Norwegian School of Economics

Bergen, Fall, 2020

Gender Differences and the Effect of Remote Working During the COVID-19 Pandemic

An Empirical Analysis of Remote Working and Its Impact on Performance in the American Mutual Fund Industry

Espen Claussen & Sindre Totland

Supervisor: Nataliya Gerasimova

Master thesis, Economics and Business Administrastion,

Financial Economics

NORWEGIAN SCHOOL OF ECONOMICS

Abstract

During the COVID-19 pandemic, forced remote working has been an important tool for authorities all over the world attempting to stop the spread of the virus. Looking at gender, early reports indicate a disproportionate impact of the remote working effect on performance, disfavoring women. Our thesis seek to investigate this remote working effect on gender performance for active mutual fund managers, hoping to see if it can help explain the industry's long-lasting disparity in gender representation. We use data on returns for 2695 actively managed U.S.-domiciled funds during the COVID-19 crisis to investigate the effect of remote working on fund performance. We identify each state's lockdown period, enabling us to use the staggered state-level adoption of the stay-at-home orders to conduct difference-in-differences analyses on both fund profitability and managerial skill. First, we investigate the role of gender on fund profitability by looking at both raw returns and factor-adjusted returns, hypothesizing that: (1) The effect of working from home during the COVID-19 crisis deteriorates factor-adjusted returns more for active mutual funds managed by women, compared to active mutual funds managed my men. Our results show no significant difference in the effect of working from home on mutual fund profitability between the genders. Secondly, we investigate the role of gender on managerial skill to identify how this is effected by the remote working effect, hypothesizing that: (2) The effect of working from home during the COVID-19 crisis deteriorates managerial skills more for female mutual fund managers, compared to male mutual fund managers. Our results show a significant difference between the genders, but not as anticipated. Contrary to our hypothesis, we find a relatively worse effect on managerial skill from working from home during the COVID-19 crisis for funds exclusively managed by men, compared to funds with at least one woman in the manager group. Given this, we do not find results justifying the current disparity in gender representation within the mutual fund industry, looking through the lens of performance.

Preface

This Master thesis was written as part of the Master's program in Financial Economics at the Norwegian School of Economics (NHH).

The paper uses difference-in-differences and triple difference analyses to examine the effect of remote working on different measures of active mutual fund performance, and how the effect varies with the gender of the fund manager. The choice of topic emerged due to our mutual interest in financial markets, combined with our curiosity about the COVID-19 pandemic and its consequences.

This process has been time-consuming and thought-provoking, particularly when deciding upon the regression strategy used in this thesis. Ironically, we have experienced the effect of remote working ourselves. With periods of lockdown in Bergen this fall, our communication with each other have been periodically limited to digital communication. Furthermore, the communication with our supervisor have been strictly online, which is sub-optimal when writing a Master Thesis.

We would like to gratefully thank our supervisor Nataliya Gerasimova for providing us with valuable counseling and crucial comments during this academic work. At last, we would like to thank the Norwegian School of Economics and all the wonderful people we have had the pleasure to meet and get to know over the course of our studies here in Bergen.

Espen Claussen **Sindre Totland**

Norwegian School of Economics,

Bergen, 20. December 2020

Contents

1. Introduction

Through offering more accessible risk diversification, economies of scale and professional management, the mutual fund industry has grown rapidly since its introduction. At the end of 2019 total worldwide assets invested was estimated to $54.9 trillion, and in the United States the actively managed mutual and exchange traded funds accounted for 61 percent ($13.4 trillion) of the total net assets in the fund market (Investment Company Institute, 2020, p. 2, 39). The industry has seen a rapidly growing number of mutual funds as a result of increased popularity, but are still characterized by a continuous disparity when looking at the gender of the fund managers. In a recent study Morningstar found that in 2000, women made up 13.4 percent of active fund managers of U.S.-domiciled funds, with the number dropping to 10.7 percent in 2019 (Morningstar, 2020).

Earlier research have been striving to determine whether gender differences in performance could explain this disparity, but results point to the absence of significant differences between male and female managed funds (Babalos, Caporale & Philippas, 2015; Atkinson, Baird & Frye, 2003; Niessen-Ruenzi & Ruenzi, 2019). Our thesis aims to extend this area of financial literature by investigating the role of gender in performance for the actively managed mutual fund industry, during a time of crisis. More specifically: the COVID-19 pandemic.

There are two reasons for studying fund performance by gender during this pandemic. The main reason being, with the rise of COVID-19, we have seen state- and nationwide lockdowns implemented by authorities all over the world as a measure to cope with the spread of the virus. This has forced people to start working from home (also called remote working), where Xiao (2020) found that the effect of working from home during the COVID-19 crisis led to both lower excess net returns and lower managerial skill in the mutual fund industry. Additionally, studies have also found results indicating a disproportionate impact of remote working on the genders' performance, disadvantageous to women (McKinsey & Lean In, 2020; Rogers, 2020).

Secondly, actively managed mutual funds are known for, on average, underperforming passive benchmarks net of fees (Fama & French, 2010), but still the industry is managing trillions of dollars worldwide. Tobias J. Moskowitz (2000) suggests that it may be that mutual funds add or subtract value when we care about performance the most. With skyrocketing unemployment levels, the COVID-19 pandemic certainly fits the description of a situation where performance matters more than usual. Being able to understand this newfound effect of forced remote working and if it affects the genders differently during a time of crisis, could provide new evidence to the discussion of the gender disparity we are currently witnessing in the industry.

This thesis utilizes various difference-in-differences analyses to investigate the topic of managerial performance for almost 2700 U.S.-domiciled funds during the lockdown period of the COVID-19 pandemic. Looking at different performance measures and different models, we thoroughly examine managerial skill in the mutual fund industry, seeking to understand how the genders are impacted by being forced to work from home. To identify the lockdown period we use the states' stay-at-home order issuance date, which carry the force of law.

Our first hypothesis addresses the profitability of the mutual funds, where we hypothesize the effect of remote working to decrease the return of female managed funds more, compared to the funds managed by men. Our results show similarities to previous literature, where we find insignificant differences when comparing the profitability of the funds for the genders (Atkinson et al., 2003; Babalos, Caporale & Philippas, 2015). In our second hypothesis, we hypothesize that the effect of remote working during the lockdown period will decrease female managerial skill more, compared to male managerial skill. Our results actually seem to indicate a significant difference between the genders, but not as hypothesized. Where the changes in managerial skill for the female managers seem to be insignificant, we see a significant decrease in managerial skill for the male managers. Interestingly, this contradicts the proposed disproportionate effects of newer reports (Rogers, 2020; McKinsey & Company, 2020), but shows similarities with the findings of Niessen and Ruenzi (2019), who found significant persistence in favor of women in a study done before COVID-19.

We acknowledge that this forced work from home effect may not explain the disparity we have seen up until today, but it should be an important aspect going forward. By limiting the period of interest and incorporating the effect of remote working into our analysis we are looking at mutual fund performance from a different angle, hoping to provide new and valuable insight to the question of gender disparity in the industry.

At last, we feel the need to explain our intentions, as we recognize studying performance by gender can be provocative. Researching the literature on this area we see a relative scarcity of papers examining gender disparities in finance, which also has been addressed by the Society for Financial Studies (2020). It is not in our interest to assess gender as an explanatory variable for fund performance, but to investigate whether or not the long-lasting disparity in the mutual fund industry could be justified when looking at performance. With this, we are hoping to help paving the way for more similar studies and essential discussions on this very important topic.

2. Literature Review

In this review we are aiming to summarize and provide a critical analysis of the research arguments for and against active management investing in the mutual fund industry, and if different manager and fund characteristics can predict performance. First, we are looking at research on the question of skill versus luck as an explanation for abnormal returns, before continuing with the discussion on whether differences in performance could be attributed to the gender of the fund manager or not.

Previous research on the topic of active fund managers' skill is mainly focused on whether or not actively managed mutual funds are able to return abnormal risk-adjusted returns with respect to their benchmark. Introducing the efficient market hypothesis, Fama and French (1970) suggest that any portfolio manager's excess returns derive from luck or random chance rather than skill. This is supported by studies showing that actively managed mutual funds, on average, fail to outperform their benchmarks (Fama, 2010) and are lacking persistence in performance (Carhart, 1997). Seeing this, it seems puzzling that active mutual fund managers are amongst the highest paid members of society. Against the conclusion of luck, there has been found evidence that supports the existence of managerial skill for actively managed funds (Grinblatt & Titman, 1989, 1993; Chevalier & Ellison, 1999; Berk & Binsbergen, 2015).

The methodology used to evaluate fund performance has gradually changed throughout time. Starting with the single factor evaluation model introduced by Jensen (1968), more advanced multi-factor models are now established as the conventional performance measures in the literature (Fama & French, 1993; Carhart, 1997). Aside from traditional performance measures, we have seen the development of conditional performance evaluation models, under the assumption that fund managers could change their investment strategy when the economic conditions are changing (Ferson & Schadt, 1996; Jha, Korkie & Turtle, 2009).

Though more thorough and advanced performance measures have been developed, one of the main questions the literature still seek to answer is if the abnormal returns are to be explained by managerial skill or pure luck. A common approach to this problem is to test for persistence in fund returns, that is, whether past winners continue to produce high returns and losers continue to underperform (Fama & French, 2010). Additionally, Amihud and Goyenko (2008) propose that managerial skill can be predicted by its R^2, obtained from a regression of a fund's returns on a multi-factor benchmark model.

Literature debating the question of managerial skill versus pure luck is divided, and some of the conflict lies in the way it is measured. Most prior studies use the net alpha to investors, the average abnormal return net of fees and expenses, as the measure of managerial skill and then look for persistence over time. This is the case for Carhart (1997) where he argues that short term persistence in equity mutual fund returns can be explained with common factors in stock returns and investment costs. Net alpha is also used to arrive at the same conclusion by Fama and French (2010), mentioned above. Given this evidence on non-persistent performance, abnormal returns is often acknowledged as a sign of luck rather than skill.

On the other hand, it is argued that investments with active managers do not outperform passive benchmarks as a consequence of the competitiveness of the market for capital investment. If investors compete with each other for superior returns, they end up ensuring that none exist. (Berk & Green, 2004). Based on this, Berk and Binsbergen (2015) argues that if skill is in short supply, the net return is determined in equilibrium by competition between investors, and not by the skill of managers. Therefore, they debate gross alpha as the correct measure for skill and continue using value added, defined as return before fees, minus the benchmark return multiplied by assets under management. Doing this they account for both the amount of money the manager takes home (his fee multiplied by the assets under management) plus the amount he creates or destroys for his investors (the overall dollar under- or over-performance relative to the benchmark) using the same mutual fund data compiled by Carhart (1997). The results of their study show an average value added per manager of about $2 million a year, and persistent skill.

With the growing popularity of the mutual fund industry, more studies have been conducted in an attempt to identify other aspects that can predict performance. We have also seen studies on managerial and fund characteristics. Chevalier and Ellison (1999) suggests that if ability exists, it is not obvious whether it resides in the manager or in the fund organization. However, by regressing annual excess returns above the value weighted NYSE/AMEX/NASDAQ composite index on different manager characteristics they find statistically significant relationships between excess returns and both age and SAT-scores. With the use of a generalized multifactor model, Prather, Bertin and Henker (2004) find fund characteristics such as market capitalization, expense ratio and the number of funds under management for a single team to be predictive of fund performance.

Some studies have also been conducted on another characteristic, manager style. By using the Fama-French 3-factor model, Davis (2001) group funds by ranking all funds after their factor weight. Davis' results are inconclusive, himself stating that "*Although the evidence of abnormal performance is slim, it is more than we would expect to see if the null hypothesis of no abnormal performance were absolutely true*".

The amount of research on different fund- and manager characteristics as a prediction for skill is steadily growing. However, the research on gender differences and its effect on performance is has been relatively scarce. Studies on gender and investment behavior shows reasons to believe that there might be differences. Halko, Alanko, and Kaustia (2012) found that women are more risk averse, while Barber and Odean (1998) found men to be more overconfident, leading to men trading more than women and thereby reduces returns more so than women.

Despite this, as mentioned in the introduction, results from the area of research points to the absence of significant differences in returns between male and female managed funds. Atkinson et al. (2003) examined the excess return before fees of 1294 fixed-income mutual funds concluding no significant differences in either performance nor risk for the female managed funds compared with the male managed funds. This is supported by a study on 747 actively managed equity funds from January 1992 to December 2009 done by Niessen and Ruenzi (2019). They utilize a combination of single-factor and multifactor models to obtain alphas net of fees and divide the observations into one portfolio with female managed funds and one with male managed funds. Testing for both factor alphas and raw returns they find no statistically significant performance difference between male and female managed funds. Interestingly, they find statistically significant differences in persistence in favor of women managers. The findings show more stable performance as well as more stable investment styles of female managers. Given the findings of less overconfidence and a higher risk aversion for women in the studies mentioned in the above paragraph, this is perhaps not so surprising.

Babalos, Caporale and Philippas (2015) compare the performance of 358 European male and female managed equity funds, excluding funds that are team managed. Using a combination of single- and multi-factor models augmented with a fixed-income securities index to account for funds' non-stock holding, the authors investigate returns not adjusted for sales charges in 14 different investment categories. The results show statistically significant alphas for both genders in one of the investment categories, Eurozone Large-Cap. Overall, female managers appear to be slightly superior to their male counterparts in terms of their alphas, but the difference is not significant in this study.

Our thesis contributes to the financial literature in several areas. First, we extend the scope of research by incorporating the effect of working from home as a possible explanation of the gender disparity in the mutual fund industry. Making use of several factor models and measures of skill we are able to present a thorough performance study using the standard methodology. The United States stay-at-home orders during the COVID-19 pandemic carry the force of law, which reduces measuring errors as a result of not knowing whether or not the managers work remotely. Secondly, the on average, long lasting underperformance of mutual funds has been justified by the belief that active mutual funds are adding value in periods when we care about performance the most (Moskowitz, 2000). Doing a gender performance study of the COVID-19 lockdown period would allow us to see if this is the case for both, non or one of the genders. Lastly, in our sample period the pandemic has created an environment with high volatility and unusually large price dislocations in the financial markets. In such an environment, assuming no other measurement errors, it might be easier to separate the skilled investors from the lucky ones, indicating a possibility of clearer results no matter the direction of our conclusion.

3. Hypothesis Development and Theoretical Background

In the following we will address important concepts that form the basis of our hypotheses. First, we will give some context on the effects of remote working and gender inequality, before looking at the theoretical background for evaluating mutual fund manager performance and how the effect of forced remote work impacts the genders.

3.1 Remote working and gender inequality

Remote work is a working style that allows employees to work outside of a traditional office environment, based on the concept that work does not need to be done in a specific place to be executed successfully (Remote Year, 2020). The concept has gained traction during this year's pandemic, with huge companies such as Microsoft and Google now planning to offer remote working indefinitely (Hadden et al., 2020).

Looking at the effects of remote working, we are witnessing a great conflict between early academic results and newer reports. Studies of remote working before the pandemic shows results of induced performance, productivity and job satisfaction (Bloom et. al., 2015, Golden & Gahendran, 2019). However, throughout this pandemic it seems like the perception of the concept has changed. In an article from March this year, the same Nicholas Bloom who presented results of induced performance as a result of remote working, states that "Working from home with your children is a productivity disaster". He goes on to explain this changed perception with four significant factors: children, space, privacy and choice (Gorlick, 2020).

As more research on forced remote working is being conducted, there seems to be evidence that support a disproportionate impact on the genders. During the pandemic we have seen an expansion of domestic activities such as housework and caregiving. Because of our society's pre-existing views on traditional gender roles and gender responsibilities, remote working can potentially traditionalize gender roles, leaving men with less family time and women with more unpaid work (Lott, 2014).

This is supported by the report from McKinsey and Lean In (2020) mentioned in the introduction, which found evidence that housework and caregiving burdens are more likely to push women out of the workforce, with mothers being more likely than fathers to worry that their performance is being negatively judged due to their caregiving responsibilities. A survey by Qualitrics and theBoardlist also found that men are 2,3 times more likely than women to say that working from home for an extended period of time would positively affect their career progression, and men are also nearly twice as likely to say that the amount they are able to work from home during the pandemic has positively affected their career (Rogers, 2020).

In addition to being a social problem, this could also have massive economic consequences. Looking at another report by McKinsey (2020), they studied the unemployment trends of the COVID-19 crisis and found that women made up 39 percent of global employment, but accounted for 54 percent of overall job losses. They portray three different scenarios of actions to counter this higher negative impact on women to see how it would influence the global GDP in 2030. Their worst case scenario assumes that the disproportionate impact on women remains unaddressed and their best case scenario implies policy makers to take decisions immediately and further on, that would significantly improve gender equality over the next decade. Comparing the two, they find the best scenario to give a global GDP in 2030 of $14 trillion higher than for the worst case scenario.

This shows the continued importance of addressing gender inequality during and after the COVID-19 crisis. The stakes are high and could have detrimental consequences if ignored. We hope to contribute to the topic by investigating the subject in the mutual fund industry.

3.2 Active Portfolio Management and Fund Performance

Active portfolio management is the attempt to achieve portfolio returns more than commensurate with risk, either by forecasting broad market trends or by identifying mispriced sectors of a market or particular securities (Bodie, Kane & Marcus, 2011). Conversely, by the definition of the Efficient Market Hypothesis, Fama (1970) proposes that it should not be possible for active fund managers to earn risk-adjusted abnormal returns trading on publicly available information. For mutual fund performance, Fama and French (2010) found that mutual fund investors in aggregate realize net returns that underperform their factor-adjusted benchmark, but that there is evidence of managerial skill, negative as well as positive.

In addition, the efficient market hypothesis assumes perfectly efficient markets. There are many studies that find evidence of market inefficiency (Dharan & Ikenberry, 1995; Desai & Jain, 1997; Frazzini, 2006). Even looking at the COVID-19 crisis it has been shown a strong loss of efficiency for the S&P index (Ammy-Driss & Garcin, 2020). The literature suggests that it should be possible for a skilled fund manager to earn risk-adjusted abnormal returns during the pandemic. Combined with the results that implies women are taking the bigger burden of increased domestic activities, we expect the remote working effect during the pandemic, on average, to induce worse performance for funds managed by women compared to men. Thus, we propose our first hypothesis:

Hypothesis 1: The effect of working from home during the COVID-19 crisis deteriorates factor-adjusted returns more for active mutual funds managed by women, compared to active mutual funds managed my men.

3.3 Manager Performance – Selective skills

While the first hypothesis investigates the remote working effect on the profitability of the mutual funds, our second hypothesis addresses the skill of the managers more directly. As mentioned in the literature review, earlier research are divided on the topic of managerial skill. Abnormal returns could be due to pure luck, so fund profitability need not imply the presence of managerial skill.

A common approach to test for skill is to measure persistence in fund returns (Fama & French, 2010). However, we believe looking at persistence in returns over a period of 3 months to separate skill from luck would be contradictory. Instead, Xiao (2020) uses an alternative measure for managerial skill; the Amihud-Goyenko 1–R^2 measure, which he defines as a measure on selective skill (Amihud & Goyenko, 2013). With this, he finds that managerial skill decreases when mutual fund managers work from home, suggesting that we should expect reduced managerial skill from both genders. Nonetheless, again considering the disproportionate impact of the remote working effect on the genders, we believe we will see an on average, bigger decrease in managerial skill for female managers compared to male managers. With this, we present our second hypothesis:

Hypothesis 2: The effect of working from home during the COVID-19 crisis deteriorates managerial skills more for female mutual fund managers, compared to male mutual fund managers.

4. Data

In this part we will describe the data that will be used to answer the two hypotheses in the thesis. First, we will present the fund data, followed by gender data, factor models and the stay-at-home-orders.

4.1 Fund Data

The fund data is retrieved from Morningstar's fund database, using their platform software, Morningstar Direct (Morningstar Direct, 2020). We collected data from the period February 1, 2020, to April 30, 2020. The data includes daily net returns, monthly total net assets under management (TNA), fund characteristics, and fund information including name of managers and home state.

Morningstar's database contains thousands of funds on a global basis. In order to sample our data to be relevant for the analysis, we used Morningstar's screening function. Our fund screening is mainly based on Xiao's (2020) fund environment. First, we allow only open-end mutual funds that are domiciled in the U.S. as we need the managers to work in the U.S. Secondly, we focus on equity funds, excluding index funds, exchange traded funds (ETFs) and non-equity funds. To strengthen our sample, we include funds investing in both equity, sector equity and international equity. Further, we only include funds with non-missing returns in the period. According to Evans (2010), funds within the incubation period outperforms non-incubated funds. After the incubation period, the effects reverses and disappears. Mutual funds with a tenure of less than 18 months before the start of our sample period are therefore excluded. Additionally, we exclude all funds with TNA of less than $15 million, as they tend to be biased upwards, according to Elton, Gruber, and Blake (2001).

After applying our screening criteria, we are left with 9513 mutual funds. However, the fund sample are now divided by share classes, and many of the funds are therefore based on the same pool of assets, where the fees and target investors are the only thing that differs. In order to aggregate the share classes of each fund into single funds, we have followed Pástor, Stambaugh, and Taylor's (2015) approach, by using the FundID variable. FundID is the same for all the share classes that belongs to a specific fund. By using the variable as a common key, we were able to use the sum of lagged monthly TNA for each share class to value weight returns and expense ratio from each share class and aggregate them to the fund level. This reduces our sample to 2867 mutual funds.

The daily net mutual fund returns are net of fees and are calculated by Morningstar, explained as: "The total returns do account for management, administrative, 12b-1 fees and other costs taken out of fund assets" (Morningstar Direct, 2020). Hence, the analysis shows how much an investor would have gotten in return if he were to invest in the funds.

For our analysis we needed to find the U.S. home state for each of the funds. Morningstar had location data available for most of the funds. Due to the lack of location information for some of the funds in the database, it required us to manually look up the information. Most of this was done through the funds' websites and the managers' LinkedIn profiles. 36 funds were dropped due to uncertain location data.

4.2 Gender data

The number of managers of each fund varies from 1 to 35, according to the data retrieved from Morningstar (Morningstar Direct, 2020). In order to look at the manager group's gender composition for each fund, we had to identify the gender of all the managers. For this, we used an algorithm provided in the *gender* package in R. The package is described as the following:

> *Infers state-recorded gender categories from first names and dates of birth using historical datasets. By using these datasets instead of lists of male and female names, this package is able to more accurately infer the gender of a name, and it is able to report the probability that a name was male or female. (Mullen et.al., 2020)*

We ran all the managers' first names against the "*ssa*" database, which consist of baby name data from the U.S. Social Security Administration. In addition, we delimited the range of birthyears to go from 1930 to 2000. This returned a data frame of 939 names and gender to each of the names, based on the probability of male or female. A considerable amount of the manager names did not have any match in the database and required us to manually identify the managers' gender. This was completed by online research, including LinkedIn-profiles, professional photos and funds' websites. Additionally, we did as Sargis and Lutton (2016), and used titles and pronouns such as *Mr., Mrs., he, she*, etc. from bio descriptions and other articles from reliable sources to determine manager gender. Funds where we could not certainly determine the manager's gender were dropped. Lastly, we put them into two different manager team categories: 1) funds managed exclusively by men and 2) funds with at least one female manager. The latter category includes both funds exclusively managed by women and funds with a mixed gender manager group. For the rest of the thesis this will be referred to as the *female group*.

After the screening process, we are left with a final sample consisting of 2695 mutual funds. Table 1 shows descriptive statistics of the fund sample. The average fund in our sample is 16.3 years old, has USD 24.8 billions in TNA, longest sitting manager tenure of 10,6 years and has an annual expense ratio of 0.98 percent. The average daily return over the sample period was -0.17 percent. The maximum and minimum daily return observed in the period was respectively 30.75 percent and -34.11 percent, which indicated that the sample data probably do not have any salient errors. Further, funds managed by men dominate the sample, with a 74 percent part of the total funds. The other 26 percent are managed by at least one woman. Of the 2695 funds in the sample, only 72 funds are managed by one woman or a team consisting purely of women.

Table 1: Descriptive statistics for the fund groups

Variable	Group	Mean	St. Dev	Min	Median	Max	Obs	T-stat
Daily return %	Male	-0.17	3.85	-34.11	0.06	30.75	123 008	-0.19
	Female	-0.16	3.75	-24.39	0.05	15.41	44 082	
	Total	-0.17	3.83	-34.11	0.06	30.75	167 090	
Fund Size (Mill)	Male	2030	5660	15,1	434	121000	1 984	-34.90
	Female	3770	14700	18,1	624	201000	711	
	Total	2490	9010	15,1	482	201000	2 695	
Tenure longest	Male	10.67	6.89	0.67	9.00	52.83	1 984	7.28
	Female	10.39	6.86	0.67	8.83	48.08	711	
	Total	10.59	6.88	0.67	8.92	52.83	2 695	
Age of fund	Male	16.34	11.80	1.51	14.39	92.04	1 984	5.28
	Female	16.00	10.72	1.55	14.34	89.82	711	
	Total	16.25	11.52	1.51	14.39	92.04	2 695	
Annual expense ratio	Male	1.00	0.37	0.00	1.00	5.04	1 984	44.86
	Female	0.91	0.34	0.00	0.94	2.30	711	
	Total	0.98	0.36	0.00	0.98	5.04	2 695	
Number of managers	Male	2.53	1.89	1.00	2.00	22.00	1 984	-0.02
	Female	5.07	4.75	1.00	3.00	35.00	711	
	Total	3.20	3.13	1.00	2.00	35.00	2 695	
Team	Male	0.74	0.44	0.00	1.00	1.00	1 984	-83.42
	Female	0.92	0.27	0.00	1.00	1.00	711	
	Total	0.79	0.41	0.00	1.00	1.00	2 695	

Note: The table presents descriptive statistics for the male and the female/mixed fund groups, as well as the total fund sample. Daily returns are percentage returns for the 62 market days in the sample. Fund Size are presented in $millions. Tenure is the longest sitting manager's tenure. Age of fund is years since inception date. Annual expense ratio is the funds' annual net expenses. Number of managers displays the number of managers in the manager groups. Team is an indicator that is 1 if the manager group is a team, and zero if the fund is managed by a solo manager. Hence, the mean represents the percentage of funds in each group that are managed by a team. T-stats are from t-tests for the means from the male and the female group.

To examine the empirical questions, we have used three factor models: the Fama French 3- and 5-factor models and Carhart 4-factor model. We retrieved daily return data for the 3 factor models from K. French's website (French, K. R., 2020a-b), which includes the market excess return, the risk-free rate, the Fama-French factors (SMB, HML, RMW, CMA), and the momentum factor (MOM). The excess return is based on the CRSP value-weighted market index, and by adding the daily risk free rate to that variable, we define the market returns that will be used in this thesis.

4.3 Stay-at-home-orders

To find the correct date and time for when each state's stay-at-home order went in to effect we used The New York Times' COVID-19 Restrictions State-level tracker website (Mervosh, Lu & Swales, 2020). For most states this tracker presented a link to the actual order, and for the states without a link we found the orders through the states' federal government homepage or through the governors' official social media accounts. To make sure our analysis capture the correct effects we had to modify the effective dates in some cases. One example is Pennsylvania, where their order went into effect April the 1st at 8 p.m. Then, the first day with remote work was April the 2nd, which is our modified effective date. Both the original and the modified dates can be seen in table 11 in the appendix. The first stay-at-home order were released by California on March the 19th and the last by South Carolina on April the 7th, while Arkansas, Iowa, Nebraska, North Dakota, Oklahoma, South Dakota, Utah and Wyoming had none or limited restrictions.

5. Methodology

In this section we will present the methods and models applied to examine the empirical analysis in this thesis. First, we will present the outcome estimators used to examine the hypotheses. Second, we will present the factor models and define the regression models used as input into the outcome regressions. In the last part, we will discuss some assumptions that need to hold if the analysis is to have a causal interpretation.

5.1 Difference-in-differences

The framework used to examine the empirical questions from our two hypotheses is the difference-in-differences (DID) research design. The DID research design has become a widely used framework to study policy questions (Imbens & Wooldridge, 2007). The basic setup consists of two groups and two periods. It estimates the effects of a policy before and after, where one group gets treated by the policy – *the treatment group,* while the other group does not get treated – *the control group.* In our case, the stay-at-home orders are defined as the policy, and mutual funds located in a state that issues a stay-at-home order during the period is the defined as the treated group. This leaves the funds located in a state that does not issue a stay-at-home-order at any point during the period as the control group. The reason for including a control group, is to establish a counterfactual for how the funds would have developed if there were not any stay-at-home orders. Unlike the basic 2x2 DID model, the stay-at-home orders get implemented in states at different points in time. Hence, we cannot use the basic difference-in-differences design. In order to be able to include multiple treatment periods in the estimations, we can use a more general difference-in-differences method (Imbens & Wooldridge, 2007).

The generalized difference-in-difference estimator can include multiple treatment periods and multiple treatment groups. Similarly to Xiao (2020), we have used aggregated stay-at-home announcements, which means that funds located in states that announce stay-at-home orders at the same time, are treated as the same group. As mentioned, the treated group in the DID models consists of funds that are located in a state that has issued a stay-at-home order. We use staggered adoption for the treated states, hence if a state first issues a stay-at-home order and joins the treated group, it stays there for all remaining periods. This is does not entirely reflect the reality, because some of the states are reopening before our end date, April 30. This is the case for Alaska, Colorado, Georgia, Mississippi, Montana and South Carolina, with South Carolina being the first to reopen with the effective date of April 21. We do not know if the fund managers in the reopening states continue with remote work or if they are going back to office, but at most this accounts for 8 trading days and 46 funds, which is believed to be insignificant for the outcome of our analysis. Also, the generalized DID deviates from the standard DID design, as funds in the control group in the generalized DID consists of funds located in a state that has not yet issued a stay at home order and funds located in a state that never issued a stay-at-home order. Thus, if a fund gets treated at time *t,* it shifts over in the treatment group and stays there for the rest of the period. The regression is presented below:

$$Outcome_{i,s,t} = \gamma_s + \lambda_t + \beta Post_{s,t} + \theta X_i + \varepsilon_{ist} \tag{1}$$

where $Outcome_{i,s,t}$ is the dependent variable, and will be explained in the following sections. γ_s and λ_t are respectively state and time fixed effects. $Post_{s,t}$ is the policy indicator, and equals 1 if fund *i* is in a state *s* which has announced a stay-at-home-order at time *t,* and zero otherwise. X_i is the time invariant covariate vector for fund *i,* which includes the tenure of the fund's longest sitting manager, the age of the fund, the fund size, the annual expense ratio, and a dummy for team, which equals 1 if there are more than one manager of the fund, and 0 if there is only one manager. Since the coefficients of interest are varying from state to state as a factor of whether or when the state announced stay-at-home-orders, the standard errors are clustered at the state level.

Before running the difference-in-differences regressions, we divide the funds into the two subgroups: the male group and the female group. The two groups' regressions will return the effects of working from home as the coefficient β of the policy variable $Post_{s,t}$. However, when running the regression for the gender groups separately, the policy indicators do not explicitly tell us anything about how the two groups of funds are affected by the policy *relative* to each other. Following Olden and Møen (2020), the $\hat{\beta}$ to our policy variable can be defined as:

$$\hat{\beta} = \left[\left(\bar{Y}_{T,Post} - \bar{Y}_{T,Pre}\right) - \left(\bar{Y}_{C,Post} - \bar{Y}_{C,Pre}\right)\right] \tag{2}$$

which is the change in outcome (post-treatment minus pre-treatment) for the treatment group, minus the change in outcome for the control group. Because we are running a regression for both the male and the female group, we get one β for each group. This allows us to examine the relative outcomes of the two groups, by subtracting the outcome of the one group from the other group. Since we are interested in the effect of the female group relative to the male group, we get the following:

$$\hat{\beta}_{Triple} = \left[\left(\bar{Y}_{F,T,Post} - \bar{Y}_{F,T,Pre}\right) - \left(\bar{Y}_{F,C,Post} - \bar{Y}_{F,C,Pre}\right)\right] - \left[\left(\bar{Y}_{M,T,Post} - \bar{Y}_{M,T,Pre}\right) - \left(\bar{Y}_{M,C,Post} - \bar{Y}_{M,C,Pre}\right)\right] \quad (3)$$

where we get the average daily post-treatments effects for the treated group *T* over the non-treated group *C* in the female group *F,* relative to the daily post-treatment effects for the treated group over the control group in the male group *M*. $\hat{\beta}_{Triple}$ is then equivalent to the coefficient of interest in a *difference-in- differences-in-differences* (DDD or triple difference) estimation.

Again, examining treatment which starts at different points in time for different groups makes it more complicated. Strumpf (2011) applies a DDD estimator when examining the impact of Medicaid on the labor supply among single women with and without kids in the 1960s and 1970s. We have constructed a DDD estimator similar to the one used in her study. It has many similarities to the DID explained above, but the policy indicator from the DID now gets interacted with a gender variable, which takes on the value 1 if the fund is in the female group, and 0 if it belongs to the male group. Additionally, we add fixed effects for the interactions state*gender and time*gender. Together with *Post*, the fixed effects controls for time-invariant gender-specific characteristics within states, time-varying changes within the gender groups and time-varying changes within states. We estimate:

$$Outcome_{i,s,t} = \gamma_s + \lambda_t + \beta_1\left(Post_{s,t}\right) + \beta_2(Female_i) + \beta_3\left(Post_{s,t} * Female_i\right) + \gamma_s * Female_i + \lambda_t * Female_i + \theta X_i + \varepsilon_{ist} \quad (4)$$

where $Post_{s,t} * Female_i$ is the stay-at-home order indicator, where $Post_{s,t}$ equals 1 if state *s* has issued a stay-at-home-order in time *t*, and 0 otherwise. $Female_i$ indicates whether the funds are in the male or the female group. So, the interaction between *Post* and *Female* equals

1 if the fund is in a state *s* with a stay-at-home order at time *t* and the fund is in the female group. Thus, the coefficient of interest is β_3, and is the coefficient of the DDD effect. As in the DID estimator, the standard errors are clustered at state level.

The female fund group includes funds both exclusively managed by women and funds with a gender diverse manager group. To try to isolate the effects of the women's impact on the fund performance, we modify the DID from equation 1, by adding a continuous variable with the ratio of women in the manager group. To test for the effect of the female ratio, we will use following regression:

$$Outcome_{i,s,t} = \gamma_i + \lambda_t + \beta_1(Post_{s,t}) + \beta_2(Post_{s,t} * Female\ Ratio\ _i) + \varepsilon_{ist} \tag{5}$$

where the variables are the same as in equation 1, but an interaction term *Female Ratio*Post* is added, where *Female Ratio* is continuous on the interval $\in [0,1]$, and 1 indicates a fund exclusively managed by women. State fixed effects are replaced by fund fixed effects, as the ratio varies at the fund level.

5.2 Factor models and 1-R^2

To examine the empirical questions from our two hypotheses, we use different factor models and the Amihud-Goyenko 1-R^2 as the outcome variable. First, we examine whether the returns from mutual funds managed by at least one woman differ from the mutual funds managed by men under the lockdown period. Second, we will examine whether there are differences in the managerial skills across fund managers' gender during the work-from-home period.

5.2.1 Fund Performance

To estimate the empirical question in hypothesis 1, we use the alphas from four different factor models as the outcome variables in equation 1 and 5. The first one is the CAPM alpha (α). The CAPM alpha is a measure of the abnormal return over the market portfolio, hence it represents the risk-adjusted excess return, as seen in equation 6. As Xiao (2020) points out, the CAPM alpha is measured as the average of a window of daily returns, and the effect of COVID-19 and the stay-at-home-orders may not be properly inherited in the α's. We have used rolling window regression to calculate the CAPM alphas, and unlike Xiao (2020) we have used a time window of the last 90 daily returns. The thought is that reducing the time window from 180 to 90 daily observations will inherent and reflect an increased explanation of the stay-at-home orders.

$$R_{it} - Rf_t = \alpha_i + \beta(R_{CRSP,t} - Rf_t) + e_{it} \tag{6}$$

The second factor-model used is the Fama-French 3-factor model, which include three different factors: a market factor net of risk-free rate, a size factor and a value factor. The third factor-model is the Carhart 4-factor model, which adds a momentum factor to the Fama-French 3-factor model. The last one is the Fama-French 5-factor model, and includes the factors from the 3-factor model in addition to an investment and a profitability factor. Factor models are widely accepted as a measure of fund performance (Fama and French, 1993; Carhart, 1997). The factor models are applied to a regression framework, where we have used the fund returns over the risk-free rate on the left-hand side, and the factor models on the right-hand side as the explanatory variables. The regressions can be expressed as:

$$R_{it} - Rf_t = \alpha_i + \beta(R_{CRSP,t} - Rf_t) + s_i SMB_t + h_i HML_t + e_{it} \tag{7}$$

$$R_{it} - Rf_t = \alpha_i + \beta(R_{CRSP,t} - Rf_t) + s_i SMB_t + h_i HML_t + m_i MOM_t + e_{it} \tag{8}$$

$$R_{it} - Rf_t = \alpha_i + \beta(R_{CRSP,t} - Rf_t) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + e_{it} \tag{9}$$

where R_{it} is return for fund *i* at time *t,* Rf_t is risk-free rate at time *t*, *SMB* is the size factor at time *t*, *HML* is the value factor at time *t, RMW* is the profitability factor at time *t*, *CMA* is the investment factor at time *t,* and at last *MOM* is the momentum factor at time *t.*

Similar to the CAPM regression, we used 90-day windows of fund returns and factor returns as the input in the rolling window regressions. We used a relatively short time horizon in the regression, as Bollen and Busse (2005) suggests that superior performance of mutual funds are short lived and not persistent over longer periods. The output of interest from the regressions is the intercepts, which is the alphas (α_i) for fund *i.* The alphas are the returns for fund *i* at time *t*, and is the average return over the last 90-day period that are not explained by the benchmark factor models. Further, we will use the daily alphas to empirically examine our hypotheses.

5.2.2 Managerial skills

To examine our hypothesis 2, we had to find an appropriate way to measure the fund managers' skills. Managerial skills are the fund managers ability to achieve risk-adjusted abnormal returns net of fees above their benchmark, by either forecasting broad market trends or by identifying mispriced securities (Bodie, Kane & Marcus, 2011). As mentioned in the literature review, using persistence measures to account for managerial skill in such a short time period will not be feasible. We have chosen Amihud and Goyenko's (2013) 1-R^2 skill measure, as in accordance with the one used in Xiao's (2020) paper. We calculated the 1- R^2 by regressing funds' net daily return over the risk free rate, on two factor models. As explained in the fund performance part above, we used rolling window regressions with a 90 days-behind window. As the 1- R^2 skill measure is robust to several factor models, we have regressed the returns on both the Carhart 4-factor model and the CAPM to find the R^2s, which is similar to equation 6 and 8. Thus, the regression is equal to the ones used when measuring fund performance, but the coefficient of interest is now the R^2. It is the proportion of the variance in funds' net excess returns over market return that can be explained by the variance of the explanatory variables from the factor models. Hence, a higher 1-R^2 reflects a higher managerial skill. According to Amihud and Goyenko (2013), a higher 1-R^2 is associated with a higher alpha and better selective skills, and they define the 1-R^2 measure as:

$$1 - R^2 = \frac{RMSE^2}{SystematicRisk^2 + RMSE^2} \tag{10}$$

where *RMSE* is the idiosyncratic volatility, and the *SystematicRisk* is the return variance that is due to the factor models' risk. Hence, selectivity is better when the idiosyncratic volatility is higher, relative to its total variance. This means that the funds volatility is less driven by the systematic volatility from the factors.

5.3 Parallel trend assumption

According to Olden and Møen (2020), an extensive part of the published studies using DID and DDD estimators relies heavily on intuition, and do not include a formal discussion of the identifying assumptions. In order for a difference-in-differences estimator to be causally interpreted, the control group and the treatment group must share a parallel trend – *the parallel trend assumption*. Thus, in the absence of stay-at-home orders, the treatment and control groups should follow the same trend.

One way to provide evidence for the assumption, is to examine it graphically. We will graph the raw data of the outcome variables and compare the trends of both the male group of funds and the female group of funds. While this gives a rough estimate of whether or not there are parallel trends before the treatment period, it may be misleading in our analysis. This approach will be sufficient for the cases where there are only two periods, one pre- and one post-treatment. However, we have states that implement the stay-at-home orders at different points in time. Thus, it may give us an indication, but it would be hard to conclude anything just by looking at the plot of the raw data.

In his lecture notes on Empirical Methods in Applied Economics at the London School of Economics, Pischke (2005) presents an alternative method to examine the parallel trend assumption for staggered adoption. His approach is to include leads and lags of the treatment in the following regression:

$$Outcome_{i,s,t} = \gamma_s + \lambda_t + \sum_{k=-5}^{5} \eta_k \, Window_{s,t+k} + \theta X_i + \varepsilon_{ist} \tag{11}$$

where we have included 5 leads and 5 lags of treatment to the outcome regression presented earlier. The *Window* are dummy variables that equal 1 at day *t+k* (before and after stay-at-home-orders), and zero otherwise. The coefficients of the leads should not be significantly different from zero in order for the parallel trend assumption to hold. In other words, this means that the states that have not yet issued a stay-at-home-order, and/or never will, share the same trend as those who are to be treated within the next one to five days. The lags are included to examine whether or not there might be a delayed reaction to the treatments.

As mentioned, a DID needs to have parallel trends between the counterfactual (control group), and the treated group to be valid. However, a DDD estimator is basically constructed of two DID estimators, and Olden and Møen (2020) states that in a DDD estimator, the two DIDs can actually be biased. This way, the DDD is the only estimator that needs to have a valid parallel trend assumption. In order for this to be the case, the two DID estimators need to be biased in the same direction, hence the ratio between their trends are parallel. Using equation 11 to graph DIDs for both the female and male fund groups, we can examine the ratios between the trends.

The estimations also rely on there being an exogenous relationship between mutual fund profits and the stay-at-home-orders. Following Xiao's working paper (2020), there are three reasons to why stay-at-home orders are exogenous to mutual fund performance. First, mutual fund managers do not have any knowledge of where the pandemic are to break out and hence mutual funds cannot hedge against their physical location before the pandemic breaks out, and the working from home effects that follows. Second, stay-at-home orders were given to mutual funds exogenously and were based on local pandemic conditions, rather than mutual fund performance. Lastly, stay-at-home orders are given strictly by authorities, and are therefore not optional. Additionally, for the DDD estimations to be valid, there should not be any shocks during the period which would affect the relative mutual fund performance of funds managed by women versus the funds managed by men.

A challenge in the analysis of the remote working effects is the fact that the Federal Reserve unexpectedly decided to announce a cut in interest rate, and open market purchase of fixed income securities only days before the first stay-at-home order was announced. Mutual funds would normally benefit from a cut in interest rates, which would potentially make an impact on our analysis. However, Xiao (2020) compared the mutual funds' returns on the shock from the Feds news, and he finds that there is an insignificant difference in performance across funds. Hence, it should not impact our analysis.

6. Results

In this section, we will estimate the effects of stay-at-home orders, with regards to our hypotheses. We will use the methods and estimators as explained above to measure the effects on both fund performance and managerial skills, and its relative effects on the fund managers' gender. Furthermore, we will discuss the results and examine the assumptions for our models.

6.1 Hypothesis 1: Fund Performance

As there is an extensive number of ways to measure fund performance, we have chosen to use four different factor models in our analysis to get an extensive view of the performance: 1) The CAPM model, 2) the Fama-French 3-factor model, 3) the Carhart 4-factor model, and 4) the Fama-French 5-factor model. Descriptive statistics of the funds' daily alphas from the rolling window regressions are presented in table 2. Both the mean and the median alphas are negative through the sample period. This indicates that fund managers in general have found it hard to outperform the factor models in our sample period, which is substantiated by the fact that the 75 percent quartile is barely positive for all of the four models. It is also important to bear in mind that the fund returns used in our data is net of fees, hence the alphas are presented from an investor's perspective.

Table 2: Descriptive statistics for the daily alphas from the factor models for the whole sample period.

Alpha	Mean	St.Dev	Min	Q25	Median	Q75	Max	Obs.
CAPM	-0,057	0,0973	-0,9309	-0,1077	-0,0446	0,0025	0,3984	167 090
Fama-French 3	-0,015	0,0525	-0,4636	-0,0389	-0,0147	0,0083	0,5411	167 090
Carhart 4	-0,0144	0,0538	-0,4864	-0,0383	-0,0135	0,0093	0,5744	167 090
Fama-French 5	-0,0184	0,0511	-0,4635	-0,0397	-0,0151	0,0059	0,5492	167 090

Table 12 in the appendix shows the unconditional regression output for the four factor models for the whole sample period. We notice that the coefficient for *Female* is positive for alpha, although not significantly different from zero. However, when adjusting for more factors in the Fama-French 3- and 5-factor model, and the Carhart's 4-factor model, the coefficient for female drops to negative. It is statistically different from zero for the Fama-French 3 factor model and when adding the Momentum factor. In the whole sample period, the daily alpha generated from these two models are 0,4 bps lower for funds with at least one female manager, relative to the funds managed exclusively by men. Additionally, we notice that fund size and the tenure of the longest sitting manager have a significant impact on the generated alphas during the whole period. Annual expense ratio and the age of the funds also impact the funds' CAPM alphas during the period. For that reason, the male funds seem to perform slightly better in our sample period. From the methodology section we put forward the outcome regressions to be applied in the analysis. In the following section we will present and discuss results obtained by implementing the four different factor models and net excess returns as the *Outcome* variables.

Table 3 presents the estimated results from the difference-in-differences and the difference-in-differences-in-differences strategy, respectively using equation 1 and equation 4, and the alphas from the CAPM model as the *Outcome* variable. Column 1 and 2 show the results from the DID estimation with time and state fixed effects, while column 3 shows the results from the DDD estimation with time, state, gender*state and gender*time fixed effects. The full set of controls are included in all of the three columns.

We find that the coefficients of interest from DID estimations on both the male and the female subgroups are not significant. The results suggest that treated funds, funds that are located in a state that issues a stay-at-home-order at any point in time during the period, have a 0,10 bps larger daily CAPM alpha during the remote working period if it has no women in the manager group. Funds managed by at least one female suffers a reduction in daily CAPM alpha by 0.13 bps. However, since there are no evidence of a significant effect, we cannot conclude on funds having a negative effect on working from home. The same applies to the DDD-estimation; we find no significant effect of working from home for funds managed by at least one female manager, relative to the male group. Interestingly, the coefficient of the DDD term is equal to the difference between the two DID coefficient from the two subgroups. This finding matches the theory presented by Olden and Møen (2020), where they state that the triple difference estimator also can be calculated as the difference between two DIDs estimators.

Table 3: Outcome regressions using the funds' CAPM alphas

	α_{CAPM}		
	(1)	(2)	(3)
	Male	Female	DDD
Post*Female			-0.0023
			(0.0045)
Female			0.0055
			(0.0044)
Post	0.0010	-0.0013	0.0010
	(0.0037)	(0.0064)	(0.0037)
Fund Size	7.44e-13***	1.13e-12***	7.88e-13***
	(1.22e-13)	(3.60e-13)	(2.14e-13)
Expense ratio	-0.0126	-0.0104	-0.0134**
	(0.0082)	(0.0065)	(0.0063)
Age of fund	0.0006***	0.0009**	0.0006***
	(0.0002)	(0.0004)	(0.0002)
Tenure (longest)	0.0001	-0.0009	-0.0000
	(0.0003)	(0.0006)	(0.0003)
Team	-0.0127	-0.0049	-0.0119
	(0.0081)	(0.0114)	(0.0072)
Constant	-0.0928***	-0.0192*	-0.0829***
	(0.0113)	(0.0098)	(0.0101)
N	123008	44082	167090
R^2	0.142	0.183	0.149
State FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
State*Female	No	No	Yes
Time*Female	No	No	Yes

Note: The table presents the outcome regression results with daily CAPM alphas as the dependent variable. Column 1 and 2 presents the difference-in-differences regression using equation 1 with respectively the male and the female subgroups of funds. Colum 3 displays the difference-in-differences-in-differences regression using equation 4. Post is the independent variable of interest in column 1 and 2, while Post*Female is the independent variable of interest of the DDD regression in column 3. All columns include state and time fixed effects, while column 3 adds state*female and time*female fixed effects. The robust standard errors are clustered at state level, and are shown in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

For the remaining three factor models, we use the same procedure as we described above for the CAPM model. Table 13 in the appendix shows that the estimates where we use the Fama-French 3-factor model as the outcome variable. Similar to the CAPM estimation, none of the estimated coefficients of interest are statistically significant, which indicates a rejection of hypothesis 1. The triple difference estimate is now negative -0.27 bps, and is close to the one we got from the CAPM model. However, the DID-estimates are lower, with -0,15 bps and -0,42 bps for respectively the male and the female subgroups. This suggests that both funds managed by female and male managers perform worse during the period of remote working. Nonetheless, the results are not significant and we cannot conclude on there being a difference between working from home or not.

In table 14 in the appendix, the results from the estimation using the Carhart 4-factor model are displayed. As mentioned earlier, the Carhart 4-factor model is similar to the Fama-French 3-factor model, but with an additional momentum factor. Consistent with the Fama-French 3-factor model, the estimates of the post period are both negative. However, the momentum factor impacts the two subgroups differently, as the female fund group have a coefficient of -0.26 bps, which is larger than in the Fama-French 3-factor estimation, and the coefficient for the male subgroup are lower, at 0,24 bps. Hence, the estimated triple difference is small, at 0.01 bps. Given the fact that there is no significant difference in the post treatment estimations from the Carhart 4-factor model, and that the triple difference estimation are very low, summarized this indicates that there are not a difference between the two fund groups in remote working generated Carhart 4-factor alpha.

The last factor model of which we have used alphas as the outcome variable, is the Fama-French 5-factor model. The results are presented in table 15 in the appendix. It adds a profitability factor and an investment factor to the 3-factor model. The estimated response to working from home are -0.08 bps for the male fund group and -0.38 bps for the female fund group in generated daily alpha. The triple difference estimate is a 0.30 bps in daily alpha. The coefficients for the post treatment period are still not significant in neither of the columns.

As the daily alphas from the rolling window regressions are calculated from the past 90-days observations, we also run the outcome regressions with net excess returns as the dependent variable, since the factor models might not fully inherit the effect of the stay-at-home orders.

However, the results presented in table 16 are not significant, and thus there are no evidence of there being any difference among the funds pre and post stay-at-home orders.

Fund size has a significant impact on the alphas and the net excess returns in all of the estimations. Moreover, the coefficients of the age of funds, the team indicator and annual fund expense ratio are significant in several of the estimations. To control for how the covariates impact our estimations, we run the triple difference for the factor models without any controls. The results are presented in table 17 in the appendix. As expected, the coefficients of the triple difference term *DDD* is not affected by the covariates. Our covariates are not time-varying, thus should they not affect the coefficient. As stated in the methodology part, they are included to control for confounding trends. Pischke (2005) shows that the time-invariant covariates also can help reduce the standard error of the policy coefficient. However, we do not find this effect in our estimation.

The findings from the estimations with regard to the factor models coincide; there are not any significant differences on the effect of working from home on fund performance between funds with only male managers and funds with at least one female manager. Hence, we can reject our hypothesis 1, that the effect of working from home during the COVID-19 crisis deteriorates factor-adjusted returns more for active mutual funds managed by women, compared to active mutual funds managed my men.

However, as previously mentioned the female group of funds consist of funds with at least one female manager. Within this group, 639 of the funds have a gender diverse manager group, while only 72 of the funds are exclusively managed by women. Hence, the results may be misguiding, as we do not account for the ratio between women and men in the funds, as well as the power distribution between them. It is plausible to believe that a higher ratio of women to men reflects a higher ratio of decisions made by women. To investigate this, we run regressions using net excess returns and the four factor models as the outcome variables with equation 5 for the female subgroup. The results are presented in table 4. The findings from the regression indicate that there is a significant positive effect on generated alpha for CAPM, Fama-French 3-factor and Carhart 4-factor if the ratio of female to men in the manager group is higher. Thus, the effect of working from home on mutual funds' alphas from the factor models are better for funds with a higher ratio of women in the manager group.

Table 4: Difference-in-differences with female ratio for the female/mixed subgroup, using equation 5.

	(1) Excess returns	(2) α_{CAPM}	(3) α_{FF3}	(4) $\alpha_{Carhart\,4}$	(5) α_{FF5}
Post*Female Ratio	-0.0589	0.0291**	0.0140*	0.0136*	0.0125
	(0.0520)	(0.0124)	(0.0077)	(0.0070)	(0.0092)
Post	0.0047	-0.0117	-0.0093	-0.0075	-0.0082
	(0.0348)	(0.0079)	(0.0088)	(0.0074)	(0.0092)
Constant	-0.1919***	-0.0184***	-0.0130***	-0.0124***	-0.0126***
	(0.0262)	(0.0037)	(0.0035)	(0.0035)	(0.0032)
N	44082	44082	44082	44082	44082
R^2	0.236	0.801	0.605	0.603	0.577
Fund FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes

Note: The table presents the DID outcome regression results when using equation 5, with net excess returns and daily alphas from the four factor models as dependent variables. Post*Female Ratio is the variable of interest. All columns include fund and time fixed effects. The robust standard errors are clustered at state level, and are shown in parentheses. Significance levels: $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

6.2 Hypothesis 2 – Managerial skill:

To estimate and verify our hypothesis 2 we use equation 4, with Amihud-Goyenko's 1-R^2 as the *Outcome* variable. To obtain our estimates of 1-R^2, we regress the fund net excess returns on both the CAPM and the Carhart 4-factor model. The mean selective skill during the whole period from the CAPM and the Carhart 4-factor model was 16.76 percent and 12.05 percent, as displayed in table 5. At first, we run the unconditional estimation, presented in table 18 in the appendix. There is not a significant difference between the 1- R^2 measured for funds in the male and female group, which indicates no differences in managerial skill between the genders during the whole sample period.

Table 5: Descriptive statistics for the daily 1-R^2 in the sample period.

1-R2	Mean	St.Dev	Min	Q25	Median	Q75	Max	Obs.
CAPM	0.1671	0.1816	0.0002	0.0413	0.1080	0.2186	1.000	167 090
Carhart 4	0.1205	0.1620	0.0001	0.0183	0.0586	0.1479	0.9976	167 090

However, as our interest lays in the effect of remote working on managerial skill, we continue our analysis by looking at the post-lockdown period compared to the pre-lockdown period. The results from the estimation when using the 1- R^2 from the CAPM is displayed in table 6. The skill measure for the male funds get 1.11 percent lower for the treated funds in the treatment period, and is statistically significant at the 1 percent level. The corresponding value for the female group of funds is an increase in managerial skill measure by 0.43 percent, but contrary to the male group it is not significant. The coefficient of interest is the DDD term, and it shows that female funds get an increase in managerial skill by 1.54 percent per day, relative to the male fund group in the period of remote working. The coefficient is significant at the 5 percent level, hence the result is the opposite of our hypothesis 2; that the effect of working from home during the COVID-19 crisis deteriorates managerial skills more for female mutual fund managers, compared to male mutual fund managers.

Table 7 presents the estimates of selective skills using 1-R^2 from the Carhart 4-factor model as the *Outcome* variable. Similar to the results in table 6, working from home had a negative effect on selective skills in the male subgroup, with a 0.78 percent decrease, while in the female subgroup there was an increase of 0.19 percent. Thus, the DDD estimate of selective skills indicates a positive effect on the female group's managerial skill by 0.97 percent relative to the male subgroup. Again, the result is the opposite of hypothesis 2.

Table 6: Outcome regressions using the funds' CAPM 1-R^2

	$1 - R^2_{CAPM}$		
	(1) Male	(2) Female	(3) DDD
Post*Female			0.0154**
			(0.0059)
Female			-0.0173**
			(0.0081)
Post	-0.0111***	0.0043	-0.0111***
	(0.0041)	(0.0042)	(0.0041)
Fund Size	-6.93e-13	-9.55e-13***	-8.58e-13***
	(5.91e-13)	(1.19e-13)	(1.79e-13)
Expense ratio	0.1273***	0.1186***	0.1254***
	(0.0172)	(0.0181)	(0.0154)
Age of fund	-0.0014***	-0.0005	-0.0012***
	(0.0005)	(0.0004)	(0.0003)
Tenure (longest)	0.0002	-0.0004	0.0001
	(0.0003)	(0.0010)	(0.0004)
Team	-0.0207	-0.0231	-0.0212
	(0.0150)	(0.0225)	(0.0130)
Constant	0.2559***	0.1709***	0.2581***
	(0.0308)	(0.0241)	(0.0274)
N	123008	44082	167090
R^2	0.355	0.357	0.356
State FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
State*Female	No	No	Yes
Time*Female	No	No	Yes

Note: The table presents the outcome regression results with daily 1-R^2 from the CAPM as the dependent variable. Column 1 and 2 presents the difference-in-differences regression using equation 1 with respectively the male and the female subgroups of funds. Colum 3 displays the difference-in-differences-in-differences regression using equation 4. Post is the independent variable of interest in column 1 and 2, while Post*Female is the independent variable of interest of the DDD regression in column 3. All columns include state and time fixed effects, while column 3 adds state*female and time*female fixed effects. The robust standard errors are clustered at state level, and are shown in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Outcome regressions using the funds' Carhart 4-factor model 1-R^2

	$1 - R^2_{Carhart\ 4}$		
	(1) Male	(2) Female	(3) DDD
Post*Female			0.0097**
			(0.0045)
Female			0.0004
			(0.0070)
Post	-0.0078***	0.0019	-0.0078***
	(0.0022)	(0.0042)	(0.0022)
Fund Size	-2.36e-13	-6.05e-13***	-5.25e-13***
	(5.42e-13)	(9.46e-14)	(1.87e-13)
Expense ratio	0.1112***	0.1051***	0.1098***
	(0.0134)	(0.0181)	(0.0122)
Age of fund	-0.0013***	-0.0005	-0.0011***
	(0.0004)	(0.0004)	(0.0003)
Tenure (longest)	-0.0003	-0.0012	-0.0005
	(0.0003)	(0.0008)	(0.0003)
Team	-0.0179	-0.0109	-0.0173*
	(0.0113)	(0.0198)	(0.0100)
Constant	0.2020***	0.1126***	0.2039***
	(0.0227)	(0.0263)	(0.0201)
N	123008	44082	167090
R^2	0.278	0.285	0.279
State FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
State*Female	No	No	Yes
Time*Female	No	No	Yes

Note: The table presents the outcome regression results with daily 1-R^2 from the CAPM as the dependent variable. Column 1 and 2 presents the difference-in-differences regression using equation 1 with respectively the male and the female subgroups of funds. Colum 3 displays the difference-in-differences-in-differences regression using equation 4. Post is the independent variable of interest in column 1 and 2, while Post*Female is the independent variable of interest of the DDD regression in column 3. All columns include state and time fixed effects, while column 3 adds state*female and time*female fixed effects. The robust standard errors are clustered at state level, and are shown in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Similarly to what we did with the alphas in the performance part, we utilize equation 5 to test the female subgroup of mutual funds on female ratio. From the results above, we have indications of better selective skills in the female subgroup when forced to work from home, relative to the male subgroup. However, the results from table 8 presents a significant negative effect of working from home if the female ratio is higher. Consequently, the results are somehow contradictory. We will address this further in the discussion section.

Table 8: Difference-in-differences with female ratio for the female/mixed subgroup, using equation 5.

	(1) $1 - R^2_{CAPM}$	(2) $1 - R^2_{Carhart\ 4}$
Post*Female Ratio	-0.0380**	-0.0321**
	(0.0166)	(0.0155)
Post	0.0179***	0.0134**
	(0.0056)	(0.0060)
Constant	0.2915***	0.2231***
	(0.0042)	(0.0043)
N	44082	44082
R^2	0.849	0.823
Fund FE	Yes	Yes
Time FE	Yes	Yes

Note: The table presents the DID outcome regression results when using equation 5, with 1-R^2 from CAPM and Carhart 4-factor model as dependent variables. Post*Female Ratio is the variable of interest. All columns include fund and time fixed effects. The robust standard errors are clustered at state level, and are shown in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

6.3 Robustness

To investigate the robustness of our DDD-regression model, we have examined the effects of applying different levels of clustering to the 1-R^2 estimations. The results are presented in table 9, where we use clusters at the fund and state-time level. The estimated coefficients are the same as the original estimations with clustering at state level, and the significance remains the same at the 5 percent level. This shows that our estimations of selective skills are robust to multiple cluster levels.

Table 9: Robustness test of DDD using 1-R^2 and clusters at different levels

	$1 - R^2_{CAPM}$		$1 - R^2_{Carhart\ 4}$	
	(1) Fund cluster	(2) State-time cluster	(3) Fund cluster	(4) State-time cluster
Post*Female	0.0154**	0.0154**	0.0097**	0.0097**
	(0.0051)	(0.0062)	(0.0049)	(0.0044)
Female	-0.0173	-0.0173**	0.0004	0.0004
	(0.0251)	(0.0069)	(0.0250)	(0.0060)
Post	-0.0111***	-0.0111**	-0.0078***	-0.0078***
	(0.0029)	(0.0043)	(0.0028)	(0.0025)
N	103168	36972	103168	36972
R^2	0.354	0.355	0.276	0.283
Controls	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
State*Female	Yes	Yes	Yes	Yes
Time*Female	Yes	Yes	Yes	Yes

Note: We test for robustness by clustering the at different levels. Column 1 and 3 uses clustering at fund level, while column 2 and 4 uses clustering at state-time level. All columns include state, time, state*female and time*female fixed effects and all controls. The robust standard errors are clustered at the state level, and are shown in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Additionally, we add state-time fixed effects to the DDD-estimation to test for state-specific time trends and shocks. Table 10 presents the results, where we notice that the coefficients for 1-R^2 for the two factor models does not change much, and are still highly significant at the 1 percent level. This suggests that the model is robust against any shock within states in the sample period, and supports the findings from the DDD-estimations.

Table 10: DDD-estimation with state-time fixed effects.

	(1) $1-R^2_{CAPM}$	(2) $1-R^2_{Carhart\ 4}$
Post*Female	0.0167***	0.0125***
	(0.0056)	(0.0041)
Female	-0.0154*	0.0025
	(0.0086)	(0.0073)
Constant	0.0936***	0.0529**
	(0.0324)	(0.0244)
N	166842	166842
R^2	0.358	0.282
Controls	Yes	Yes
State*Female	Yes	Yes
Time*Female	Yes	Yes
Time*State	Yes	Yes

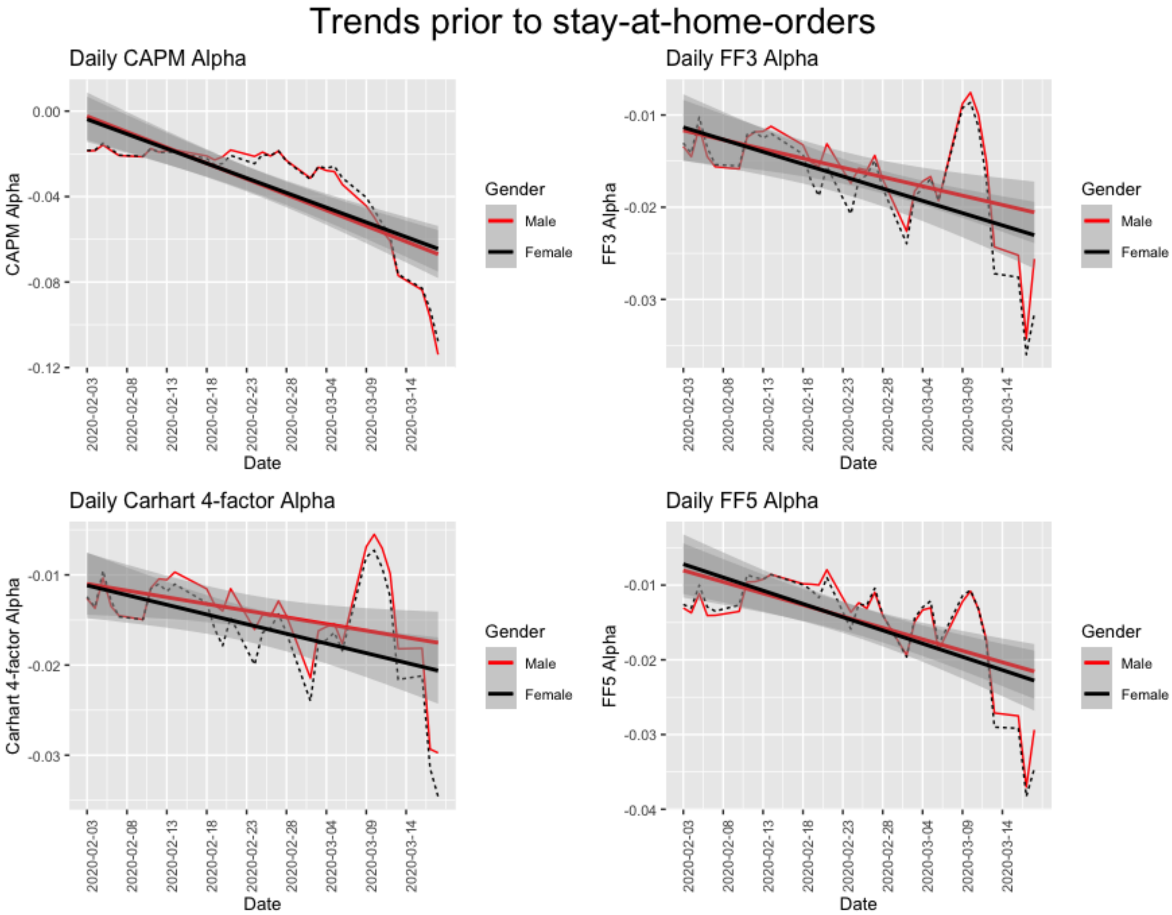
Note: The table presents the regression results from equation 4, but replacing State and time fixed effects with state*time fixed effects. The regression includes all controls. The post term gets dropped due the the time*state fixed effects. The robust standard errors are clustered at the state level, and are shown in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

6.4 Parallel trend assumption

The main assumption when using the difference-in-differences and the triple difference estimators, is the parallel trend assumption. If the assumption is violated, the estimations may be biased, and one cannot conclude on a causal treatment effect. In this part we will investigate the trends in the models used above. As mentioned in the methodology part, we only need the triple difference estimator to be unbiased. Consequently, we will look at the trends between the male funds' subgroup and the female funds' subgroup.

At first, we will examine the raw data by subgroups, presented in figure 1. It shows the linear trendlines from the start of our sample period until the last day before California announces the first U.S. stay-at-home-order. From the figure, we can see that the CAPM alpha and the Fama-French 5-factor model seems to indicate parallel trends in the period before the first stay-at-home-order is issued. On the contrary, daily alphas from the Fama-French 3-factor model and the Carhart 4-factor model do not share parallel trends prior to the stay-at-home-orders. They are similar at the start of April, but the female funds have a steeper slope compared to the male funds.

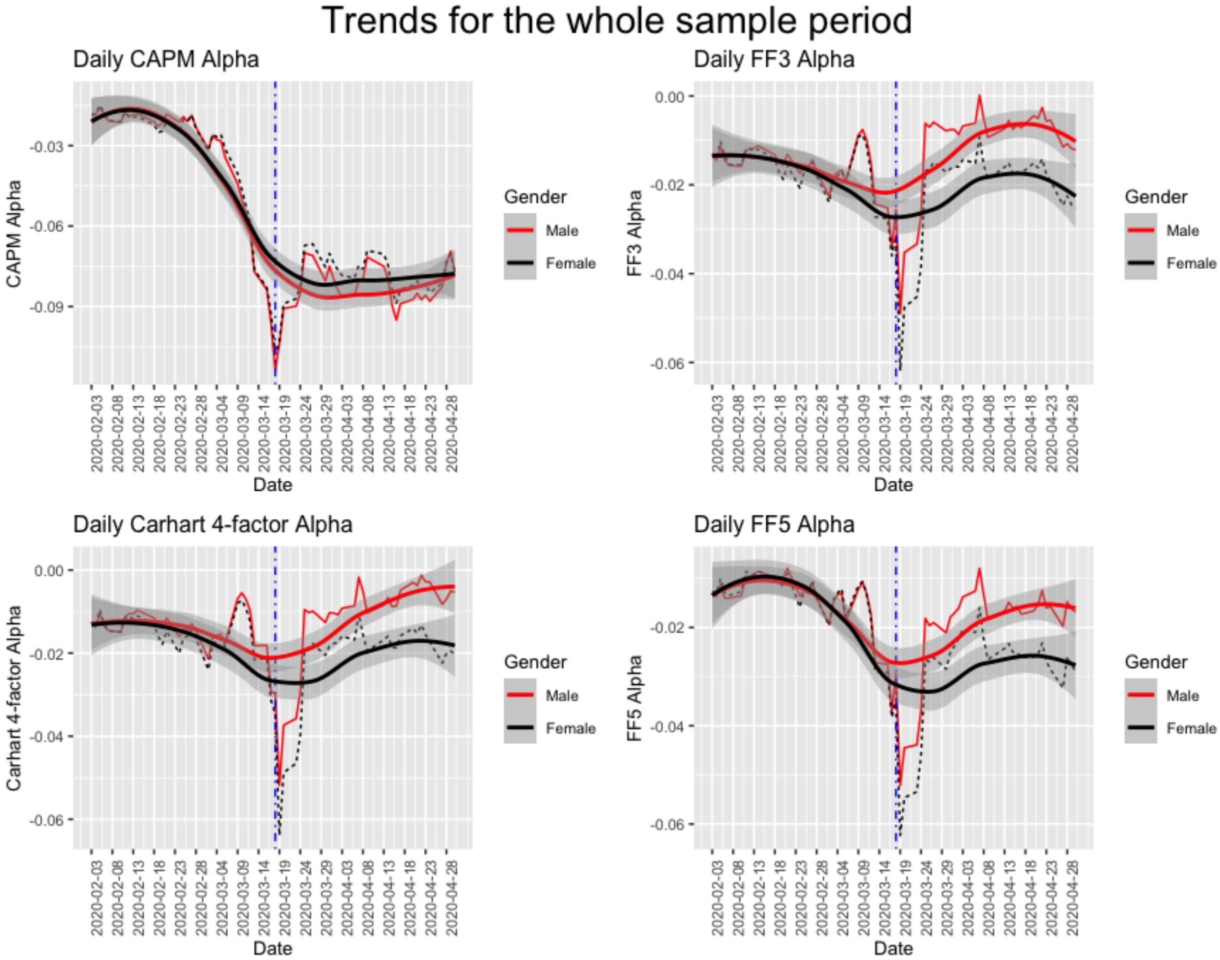
Figure 1: Unconditional linear time trends for the funds alphas from the factor models before the first stay-at-home orders.



Note: The linear trend lines present the linear trends for the two subgroups: red and black lines represents respectively the male and female subgroups. The grey areas are the 95 percent confidence intervals. The graphs in the background are daily alphas.

Additionally, we have graphically examined the smoothed trends throughout our whole sample period in figure 2. The blue vertical line represents the day of the California stay-at-home order. Again, we can see that the CAPM seems to have parallel trends. The two groups have a similar trend until the day of the announcement of the first stay-at-home-order, before we can see a shift in trend for the female group, compared to the male group. The trends in the Fama-French 3-factor model and the Carhart 4-factor model start to deviate too early, thus the parallel trend assumption seems unlikely to hold. As for the Fama-French 5-factor, the trends between the two groups deviates a little before the treatment period, but they still seem likely to satisfy the assumption.

Figure 2: Unconditional smoothed time trends for the funds alphas from the factor models for the whole sample period.

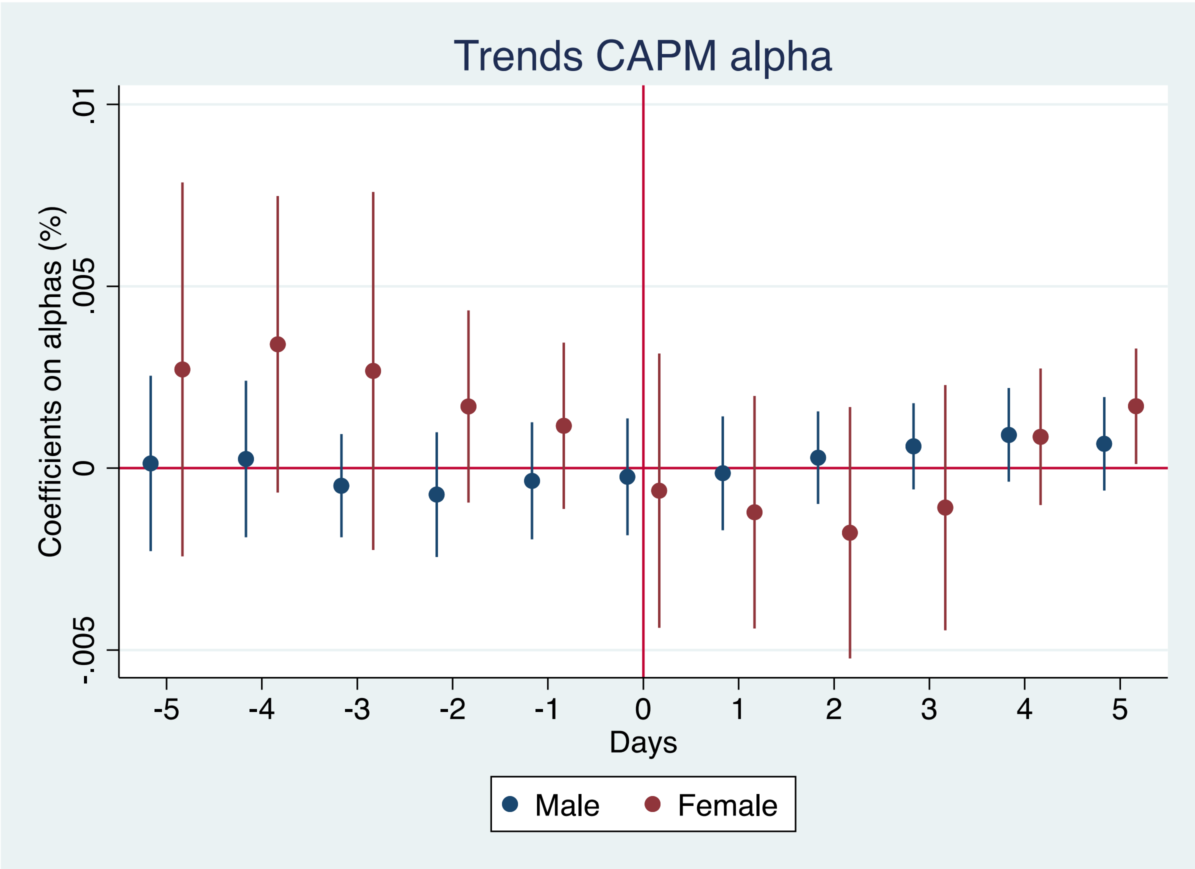


Note: The smoothed trend lines present the linear trends for the two subgroups: red and black lines represents respectively the male and female subgroups. The grey areas are the 95 percent confidence intervals. The graphs in the background are daily alphas.

In our study we have treatment at different points in time, and we should not make a final conclusion of the parallel trends from the graphs above. Since we have treatment at multiple points in time, we will use equation 10 to further examine the parallel trends assumption. Figure 3 shows the plotted coefficients from the equation, using the two gender subgroups and CAPM alpha. Similar plots for the other factor models' alphas are presented in figure 6-8 in the appendix. The coefficient for each subgroup represents the difference between the counterfactual and the treated group, and are presented in table 19 and 20 in the appendix. Thus, the intuition is that the coefficients of the leads should not be significantly different from zero. The graphs confirm what we saw from the unconditional data plots; the CAPM alpha and Fama-French 5-factor models have a consistent ratio between the two groups' coefficients, and seems to satisfy the parallel trend assumption. Equally, the Fama-French 3- and Carhart

4-factor models have a similar trend the last three days before the stay-at-home-orders, but there's a shift between lead 5 and 4. Thus, there might be bias in the estimations.

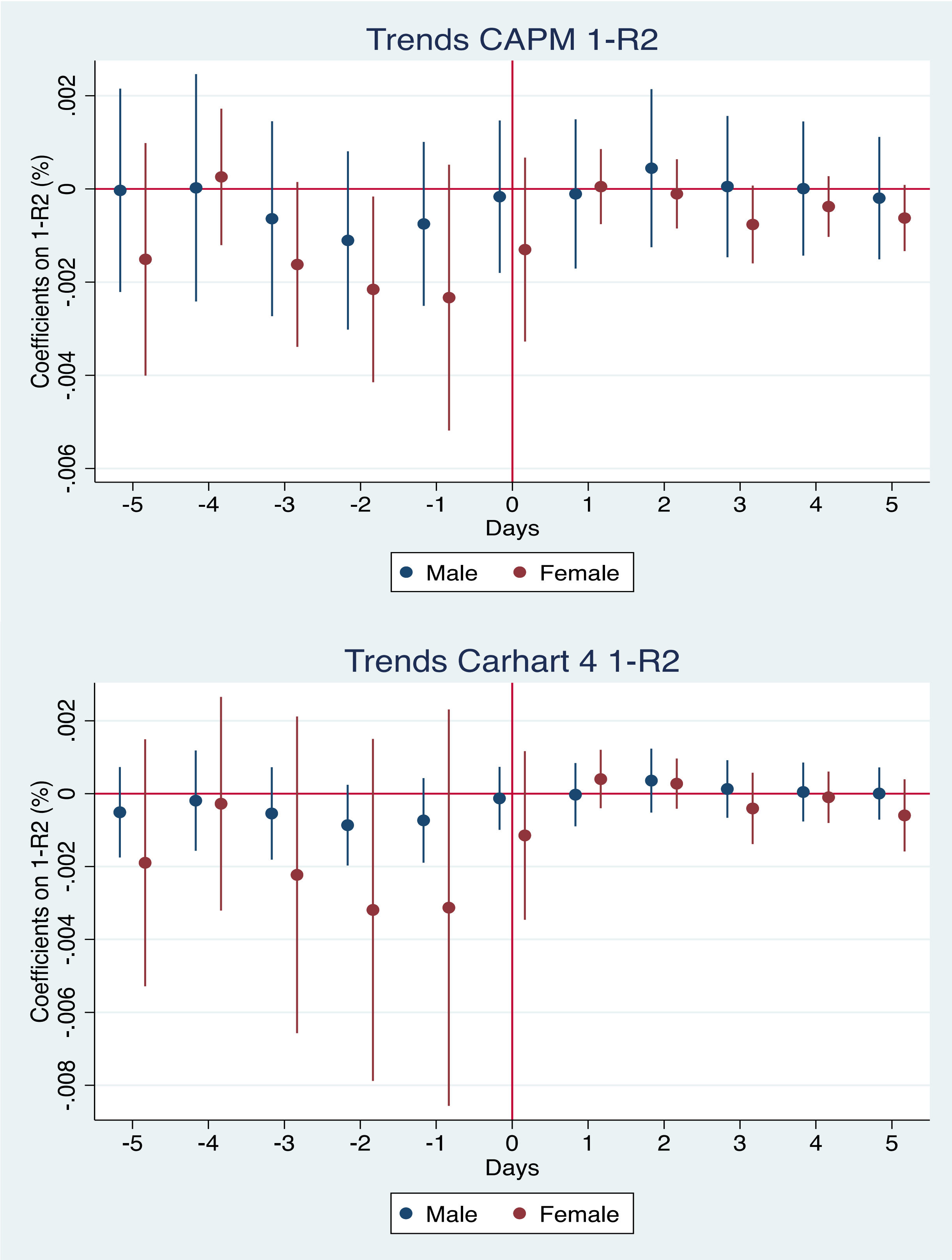
Figure 1: Leads and lags around the stay-at-home orders for the funds' CAPM alphas



Note: The figure presents the coefficients of leads and lags around the stay-at-home order announcements, using equation 11. The main interest is that leads should be close to zero, or the ratio between male and female should be parallel for the leads.

As for the measuring of managerial skills, the trends plot of the 1-R^2 from the two factor models are presented in figure 4 and 5. The trend plots are not convincing for neither of the models, as the two groups does not seem to have parallel trends in the days before the stay-at-home order announcements. However, none of the coefficient of the leads are statistically different from zero for the 1-R^2 from the Carhart 4-factor model. This can be seen in Table 20 in the appendix. Hence, we the parallel trend assumption seems likely to hold when using the Carhart 4-factor model to measure selective skill.

Figure 4 and 5: Leads and lags around the stay-at-home orders for the CAPM and the Carhart *4-factor models 1-R^2*



Note: The figures presents the coefficients of leads and lags around the stay-at-home order announcements, using equation 11. The main interest is that leads should be close to zero, or the ratio between male and female should be parallel for the leads.

7. Discussion

In this part we will discuss the results from our analysis. First, we will discuss the regression results and the two hypotheses. Second, we will compare our findings with regards to recent literature. At last, we will discuss potential limitations of our methods and the implication of these for the interpretation of our analysis.

7.1 Hypothesis 1 – Fund performance

Our results from the analysis on fund performance are unambiguous. For the unconditional estimation, none of our models show evidence of significant difference in performance. Though all models but the CAPM estimates a negative coefficient for the female dummy, none of the coefficients are statistically significant, implying no evidence of different performance between the genders.

Going on with the difference-in-differences analysis the results stay the same. The coefficients remain unsignificant looking at the subgroups, but it is worth noting how the coefficient for *Post* is negative for the female subgroup in all factor-models, and always more negative than the same coefficient for the male sub group. Given this, as explained in Olden and Møen (2020), it is expected that the triple difference coefficient is negative in all models, indicating that the female managed funds deliver lower returns than the male managed funds. However, the coefficient of the triple difference estimator is unsignificant as well, implying that there is no significant difference in performance when comparing the genders. Not considering the remote working effect, these wouldn't have been surprising results. As discussed in the literature review, these results are consistent with the findings of Babalos, Caporale and Philippas (2015) and Niessen and Ruenzi (2019).

Interestingly, we find significant evidence of funds with a gender diverse management benefiting from having a higher ratio of female managers in their manager group during the period of remote working. A possible explanation is that the working environment is less hostile when the managers are working from their homes, and therefore lead to female managers having more room to make investment decisions. Another possible explanation could be that female managers in general are less risky in their investment strategies, which concurs with the findings of Niessen and Ruenzi (2019).

7.2 Hypothesis 2 – Managerial skill

Looking at the results from the unconditional model, we find no significant difference in the change of managerial skill between the female subgroup compared to the male subgroup. We can question the importance of these findings, but given the results of no significant difference on fund performance between the two subgroups, combined with the literature concluding with the same, it is not unexpected that the managerial skill shows the same results for our sample.

Conversely, when looking at both 1-R^2 derived from the CAPM and the Carhart 4-factor model we are witnessing significant differences, but not as we hypothesized. The regression of the CAPM 1-R^2 gives us a statistically significant negative coefficient of -0,0111 for the Post-variable in the male subgroup, indicating significant change in managerial skill for the male managers. As opposed to this, the Post-variable for the female subgroup is positive, but insignificant. The triple difference coefficient is significant and equals the difference between the coefficients of the subgroups, 0,0154, and confirm the presence of a decrease in managerial skill for male managers compared to female managers. The triple difference coefficient of the 1-R^2 from the Carhart 4-factor is consistent with what we see for the CAPM, indicating a significant positive effect of managerial skill when working from home for the female group relative to the male group.

Similar to what we did in the performance section, we added female ratio to see if it impacted the results for the female group, which consists of funds with at least one female manager. The results were somehow surprising; a higher ratio of female managers to male managers in funds indicates a significant negative effect of working from home. Thus, the results are contradictory; the female fund group have a better effect of stay-at-home orders on selective skills than the male group, but a higher ratio of the female managers in the funds with a gender diverse management indicates a negative effect of stay-at-home orders.

Our results are not in line with our hypothesis, it is actually indicating a negative remote working effect for men relative to women, and not the other way around as we expected with our hypothesis.

7.3 Limitations

From our hypotheses, we expected that the disproportionate impact of the remote working effect on genders would yield results where the returns of female managed mutual funds where inferior those whom were managed by men. Our results suggest otherwise, and one possible explanation is that our study is limited by the fact that we do not have more available data on managerial characteristics. The difference in time spent on caregiving responsibilities might vary in great extent from manager to manager depending on how many kids a manager has, if they have sole custody or not, if they have a nanny and possibly what their significant other's occupation and education. The reports we discuss in the hypothesis development argued that women were expected to be responsible for the care-giving activities. A report by Modern Fertility (2020) released this year showed that among the top reasons why people are putting of kids, 51 percent answered that they wanted to earn a higher salary first, and 35 percent answered that they wanted to reach a certain title or level in the career first. With this in mind, it is not unreasonable to believe that women working around the clock are putting of having kids. If this is the case in the mutual fund industry, the care-giving activities would be fewer and being able to control for this should increase the precision of our results. Thus, the effect of remote work might not be as disproportionate as we have assumed.

The daily input we have used in all the outcome regressions in this thesis are based on factor models from rolling regression windows of the last 90 days. Thus, a possible challenge to the results from our outcome regressions are in what degree the alphas and $1\text{-}R^2$ variables actually inherits the effect of working from home. Xiao (2020) mentions the same problem in his paper, however he used a 180-day regression window. The stay-at-home orders are announced between March 19 and April 7, which leaves about a month of observation for the factors to inherit the effects of stay-at-home orders. Thus, there should be enough time for the factors to inherit the effects of working from home. However, it is likely that the effects increase during the post-treatment period, as the number of days of the 90-days window that are post-treatment will increase with time. As a result, our estimations may show the effect of stay-at-home orders as weaker than they actually were.

8. Conclusion

Analyzing the effect of stay-at-home orders on mutual fund performance and managerial skill with regards to the gender of managers is challenging. The situation is a consequence of the COVID-19 pandemic, and are not similar to anything we have seen in the modern times. The time period of our study was characterized by volatile markets and uncertainty in the macro economy. Hence, there are a lot of factors that may have impacted the mutual funds beyond remote working. By using the difference-in-differences and difference-in-differences-in-differences strategy, we have tried to isolate the effects of working from home.

First, we used the regression design with factor models' alphas and net excess returns as input to investigate the effect of working-from-home on mutual fund performance. We found no significant difference in the effect of working from home on mutual fund profitability between the funds managed by at least one woman and the funds managed exclusively by men. Second, we used the same strategy to test for differences in managerial skill. The results suggest that mutual funds managed by at least one woman show a positive effect on managerial skill from remote working, relative to the male group. When using the 1-R^2 from the CAPM as a skill measure, the parallel trend assumption is violated, which indicates a possible biased estimation. However, when using the Carhart 4-factor model, the parallel trend assumption seems to hold, indicating a causal interpretation. The DDD regressions are robust when testing for multiple cluster levels and state specific shocks, which supports our findings.

Because the funds in the female subgroup mainly includes manager groups with both genders, we wanted to test for the effect of the ratio of women to men within the subgroup. When testing for profitability, the findings suggest a significant positive effect of having a higher female ratio. On the other hand, we find that a higher female ratio suggests a significant negative effect on managerial skill. The latter result contradicts with our findings from the DDD regression on managerial skill.

These contradictory results show that there may be limitations to our analysis. The time period is relatively short, and the amount of funds managed exclusively by women is low. These are factors that may affect our analysis, and makes it harder to conclude with causal relationships. For further research, we would advise including more managerial characteristics, such as family situation and educational background. To increase the amount of all-female managed funds we suggest including funds from other parts of the world whom have experienced forced remote working. As the pandemic is still ongoing, we are witnessing new lockdowns in countries that have re-opened once already. Extending the sample period to include these recent lockdowns would also be recommended.

References

Ammy-Driss, A. & Garcin, M. (2020). Efficiency of the financial markets during the COVID-19 crisis: time-varying parameters of fractional stable dynamics.

Amihud, Y. & Goyenko, R. (2013). Mutual Fund's R2 as Predictor of Performance, *The Review of Financial Studies*, Volume 26, Issue 3, March 2013, Pages 667–694, https://doi.org/10.1093/rfs/hhs182Atkinson, S., Baird, S., Frye, M. (2003). Do Female Mutual Fund Managers Manage Differently? *Journal of Financial Research, 2003 (26).* https://doi.org/10.1111/1475-6803.00041

Babalos, V., Caporale, G. & Philippas, N. (2015). Gender, style diversity, and their effect on fund performance. *Research in International Business and Finance, 2015 (35),* 57-74. https://doi.org/10.1016/j.ribaf.2015.02.020

Barber, B. M. & Odean, T. (1998). Boys Will Be Boys: Gender, Overconfidence, and Common Stock Investment (November 1998). Available at SSRN: https://ssrn.com/abstract=139415 or http://dx.doi.org/10.2139/ssrn.139415

Berk, J. B., & Binsbergen, J. v. (2015). Measuring Skill in the Mutual Fund Industry. Journal of Financial Economics, 118 (1), 1-20. http://dx.doi.org/10.1016/j.jfineco.2015.05.002

Bloom, N., Liang, J., Roberts, J. & Ying, Z. (2013). Does Working from Home Work? Evidence from a Chinese Experiment. *The Quarterly Journal of Economics*. 130. 10.1093/qje/qju032.

Bodie, Z., Kane, A., & Marcus, A. J. (2011). Investments. New York: McGraw-Hill/Irwin.

Fama, E. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. The Journal of Finance, 25(2), 383-417. doi:10.2307/2325486

Bollen, N. P. B. & Busse, J. A. (2005). Short-Term Persistence in Mutual Fund Performance, *The Review of Financial Studies 18*(2): 569–597, https://doi.org/10.1093/rfs/hhi007

Carhart, M. (1997). On Persistence in Mutual Fund Performance. The Journal of Finance, 52(1), 57-82. doi:10.2307/2329556

Chevalier, J. and Ellison, G. (1999), Are Some Mutual Fund Managers Better Than Others? Cross-Sectional Patterns in Behavior and Performance. The Journal of Finance, 54: 875-899. https://doi.org/10.1111/0022-1082.00130

Davis, J. (2001). Mutual Fund Performance and Manager Style. Financial Analysts Journal, 57(1), 19-27. Retrieved December 13, 2020, from http://www.jstor.org/stable/4480292

Desai, H. & Jain, P. (1997). Long-Run Common Stock Returns Following Stock Splits and Reverse Splits. The Journal of Business. 70. 409-33. 10.1086/209724.

Dharan, B., & Ikenberry, D. (1995). The Long-Run Negative Drift of Post-Listing Stock Returns. The Journal of Finance, 50(5), 1547-1574. doi:10.2307/2329326

Elton, E. J., Gruber, M. J. & Blake, C. R. (2001), A First Look at the Accuracy of the CRSP Mutual Fund Database and a Comparison of the CRSP and Morningstar Mutual Fund Databases. *The Journal of Finance, 56*(6): 2415-2430. https://doi.org/10.1111/0022-1082.00410

Evans, R. B. (2010), Mutual Fund Incubation. *The Journal of Finance*, *65*(4): 1581-1611. https://doi.org/10.1111/j.1540-6261.2010.01579.x

Fama, E. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, 25(2), 383-417. doi:10.2307/2325486

Fama, E. F. & French, K. R. (1993). Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics,* Volume 33, Issue 1,1993,Pages 3-56,ISSN 0304-405X,https://doi.org/10.1016/0304-405X(93)90023-5.(http://www.sciencedirect.com/science/article/pii/0304405X93900235)

Fama, E. F. & French, K. R. (2010). Luck versus skill in the cross section of mutual fund returns. *Journal of Finance, 2010 (65),* 1915–47.

Ferson, W., & Schadt, R. (1996). Measuring Fund Strategy and Performance in Changing Economic Conditions. The Journal of Finance, 51(2), 425-461. doi:10.2307/2329367

Frazzini, A. (2006). The Disposition Effect and Under-Reaction to News. Journal of Finance. 61. 2017-2046. 10.1111/j.1540-6261.2006.00896.x.

French, K. R. (2020a) *Fama/French 5 Factors (2x3) [Daily].* [Data file]. Retrieved 26.09.2020 from https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

French, K. R. (2020b) *Momentum factor (MOM) [Daily].* [Data file]. Retrieved 26.09.2020 from https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

Golden, T. D., Gajendran, R. S. (2019). Unpacking the Role of a Telecommuter's Job in Their Performance: Examining Job Complexity, Problem Solving, Interdependence, and Social Support. *J Bus Psychol* 34, 55–69 (2019). https://doi.org/10.1007/s10869-018-9530-4

Gorlick, A. (2020, March 30). The productivity pitfalls of working from home in the age of COVID-19. *Standford University*. Retrieved from: https://news.stanford.edu/2020/03/30/productivity-pitfalls-working-home-age-covid-

19/#:~:text=Nicholas%20Bloom%20is%20widely%20known,benefits%20of%20working%20from%20home.&text=The%20experiment%20revealed%20that%20working,drop%20in%20employee%2Dquit%20rates

Green, R. C. & Berk, J. B. (2002). Mutual Fund Flows and Performance in Rational Markets (December 9, 2002). Available at SSRN: https://ssrn.com/abstract=383061 or http://dx.doi.org/10.2139/ssrn.383061

Grinblatt, M., & Titman, S. (1989). Mutual Fund Performance: An Analysis of Quarterly Portfolio Holdings. The Journal of Business, 62(3), 393-416. Retrieved December 13, 2020, from http://www.jstor.org/stable/2353353

Grinblatt, M., & Titman, S. (1993). Performance Measurement without Benchmarks: An Examination of Mutual Fund Returns. The Journal of Business, 66(1), 47-68. Retrieved December 13, 2020, from http://www.jstor.org/stable/2353341

Hadden, J., Casado, L., Sonnemaker, T. & Borden, T. (2020, October 13). 20 major companies that have announced employees can work remotely long-term. *Business Insider*. Retrieved from: https://www.businessinsider.com/companies-asking-employees-to-work-from-home-due-to-coronavirus-2020?r=US&IR=T

Halko, M., Alanko, E. & Kaustia, M. (2012). The gender effect in risky asset holdings,Journal of Economic Behavior & Organization,Volume 83, Issue 1,2012,Pages 66-81,ISSN 0167-2681, https://doi.org/10.1016/j.jebo.2011.06.011.(http://www.sciencedirect.com/science/article/pii/S0167268111001569)

Imbens, G. M. & Wooldrigde J. M. (2007). *Difference-in-differences estimation.* [lecture notes]. National Bureau of Economic Research. http://www2.nber.org/WNE/lect_10_diffindiffs.pdf

Investment Company Institute. (2020, May 6). 2020 Investment Company Factbook. Retrieved from: https://www.ici.org/pdf/2020_factbook.pdf

Jensen, M. C. (1968): “The Performance of Mutual Funds in the Period 1945–1964,” Journal of Finance, 23, 389–416.

Jha, R., Korkie, B., Turtle, H. J. (2009) Measuring performance in a dynamic world: Conditional mean–variance fundamentals,Journal of Banking & Finance,Volume 33, Issue 10,2009,Pages 1851-1859,ISSN 0378-4266, https://doi.org/10.1016/j.jbankfin.2009.04.007.(http://www.sciencedirect.com/science/article/pii/S0378426609000818)

Lott, Y. (2014). Working time flexibility and autonomy: Facilitating time adequacy? A European perspective. *WSI Discussion paper.*

McKinsey & Company. (2020, July 15). COVID-19 and gender equality: Countering the regressive effects. Retrieved from: https://www.mckinsey.com/featured-insights/future-of-work/covid-19-and-gender-equality-countering-the-regressive-

McKinsey & Lean In. (2020, October 1). Women in the Workplace. Retrieved from: https://wiw-report.s3.amazonaws.com/Women_in_the_Workplace_2020.pdf

Mervosh, S., Lu, D & Swales, V. (n.d.). See Which States and Cities Have Told Residents to Stay at Home. *The New York Times.* Retrieved from: https://www.nytimes.com/interactive/2020/us/coronavirus-stay-at-home-order.html

Modern Fertility & SoFi. (2020). Modern State of Fertility 2020: Career & Money. Retrieved from: https://modernfertility.com/modern-state-fertility-2020-sofi-career-money#takeaway-1)

Morningstar. (2020, March 2). Women in Investing: Morningstar's View. Retrieved from: https://www.morningstar.com/articles/967691/women-in-investing-morningstars-view

Morningstar Direct [Online]. (September 17, 2020). Available: Morningstar.

Moskowitz, T. J. (2000), Discussion. *The Journal of Finance, 2002 (55)*. 1695-1703. https://doi.org/10.1111/0022-1082.00264

Mullen, L., Blevins, C. & Schmidt, B. (May 15, 2020). Package 'gender'. [Fact sheet]. R-projects. https://cran.r-project.org/web/packages/gender/gender.pdf

Niessen-Ruenzi, A. & Ruenzi, S. (2019). Sex Matters: Gender Bias in the Mutual Fund Industry. *Management Science, 65(7),* 3001-3025. https://doi.org/10.1287/mnsc.2017.2939

Olden, A & Moen, J. (2020) (April 22, 2020). The Triple Difference Estimator. *NHH Dept. of Business and Management Science Discussion Paper No. 2020/1*, http://dx.doi.org/10.2139/ssrn.3582447

Pastor, L., Stambaugh, R. F. & Taylor, L. A. (2015), Scale and skill in active management. *Journal of Financial Economics 116*(1): 23-45. https://doi.org/10.1016/j.jfineco.2014.11.008

Pischke, J. S. (2005). Empirical Methods in Applied Economics Lecture Notes. [lecture notes]. London School of Economics. Retrieved from http://econ.lse.ac.uk/staff/spischke/ec524/evaluation3.pdf

Prather, L., Bertin, W. J., Henker, T. (2004). Mutual fund characteristics, managerial attributes, and fund performance, Review of Financial Economics, Volume 13, Issue 4, 2004, Pages 305-326, ISSN 1058-3300, https://doi.org/10.1016/j.rfe.2003.11.002.

Rogers, B. (2020, August 26). Not in the same boat: Career progression in the pandemic. Retrieved from: https://www.qualtrics.com/blog/inequitable-effects-of-pandemic-on-careers/

Remote Year Inc. (2020, n.d.). What is Remote Work? Retrieved from: https://www.remoteyear.com/blog/what-is-remote-work

Sargis, M. & Lutton, L. P. (2016). Fund managers by gender The Global landscape. Morningstar Research. https://www.morningstar.com/lp/fund-managers-by-gender-the-global-landscape

Society for Financial Studies. (2020, n.d.). Call for Proposals on "Discrimination, Disparities, and Diversity in Finance". Retrieved from: http://sfs.org/wp-content/uploads/2020/10/RCFS-Registered-Reports-2021.pdf

Strumpf, E. (2011). Medicaid's Effect on Single Women's Labor Supply: Evidence from the Introduction of Medicaid. *Journal of Health Economics 30*(3), 2011, https://doi.org/10.1016/j.jhealeco.2011.02.002

Xiao, H. (2020). Does Working from Home Decrease Profitability and Productivity? Evidence from the Mutual Fund Industry. Available at SSRN: https://ssrn.com/abstract=3743993

World Health Organization. (2020, October 1). *The best time to prevent the next pandemic is now: countries join voices for better emergency preparedness.* Retrieved from: https://www.who.int/news/item/01-10-2020-the-best-time-to-prevent-the-next-pandemic-is-now-countries-join-voices-for-better-emergency-preparedness

Appendix

Table 11: Dates of stay-at-home order announcements across states

A: States with stay-at-home-orders			
State	Effective date	Effective time (local)	Modified effective date
Alabama	April 4	5 p.m.	April 5
Alaska	March 28	5 p.m.	March 29
Arizona	March 31	5 p.m.	April 1
California	March 19	5:30 p.m.	March 20
Colorado	March 26	6 a.m.	March 26
Connecticut	March 23	8 p.m.	March 24
Delaware	March 24	8 a.m.	March 24
District of Columb	April 1	12:01 a.m.	April 1
Florida	April 3	12:01 a.m.	April 3
Georgia	April 3	6 p.m.	April 4
Hawaii	March 25	12:01 a.m.	March 25
Idaho	March 25	1:30 p.m.	March 26
Illinois	March 21	5 p.m.	March 22
Indiana	March 24	11:59 p.m.	March 25
Kansas	March 30	12:01 a.m.	March 30
Kentucky	March 26	8 p.m.	March 27
Louisiana	March 23	5 p.m.	March 24
Maine	April 2	12:01 a.m.	April 2
Maryland	March 30	8 p.m.	March 31
Massachusetts	March 24	12 p.m.	March 25
Michigan	March 24	12:01 a.m.	March 24
Minnesota	March 27	11:59 p.m.	March 28
Mississippi	April 3	5 p.m.	April 4
Missouri	April 6	12:01 a.m.	April 6
Montana	March 28	12:01 a.m.	March 28
Nevada	April 1	10 a.m.	April 2
New Hampshire	March 27	11:59 p.m.	March 28
New Jersey	March 21	9 p.m.	March 22
New Mexico	March 24	8 a.m.	March 24
New York	March 22	8 p.m.	March 23
North Carolina	March 30	5 p.m.	March 31
Ohio	March 23	11:59 p.m.	March 24
Oregon	March 23	10:30 a.m.	March 24
Pennsylvania	April 1	8 p.m.	April 2
Rhode Island	March 28	3:30 p.m.	March 29
South Carolina	April 7	5 p.m.	April 8
Tennessee	March 31	11:59 p.m.	April 1
Texas	April 2	12:01 a.m.	April 2
Vermont	March 25	5 p.m.	March 26
Virginia	March 30	5:30 p.m.	March 31
Washington	March 23	5 p.m.	March 24
West Virginia	March 24	8 p.m.	March 25
Wisconsin	March 25	8 a.m.	March 25
B: States without a stay-at-home order			
Arkansas	-	-	-
Iowa	-	-	-
Nebraska	-	-	-
North Dakota	-	-	-
Oklahoma	-	-	-
South Dakota	-	-	-
Utah	-	-	-
Wyoming	-	-	-

Note: States announced stay-at-home orders at different times. The effective date presents the date where the stay-at-home orders went into

Table 12: Unconditional regressions factor models' alphas for the whole sample period.

	(1)	(2)	(3)	(4)
	α_{CAPM}	α_{FF3}	$\alpha_{Carhart\,4}$	α_{FF5}
Female	0.0041	-0.0039*	-0.0040*	-0.0031
	(0.0043)	(0.0022)	(0.0023)	(0.0022)
Fund Size	6.39e-13***	2.09e-13***	1.86e-13***	1.80e-13***
	(2.35e-13)	(3.96e-14)	(3.65e-14)	(4.39e-14)
Expense ratio	-0.0135**	0.00778	0.00803	0.00361
	(0.00652)	(0.00498)	(0.00496)	(0.00426)
Age of fund	-0.0135**	0.0078	0.0080	0.0036
	(0.0065)	(0.0050)	(0.0050)	(0.0043)
Tenure (longest)	0.0007***	0.0002***	0.0002***	0.0002***
	(0.0002)	(0.0001)	(0.0001)	(0.0001)
Team	0.0000	0.0001	0.0001	0.0001
	(0.0003)	(0.0002)	(0.0002)	(0.0002)
Constant	-0.0122*	-0.0058**	-0.0055**	-0.0049**
	(0.0072)	(0.0022)	(0.0021)	(0.0021)
N	167090	167090	167090	167090
R^2	0.138	0.054	0.053	0.057
State FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes

Note: The table presents the regression results from using the factor models' alphas as dependent variables in a unconditional version of equation 1 – the post term is dropped. Hence, the independent variable presents the difference between the female funds and the male funds. All columns include state and time fixed effects. The robust standard errors are clustered the state level, and are shown in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 13: Outcome regressions using the funds' Fama-French 3-factor model alphas

	α_{FF3}		
	(1) Male	(2) Female	(3) DDD
Post*Female			-0.0027
			(0.0051)
Female			0.0030
			(0.0036)
Post	-0.0015	-0.0042	-0.0015
	(0.0029)	(0.0069)	(0.0029)
Fund Size	2.25e-13**	3.03e-13***	2.84e-13***
	(9.08e-14)	(3.09e-14)	(4.26e-14)
Expense ratio	0.0073	0.0085	0.0078
	(0.0051)	(0.0064)	(0.0050)
Age of fund	0.0001*	0.0003	0.0002***
	(0.0001)	(0.0002)	(0.0001)
Tenure (longest)	0.0001	0.0001	0.0001
	(0.0002)	(0.0002)	(0.0002)
Team	-0.0059**	-0.0060	-0.0059**
	(0.0025)	(0.0062)	(0.0023)
Constant	-0.0622***	-0.0246**	-0.0632***
	(0.0085)	(0.0102)	(0.0079)
N	123008	44082	167090
R^2	0.050	0.100	0.064
State FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
State*Female	No	No	Yes
Time*Female	No	No	Yes

Note: The table presents the outcome regression results with daily Fama-French 5-factor model alphas as the dependent variable. Column 1 and 2 presents the difference-in-differences regression using equation 1 with respectively the male and the female subgroups of funds. Colum 3 displays the difference-in-differences-in-differences regression using equation 4. Post is the independent variable of interest in column 1 and 2, while Post*Female is the independent variable of interest of the DDD regression in column 3. All columns include state and time fixed effects, while column 3 adds state*female and time*female fixed effects. The robust standard errors are clustered at state level, and are shown in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 14: Outcome regressions using the funds' Carhart 4-factor model alphas

	$\alpha_{Carhart\ 4}$		
	(1) Male	(2) Female	(3) DDD
Post*Female			-0.0001
			(0.0042)
Female			0.0034
			(0.0035)
Post	-0.0024	-0.0026	-0.0024
	(0.0032)	(0.0056)	(0.0032)
Fund Size	1.86e-13**	2.99e-13***	2.68e-13***
	(8.39e-14)	(2.98e-14)	(4.18e-14)
Expense ratio	0.0076	0.0084	0.0080
	(0.0050)	(0.0065)	(0.0049)
Age of fund	0.0002*	0.0003	0.0002***
	(0.0001)	(0.0002)	(0.0001)
Tenure (longest)	0.0001	0.0001	0.0001
	(0.0002)	(0.0002)	(0.0002)
Team	-0.0056**	-0.0058	-0.0056**
	(0.0025)	(0.0061)	(0.0022)
Constant	-0.0637***	-0.0225**	-0.0645***
	(0.0081)	(0.0105)	(0.0077)
N	123008	44082	167090
R^2	0.047	0.102	0.062
State FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
State*Female	No	No	Yes
Time*Female	No	No	Yes

Note: The table presents the outcome regression results with daily Carhart 4-factor model alphas as the dependent variable. Column 1 and 2 presents the difference-in-differences regression using equation 1 with respectively the male and the female subgroups of funds. Colum 3 displays the difference-in-differences-in-differences regression using equation 4. Post is the independent variable of interest in column 1 and 2, while Post*Female is the independent variable of interest of the DDD regression in column 3. All columns include state and time fixed effects, while column 3 adds state*female and time*female fixed effects. The robust standard errors are clustered at state level, and are shown in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 15: Outcome regressions using the funds' Fama-French 5-factor alphas

	α_{FF5}		
	(1) Male	(2) Female	(3) DDD
Post*Female			-0.0030
			(0.0055)
Female			0.0020
			(0.0035)
Post	-0.0008	-0.0038	-0.0008
	(0.0028)	(0.0070)	(0.0028)
Fund Size	1.79e-13	2.86e-13***	2.56e-13***
	(1.14e-13)	(2.88e-14)	(4.85e-14)
Expense ratio	0.0026	0.0057	0.0035
	(0.0044)	(0.0060)	(0.0042)
Age of fund	0.0002**	0.0003*	0.0002***
	(0.0001)	(0.0002)	(0.0001)
Tenure (longest)	0.0001	0.0001	0.0001
	(0.0003)	(0.0001)	(0.0002)
Team	-0.0052**	-0.0042	-0.0050**
	(0.0024)	(0.0059)	(0.0022)
Constant	-0.0594***	-0.0231**	-0.0609***
	(0.0074)	(0.0108)	(0.0071)
N	123008	44082	167090
R^2	0.049	0.106	0.065
State FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
State*Female	No	No	Yes
Time*Female	No	No	Yes

Note: The table presents the outcome regression results with daily Fama-French 5-factor model alphas as the dependent variable. Column 1 and 2 presents the difference-in-differences regression using equation 1 with respectively the male and the female subgroups of funds. Colum 3 displays the difference-in-differences-in-differences regression using equation 4. Post is the independent variable of interest in column 1 and 2, while Post*Female is the independent variable of interest of the DDD regression in column 3. All columns include state and time fixed effects, while column 3 adds state*female and time*female fixed effects. The robust standard errors are clustered at state level, and are shown in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 16: Outcome regressions using the funds' net excess returns

	$R_{i,t} - R_{CRSP,t}$		
	(1)	(2)	(3)
	Male	Female	DDD
Post*Female			-0.0415
			(0.0379)
Female			0.0143
			(0.0216)
Post	0.0251	-0.0164	0.0251
	(0.0234)	(0.0278)	(0.0234)
Fund Size	1.65e-12***	1.17e-12***	1.23e-13***
	(6.23e-13)	(1.63e-13)	(3.17e-13)
Expense ratio	-0.0179	-0.0094	-0.0164
	(0.0121)	(0.0122)	(0.0105)
Age of fund	0.0007***	0.0015*	0.0009***
	(0.0003)	(0.0007)	(0.0003)
Tenure (longest)	0.0003	-0.0014**	-0.0001
	(0.0005)	(0.0007)	(0.0004)
Team	-0.0175	-0.0122	-0.0170
	(0.0132)	(0.0242)	(0.0119)
Constant	-0.3664***	-0.1938***	-0.3661***
	(0.0230)	(0.0335)	(0.0218)
N	123008	44082	167090
R^2	0.194	0.228	0.202
State FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
State*Female	No	No	Yes
Time*Female	No	No	Yes

Note: The table presents the outcome regression results with net excess returns as the dependent variable. Column 1 and 2 presents the difference-in-differences regression using equation 1 with respectively the male and the female subgroups of funds. Colum 3 displays the difference-in-differences-in-differences regression using equation 4. Post is the independent variable of interest in column 1 and 2, while Post*Female is the independent variable of interest of the DDD regression in column 3. All columns include state and time fixed effects, while column 3 adds state*female and time*female fixed effects. The robust standard errors are clustered at state level, and are shown in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 17: DDD-regression using the funds' factor model alphas without controls

	(1)	(2)	(3)	(4)
	α_{CAPM}	α_{FF3}	$\alpha_{Carhart\ 4}$	α_{FF5}
Post*Female	-0.0023	-0.0027	-0.0001	-0.0030
	(0.0045)	(0.0051)	(0.0042)	(0.0055)
Female	0.0059*	0.0003	0.0006	0.0001
	(0.0029)	(0.0028)	(0.0028)	(0.0027)
Post	0.0010	-0.0015	-0.0024	-0.0008
	(0.0037)	(0.0029)	(0.0032)	(0.0028)
Constant	-0.1014***	-0.0544***	-0.0550***	-0.0570***
	(0.0022)	(0.0012)	(0.0013)	(0.0010)
N	167090	167090	167090	167090
R^2	0.133	0.055	0.054	0.058
State FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
State*Female	Yes	Yes	Yes	Yes
Time*Female	Yes	Yes	Yes	Yes

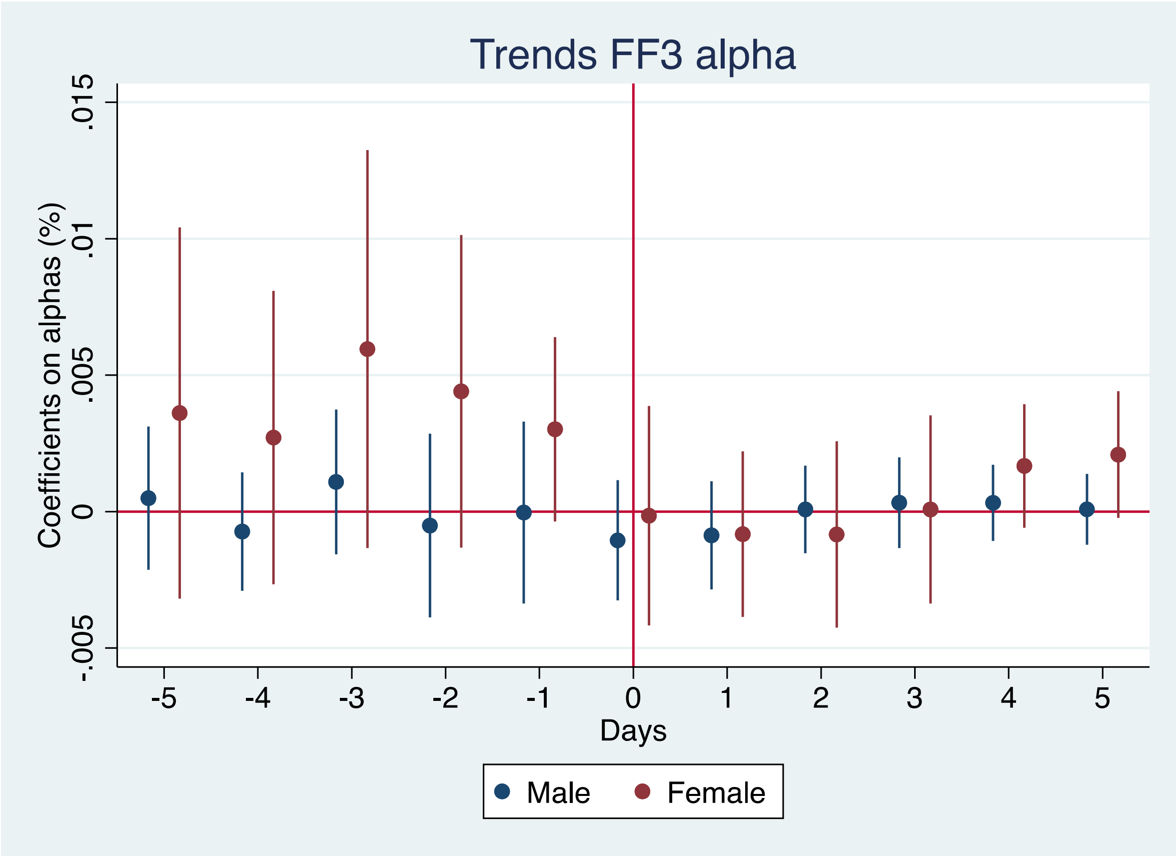
Note: The table presents the regression results from using the factor models' alphas as dependent variables in a equation 4, but the controls are excluded. All columns include state and time, state*female and time*female fixed effects. The robust standard errors are clustered at state level, and are shown in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 18: Unconditional regressions using the 1-R^2 for the whole sample period.

	(1)	(2)
	$1 - R^2_{CAPM}$	$1 - R^2_{Carhart\ 4}$
Female	0.0037	0.0093
	(0.0086)	(0.0083)
Fund Size	-7.74e-13***	-3.89e-13**
	(1.56e-13)	(1.48e-13)
Expense ratio	0.1229***	0.1072***
	(0.0152)	(0.0121)
Age of fund	-0.0012***	-0.0012***
	(0.0003)	(0.0003)
Tenure (longest)	0.0001	-0.0005
	(0.0004)	(0.0003)
Team	-0.0226*	-0.0187*
	(0.0133)	(0.0102)
Constant	0.2622***	0.2087***
	(0.0265)	(0.0193)
N	167090	167090
R^2	0.349	0.271
State FE	Yes	Yes
Time FE	Yes	Yes

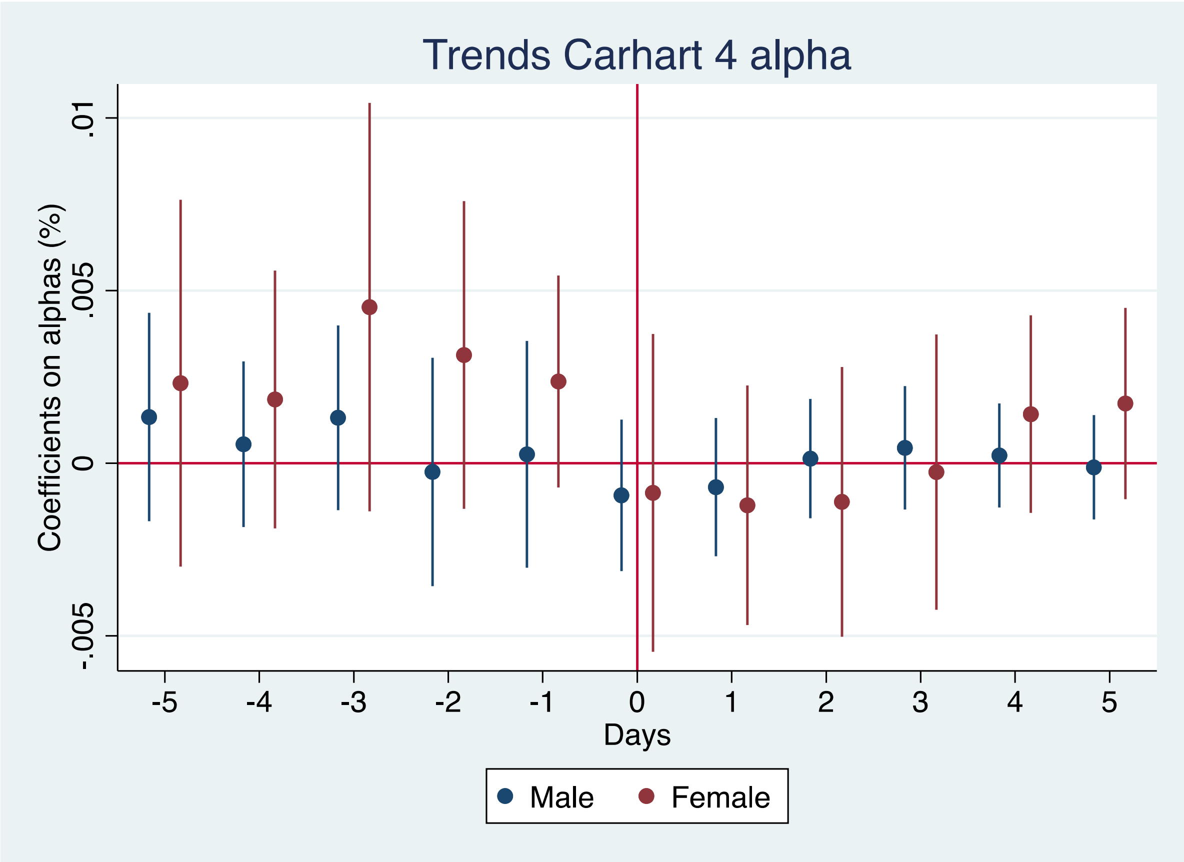
Note: The table presents the regression results from using the factor models' 1-R^2 as dependent variables in a unconditional version of equation 1 – the post term is dropped. Hence, the independent variable presents the difference between the female funds and the male funds. All columns include state and time fixed effects. The robust standard errors are clustered at state level, and are shown in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure 6: Leads and lags around the stay-at-home orders for the Fama-French 3-factor model



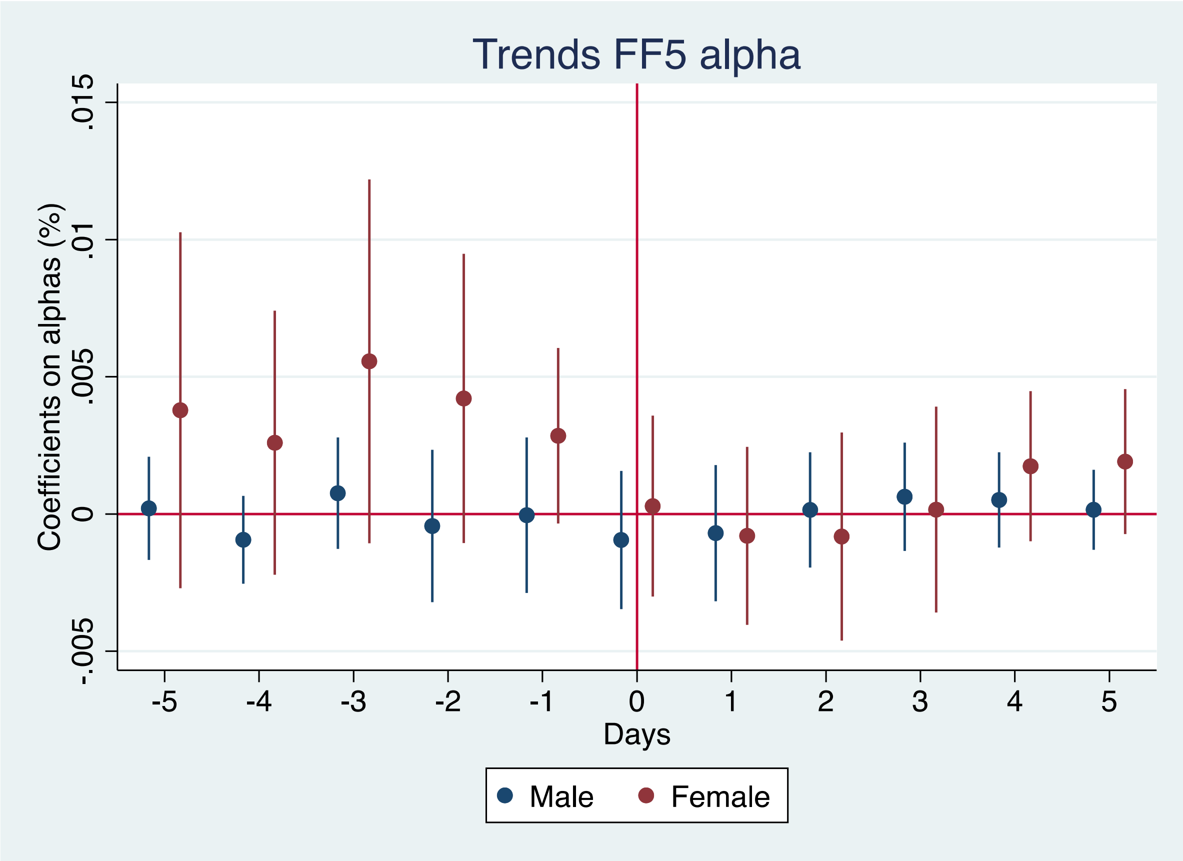
Note: The figure presents the coefficients of leads and lags around the stay-at-home order announcements, using equation 11. The main interest is that leads should be close to zero, or the ratio between male and female should be parallel for the leads.

Figure 7: Leads and lags around the stay-at-home orders for the Carhart 4-factor model



Note: The figure presents the coefficients of leads and lags around the stay-at-home order announcements, using equation 11. The main interest is that leads should be close to zero, or the ratio between male and female should be parallel for the leads.

Figure 8: Leads and lags around the stay-at-home orders for the Fama-French 5-factor model



Note: The figure presents the coefficients of leads and lags around the stay-at-home order announcements, using equation 11. The main interest is that leads should be close to zero, or the ratio between male and female should be parallel for the leads.

Table 19: Regression output from leads and lags using alphas.

	α_{CAPM}		α_{FF3}		$\alpha_{Carhart\ 4}$		α_{FF5}	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Male	Female	Male	Female	Male	Female	Male	Female
Lead (-5)	0.000131	0.00271	0.000492	0.00361	0.00134	0.00232	0.000205	0.00378
	(0.00119)	(0.00251)	(0.00130)	(0.00333)	(0.00149)	(0.00260)	(0.000929)	(0.00317)
Lead (-4)	0.000253	0.00341*	-0.000730	0.00271	0.000549	0.00185	-0.000940	0.00260
	(0.00107)	(0.00199)	(0.00107)	(0.00263)	(0.00119)	(0.00182)	(0.000793)	(0.00235)
Lead (-3)	-0.000483	0.00267	0.00109	0.00596	0.00132	0.00452	0.000759	0.00556*
	(0.000701)	(0.00241)	(0.00131)	(0.00357)	(0.00132)	(0.00289)	(0.00100)	(0.00324)
Lead (-2)	-0.000725	0.00170	-0.000511	0.00441	-0.000253	0.00313	-0.000438	0.00421
	(0.000848)	(0.00129)	(0.00167)	(0.00280)	(0.00163)	(0.00218)	(0.00137)	(0.00258)
Lead (-1)	-0.000348	0.00117	-0.0000330	0.00302*	0.000257	0.00237	-0.0000420	0.00285*
	(0.000795)	(0.00112)	(0.00165)	(0.00165)	(0.00162)	(0.00150)	(0.00140)	(0.00156)
Orders (0)	-0.000237	-0.000618	-0.00105	-0.000149	-0.000931	-0.000860	-0.000946	0.000288
	(0.000793)	(0.00184)	(0.00109)	(0.00196)	(0.00108)	(0.00225)	(0.00125)	(0.00161)
Lag (1)	-0.000138	-0.00121	-0.000870	-0.000827	-0.000695	-0.00122	-0.000696	-0.000795
	(0.000774)	(0.00156)	(0.000981)	(0.00148)	(0.000989)	(0.00170)	(0.00123)	(0.00159)
Lag (2)	0.000289	-0.00177	0.0000812	-0.000836	0.000132	-0.00112	0.000150	-0.000821
	(0.000628)	(0.00169)	(0.000794)	(0.00167)	(0.000854)	(0.00191)	(0.00104)	(0.00185)
Lag (3)	0.000602	-0.00108	0.000326	0.0000785	0.000444	-0.000256	0.000628	0.000160
	(0.000587)	(0.00165)	(0.000823)	(0.00169)	(0.000883)	(0.00195)	(0.000975)	(0.00183)
Lag (4)	0.000914	0.000863	0.000322	0.00167	0.000224	0.00142	0.000512	0.00174
	(0.000637)	(0.000917)	(0.000691)	(0.00111)	(0.000745)	(0.00140)	(0.000859)	(0.00134)
Lag (5)	0.000671	0.00170**	0.0000837	0.00209*	-0.000121	0.00173	0.000156	0.00191
	(0.000635)	(0.000777)	(0.000642)	(0.00114)	(0.000746)	(0.00135)	(0.000721)	(0.00129)
N	103168	36972	103168	36972	103168	36972	103168	36972
R^2	0.137	0.180	0.053	0.107	0.051	0.111	0.053	0.112
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: The table presents the regression results from equation 10. The columns presents the male and female subgroups and the four factor models' alphas as dependent variables. All columns include state, time, state*female and time*female fixed effects and all controls. The robust standard errors are clustered at state level, and are shown in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 20: Regression output from leads and lags using 1-R^2.

	$1-R^2_{CAPM}$		$1-R^2_{Carhart\,4}$	
	(1) Male	(2) Female	(5) Male	(6) Female
Lead (-5)	-0.0000	-0.0015	-0.0005	-0.0019
	(0.0011)	(0.0012)	(0.0006)	(0.0017)
Lead (-4)	0.0000	0.0003	-0.0002	-0.0003
	(0.0012)	(0.0007)	(0.0007)	(0.0014)
Lead (-3)	-0.0006	-0.0016*	-0.0005	-0.0022
	(0.0010)	(0.0009)	(0.0006)	(0.0021)
Lead (-2)	-0.0011	-0.0022**	-0.0009	-0.0032
	(0.0009)	(0.0010)	(0.0005)	(0.0023)
Lead (-1)	-0.0008	-0.0023	-0.0007	-0.0031
	(0.0009)	(0.0014)	(0.0006)	(0.0027)
Orders (0)	-0.0002	-0.0013	-0.0001	-0.0011
	(0.0008)	(0.0010)	(0.0004)	(0.0011)
Lag (1)	-0.0001	0.0000	-0.0000	0.0004
	(0.0008)	(0.0004)	(0.0004)	(0.0004)
Lag (2)	0.0004	-0.0001	0.0004	0.0003
	(0.0008)	(0.0004)	(0.0004)	(0.0003)
Lag (3)	0.0001	-0.0008*	0.0001	-0.0004
	(0.0007)	(0.0004)	(0.0004)	(0.0005)
Lag (4)	0.0000	-0.0004	0.0000	-0.0001
	(0.0007)	(0.0003)	(0.0004)	(0.0003)
Lag (5)	-0.0002	-0.0006*	0.0000	-0.0006
	(0.0006)	(0.0003)	(0.0004)	(0.0005)
N	103168	36972	103168	36972
R^2	0.354	0.355	0.276	0.283
Controls	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes

Note: The table presents the regression results from equation 10. The columns presents the male and female subgroups and the 1-R^2 as dependent variables. All columns include state, time, state*female and time*female fixed effects and all controls. The robust standard errors are clustered at state level, and are shown in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$