among U.S. equity funds, it is not appropriate for all funds given its large-cap focus. We consider three types of benchmarks that are tailored to each fund's investment style: Morningstar-designated FTSE/Russell benchmarks, fund-designated prospectus benchmarks, and factor-model benchmarks. We find that active funds also underperform these fund-specific benchmarks, although by lower margins. For example, 57.6% of funds underperform their FTSE/Russell benchmarks and 54.2% of funds underperform their prospectus benchmarks. The average fund underperformance relative to the FTSE/Russell benchmark is −2.1% (t = −3.90) during the crisis period, or −11% on an annualized basis. Relative to the prospectus benchmark, the average underperformance is −1.5% (t = −2.49) during the crisis, or −7.7% annualized.

Besides benchmark-adjusted fund returns, we also examine factor-adjusted returns by computing fund alphas relative to five different factor models. All five alphas are significantly negative on average during the crisis period, ranging from −7.6% annualized (t = −3.25) for the six-factor model that includes the five factors of Fama and French (2015) plus momentum, to −29.1% annualized (t = −7.02) for the CAPM. The fraction of funds with negative alphas ranges from 60.4% for the four-factor Carhart (1997) model to a stunning 80.2% for the CAPM. In short, active funds perform poorly during the COVID-19 crisis.

Prior tests of the same hypothesis arrive at a different conclusion. Moskowitz (2000)

shows that active funds' returns from 1975 to 1994 are higher during recessions by 6% per year, on average. Kosowski (2011) analyzes the period from 1962 to 2005 and finds that mutual fund alphas in recessions exceed those in expansions by 3% to 5% per year, on average. Glode (2011) reports that funds with poor unconditional performance generate countercyclical risk-adjusted returns in 1980 through 2005. Kacperczyk, van Nieuwerburgh, and Veldkamp (2016) find that fund alphas are 1.6% to 4.6% per year higher in recessions over the 1980–2005 period. Unlike our study, all of these studies examine periods in which recessions are substantially milder than the COVID-19 crisis.

While active funds as a whole underperform, their performance during the COVID-19 crisis exhibits substantial heterogeneity. One of the strongest predictors of performance is the sustainability rating from Morningstar. Morningstar assigns between one and five sustainability "globes" to each fund, with more globes denoting higher sustainability. We find that funds with more globes as of January 31, 2020 have higher benchmark-adjusted returns between February 20 and April 30, 2020. Remarkably, the relation is monotonic across the globe categories: five-globe funds outperform four-globe funds, which in turn outperform three-globe funds, etc. High-globe funds (those with four or five globes) significantly outperform the remaining funds within the same investment style by 14.2% per year ($t = 4.85$) in terms of FTSE/Russell benchmark-adjusted returns. This result is driven largely by environmental sustainability—funds with higher environmental ratings outperform those with lower ratings.

Our findings linking fund performance to sustainability resemble those of Nofsinger and Varma (2014) who find that socially responsible mutual funds tend to outperform during periods of market crises. Their findings are based on a sample of 240 U.S. domestic equity mutual funds in the period of 2000 through 2011, which includes two recessions (2001 and 2007–2009). We examine only one recession (2020) but many more funds. Another related study, Albuquerque et al. (2020), finds that U.S. firms with high environmental and social ratings earn comparatively high stock returns in the first quarter of 2020. Ding et al. (2020) report a similar finding based on corporate social responsibility ratings of firms in 56 countries.[3] Our fund-level evidence complements their stock-level evidence in highlighting the role of sustainability during the COVID-19 crisis. The high returns of sustainable funds and stocks suggest that market participants' tastes continue to shift toward green assets and green products during this crisis (Pástor, Stambaugh, and Taylor, 2020).

Besides sustainability ratings, another strong predictor of fund performance during the

[3]These findings echo those based on the 2008–2009 recession. Lins, Servaes, and Tamayo (2017) show that U.S. firms with higher environmental and social ratings perform better during that recession.

COVID-19 crisis is the fund's star rating from Morningstar. Star ratings assigned as of January 31, 2020 predict performance between February 20 and April 30, 2020 positively and significantly. Similar to Morningstar globes, the relation is monotonic: five-star funds outperform four-star funds, which outperform three-star funds, etc. Five-star funds significantly outperform one-star funds in terms of cumulative benchmark-adjusted returns. One additional star is associated with an increase in performance of 5.78% per year ($t = 2.84$) in terms of FTSE/Russell benchmark-adjusted returns. That is, a five-star fund outperforms a one-star fund of the same style by about 23% per year, on average.

Finally, we find that growth funds outperform value funds. This finding is only partly driven by the well-known fact that the growth style outperforms the value style during the crisis because we measure fund performance net of the fund's style. In other words, we find that growth funds beat value funds on a style-adjusted basis. This result is strong when the style adjustment is performed through a factor model but it is insignificant when the adjustment is based on the style benchmark. The mixed nature of this evidence suggests roles for both active management and the superior performance of the growth style in explaining the different performance of growth and value funds during the crisis.

In addition to fund performance, we analyze capital flows in and out of active mutual funds. During the COVID-19 crisis, active funds experience steady outflows that largely continue long-term trends. The outflows are rapid during the market crash but they continue, albeit at a slower pace, during the market rebound after March 23, 2020.

Fund flows vary substantially across funds. Similar to performance, crisis-period flows are predictable by funds' pre-crisis sustainability ratings. Flows are near-monotonic across the five globe categories, with five-globe funds having the largest net flows and one-globe funds having the lowest flows between February 20 and April 30, 2020. In particular, one-globe funds suffer outflows of 2.6% of assets under management over the ten-week period, whereas five-globe funds' net flows are roughly zero. This difference, which is statistically significant, is driven especially by environmental concerns. Furthermore, funds that apply exclusion criteria in their investment process receive net inflows during the crisis, whereas funds that do not apply exclusions experience outflows. It is well known that mutual fund investors have come to favor sustainability-oriented funds in the 2010s (e.g., Bialkowski and Starks, 2016, and Hartzmark and Sussman, 2019). We find that this pre-crisis trend toward sustainability continues during the COVID-19 crisis.

A popular perspective in traditional neoclassical economics is that sustainability issues, such as environmental quality, are "luxury goods" that are likely to be of concern only

to those whose more basic needs for food, housing, and survival are adequately met (e.g., Baumol and Oates, 1979).[4] This perspective predicts that interest in sustainability should subside during a major economic and health crisis. In contrast, we find that investors retain their commitment to sustainability during the COVID-19 crisis. This finding suggests that investors have come to view sustainability as a necessity rather than a luxury good.

The performance hypothesis rejected by our evidence—that active funds outperform in recessions—is not the only possible explanation for why active management remains popular despite its poor track record. Gruber (1996) suggests that some investors suboptimally rely on active management because they are influenced by advertising, brokers, institutional arrangements, or tax considerations. Pástor and Stambaugh (2012) argue that a large active management industry can be rationalized if investors believe that active managers face decreasing returns to scale. In their model, rational investors respond to past underperformance of active funds by withdrawing money, which improves those funds' future performance to the point where investors are indifferent between investing actively or passively.

Our focus on crisis-period fund performance is also related to the literature on time-varying fund manager skill. An important early contribution is Ferson and Schadt (1996). More recently, Kacperczyk, van Nieuwerburgh, and Veldkamp (2016) develop a model of optimal attention allocation over the business cycle. In their model, fund managers allocate more attention to idiosyncratic shocks in expansions and to aggregate shocks in recessions. Similarly, Kacperczyk, van Nieuwerburgh, and Veldkamp (2014) find that fund managers exhibit better stock picking in expansions and better market timing in recessions. We do not attempt to separate stock selection from market timing during the COVID-19 crisis because such an exercise would require time series of fund holdings, which are widely available only on a quarterly basis. Data availability also limits our ability to test the hypothesis that the profit opportunities created by COVID-19 lead active funds to trade more, improving their future performance (Pástor, Stambaugh, and Taylor, 2017). Whether funds increase their turnover in 2020, and whether this turnover causes better future fund performance, remains to be seen because the turnover data from the SEC are only annual.

The paper is organized as follows. In Section 2, we describe our data. In Section 3, we analyze fund performance and its determinants. In Section 4, we discuss fund flows and their determinants. Section 5 concludes. Additional empirical results and database construction details are located in the Appendix, which is available on the authors' websites.

[4]An example of this common view is the controversial "Summers memo" from 1991, in which the World Bank's Chief Economist suggests that the Bank should be encouraging more migration of dirty industries to the least-developed countries. One of the reasons given in the memo is that "the demand for a clean environment for aesthetic and health reasons is likely to have very high income elasticity."

2. Data

We use daily data from Morningstar Direct covering the period from January 1, 2017 to April 30, 2020. Our original sample covers 4,292 actively managed U.S. equity mutual funds, although we primarily focus on the 3,626 funds with at least one non-missing net return between February 20 and April 30, 2020. The latter sample represents $4.9 trillion of total net assets as of January 31, 2020.

Our fund universe is constructed largely following Pástor, Stambaugh, and Taylor (2015), with two main differences. First, we also include international and sector equity funds domiciled in the U.S. Second, we do not require funds to appear in both CRSP and Morningstar; we use only Morningstar data. As in Pástor, Stambaugh, and Taylor (2015), we use the Morningstar FundID variable to aggregate share classes to the fund level.[5]

We use keywords in the Morningstar Category variable and the prospectus benchmark to exclude bond funds, money market funds, real estate funds, target retirement funds, and other non-equity funds. We also exclude funds identified by Morningstar as passive index funds and funds whose name contains the word "index." In our baseline results, we also use a fund size filter to include only funds with at least $15 million of net assets on January 31, 2020. Excluding the smallest funds has been advocated by Elton, Gruber, and Blake (2001), among many others. This screen is particularly relevant for fund flows because modest dollar flows can translate into extreme percentage flows for the smallest funds. This subsample covers 2,764 funds and $4.891 trillion of total net assets.

Throughout our analysis, we use funds' returns net of the expense ratio because our goal is to measure the return delivered to clients after fees. Despite being very fresh (we downloaded data through April 30, 2020 in May 2020), the data appear to be free of salient errors. For example, none of our sample's 2,692,799 fund-level daily net return observations are below -90% and only one is larger than 100%. We adjust fund returns for daily benchmark returns, also obtained from Morningstar, and for daily factor returns, which we obtain from Ken French's data library along with the risk-free rate.

We rely on three main Morningstar categorization variables throughout our analysis: the

[5]Many funds have multiple share classes, which are tied to the same pool of assets but have different fee structures. Since different share classes of the same fund have the same Morningstar FundID value, we can use the FundID variable to aggregate the share classes up to the fund level. Specifically, we compute a fund's total net assets by summing total net assets across the fund's share classes, setting the fund-level variable to missing if total net assets are missing for any of the share classes on that date. The fund's net returns, net expense ratio, and turnover ratio are averaged (lag-asset-weighted) across all share classes with non-missing values.

Morningstar Category, the Morningstar Institutional Category, and the Global Category. In our full sample of 3,626 funds, there are 39 Global Categories, 52 Morningstar Categories, and 93 Morningstar Institutional Categories. Each of these variables classifies a fund based on its investment style, sector, and geographical orientation. The Global Categories are the coarsest classification system and are used by Morningstar as groupings within which sustainability can be ranked. These categories include U.S. equity large-cap blend, U.S. equity small-cap, energy sector equity, and Latin America equity. We use the Global Categories for style fixed effects. The Morningstar Category variable is built on the 3-by-3 style box of size tilts (large-cap vs. small-cap) and growth vs. value style tilts. We use this variable to follow Pástor, Stambaugh, and Taylor (2015) in their classification of funds into equity and non-equity categories. Morningstar uses these groupings to rank performance in terms of star ratings. Last, the Morningstar Institutional Category variable is built on an extended version of the 3-by-3 style box with size tilts including micro-cap and giant and style tilts including deep value and high growth. We use this finest classification system for clustering our standard errors. While this is conservative relative to the more standard treatment of clustering at the fund level, we believe this appropriately accounts for how the health crisis shock may generate residual correlation among funds with similar strategies. For further details of our data construction, see the Appendix.

3. Fund Performance

Figure 1 provides a preliminary look at the performance of active funds during the COVID-19 crisis, along with the performance of the most popular passive benchmark: the S&P 500 index. We normalize the levels of both the S&P 500 and each fund's net asset value to 100 as of February 19, 2020. For each day t after February 19, we compute the price indices for each fund as well as the S&P 500 by compounding the corresponding daily returns:

$$F_t = 100(1+r_1^F)(1+r_2^F)...(1+r_t^F) \quad (1)$$

$$B_t = 100(1+r_1^B)(1+r_2^B)...(1+r_t^B)\ , \quad (2)$$

where F_t is the fund price index, B_t is the price index for the passive benchmark, r_t^F is the fund's net return on day t, and r_t^B is the benchmark return. Figure 1 plots both B_t and the average value of F_t across all funds. The figure also plots a 95% confidence interval around average F_t. Standard errors are clustered on the Morningstar Institutional Category, both here and in all subsequent figures reporting confidence intervals.

Computing F_t in equation (1) requires all of the fund's daily returns starting February

20 through day t. Any gap in the fund's return series, however short, would necessitate the fund's deletion from the average calculation. To avoid deleting too many funds, we replace any missing returns by the average return across all funds with the same FTSE/Russell benchmark on the same day, thus preserving the average level of performance across funds. We only replace missing returns for which there exists a non-missing return later in the fund's history by April 30. That is, we do not replace any missing returns at the end of a fund's history because funds that stop reporting returns to Morningstar may no longer be alive. Altogether, we replace 19,124 missing returns, which account for 13.9% of our sample. We apply this "patch" not only in Figure 1 but also in Figures 2 through 8.[6] We do not replace missing returns in Tables 1 through 3 because the analysis behind those tables does not require funds to have continuous return series.

Figure 1 shows that the S&P 500 loses 34% of its value between February 19 and March 23, before gaining 30% by April 30, 2020. The average active fund performs similarly, but it significantly underperforms the S&P 500 during the crisis. The April 30 price index levels are 86.01 for the S&P 500 but only 82.45 for the average active fund.

Given its focus on large-cap stocks, the S&P 500 is not the most appropriate benchmark for every fund. Several large-cap technology stocks performed well during the crisis, making the S&P 500 hard to beat. We thus compare each fund's return also to the returns of two benchmarks tailored to the fund's investment style: the prospectus benchmark and the FTSE/Russell benchmark. The prospectus benchmark is chosen by the fund itself (with some potential for strategic choice, as discussed by Sensoy, 2009), whereas the FTSE/Russell benchmark is assigned to each fund by Morningstar based on the fund's holdings.

Figure 2 compares fund performance to the FTSE/Russell benchmark (Panel A), the prospectus benchmark (Panel B), and the S&P 500 (Panel C). Unlike Figure 1, which plots index levels, Figure 2 plots the cumulative performance of the average active fund relative to the benchmark. Specifically, at each date t after February 19, 2020, the figure plots the average value of $\log(F_t) - \log(B_t)$, where F_t and B_t are defined in equations (1) and (2).

Figure 2 shows that active funds significantly underperform their benchmarks, on average. As of April 30, 2020, the average underperformance over the ten-week period is 1.53% relative to the FTSE/Russell benchmark, 0.94% relative to the prospectus benchmark, and 4.77% relative to the S&P 500. This underperformance is highly statistically significant in Panel C, and marginally significant in Panels A and B. Moreover, Figure 2 underestimates the actual underperformance due to a mild survivorship bias because average F_t can only be computed

[6] All figures look virtually identical if we replace missing returns not by average fund returns but rather by the returns on the fund's FTSE/Russell benchmark on the same day.

across funds that have survived through time t. During the ten-week period, 22 funds drop out of our sample, so their returns are excluded from the plot as of April 30.

Table 1 reports average benchmark-adjusted fund performance in a way that is immune to the survivorship bias. For each fund, live or dead, we take all of the fund's available daily returns in the given time period and subtract the same days' returns on the corresponding benchmark. We report annualized averages of those benchmark-adjusted returns in the first three columns of Panel A of Table 1. The average fund underperforms its FTSE/Russell benchmark by 11.02% per year, with a t-statistic of −3.90. Average underperformance relative to the prospectus benchmark is slightly smaller, 7.70% per year, but still significant (t = −2.49). The average fund underperforms the S&P 500 by a whopping 29.12% per year (t = −5.37). We observe underperformance not only during the full ten-week crisis period but also during both subperiods, the first of which captures the market crash and the second the recovery, and also in the pre-crisis period (October 1, 2019 to January 31, 2020).

The remaining columns of Panel A of Table 1 report average fund alphas from five multifactor models: the capital asset pricing model (CAPM), the three-factor model of Fama and French (1993), the four-factor model of Carhart (1997), the five-factor model of Fama and French (2015), and a six-factor model that includes those five factors plus momentum. For a fund's alpha to be included in the average, the fund must have at least 15 non-missing net returns for the time period of interest. All five alphas are significantly negative during the crisis period, ranging from −7.62% per year (t = −3.25) for the six-factor model to −29.11% per year (t = −7.02) for the CAPM. The alphas are particularly negative during the crash period (February 20 to March 23, 2020), ranging from −37.75% to −80.94% per year across the five models.

Panel B of Table 1 shows that 57.6% of funds underperform their FTSE/Russell benchmarks during the crisis. Additionally, 54.2% of funds underperform prospectus benchmarks and 74.2% of funds—almost three quarters!—underperform the S&P 500. More than 80% of funds have negative CAPM alphas during the crisis period. The fraction of funds with negative alphas ranges from 60.4% to 80.2% across the five models. Regardless of how we look at the data, we see active funds underperforming during the crisis.

3.1. Sustainability

We find that funds with higher sustainability ratings perform better during the crisis. For each fund, Morningstar evaluates how well the fund's holdings perform on ESG issues relative

to the fund's peer group (i.e., Morningstar Global Category). Morningstar uses company-level ESG scores from Sustainalytics to determine each fund's asset-weighted average unmanaged ESG risk exposure. Then, within each peer group, these scores are fitted to an approximate normal distribution to award 1, 2, 3, 4, or 5 sustainability globes to each fund.[7] Funds with 5 globes are the most sustainable and funds with 1 globe are the least sustainable. We find that funds with more globes perform better during the crisis.

Figure 3 shows the distributions of cumulative returns during the crisis across funds with different sustainability ratings, which are assigned as of January 2020. We collect funds in two groups: funds with 4 or 5 globes ("high sustainability") and funds with 1 or 2 globes ("low sustainability"). Panel A shows the distributions of cumulative total fund returns, whereas Panel B shows cumulative benchmark-adjusted returns, which are adjusted using FTSE/Russell benchmarks. Specifically, Panel A shows $\log(F_t)$, where F_t is in equation (1) and t corresponds to April 30, 2020, while Panel B shows $\log(F_t) - \log(B_t)$, where B_t is in equation (2) for the fund's FTSE/Russell benchmark. Both panels clearly show that more sustainable funds perform better in the crisis.

Figure 4 presents the sustainability result from a perspective similar to Figure 2, plotting cumulative fund performance relative to the benchmark, or $\log(F_t) - \log(B_t)$. We consider the same three benchmarks as before: FTSE/Russell (Panels A and B), prospectus (Panels C and D), and the S&P 500 (Panels E and F). In the left panels (A, C, and E), we plot the average cumulative performance differences for each of the five globe groups. In the right panels (B, D, and F), we report 95% confidence intervals for high-sustainability funds (4 or 5 globes) and low-sustainability funds (1 or 2 globes).

Remarkably, Figure 4 shows a monotonic relation between benchmark-adjusted fund performance and sustainability globes: five-globe funds outperform four-globe funds, which outperform three-globe funds, which in turn outperform two-globe funds, which beat one-globe funds. This monotonicity is present for all three benchmarks. The performance difference between high-sustainability funds and low-sustainability funds is marginally statistically significant.[8] The significance is stronger in the subsequent regression analysis in Table 2.

Given the important role of sustainability in determining fund performance, we investigate which dimensions of sustainability—E, S, or G—matter the most during the crisis. After sorting funds by their individual E, S, and G scores from Morningstar, we separate

[7]Within each peer group, the top 10% of funds receive 5 globes, the next 22.5% receive 4 globes, the next 35% receive 3 globes, the next 22.5% receive 2 globes, and the bottom 10% receive 1 globe.

[8]This is a difference-in-difference type of calculation, where one difference is between the fund and its benchmark and the other difference is between funds with different numbers of globes.

funds into two groups, "greener" (top 30%) and "browner" (bottom 30%), for each of the three scores. We do the same for the composite historical sustainability score, based on which Morningstar assigns globes to each fund.[9] We perform the greener-versus-browner comparisons in Figure 5, whose four panels are analogous to Panel B of Figure 4, except that sustainability globes are replaced by the four metrics described above. In all four panels, we benchmark funds against FTSE/Russell.

Figure 5 shows that funds with high sustainability scores outperform those with low scores. This result from Panel A is not surprising, given the prior results from Figure 4. More interesting, funds with high environmental (E) scores outperform those with low E scores (Panel B), whereas funds with high social (S) scores underperform those with low S scores (Panel C). According to Panel D, funds' governance (G) scores have no effect on performance. To make the figure easy to read, we do not show confidence intervals, but we do show them in the Appendix. Only the pattern in Panel B is statistically significant.

Figures 3 through 5 demonstrate that more sustainable funds perform better during the crisis. We further examine this result by conducting regression analysis, with two benefits. First, regressions allow us to see whether the result survives the inclusion of many control variables. Second, we remove the slight survivorship bias discussed earlier.

Table 2 analyzes the determinants of crisis-period fund performance in cross-sectional regressions with large numbers of controls. Panel A focuses on benchmark-adjusted performance, using FTSE/Russell benchmarks. Panel B considers factor-adjusted performance, using the four-factor Carhart (1997) model. The right-hand-side variables include indicators for sustainability, exclusions, and the growth investment style, as well as the Morningstar star rating. Fund-level controls include the log of fund age, the log of the fund's total net assets, fund turnover, expense ratio, cash position, the Morningstar medal rating, and market beta (in Panel B only because there is no beta estimation in Panel A). Industry controls are the fund's net investment position as a percentage of net assets in industries including energy, healthcare, and technology, among others. All regressions include style fixed effects, where style is measured at the level of the Morningstar Global Category. As a result, the relevant comparisons are across funds within the same investment style.

Table 2 confirms that funds with higher sustainability ratings perform better during the crisis. As before, we define high-sustainability funds as those with 4 or 5 globes. Column 1 of the table includes no controls other than style fixed effects. In column 1 of Panel A, the slope on the high-sustainability indicator is 14.21 (t = 4.85), indicating that high-

[9] A fund's individual E, S, and G scores do not simply add up to the fund's historical sustainability score. There does not appear to be a simple direct mapping between the two sets of scores.

sustainability funds outperform the remaining funds within the same style by 14.21% per year during the crisis. The high-sustainability indicator is also highly significant in column 1 of Panel B (t = 4.25), where fund returns are factor-adjusted rather than benchmark-adjusted. Sustainability thus remains a significant determinant of performance even after style fixed effects are included. This is not surprising, given our prior results, because both sustainability ratings and fund returns are style-adjusted, though in slightly different ways—sustainability ratings by Morningstar, with respect to the Morningstar Global Category, and returns by us, with respect to the fund's FTSE/Russell benchmark.

More interesting, sustainability remains significantly associated with fund performance after the inclusion of fund-level and industry controls. The slope on the high-sustainability indicator decreases as controls are added, but it remains both statistically and economically significant even when all controls are included: 9.76 (t = 2.60) in Panel A and 3.47 (t = 3.15) in Panel B. To summarize, we find that funds with high sustainability ratings perform better during the crisis.

A subset of funds employ exclusions in their investment process. These funds exclude from their portfolios stocks of firms such as tobacco producers or gun manufacturers whose business they deem unacceptable. Exclusions represent one possible approach to sustainability, one that discards the opportunities to engage with the firm as well as to benefit from the potential mispricing of the firm's stock. 107 of our funds employ exclusions, representing 3.9% of our TNA-screened sample.

Table 2 shows that funds that employ exclusions outperform same-style funds that do not employ exclusions by 8.61% per year (t = 3.26) in terms of benchmark-adjusted returns. However, this result weakens, and eventually loses significance, after adding enough control variables. The result does not obtain on a factor-adjusted basis; if anything, it goes the other way (Panel B). The exclusion aspect of sustainability therefore does not have a robust association with fund performance during the crisis.

3.2. Star Ratings

Besides sustainability, the most important determinant of active fund performance during the crisis is the fund's star rating as of January 31, 2020. To calculate star ratings, Morningstar computes each fund's risk-adjusted performance over the prior three, five, and ten years relative to the fund's peer group. Averaging across the three periods, Morningstar awards

1, 2, 3, 4, or 5 stars to each fund, with 5 stars going to the best-performing funds.[10] We find that funds with higher star ratings perform better during the crisis.

Figure 6 shows the distributions of cumulative returns during the crisis across funds with different star ratings, similar to Figure 3. We collect funds in two groups: funds with 4 or 5 stars ("high") and funds with 1 or 2 stars ("low"). Panel A shows the distributions of cumulative total fund returns whereas Panel B shows cumulative returns adjusted for FTSE/Russell benchmark returns. Both panels clearly show that funds with more stars perform better during the crisis.

Figure 7 shows the same result from a different perspective. Similar to the layouts from Figures 2 and 4, Figure 7 plots cumulative benchmark-adjusted fund performance for groups of funds with different star ratings. The relation between benchmark-adjusted performance and star ratings is monotonic across the five star groups, with five-star funds performing the best and one-star funds performing the worst. This striking monotonicity is observed for all three benchmarks. The figure also shows that five-star funds outperform one-star funds by a significant margin for all three benchmarks.

Table 2 confirms the important role of the star rating in our regression setting with controls and style fixed effects. The star rating significantly predicts both benchmark-adjusted and factor-adjusted returns, with t-statistics ranging from 2.42 to 3.50 in Panel A and from 2.79 to 5.92 in Panel B. This is a surprising result—it is not clear a priori why Morningstar star ratings, which are computed before the crisis from historical risk-adjusted returns, should have such strong predictive power for fund performance during the crisis. The result is significant not only statistically but also economically. For example, the slope coefficient of 5.78 in column 3 of Panel A indicates that one extra star is associated with a higher crisis-period benchmark-adjusted return of 5.78% per year. Therefore, a five-star fund outperforms a one-star fund of the same style by four times that amount, 23.1% per year, on average.

3.3. Value versus Growth

Sustainability and stars are the most robust predictors of active fund performance during the crisis. Next in line, though less robust, is the value/growth investment style. We find that growth funds tend to outperform value funds. Importantly, we are not saying that the

[10] As with the sustainability globes, within each peer group, the top 10% of funds receive 5 stars, the next 22.5% receive 4 stars, the next 35% receive 3 stars, the next 22.5% receive 2 stars, and the bottom 10% receive 1 star. A fund must have at least three years of performance to be considered for a rating, and depending on its age, a combination of three-year, five-year, and ten-year performance measures are averaged to construct the fund's raw performance score.

growth style outperforms the value style during the crisis—that is well known (e.g., HML's crisis-period return is −18%). What we are saying is that growth funds deliver higher returns than value funds on a style-adjusted basis.

To decide which funds follow the value and growth investment styles, we use the equity style box variable from Morningstar. We define value funds as funds classified as large-cap value, mid-cap value, or small-cap value. We define growth funds as funds classified as large-cap growth, mid-cap growth, or small-cap growth.

Figure 8 shows that growth funds outperform value funds, on average, for all three benchmarks. The outperformance is statistically significant when measured against the S&P 500 and prospectus benchmarks, but not against the FTSE/Russell benchmarks.

Table 2 finds the same outperformance in our regression setting with controls and style fixed effects. In Panel B, where fund performance is factor-adjusted, the indicator variable for the growth tilt is always positive and significant, with t-statistics exceeding 3.7 in all specifications. For example, using the estimate from column 4, growth funds outperform non-growth funds by 10.62% per year (t = 5.58) during the crisis. In Panel A, the estimated slopes on the growth indicator are similar in magnitude to those in Panel B but their statistical significance is substantially weaker, with t-statistics ranging from 0.75 to 2.35 across the four specifications.

3.4. Robustness

Our main results are robust to a variety of methodological modifications. As noted earlier, our sample is restricted to active equity funds with at least $15 million in total net assets as of January 31, 2020. However, we show in the Appendix that the results from Table 1 are extremely similar even if we do not impose this size screen.

Another screen that is commonly imposed on mutual fund samples is an age screen. Researchers often exclude young funds because of a concern about the incubation bias (Evans, 2010). This bias can appear if researchers analyze historical fund data with a delay that would allow the bias to creep in. The bias is not a concern in our study because there is no such delay—we constructed our fund sample in May 2020, shortly after the end of our sample period. Nonetheless, we show in the Appendix that our main results are extremely similar also when we exclude funds less than two years old from the sample.

Our tables report evidence based on simple returns. Our plots of cumulative performance

are based on log (i.e., continuously compounded) returns because those returns cumulate over time in a tractable manner. This distinction is immaterial—our main table results are very similar if we replace simple returns by log returns, as we show in the Appendix.

To remove the effects of investment style, we include style fixed effects in our regressions. Nonetheless, we show in the Appendix that the regression results from Table 2 are similar if style fixed effects are excluded. The Appendix also reports subperiod results for Table 2 and its variations, dividing the full crisis period into the crash period (February 20 to March 23, 2020) and the recovery period (March 24 to April 30, 2020).

Recall that Figure 3 shows the distributions of crisis-period returns across two groups of funds, those with high sustainability ratings (4 or 5 globes) and low sustainability ratings (1 or 2 globes). In the Appendix, we present analogous plots showing three distributions corresponding to funds with 1, 3, and 5 globes, and also five distributions, one for each possible number of globes. Those plots are more cluttered but they convey the same message as Figure 3—that more sustainable funds perform better in the crisis.

Similarly, Figure 6 shows the distributions of returns across funds with high star ratings (4 or 5 stars) and low star ratings (1 or 2 stars). In the Appendix, we present analogous plots showing three distributions corresponding to funds with 1, 3, and 5 stars, and also five distributions, one for each possible number of stars. Again, those plots convey the same message as Figure 6: funds with more stars perform better during the crisis.

4. Fund Flows

Our key measure for assessing fund flows is the cumulative net fund flow percentage. Daily net fund flows, in dollars, are computed following Barber, Huang, and Odean (2016) as

$$FD_{i,t} = TNA_{i,t} - (1 + R_{i,t})TNA_{i,t-1} \ , \tag{3}$$

where $TNA_{i,t}$ is the total net assets of fund i on date t and $R_{i,t}$ is the net return of fund i on date t. To convert the dollar values of net fund flows into a cumulative percentage, the values of $FD_{i,t}$ are accumulated across the time period of interest and divided by the total net assets of fund i on the day before the period of interest begins. Given the sensitivity of the cumulative net flow percentage to missing values, we restrict consideration to funds with entirely non-missing daily net fund flows.[11]

[11] In our baseline fund size filtered sample, we retain 2,082 funds (75.3%) over the full February 20 to April 30, 2020 time period, 2,137 funds (77.3%) over the February 20 to March 23, 2020 time period, and 2,219 funds (80.3%) over the March 24 to April 30, 2020 time period.

Figure 9 shows the time series of cumulative net fund flows into active equity mutual funds, both in dollar terms and in percentage terms. Panel A shows that active funds experience steady outflows during the COVID-19 crisis of about 43 billion dollars, or 1.3% of assets under management. The pace of outflows is fairly rapid during the market crash between February 20 and March 23, 2020. Outflows continue, albeit at a slower pace, after the market rebound. Panel B shows that between January 2017 and April 2020, active funds experience outflows amounting to about 5% per year as a fraction of their initial assets. These steady outflows reflect the well-known ongoing trend toward passive investment management. Year 2020 does not stand out relative to prior years, indicating that crisis-period outflows largely extend their long-term pre-crisis trend.

4.1. Sustainability

We find that funds with higher sustainability ratings (i.e., more Morningstar globes) receive larger net flows during the crisis. Figure 10 plots cumulative net fund flows across funds with different globe ratings over the February 20 to April 30, 2020 time period. Panel A plots aggregate cumulative flows for each globe category, which we compute as total cumulative net flows into that category divided by that category's total net assets on February 19, 2020. The panel shows a near-monotonic relation between those ratings and net fund flows, with five-globe funds having the largest flows and one-globe funds the lowest flows. Panel B focuses on the five-globe and one-globe categories. For both of them, the panel plots average fund-level cumulative flows across all funds in the given category, scaled as a percent of the fund's February 19, 2020 total net assets, along with the 95% confidence intervals. The panel shows that five-globe funds receive significantly larger net flows than one-globe funds. In short, investors favor sustainable funds when moving their money during the crisis.

Figure 11 shows a similar pattern based on a different measure of sustainability: an indicator of whether the fund employs exclusions in its investment process. Funds that do not employ exclusions, which account for the vast majority of funds, experience net outflows during the crisis. However, funds that do employ exclusions experience net inflows, and the difference between the two groups is statistically significant.

Figure 12 unpacks sustainability into its E, S, and G dimensions. As before, we separate funds into high-E, low-E, high-S, low-S, high-G and low-G, where the high (low) group always denotes the top (bottom) 30% of funds. Figure 12 shows that cumulative aggregate net flows during the crisis are larger for high-E funds than for low-E funds. Low-E funds experience substantial outflows of 2.7% of assets, whereas the outflows from high-E funds

are only 0.3%. Net flows are also larger for high-G funds than for low-G funds: low-G funds have outflows of 2.0% whereas high-G funds' outflows are only 1.0%. High-S funds actually experience larger outflows than low-S funds, but the difference between the two categories' total flows is relatively small, only 0.7%, and the average outflow is in fact slightly larger for the low-S category. The effect of sustainability on fund flows thus seems driven by E and, to a lesser extent, G. Investors seem to have retained their focus on environmental issues even during the health crisis of 2020.

Table 3 reports results from cross-sectional regressions of crisis-period net fund flows on the sustainability variables as well as style fixed effects and a large number of fund and industry controls. The sustainability variables remain significant even with these controls. The exclusions variable is the most robust, with t-statistics ranging from 2.50 to 4.02 across the seven different specifications. The indicator for high-E (i.e., green) funds is also generally significant, with t-statistics ranging from 1.89 to 3.09. The five-globe indicator is significant when included on its own as well as in several other specifications, but it loses significance when the exclusions indicator is included. Overall, these results indicate that investors favor sustainable funds while reallocating money during the pandemic of 2020.

4.2. Other Determinants of Fund Flows

Figure 13 plots cumulative net fund flows for funds with different star ratings as of January 31, 2020. Panel A shows a monotonic flow-star relation, with higher-star funds receiving higher net flows. Panel B shows that the differences in average cumulative net flows are statistically significant: five-star funds receive significantly larger average net flows than three-star funds, whose average flows are significantly larger than those of one-star funds. The positive flow-star relation is highly significant also after controlling for style fixed effects and the other controls (see Table 3). This relation is not surprising, however, because star ratings are constructed based on past returns, which are well known to have predictive power for fund flows. Star ratings essentially perform the role of catch-all controls for past returns in our regressions.

Figure 14 shows that growth funds receive significantly larger net flows than value funds. This relation holds largely at the style level because it vanishes when we run regressions that include style fixed effects. We include the growth indicator variable among the fund-level controls in Table 3 but we suppress it because it is never statistically significant.

As in our analysis of fund performance, we also report our regression results in specifi-

cations without style fixed effects. The results are quite similar to those reported in Table 3, as we show in the Appendix. The Appendix also reports subperiod (crash and recovery) results for Table 3, as well as for its version with no style fixed effects.

5. Conclusions

We analyze the performance and flows of U.S. active equity mutual funds during the COVID-19 crisis. We find that most active funds underperform passive benchmarks during the crisis, contradicting the hypothesis that active funds outperform in recessions. This underperformance is particularly large when measured against the S&P 500 index, but it is observed also relative to fund-specific style benchmarks. Funds with high sustainability ratings and high star ratings outperform those with low sustainability ratings and low star ratings, respectively. When reallocating capital across funds, investors favor funds with high sustainability ratings and funds that apply exclusion criteria. That investors retain their focus on sustainability during a major crisis indicates that sustainability is not just a luxury good.

While this paper appears to be the first to analyze the performance and flows of active funds during the COVID-19 crisis, it leaves plenty of room for future research. For example, while we focus on equity funds, fixed income funds also deserve a careful investigation. So do the sources of active funds' underperformance during the crisis. It would also be useful to extend our work to deepen our understanding of the dependence of investors' demand for sustainability on economic conditions.

Table 1
Fund Performance

This table describes active equity mutual funds' performance against both benchmarks and factor models. Panel A reports simple averages across funds of estimated deltas and alphas, all reported in annualized percentage terms. The deltas are average differences between the fund's net returns and its benchmark returns. The benchmarks are the FTSE/Russell benchmark, the prospectus benchmark, and the S&P 500. The alphas are estimated intercepts from the regressions of excess net fund returns on factor returns. The factor models are described in the text. Panel B reports the fraction of funds that underperform (i.e., have a negative delta or alpha). The time periods are: crisis (February 20 to April 30, 2020); crash (February 20 to March 23, 2020); recovery (March 24 to April 30, 2020); and pre-crisis (October 1, 2019 to January 31, 2020). Standard errors are clustered on the Morningstar Institutional Category. t-statistics are in brackets.

	(1) $\Delta_{Bench}^{FTSE/Rus}$	(2) Δ_{Bench}^{Prosp}	(3) $\Delta_{Bench}^{S\&P500}$	(4) α^{CAPM}	(5) α^{FF3}	(6) α^{Car4}	(7) α^{FF5}	(8) $\alpha^{FF5+Mom}$
Panel A. Average Fund Performance (%)								
Crisis	-11.02	-7.70	-29.12	-29.11	-11.30	-7.84	-9.90	-7.62
	[-3.90]	[-2.49]	[-5.37]	[-7.02]	[-5.16]	[-3.22]	[-4.51]	[-3.25]
Crash	-7.91	-8.19	-64.31	-80.94	-37.75	-38.98	-49.84	-51.11
	[-1.54]	[-1.33]	[-4.77]	[-8.33]	[-4.25]	[-3.90]	[-5.91]	[-5.77]
Recovery	-12.68	-7.55	-5.81	8.47	17.49	20.76	18.75	22.83
	[-3.94]	[-2.13]	[-1.14]	[1.46]	[2.83]	[4.12]	[3.00]	[4.42]
Pre-Crisis	-2.11	-1.28	-3.03	-5.14	-3.12	-3.01	-2.65	-2.59
	[-2.86]	[-2.09]	[-1.80]	[-3.87]	[-4.37]	[-3.64]	[-3.22]	[-2.88]
Panel B. Fraction of Funds Underperforming (%)								
Crisis	57.59	54.17	74.24	80.15	69.66	60.35	67.80	60.43
Crash	51.62	48.57	63.48	83.79	78.57	80.01	80.65	81.41
Recovery	55.73	55.64	55.77	53.09	39.53	34.79	40.08	34.28
Pre-Crisis	63.58	59.85	67.26	71.20	73.31	71.54	70.02	69.51

Table 2
Determinants of Fund Performance During the Crisis

The table reports slope coefficients estimated from regressions of fund performance in February 20 to April 30, 2020 on fund characteristics and controls. In Panel A, the dependent variable is FTSE/Russell-benchmark-adjusted performance; in Panel B, it is the Carhart four-factor alpha. Both performance measures are estimated using simple returns and expressed in annualized percentage terms. Global category fixed effects are based on the Morningstar Global Category variable. Fund-level controls include the log of the fund's age in days, the log of the fund's January 31, 2020 total net assets (TNA), turnover ratio as of January 2020, net expense ratio as of January 2020, net cash position (as a percent of TNA) as of January 2020, Morningstar medal rating as of January 2020, and, in Panel B only, market beta estimated from the October 1, 2019 to January 31, 2020 period. Industry controls include the fund's net position as a percent of TNA in basic materials, communication services, consumer cyclical, consumer defensive, healthcare, industrials, real estate, technology, energy, financial services, and utilities. Standard errors are clustered on the Morningstar Institutional Category. t-statistics are in brackets.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
			Panel A. Benchmark-Adjusted Performance				
$\mathbb{I}$(4 or 5 Sustainability Globes)	14.21				11.51	8.61	9.76
	[4.85]				[3.22]	[2.26]	[2.60]
$\mathbb{I}$(Employs Exclusions)		8.61			5.47	2.03	2.79
		[3.26]			[2.44]	[1.05]	[1.24]
Star Rating			5.78		5.12	7.00	6.49
			[2.84]		[2.42]	[3.50]	[3.41]
$\mathbb{I}$(Growth Tilt)				12.43	7.24	9.39	5.15
				[2.35]	[1.16]	[1.70]	[0.75]
Global Category FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fund-Level Controls	No	No	No	No	No	Yes	Yes
Industry Controls	No	No	No	No	No	No	Yes
Observations	2,494	2,561	2,286	2,561	2,251	1,632	1,604
Adjusted R^2	0.06	0.05	0.06	0.06	0.06	0.12	0.15
			Panel B. Factor-Adjusted Performance				
$\mathbb{I}$(4 or 5 Sustainability Globes)	5.59				2.67	3.04	3.47
	[4.25]				[2.39]	[2.55]	[3.15]
$\mathbb{I}$(Employs Exclusions)		−0.89			−2.61	−3.46	−3.16
		[−0.50]			[−1.52]	[−2.12]	[−2.19]
Star Rating			3.15		2.51	3.25	3.13
			[3.35]		[2.79]	[5.92]	[5.42]
$\mathbb{I}$(Growth Tilt)				10.62	7.53	7.51	7.77
				[5.58]	[4.09]	[3.74]	[4.10]
Global Category FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fund-Level Controls	No	No	No	No	No	Yes	Yes
Industry Controls	No	No	No	No	No	No	Yes
Observations	2,233	2,363	2,104	2,363	2,020	1,522	1,494
Adjusted R^2	0.10	0.12	0.11	0.12	0.10	0.42	0.46

Table 3
Determinants of Net Fund Flows During the Crisis

The table reports slope coefficients estimated from regressions of net fund flows in February 20 to April 30, 2020 on fund characteristics and controls. A fund's net flow is expressed as a percent of the fund's February 19, 2020 total net assets. Flows are winsorized at the 2.5% and 97.5% levels. Global category fixed effects are based on the Morningstar Global Category variable. Fund-level controls include an indicator for a growth tilt, the log of the fund's age in days, the log of the fund's January 31, 2020 total net assets (TNA), turnover ratio as of January 2020, net expense ratio as of January 2020, net cash position (as a percent of TNA) as of January 2020, Morningstar medal rating as of January 2020, and market beta estimated from the October 1, 2019 to January 31, 2020 period. Industry controls include the fund's net position as a percent of TNA in basic materials, communication services, consumer cyclical, consumer defensive, healthcare, industrials, real estate, technology, energy, financial services, and utilities. Standard errors are clustered on the Morningstar Institutional Category. t-statistics are in brackets.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$\mathbb{I}$(5 Sustainability Globes)	1.76	0.92	1.36	1.44		0.70	1.16	1.22				
	[3.19]	[1.82]	[2.23]	[2.32]		[1.23]	[1.77]	[1.82]				
$\mathbb{I}$(Employs Exclusions)					2.84	2.61	2.70	2.75		2.69	2.95	3.11
					[3.37]	[3.09]	[2.50]	[2.53]		[4.02]	[3.46]	[3.43]
$\mathbb{I}$(Greener E)									1.67	1.02	1.04	1.91
									[3.09]	[2.08]	[1.89]	[2.72]
$\mathbb{I}$(Greener S)									−0.48	−0.43	−0.39	−0.88
									[−0.93]	[−0.83]	[−0.64]	[−1.22]
$\mathbb{I}$(Greener G)									0.76	0.63	1.00	1.26
									[1.39]	[1.09]	[1.31]	[1.69]
Star Rating		1.83	1.78	1.90		1.82	1.76	1.89		1.70	1.38	1.49
		[7.49]	[6.04]	[6.00]		[7.51]	[6.04]	[5.96]		[5.77]	[3.91]	[4.01]
Global Category FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fund-Level Controls	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
Industry Controls	No	No	No	Yes	No	No	No	Yes	No	No	No	Yes
Observations	2,082	1,863	1,434	1,390	2,082	1,863	1,434	1,390	1,503	1,348	1,037	1,020
Adjusted R^2	0.02	0.08	0.11	0.11	0.02	0.09	0.11	0.11	0.01	0.08	0.10	0.11

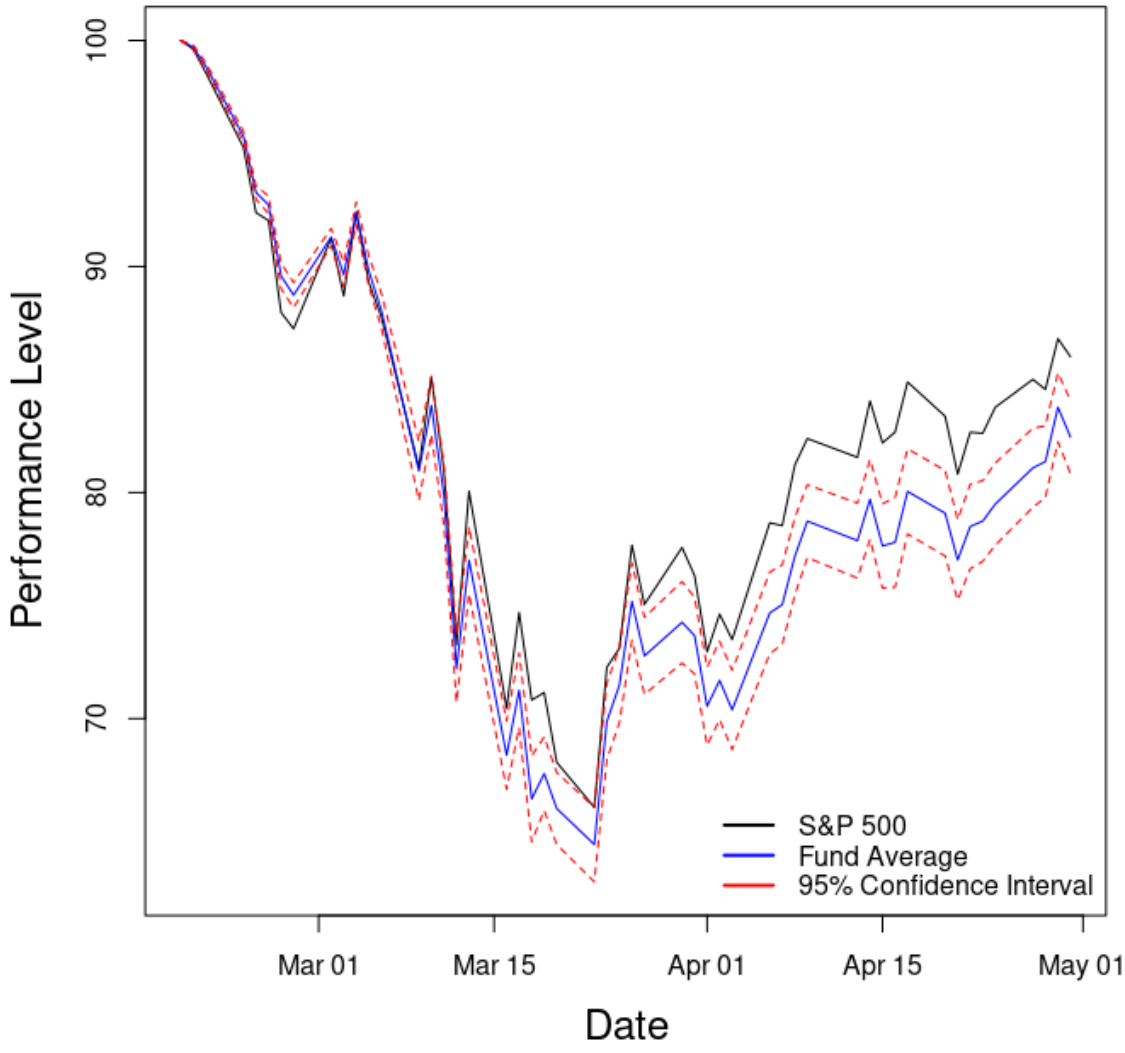

Figure 1. Average Fund Performance vs. the S&P 500 During the Crisis. This figure plots the performance of the average active equity mutual fund against the S&P 500 in February 20 through April 30, 2020. Both price indices are initialized at 100 on February 19, 2020 and computed by compounding daily returns. The fund average is computed by adding the average difference between the fund price index and the S&P 500 price index to the S&P 500 price index. Standard errors are estimated for this difference and are clustered on the Morningstar Institutional Category. 95% confidence intervals are plotted in red.

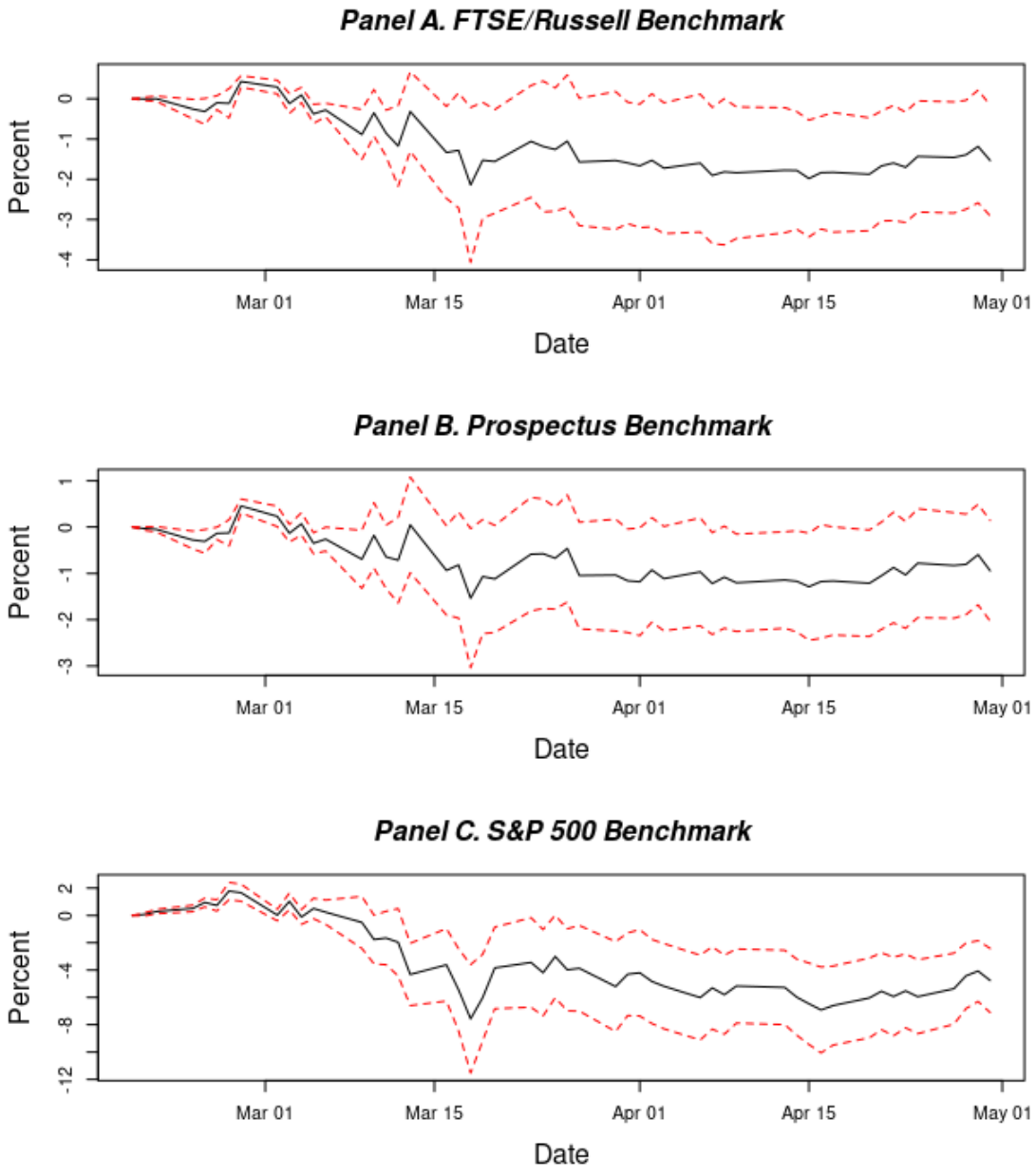

Figure 2. Average Benchmark-Adjusted Fund Performance. This figure plots the cumulative compound performance of the average active equity mutual fund in February 20 through April 30, 2020 relative to three benchmarks: the Morningstar-designated FTSE/Russell benchmark (Panel A), the prospectus benchmark (Panel B), and the S&P 500 (Panel C). Relative performance is measured by $\log(F_t) - \log(B_t)$, where F_t and B_t are the cumulative compounded daily returns of the average fund and the benchmark, respectively. Standard errors are estimated for this difference and are clustered on the Morningstar Institutional Category. 95% confidence intervals are plotted in red.

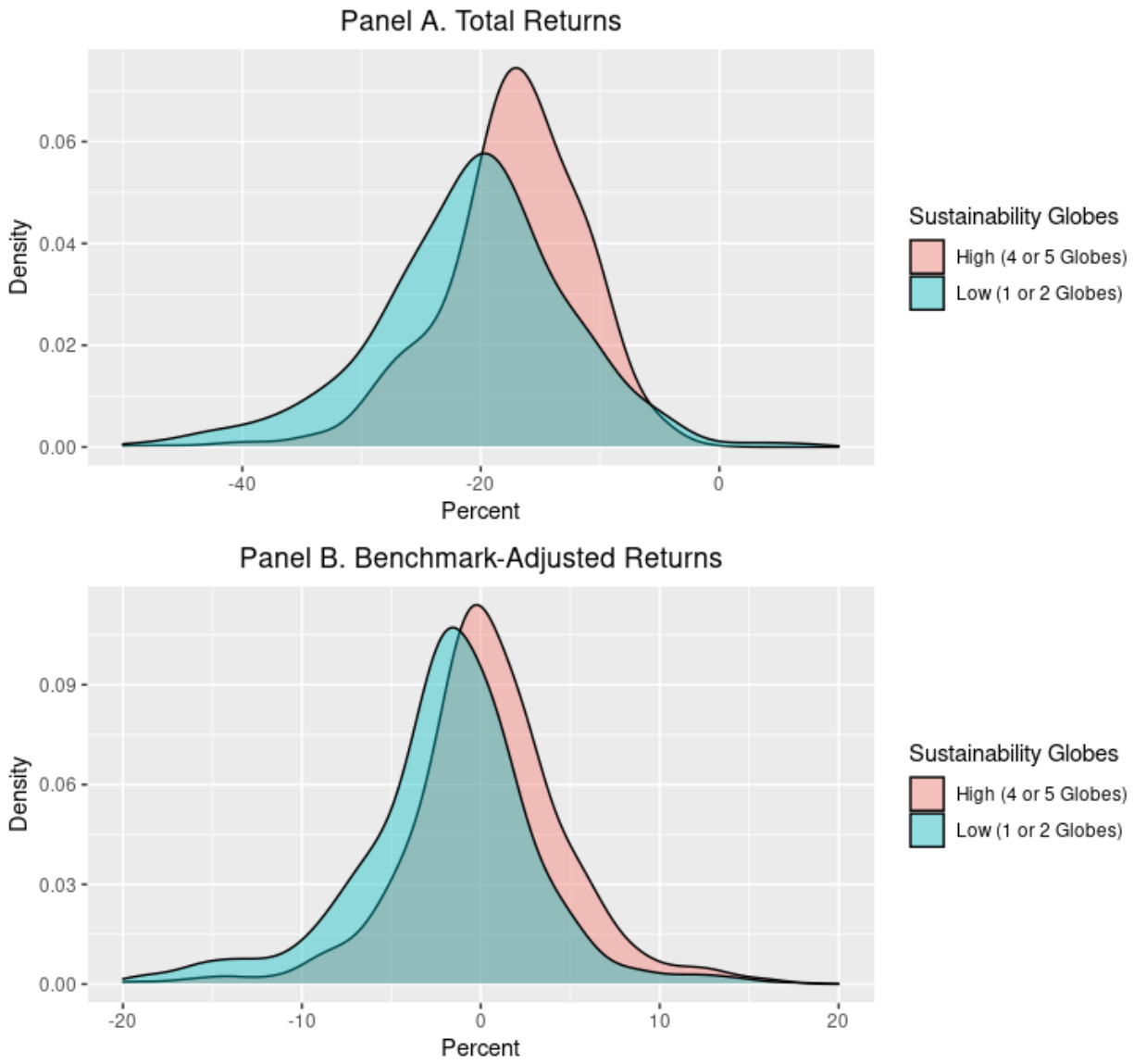

Figure 3. Cumulative Return Densities Across Sustainability Ratings. This figure plots densities of funds' cumulative returns from February 20 to April 30, 2020 for two categories of sustainability: high (four or five Morningstar globes) and low (one or two globes), both assigned as of January 2020. In Panel A, the cumulative returns are unadjusted, given by $\log(F_t)$ where $F_t = (1 + r_1^F)(1 + r_2^F)\dots(1 + r_t^F)$ is the fund's cumulative total return. In Panel B, the cumulative returns are benchmark-adjusted, given by $\log(F_t) - \log(B_t)$, where B_t is the cumulative total return of the fund's FTSE/Russell benchmark.

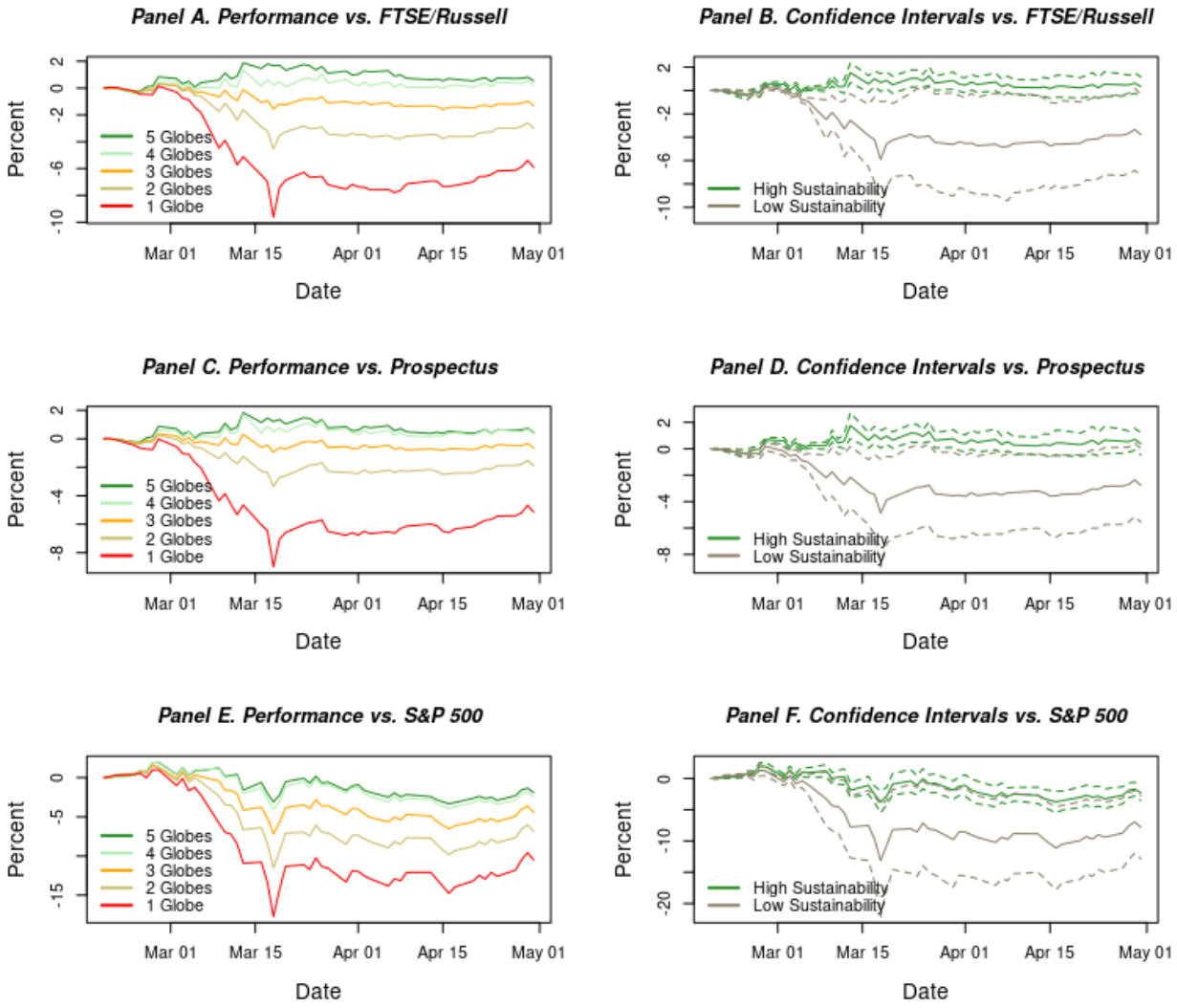

Figure 4. Benchmark-Adjusted Fund Performance: Sustainability Ratings. This figure plots the cumulative compound performance in February 20 through April 30, 2020 for fund categories with different numbers of Morningstar sustainability globes assigned as of January 2020. "High sustainability" denotes funds with 4 or 5 globes while "low sustainability" denotes funds with 1 or 2 globes. Performance is measured relative to the FTSE/Russell benchmark (Panels A and B), the prospectus benchmark (Panels C and D), and the S&P 500 (Panels E and F). Relative performance is measured by $\log(F_t) - \log(B_t)$, where F_t and B_t are the cumulative compounded daily returns of the average fund and the benchmark, respectively. Standard errors are estimated for this difference and are clustered on the Morningstar Institutional Category.

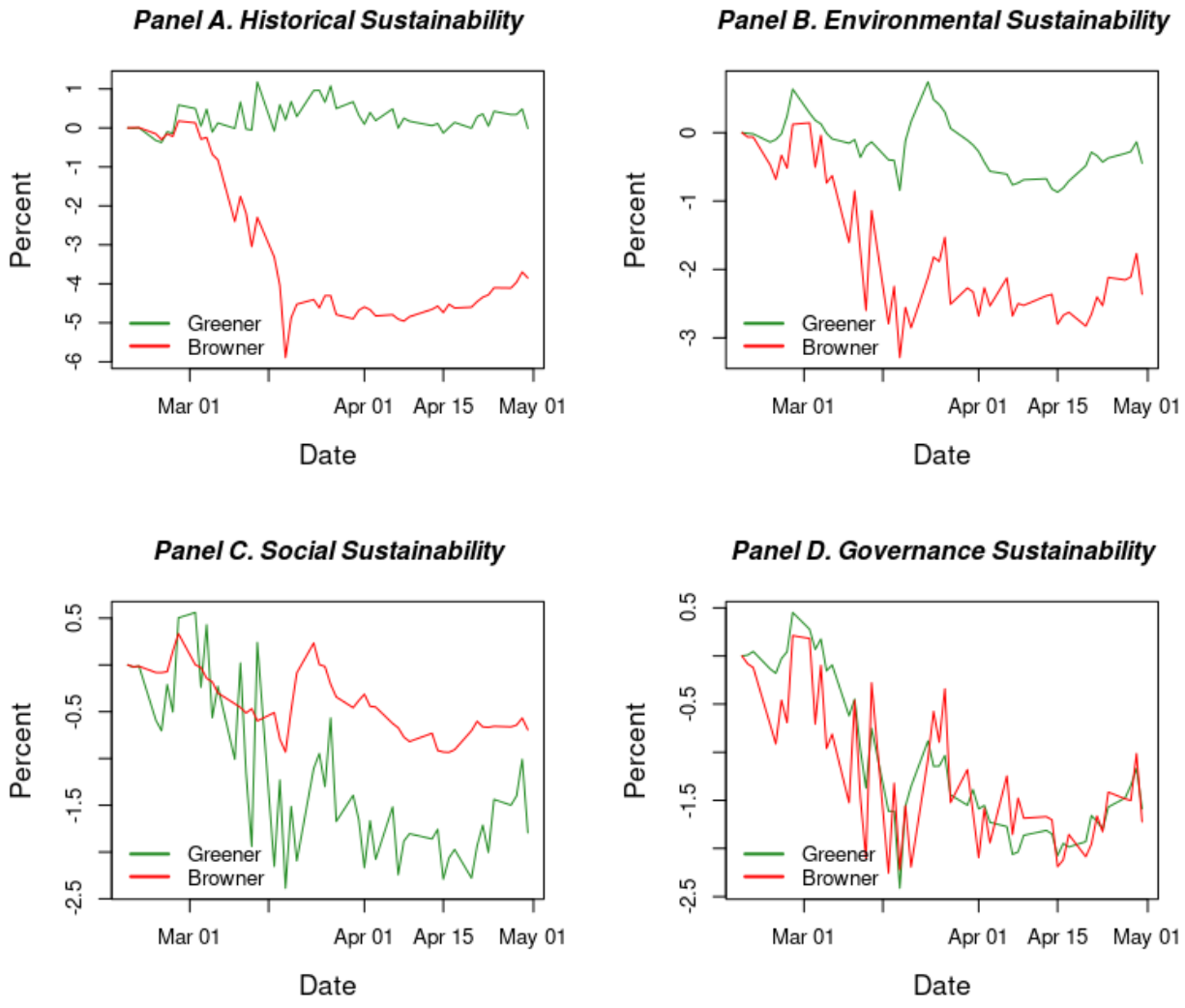

Figure 5. Benchmark-Adjusted Fund Performance: Sustainability Components. This figure plots the cumulative compound performance in February 20 through April 30, 2020 for fund categories with different Morningstar sustainability scores. These scores represent historical portfolio sustainability scores (Panel A), environmental scores (Panel B), social scores (Panel C), and governance scores (Panel D). For each of the four scores, the top 30% most sustainable funds are labeled as "greener" and the bottom 30% of funds are labeled "browner." Performance is measured relative to the FTSE/Russell benchmark by $\log(F_t) - \log(B_t)$, where F_t and B_t are the cumulative compounded daily returns of the average fund and the benchmark, respectively.

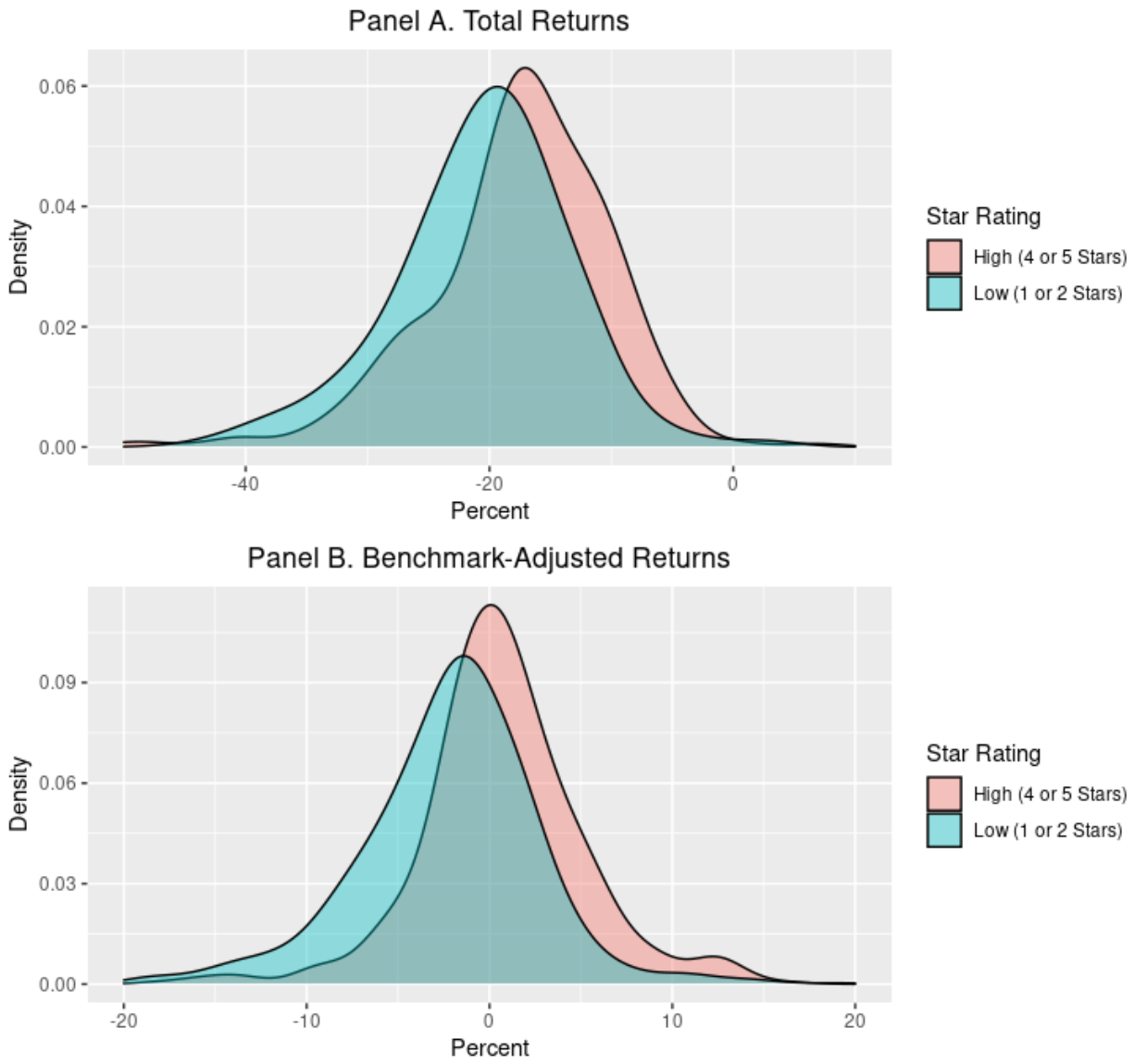

Figure 6. Cumulative Return Densities Across Star Ratings. This figure plots densities of funds' cumulative returns from February 20 to April 30, 2020 for two categories of star ratings: high (four or five Morningstar stars) and low (one or two stars), both assigned as of January 2020. In Panel A, the cumulative returns are unadjusted, given by $\log(F_t)$ where $F_t = (1 + r_1^F)(1 + r_2^F) \dots (1 + r_t^F)$ is the fund's cumulative total return. In Panel B, the cumulative returns are benchmark-adjusted, given by $\log(F_t) - \log(B_t)$, where B_t is the cumulative total return of the fund's FTSE/Russell benchmark.

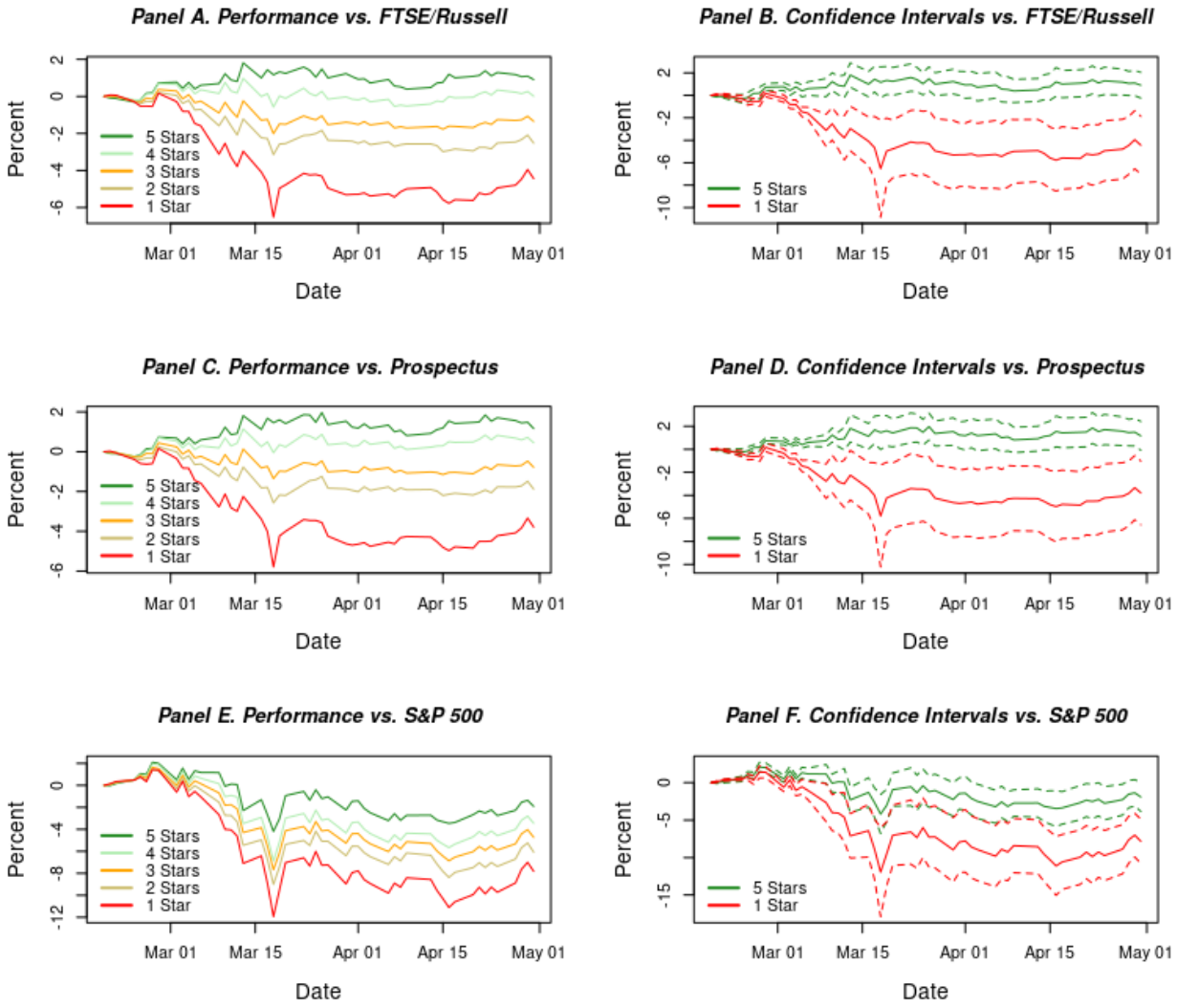

Figure 7. Benchmark-Adjusted Fund Performance: Star Ratings. This figure plots the cumulative compound performance in February 20 through April 30, 2020 for fund categories with different numbers of Morningstar stars assigned as of January 2020. Performance is measured relative to the FTSE/Russell benchmark (Panels A and B), the prospectus benchmark (Panels C and D), and the S&P 500 (Panels E and F). Relative performance is measured by $\log(F_t) - \log(B_t)$, where F_t and B_t are the cumulative compounded daily returns of the average fund and the benchmark, respectively. Standard errors are estimated for this difference and are clustered on the Morningstar Institutional Category.

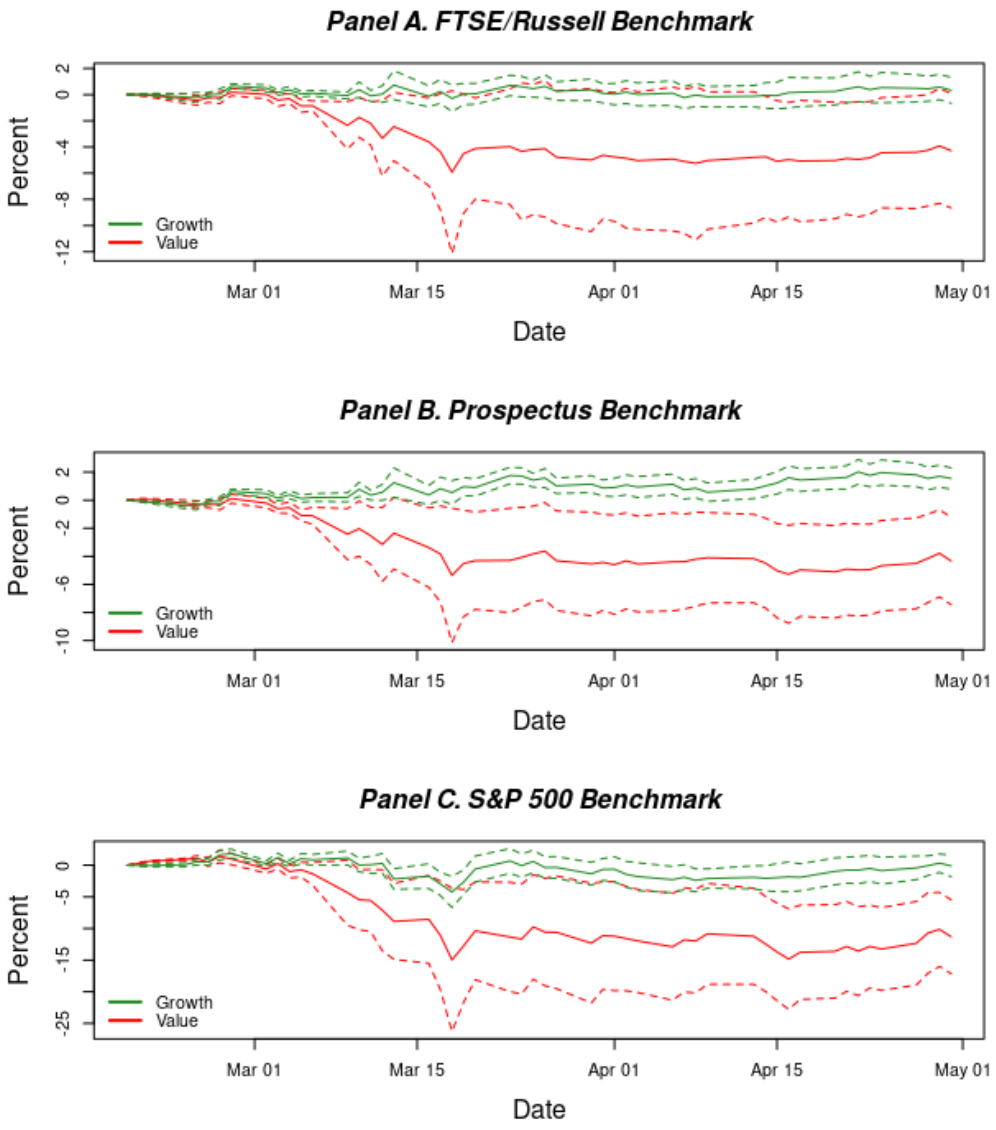

Figure 8. Benchmark-Adjusted Fund Performance: Growth vs. Value Funds. This figure plots the cumulative compound performance in February 20 through April 30, 2020 for growth vs value funds, as determined by the Morningstar equity style box. Performance is measured relative to the FTSE/Russell benchmark (Panel A), the prospectus benchmark (Panel B), and the S&P 500 (Panel C). Relative performance is measured by $\log(F_t) - \log(B_t)$, where F_t and B_t are the cumulative compounded daily returns of the average fund and the benchmark, respectively. Standard errors are estimated for this difference and are clustered on the Morningstar Institutional Category. 95% confidence intervals are shown.

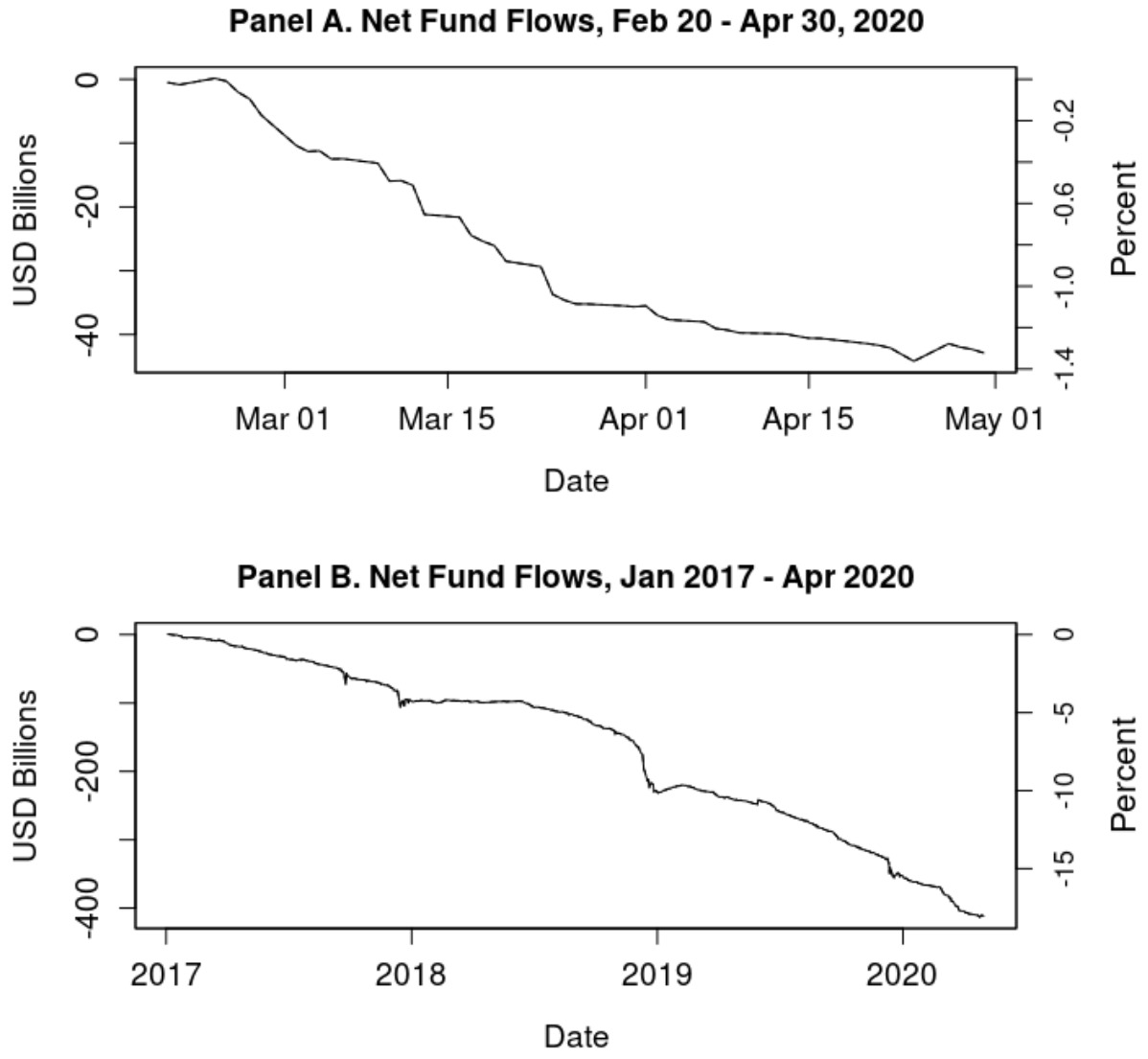

Figure 9. Aggregate Net Fund Flows. This figure plots aggregate net flows into active equity mutual funds during the crisis period (Panel A) and over the past three years (Panel B). Specifically, Panel A plots total cumulative net fund flows (in both USD billions and as a percent of February 19, 2020 aggregate total net assets) over the February 20 to April 30, 2020 period. Panel B covers the January 4, 2017 to April 30, 2020 period, and it expresses flows as a percent of January 3, 2017 total net assets.

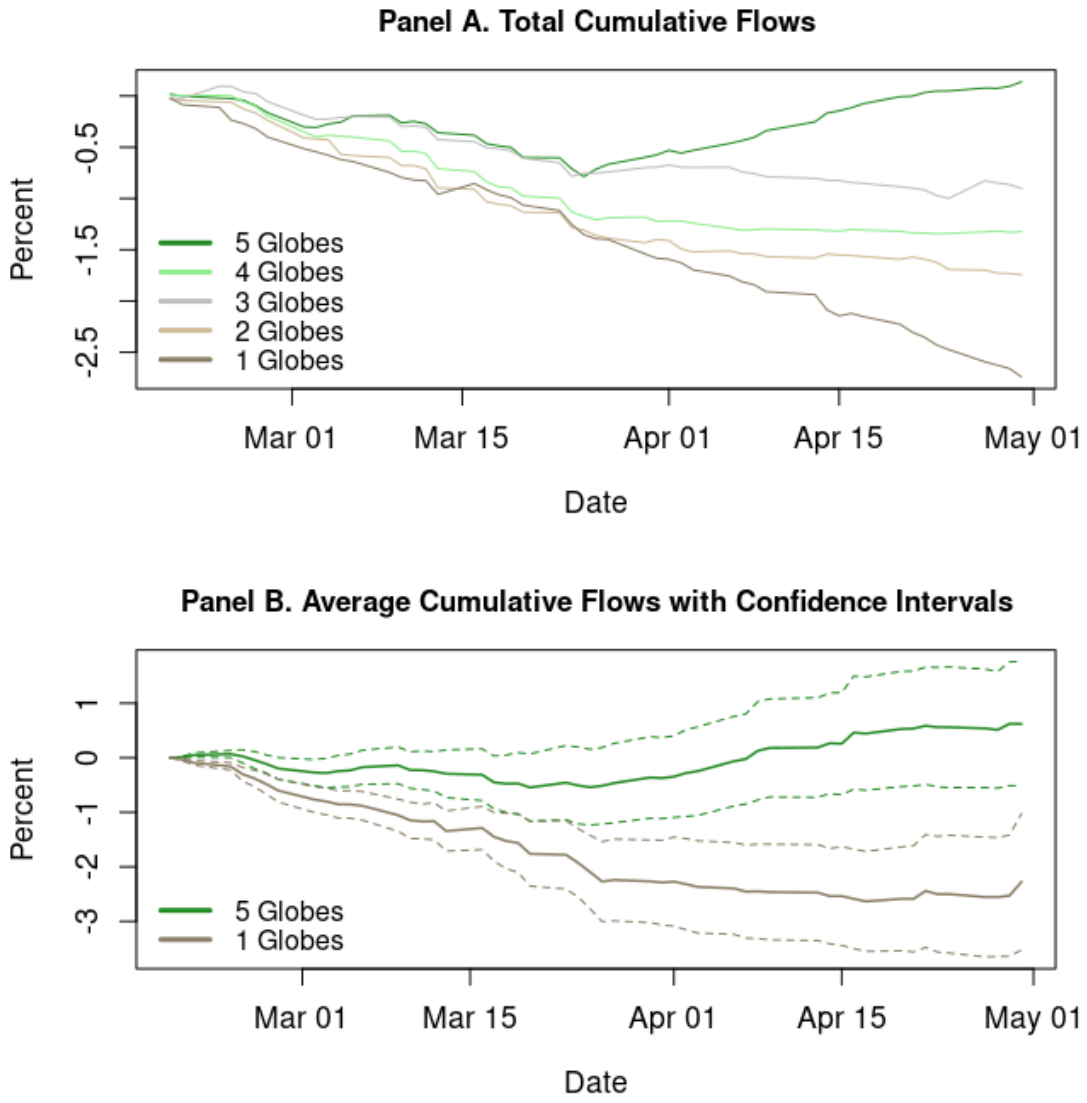

Figure 10. Fund Flows and Sustainability Ratings. This figure plots net fund flows over the February 20 to April 30, 2020 period for categories of funds sorted by Morningstar sustainability ratings. Panel A plots aggregate cumulative net flows for each of the five globe categories. Flows are aggregated within each category and accumulated over time, then scaled by the category's total net assets on February 19, 2020. Panel B plots the average across funds of cumulative net flows as a percent of the fund's February 19, 2020 total net assets, for the five- and one-globe categories only. Unlike in Panel A, the sample in Panel B is restricted to funds with at least $15 million of total net assets as of January 31, 2020 and the net fund flow percentage is winsorized at the 2.5% and 97.5% levels. Panel B also plots 95% confidence intervals, with standard errors clustered on the Morningstar Institutional Category.

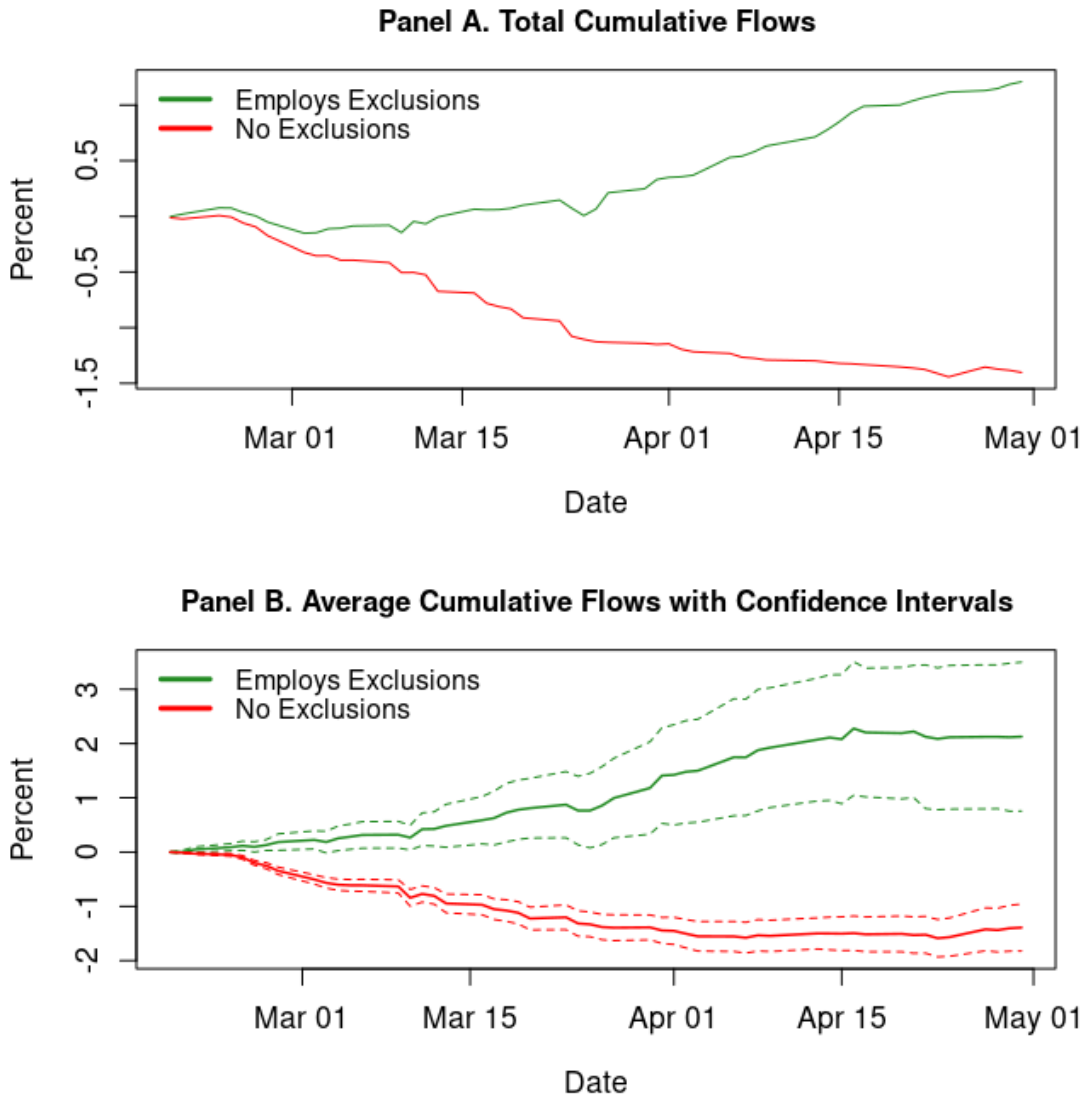

Figure 11. Fund Flows and Exclusions. This figure plots net fund flows over the February 20 to April 30, 2020 period for two categories of funds: those that do and do not employ exclusions in their investment process. Panel A plots aggregate cumulative net flows for both categories. Flows are aggregated within each category and accumulated over time, then scaled by the category's total net assets on February 19, 2020. Panel B plots the average across funds of cumulative net flows as a percent of the fund's February 19, 2020 total net assets, for both categories. Unlike in Panel A, the sample in Panel B is restricted to funds with at least $15 million of total net assets as of January 31, 2020 and the net fund flow percentage is winsorized at the 2.5% and 97.5% levels. Panel B also plots 95% confidence intervals, with standard errors clustered on the Morningstar Institutional Category.

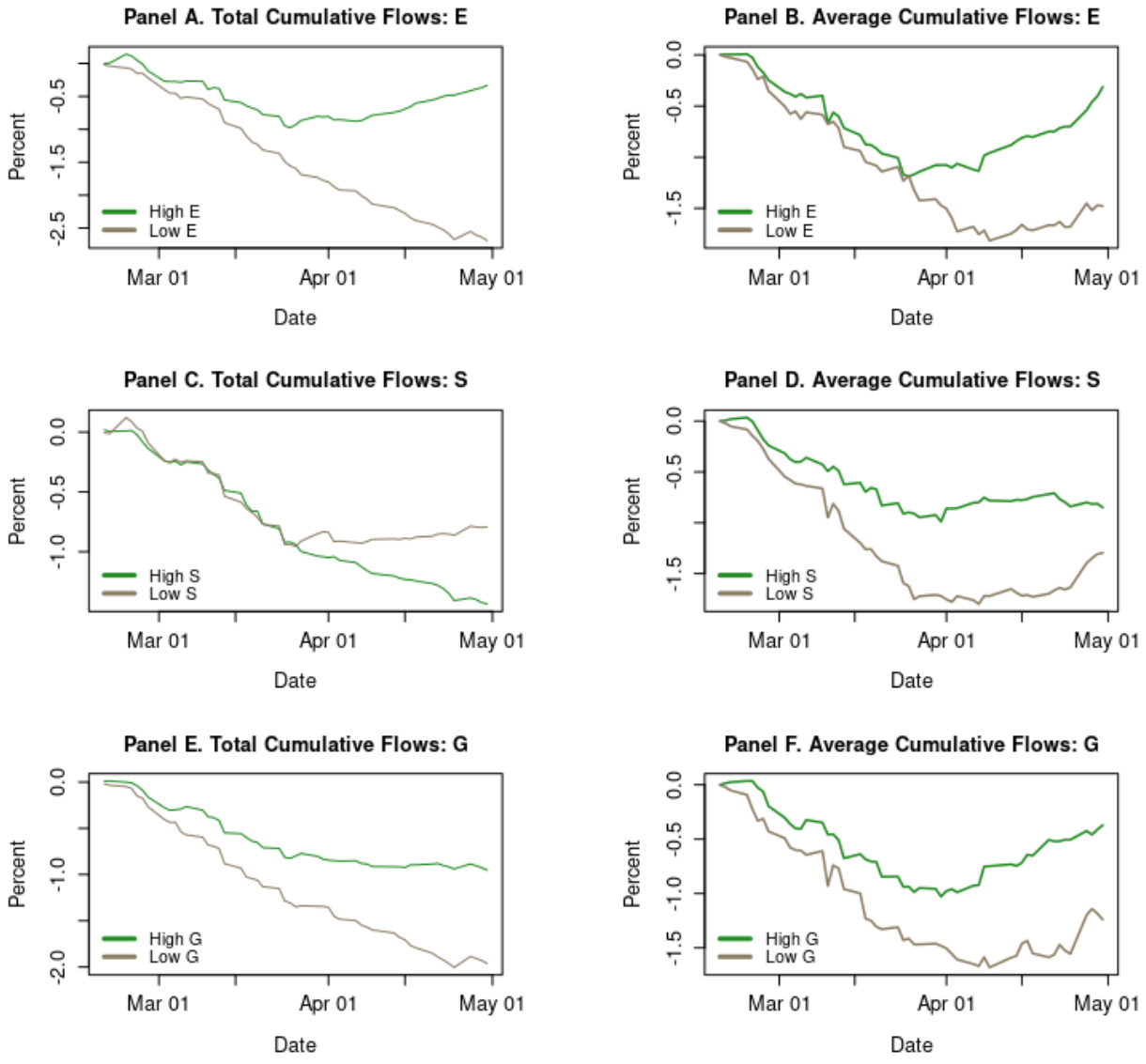

Figure 12. Fund Flows and ESG Scores. This figure plots net fund flows over the February 20 to April 30, 2020 period for funds in the top 30% ("high") and bottom 30% ("low") of environmental, social, and governance sustainability scores. The left panels plot aggregate cumulative net flows for the high and low E, S, and G categories. Flows are aggregated within each category and accumulated over time, then scaled by the category's total net assets on February 19, 2020. The right panels plot the average across funds of cumulative net flows as a percent of the fund's February 19, 2020 total net assets, for both the high and low categories. Unlike in the left panels, the samples in the right panels are restricted to funds with at least $15 million of total net assets as of January 31, 2020 and the net fund flow percentage is winsorized at the 2.5% and 97.5% levels.

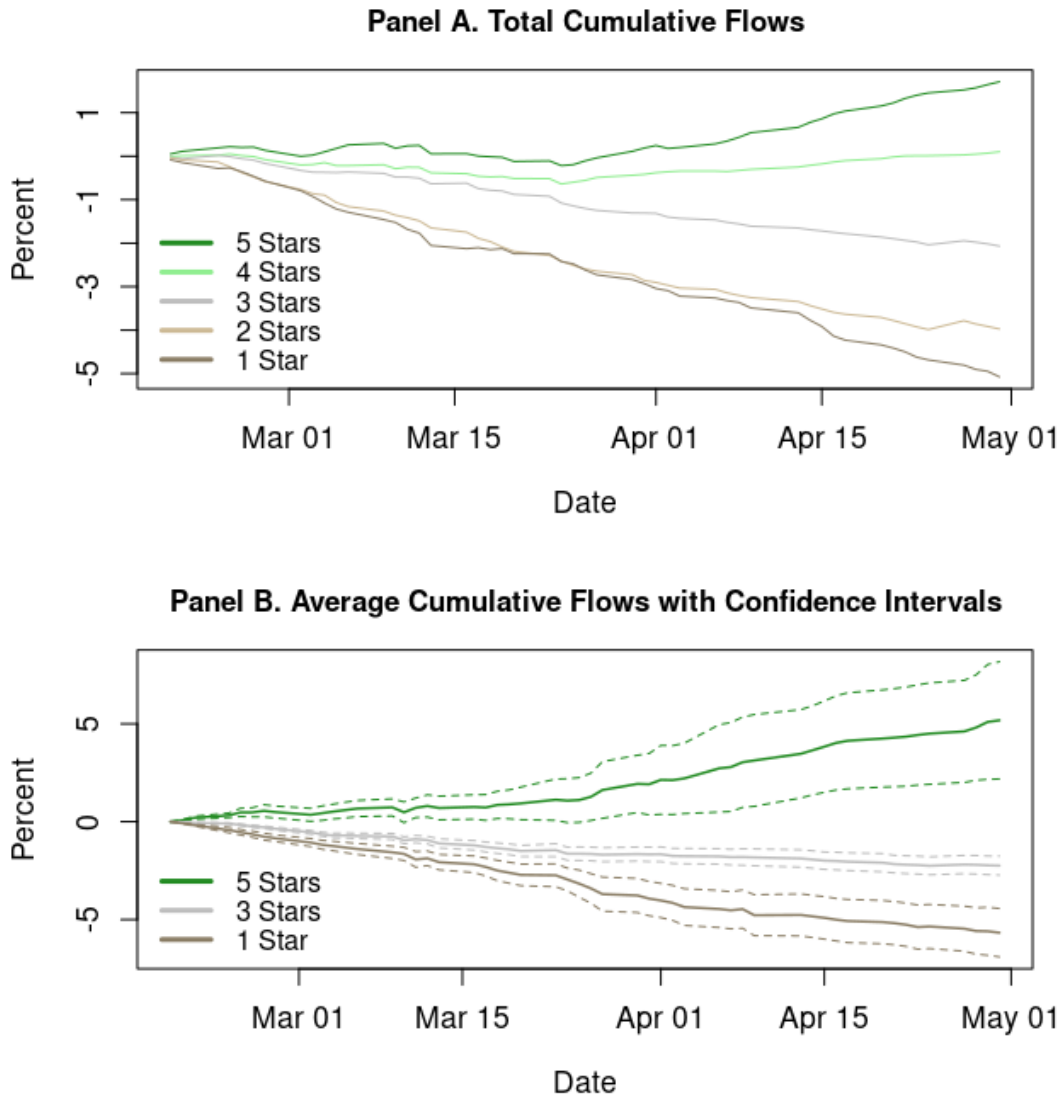

Figure 13. Fund Flows and Star Ratings. This figure plots net fund flows over the February 20 to April 30, 2020 period for categories of funds sorted by Morningstar star ratings. Panel A plots aggregate cumulative net flows for each of the five star categories. Flows are aggregated within each category and accumulated over time, then scaled by the category's total net assets on February 19, 2020. Panel B plots the average across funds of cumulative net flows as a percent of the fund's February 19, 2020 total net assets, for the five-, three-, and one-star categories only. Unlike in Panel A, the sample in Panel B is restricted to funds with at least $15 million of total net assets as of January 31, 2020 and the net fund flow percentage is winsorized at the 2.5% and 97.5% levels. Panel B also plots 95% confidence intervals, with standard errors clustered on the Morningstar Institutional Category.

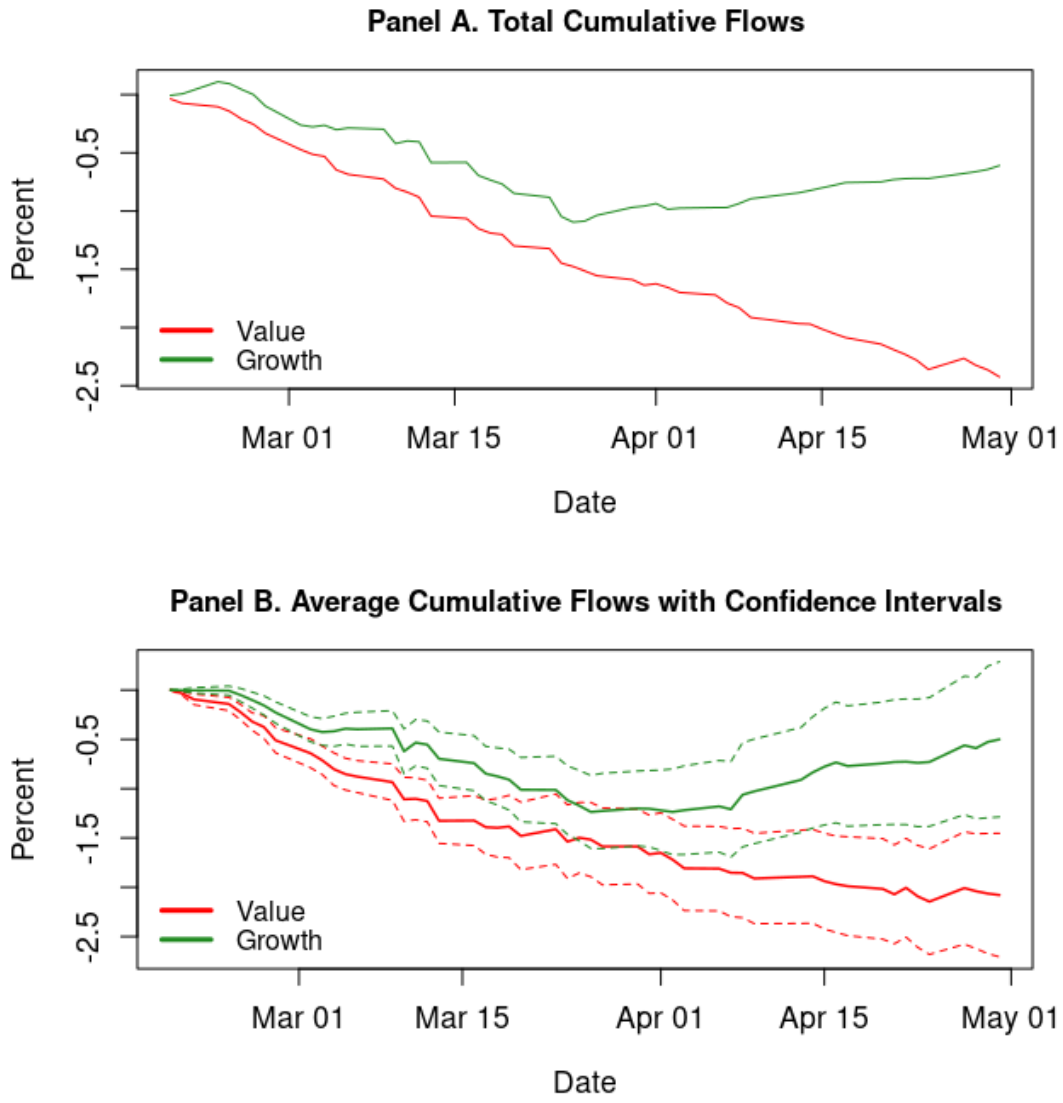

Figure 14. Fund Flows and Growth vs. Value Funds. This figure plots net fund flows over the February 20 to April 30, 2020 period for growth vs value funds, as determined by the Morningstar equity style box. Panel A plots aggregate cumulative net flows for both categories. Flows are aggregated within each category and accumulated over time, then scaled by the category's total net assets on February 19, 2020. Panel B plots the average across funds of cumulative net flows as a percent of the fund's February 19, 2020 total net assets, for both categories. Unlike in Panel A, the sample in Panel B is restricted to funds with at least $15 million of total net assets as of January 31, 2020 and the net fund flow percentage is winsorized at the 2.5% and 97.5% levels. Panel B also plots 95% confidence intervals, with standard errors clustered on the Morningstar Institutional Category.

REFERENCES

Albuquerque, Rui A., Yrjo Koskinen, Shuai Yang, and Chendi Zhang, 2020, Love in the time of COVID-19: The resiliency of environmental and social stocks, Working paper, Boston College.

Alfaro, Laura, Anusha Chari, Andrew N. Greenland, and Peter K. Schott, 2020, Aggregate and firm-level stock returns during pandemics, in real time, Working paper, Harvard University.

Barber, Brad M., Xing Huang, and Terrance Odean, 2016, Which factors matter to investors? Evidence from mutual fund flows, *The Review of Financial Studies* 29, 2600–2642.

Baumol, William J. and Wallace E. Oates, 1979, *Economics, Environmental Policy, and Quality of Life*, Inglewood Cliffs, NJ: Prentice-Hall.

Bialkowski, Jedrzej, and Laura T. Starks, 2016, SRI funds: Investor demand, exogenous shocks and ESG profiles, Working paper.

Bretscher, Lorenzo, Alex Hsu, and Andrea Tamoni, 2020, The supply channel of uncertainty shocks and the cross-section of returns: Evidence from the COVID-19 crisis, Working paper, LBS.

Carhart, Mark M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57–82.

Ding, Wenzhi, Ross Levine, Chen Lin, and Wensi Xie, 2020, Corporate immunity to the COVID-19 pandemic, Working paper, University of Hong Kong.

Elton, Edwin J., Martin Jay Gruber, Sanjiv Das, and Matthew Hlavka, 1993, Efficiency with costly information: A reinterpretation of evidence from managed portfolios, *Review of Financial Studies* 6, 1–22.

Elton, Edwin J., Martin Jay Gruber, and Christopher R. Blake, 2001, A first look at the accuracy of the CRSP mutual fund database and a comparison of the CRSP and Morningstar mutual fund databases, *Journal of Finance* 56.

Evans, Richard, 2010, Mutual fund incubation, *Journal of Finance* 65, 1581–1611.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F., and Kenneth R. French, 2010, Luck versus skill in the cross section of mutual fund returns, *Journal of Finance* 65, 1915–1947.

Fama, Eugene F., and Kenneth R. French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1–22.

Fahlenbrach, Rudiger, Kevin Rageth, and René M. Stulz, 2020, How valuable is financial flexibility when revenue stops? Evidence from the Covid-19 crisis, Working paper, Swiss Finance Institute.

Ferson, Wayne E., and Rudi W. Schadt, 1996, Measuring fund strategy and performance in changing economic conditions, *Journal of Finance* 51, 425–461.

Gerding, Felix, Thorsten Martin, and Florian Nagler, 2020, The value of fiscal capacity in the face of a rare disaster, Working paper, Bocconi University.

Glode, Vincent, 2011, Why mutual funds "underperform", *Journal of Financial Economics* 99, 546–559.

Gormsen, Niels Joachim, and Ralph SJ Koijen, 2020, Coronavirus: Impact on stock prices and growth expectations, Working paper, University of Chicago.

Gruber, Martin J., 1996, Another puzzle: The growth in actively managed mutual funds, *Journal of Finance* 51, 783–810.

Haddad, Valentin, Alan Moreira, and Tyler Muir, 2020, When selling becomes viral: Disruptions in debt markets in the COVID-19 crisis and the Fed's response, Working paper, UCLA.

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

Jensen, Michael C., 1968, The performance of mutual funds in the period 1945–1964, *Journal of Finance* 23, 389–416.

Kacperczyk, Marcin, Stijn Van Nieuwerburgh, and Laura Veldkamp, 2014, Time-varying fund manager skill, *Journal of Finance* 69, 1455–1484.

Kacperczyk, Marcin, Stijn Van Nieuwerburgh, and Laura Veldkamp, 2016, A rational theory of mutual funds' attention allocation, *Econometrica* 84, 571–626.

Kargar, Mahyar, Benjamin T. Lester, David Lindsay, Shuo Liu, and Pierre-Oliver Weill, 2020, Corporate bond liquidity during the COVID-19 crisis, Working paper, University of Illinois.

Kosowski, Robert, 2011, Do mutual funds perform when it matters most to investors? US mutual fund performance and risk in recessions and expansions, *Quarterly Journal of Finance 1*, 607–664.

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

Malkiel, Burton G., 1995, Returns from investing in equity mutual funds 1971 to 1991, *Journal of Finance* 50, 549–572.

Moskowitz, Tobias J., 2000, Mutual fund performance: An empirical decomposition into stock-picking talent, style, transactions costs, and expenses: Discussion, *Journal of Finance* 55, 1695–1703.

Nofsinger, John, and Abhishek Varma, 2014, Socially responsible funds and market crises, *Journal of Banking and Finance* 48, 180–193.

O'Hara, Maureen, and Xing Alex Zhou, 2020, Anatomy of a liquidity crisis: Corporate bonds in the COVID-19 crisis, Working paper, Cornell University.

Pagano, Marco, Christian Wagner, and Josef Zechner, 2020, Disaster resilience and asset prices, Working paper.

Pástor, Ľuboš, and Robert F. Stambaugh, 2002, Mutual fund performance and seemingly unrelated assets, *Journal of Financial Economics* 63, 315–349.

Pástor, Ľuboš, and Robert F. Stambaugh, 2012, On the size of the active management industry, *Journal of Political Economy* 120, 740–781.

Pástor, Ľuboš, Robert F. Stambaugh, and Lucian A. Taylor, 2015, Scale and skill in active management, *Journal of Financial Economics* 116, 23–45.

Pástor, Ľuboš, Robert F. Stambaugh, and Lucian A. Taylor, 2017, Do funds make more when they trade more?, *Journal of Finance* 72, 1483–1528.

Pástor, Ľuboš, Robert F. Stambaugh, and Lucian A. Taylor, 2020, Sustainable investing in equilibrium, *Journal of Financial Economics*, forthcoming.

Ramelli, Stefano, and Alexander F. Wagner, 2020, Feverish stock price reactions to COVID-19, Working paper, University of Zurich.

Schrimpf, Andreas, Hyun Song Shin, and Vladyslav Sushko, 2020, Leverage and margin spirals in fixed income markets during the COVID-19 crisis, Working paper, BIS.

Sensoy, Berk, 2009, Performance evaluation and self-designated benchmark indexes in the mutual fund industry, *Journal of Financial Economics* 92, 25–39.

Wermers, Russ, 2000, Mutual fund performance: An empirical decomposition into stock-picking talent, style, transactions costs, and expenses, *Journal of Finance* 55, 1655–1695.