Data Science

Capstone Report - Fall 2023

The Impact of the COVID-19 Pandemic on Chinese Mutual Funds Investors

Kexin Deng

supervised by
Yiqing Lü

Declaration

I declare that this senior capstone was composed entirely by myself with the guidance of my advisor, and that it has not been submitted, in whole or in part, to any other application for a degree. Except where it is acknowledged through reference or citation, the work presented in this capstone is entirely my own.

Preface

The impact of COVID-19 spans various aspects of our lives, and its influence on investment and consumption behavior, particularly for individual investors, remains to be seen by the public with limited prior research. However, understanding these dynamics is crucial for addressing the current pandemic and enhancing decision-making in the face of future shocks. In this data science capstone project, I explore consumption and investment responses during the COVID-19 pandemic and their profound influence on the behavior of Chinese mutual fund investors. This research aims to address the following research questions. Do investors' consumption and reinvestment responses differ before and after the pandemic? The question can be broken down to each type of response, like netflow, turnover, rebalancing, timing, etc. Do these responses also vary across investors of different wealth levels before and after the pandemic, and considering city-wise pandemic severity? Using retail investors' data on the Alipay platform, I connect the investment and consumption from a micro-level investor perspective that unravels the connections between financial literacy and investor behavior, considering geographical and temporal dimensions. This project makes up for the gap in the current literature landscape. It is the first to deal with such a question; thus, the results are meaningful to Alipay, policy-makers, and the broader academic and industry realm.

Acknowledgements

I could not have completed this thesis without the help of my professors and fellow research assistants in the Alipay & New York University Joint Research program. Professor Yiqing Lü dedicated her time to helping students understand research methods and processes and how to design advanced models to analyze large amounts of data. Professor Guodong Chen, who led me to the program, was available for advice to any of us who asked, and I appreciate his endless support for the program to the furthest extent. I would also like to extend my deep gratitude to the capstone thesis mentor, Professor Li Guo, for her detailed guidance on rigorous academic research, including writing the proposal, literature review, presentation, and paper writing. I would also like to appreciate my previous peers in this program, namely Runming, Lydia, and Yuexin, and my current peers, Ziming and Erin. This thesis would not have been possible without their support and encouragement. Lastly, I would like to thank my friends and family for being with me through this project.

Abstract

This research explores the impact of the COVID-19 pandemic on Chinese mutual fund investors across different wealth levels, utilizing data from the Alipay AntFin Open Research Laboratory and employing two-way fixed effects and difference-in-differences analyses. Wealthier investors were observed cashing out and reducing called investments while maintaining higher overall investment levels post-pandemic. The study delves into changes in dispositional effect, portfolio rebalancing, timing, and alpha risk. City-level severity measurements suggest increased mutual fund turnover and net flow in more severe pandemic situations, favoring stock and mixed funds in more severe cities while less favoring money market funds. The research suggests that extreme health events can prompt investor behavior shifts beyond the traditional diagrams, offering a Chinese mutual fund perspective and advocating for the consideration of inequality in financial market assessments during COVID-19 and similar health crises.

Keywords

Pre and Post-COVID-19 Pandemic, Psychology of Stock Market, Mutual Fund Investors; Market Reaction; Consumption; Capital Gains; Turnover; Netflow

Contents

1 Introduction

Following several rounds of financial turmoil, it is becoming a common observation both in academia and in industry that inequality often grows in the wake of crises. The pandemic affected over 219 countries and claimed the lives of 1.7 million people worldwide by the end of 2020 [1]. It also significantly affected global capital markets and exacerbated existing social inequality concerns. In 2020, an estimated 49 million individuals could have been plunged into extreme poverty due to COVID-19, as indicated in a report on the World Bank blog [2]. Unlike other crises like the Great Financial Crisis (GFC) of 2007–2009, the COVID-19 pandemic is unique for being exogenous, uncertain, and global [3] [Borio 2020]. Being exogenous means, it did not emerge from the usual financial imbalances that have triggered recessions since the 1980s. The COVID-19 pandemic also displays genuine uncertainty, where a wide range of potential outcomes hinges on unpredictable non-economic factors. Furthermore, it is a genuinely global phenomenon compared to previous financial crises, which did not impact many countries like those in Asia.

Studying China's response to COVID-19 is particularly significant due to unique policies, including early lockdown and quarantine measures, the Dynamic zero-COVID strategy, and a phased reopening of provinces. These policies likely produced varied effects on consumption and investments across income levels. The United Nations Conference on Trade and Development(UNCTAD) pointed out that China's implementation of a diverse set of fiscal, monetary, and financial measures, such as dedicated funds for epidemic response materials, reduced taxes and fees for affected industries, enhanced market liquidity, and deferred loan repayments for small enterprises, may impact individuals differently based on their wealth levels [4] [Zhang et al. 2022]. These implementations underscore the importance of studying Chinese mutual fund investors to see how these policies influenced their investment and consumption patterns. Such research design contributes to a better understanding of inequality dynamics before and after pandemics.

With this background in mind, this capstone aims to examine how the COVID-19 pandemic influenced the behavior of Chinese mutual fund investors across different wealth level. The capstone has obtained detailed data from the Ant Research Repository, comprising two hundred thousand investors. This data is extensive and of exceptionally high quality, making it a valuable resource to support my research. I focus on critical metrics such as investor decisions, including turnover, netflow, capital gains, and consumption decisions like sensitivity. Fundamental questions of this research revolve around the following: (1) Do investors' consumption and reinvestment responses differ before and after the pandemic? The question can be broken down to each type of response, like netflow, turnover, rebalancing, etc. (2) Do these responses also vary across investors of different wealth levels/ holdings and city-level pandemic severity before and after the pandemic? By answering these questions, the research will understand how these investors responded to shifts in capital gains and dividend distributions, delving into factors like changes in investment performance (alpha) and risk-taking (turnover) while uncovering the underlying mechanisms related to financial constraints, financial literacy, and investor behavior, taking into account geographical and temporal dimensions. Ultimately, this project uses micro-level evidence to shed light on potential macroeconomic implications, offering insights for scholars, industry leaders, and policymakers to assess the financial market's reactions to the COVID-19 pandemic and similar public health safety events.

The remainder of this paper proceeds as follows. Section 2 introduces the related work. Section 3 conducts the main empirical research and analysis, providing the solution. Section IV offers results and discussions. Section 5 is attributed to discussions, and finally, Section 6 concludes.

2 Related Work

2.1 COVID-19 Pandemic and Financial Market Impact

The COVID-19 pandemic, often described as a black swan event, came as a surprise yet profoundly influenced global financial markets. It disrupted conventional economic models and market behaviors, exemplifying the unpredictability of the financial investment arena. In regards to the epidemiology model, Antras et al. [5] [2020] contributed a novel human interactions model incorporating elements of business travel and trade, drawing from both the gravity equation and the Susceptible-Infected-Recovered model (SIR). Scholars like Lai [6][2020] have revealed that non-pharmaceutical interventions can achieve rapid impact in pandemic control; however, may introduce delays that, rather than just diminishing, also prolong the battle. These two foundational works of epidemiology and social policy view together construct a background for our study related to COVID-19 that highlights the extensive and persistence of risk.

In terms of real data concerning consumption in China, CEIC Databytes show a sharp decline in year-on-year consumer goods sales, dropping from approximately 8% in December 2019 to -1% in March 2020, followed by a gradual recovery to 4.6% by the end of 2020. Annual retail sales of consumer goods decreased by 3.9% compared to 2019 and the catering industry was most severely impacted with a sharp income decline of 16.6% in 2020 [7] [CEIE Databyte 2021]. These statistics are instrumental in understanding cautious consumer behavior at the micro-level, a key aspect of my capstone project. Moving to the academic realm, Eichenbaum's research highlights that individuals during the pandemic tend to reduce consumption and work, contributing to the mitigation of epidemics. Still, this reduction often leads to economic recessions [8] [Eichenbaum 2021]. Similarly, Galiani's findings indicate that pandemics have a dual impact on supply and demand, hindering production in non-essential sectors and reducing consumer income [9][Galiani 2021]. These studies, which provide valuable insights into global trends and sector-specific impacts, serve as important reference points. However, my capstone's contribution lies in its emphasis on implications within China's unique context.

Moving to investment during the COVID era in China, CEIC Insights suggests that the yield curve for Chinese government bonds and corporate bonds exhibited a V-shaped pattern. In 2020, yields consistently reached their lowest levels from early April to late May, with government bond yields falling to their lowest points since the 2009 financial crisis. Under the influence of central bank stimulus policies, yields for bonds with various maturities steadily increased and eventually returned to pre-pandemic levels[10]. In addition to industry data, academic research also provides valuable insights into investment dynamics during the pandemic. Al-Awadhi et al. [2020] investigated the influence of the pandemic on the stock market, revealing a significant negative correlation between stock returns and the growth of confirmed COVID-19 cases and deaths [11]. The pandemic's effect on the stock market is further exacerbated by extensive media coverage, leading to a sentiment shock and heightened market volatility, as discussed by Haroon and Rizvii [12][2020]. These findings shed light on the interconnectedness of health crises, investor sentiments, and financial markets.

2.2 COVID-19 Pandemic and Financial Market Impact

The most important literature to this capstone is *Capital in the Twenty-First Century* by Thomas Piketty, which focuses on wealth and income inequality in Europe and the United States since the 18th century. Piketty's central message emphasizes how capital endowments can exacerbate income inequality, a theme closely tied to this capstone. His work suggests that when capital returns outstrip GDP growth, wealth accumulates faster for the affluent, surpassing the rate at which the working class generates income through labor. This unequal distribution of wealth causes social and economic instability [13] [Piketty 2014]. The macro-level insights provided by Piketty's work also have implications for this capstone in understanding the micro-level behaviors of investors. This raises the question of whether similar patterns can be observed among individuals with varying levels of capital accumulation in terms of their investment and consumption.

Expanding from Piketty'sPiketty's work to various sources on inequalities, the belief that "COVID-19 does not discriminate" is a dangerous myth, as it overlooks the increased vulnerability of the socially and economically deprived [14] [Patel 2020]. Like Patel's work, my capstone examines often overlooked disparities and vulnerabilities linked to the pandemic. The economic and social ramifications of COVID-19 are extensive and intricately connected to different topics of inequity, such as disparities in income, spending, and investing. Examining the income inequality's impact on COVID-19 in a global context, scholars surprisingly found that despite wealthier countries having better healthcare and public health systems, higher Gross National Income (GNI) per capita leads to higher COVID-19 mortality rates [15] [Ferreira et al. 2021]. It was also discovered during COVID-19, international income inequality decreased on a country-by-country basis, yet population-weighted inequality increased in line with the initial expectation [16] [Deaton 2021].

Regarding consumption inequality, economic downturns have lasting effects on consumers. Individuals experiencing high unemployment remain pessimistic about their future financial prospects, resulting in reduced spending on food and overall consumption in subsequent years. This effect is particularly pronounced among younger cohorts when controlled for income, wealth, and employment status. [17] [Malmendier and Shen 2018]. My capstone builds upon and complements the existing literature on Chinese mutual fund consumers by examining how the pandemic has introduced additional complexities that could exacerbate wealth disparities and give rise to uncertain stock-driven inequality.

2.3 Mutual Fund Investing and Investor Demographics

Transitioning to a more detailed exploration of mutual fund investing and its intersection with investor demographics, I shall first focus on uncovering mutual funds. Mutual funds are financial vehicles that pool money from multiple investors and use that money to purchase a diversified portfolio of stocks, bonds, or other securities. In our dataset, the fund types include stock, bond, gold, mixed funds, monetary funds, Index funds, and DQII funds. Since mutual funds are overseen by fund managers, an area of mutual fund literature is centered around agency problems. Classical literature modeling the flow-performance relationship highlights agency conflicts between mutual fund companies and investors. Investors expect fund companies to use their judgment and maximize risk-adjusted returns, while fund companies aim to maximize profits and attract investments. If these interests do not align, inefficiencies can arise [18][Chevalier and Ellison 1997]. Another essential piece of literature employs an active portfolio management model to study abnormal patterns associated with rational beliefs based on past performance. This model provides insights into the relationship between fund inflows and performance, where performance is indeed correlated with the average skill level of fund managers. It's noteworthy that there is significant heterogeneity among these managers, which can contribute to variations in fund performance [18][Berk and Green 2004]. While mutual fund managers undoubtedly play a crucial role, this capstone will shift its focus to investor heterogeneity rather than that of mutual fund managers. Additionally, including mutual fund manager literature contributes to the discussion on inequality within this capstone because individual investors' access and ability to identify skilled fund managers are also constrained by their wealth levels.

Now that mutual fund investing has been explained, this capstone shifts its focus from average effects to a comprehensive exploration of heterogeneity. Notably, one heterogeneity aspect involves studying institutional investor and corporate responses to the COVID-19 pandemic. Ataullah, Le, and Wood [2022] found that COVID-19 influences firms' dividend and share buyback decisions via institutional investors. Firms with substantial institutional ownership actively collaborate with management. Thus, managers prioritize retaining earnings over distributing dividends to mitigate pandemic-related uncertainties and promote long-term growth [19]. This research relates to the capstone in that individual retail investors are less favored in companies with fewer institutional investors. These individual investors have limited access to opportunities and are disadvantaged compared to their counterparts.

Another dimension of heterogeneity revolves around demographic factors. Asri Nur Wahyuni and Yanti Puji Astuti [2021] studied investment decisions during the COVID-19 pandemic in Indonesia; it was revealed that certain demographic factors play a significant role in shaping these choices. Specifically, age and education were found to be influential factors. Conversely, demographic variables such as gender, income, occupation, marital status, and prior investment experience had no significant impact on investment decisions during the pandemic [20]. This literature piece serves as a reference for my research, as I will also be investigating the behavior of Chinese mutual fund investors. I will consider age and education as crucial factors and then compare the results obtained from Chinese data to those obtained among investors in Indonesia.

2.4 Investor Theory and Financial Literacy

In trading theory regarding individual investors, Grossman and Stiglitz [21] [1980] propose that within a rational expectation framework, investors will trade when the marginal benefit equals or exceeds the marginal trade cost. Conversely, Odean [22] [1998], Gervais and Odean [23] [1998] develop theoretical financial market models wherein investors exhibit overconfidence that leads to detrimental decisions. Based upon the two, B. Barber and T. Odean [2000] found that individual investors holding common stocks trade actively and result in poor performance while households exhibit a preference for investing in small, high-beta stocks, with a less apparent inclination towards value stocks [24]. More recently, researchers examined the impact of the 2007-2008 financial crisis on portfolio decisions. They found that enduring shocks and challenges, such as financial setbacks, diminished investment returns, or extended periods of joblessness, consistently lead to an increased inclination toward risk aversion [25] [Andersen et al. 2019]. These pieces of literature contribute to my hypothesis, indicating that overconfidence and gambling incentives may be more pronounced in less wealthy investors. When combined with potentially poor timing, these factors can result in differences in transaction abilities.

3 Solution

In this solution section, I have to note the overall structure before diving into the interesting discoveries I made. Section 3.1 offers a data overview and explanation of variable construction. Starting from 3.2, the following six subsections are all independent specifications and offer their independent conclusion that contribute to the overall conclusions I could draw for the research. Each subsection has its group of formulas and coding files expanding over 20 specifications, and the line plots are my production, with their points being coefficients that I derived from numerous regressions. I should also note that the table is not simply derived from a regression but combined with one or multiple coding files. Each section is an independent specification, highlighting the nature of comprehensiveness and the huge workload of this research.

3.1 Data and Variable Construction

This project obtains data across 200,000 investors from 2019 to 2020 from Alipay, including report month, fund code, account ID, fund holding amounts, capital gains, dividends, and investors' monthly consumption amounts. This project applies descriptive statistics, two-way-fixed-effect models, and first-difference models to address the research questions. The data processing and model building are conducted using Python. Data were analyzed at an aggregate level, and no participants were contacted.

3.2 Turnover and Flow Patterns

I calculated metrics on a monthly basis: turnover, unit net flow, unit inflow, unit outflow, netflow, purchase, redemption, and netflow percentage change by grouping ID and report month, aggregating at the monthly level. Turnover represents the weight of the trading activity relative to the total asset value of the portfolio. In this context, unit net flow, unit inflow, and unit outflow are all scaled proportionally based on the fund amount. Metrics such as purchases (inflow), redemptions (outflow), and net flow are portfolio characteristics not subject to scaling.

$$Turnover_{i,t} = \frac{Purchase\_amt_{i,t} + Redemption\_amt_{i,t}}{Fund\_amt_{i,t}} \tag{1}$$

$$Unit\ Net\ Flow_{i,t} = \frac{Purchase\_amt_{i,t} - Redemption\_amt_{i,t}}{Fund\_amt_{i,t}} \tag{2}$$

$$Unit\ Inflow_{i,t} = \frac{Purchase\_amt_{i,t}}{Fund\_amt_{i,t}} \tag{3}$$

$$Unit\ Outflow_{i,t} = \frac{Redemption\_amt_{i,t}}{Fund\_amt_{i,t}} \tag{4}$$

$$Netflow_{i,t} = Purchase\_amt_{i,t} - Redemption\_amt_{i,t} \tag{5}$$

Winsorization for reducing the impact of extreme observations is also applied to the panel data at the 0.01 level. In addition to metrics related to financial turnover or flow patterns, I aim to investigate how the pandemic influences turnover and its associated flow responses at a microstructural level with the introduction of a dummy variable. The report months, including and preceding January 2020, are labeled as pre-pandemic with a value of 0. The remaining period, from February 2020 to the end of 2020, is considered post-pandemic with a value of 1.

I then conducted the baseline model of the following eight groups of regressions for the five incomes. The determination of 5 divisions of income levels is based on individuals' fund amounts in the first quarter of 2019. The process involves sorting the data in ascending order and dividing it into five quartiles, thus delineating five discrete income levels during the segmentation procedure. This segmentation is grounded in numerous attempts from projects that preceded this capstone.

$$y_{i,t} = \alpha_i + \beta_1 Pandemic\_Dummy\_amt_{i,t} + e_{i,t} \tag{6}$$

Dependent variables y include Turnover, Unit net flow, Unit inflow, Unit outflow, Netflow, Purchase, Redemption, and Netflow Pct Change. In each income level, the dependent variables for investor i in month t are $Turnover\_amt_{i,t}$, $Unit\ net\ flow\_amt_{i,t}$, $Unit\ inflow\_amt_{i,t}$, $Unit\ outflow\_amt_{i,t}$, $Netflow\_amt_{i,t}$, $Purchase\_amt_{i,t}$, $Redemption\_amt_{i,t}$, as well as $Netflow\ Pct\ Change\_amt_{i,t}$. Specifically, α_i represents the entity fixed effect. The Pandemic Dummy indicates whether it is pre- or post-pandemic in month t.

Table 1: Flows pattern related variables regressed on pandemic dummy for different wealth levels

	Wealth Level				
	Level 1 (1)	**Level 2** (2)	**Level 3** (3)	**Level 4** (4)	**Level 5** (5)
Panel A: Dependent variable - Turnover					
Pandemic Dummy	-0.2755***	0.1047***	0.1030***	0.1221***	0.1344***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
R^2	0.0010	0.0002	0.0002	0.0004	0.0005
Observations	885301	912390	916490	919818	926349
Panel B: Dependent Variable - Unit net flow					
Pandemic Dummy	0.2673***	-0.0009	-0.0093	-0.0282***	-0.0367***
	(0.0000)	(0.8728)	(0.0956)	(0.0000)	(0.00006)
R^2	0.0016	2.95e-08	3.188e-06	3.12e-05	6.072e-05
Observations	885301	912390	916490	919818	926349
Panel C: Dependent Variable - Unit inflow					
Pandemic Dummy	0.0283***	0.0481***	0.0434***	0.0412***	0.0402***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
R^2	0.0025	0.0103	0.0085	0.0080	0.0089
Observations	885301	912390	916490	919818	926349
Panel D: Dependent Variable - Unit outflow					
Pandemic Dummy	-0.2822***	0.0530***	0.0570***	0.0764***	0.0881***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
R^2	0.0013	7.24e-05	8.698e-05	0.0002	0.0003
Observations	885301	912390	916490	919818	926349
Panel E: Dependent Variable - Netflow					
Pandemic Dummy	140500***	131800***	143400***	163000***	201400***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
R^2	0.0027	0.0021	0.0016	0.0012	0.0008
Observations	885301	912390	916490	919818	926349

Panel F: Dependent Variable - Purchase (Inflow)					
Pandemic Dummy	536000***	584500	682600	839700***	1121000***
	(0.0000)	(0.8728)	(0.0956)	(0.0000)	(0.0000)
R^2	0.0315	0.0343	0.0317	0.0288	0.0195
Observations	885301	912390	916490	919818	926349
Panel G: Dependent Variable - Redemption (Outflow)					
Pandemic Dummy	389900***	448800***	535100***	674300***	869200***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
% R^2 Explained	0.0166	0.0196	0.0178	0.0156	0.0087
Observations	885301	912390	916490	919818	926349
Panel H: Dependent Variable - Netflow Pct Change					
Pandemic Dummy	7.7976***	-63.598***	-83.984***	-81.270***	-254.54***
	(0.7802)	(0.0323)	(0.0088)	(0.0791)	(0.0061)
R^2	1.061e-07	6.646e-06	8.111e-06	3.93e-06	1.015e-05
Observations	578007	630534	647432	657910	674565
Time Fixed Effect	NO	NO	NO	NO	NO
Entity Fixed Effect	YES	YES	YES	YES	YES

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

3.3 Disposition Effect, Capital Gains, and Sensitivity Shift

In this section, I'll explore the specification involving capital gains and delve into disposition effect and Sensitivity Shift.

$$y_{i,t,d} = \alpha_i, d + \gamma_t, d + \beta_1 Capital\_\, Gain_{i,t,d} + e_{i,t,d} \tag{7}$$

Dependent variables y include unit net flow, unit inflow, unit outflow, netflow, purchase, and redemption. To examine the interaction of flow patterns with capital gains, the flow patterns, including the respective dependent variables $Unit\ net\ flow\_\, amt_{i,t}$, $Unit\ inflow\_\, amt_{i,t}$, $Unit\ outflow\_\, amt_{i,t}$, $Netflow\_\, amt_{i,t}$, $Purchase\_\, amt_{i,t}$, $Redemption\_\, amt_{i,t}$ are regressed on capital gain for each wealth level. Ten regressions are conducted for each specification of dependent variables, totaling five income levels and two pandemic periods. For the specified models, both time fixed effects and entity fixed effects are included. Each dependent variable represents investor i in month t, belonging to pandemic period d. α_i, d represents the entity fixed effect, and γ_t, d represents the time fixed effect, where d takes values of 0 and 1.

Table 2: Flow pattern on capital gain For different wealth levels before and after Pandemic

	Before Pandemic					After Pandemic				
	Level 1 (1)	Level 2 (2)	Level 3 (3)	Level 4 (4)	Level 5 (5)	Level 1 (1)	Level 2 (2)	Level 3 (3)	Level 4 (4)	Level 5 (5)
Panel A: Dependent Variable - Unit netflow										
Cap Gain	1.9e-06***	-9.0e-07***	-4.9e-07***	-3.0e-07***	-5.3e-08***	2.5e-08***	5.5e-08***	5.8e-08***	4.7e-08***	4.0e-08***
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.1)	(0.0)	(0.0)	(0.0)	(0.0)
R^2	0.0017	0.0011	0.0007	0.0008	0.0002	6.8e-06	5.5e-05	9.8e-05	0.0001	0.0002
Panel B: Dependent Variable - Unit inflow										
Cap Gain	-1.0e-07***	-1.2e-07***	-9.8e-08***	-6.2e-08***	-1.9e-08***	-2.3e-08***	-2.5e-08***	-2.2e-08***	-1.8e-08***	-1.0e-08***
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
R^2	0.0009	0.0031	0.0042	0.0046	0.0036	0.0007	0.0014	0.0018	0.0021	0.0019
Panel C: Dependent Variable - Unit outflow										
Cap Gain	2.1e-06***	8.7e-07***	4.4e-07***	2.5e-07***	3.8e-08***	-6.8e-08***	-9.4e-08***	-9.4e-08***	-6.9e-08***	-5.7e-08***
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
R^2	0.0014	0.0008	0.0004	0.0004	7.8e-05	3.6e-05	0.0001	0.0002	0.0002	0.0003
Panel D: Dependent Variable - Netflow										
Cap Gain	-1.9662***	-1.2491***	-1.1549***	-0.9070***	-0.3738***	-0.6957***	-0.4781***	-0.4790***	-0.3943***	-0.3161***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
R^2	0.0261	0.0139	0.0132	0.0106	0.0048	0.0163	0.0090	0.0096	0.0071	0.0055
Panel E: Dependent Variable - Purchase										
Cap Gain	0.3661***	0.1808***	-0.0379***	-0.1884***	-0.2531***	0.0366***	0.0484***	-0.0130***	-0.0596***	-0.1521***
	(0.0004)	(0.0119)	(0.3757)	(0.0000)	(0.0000)	(0.1810)	(0.0230)	(0.4408)	(0.0000)	(0.0000)
R^2	0.0010	0.0003	1.721e-05	0.0006	0.0021	4.337e-05	9.485e-05	7.422e-06	0.0002	0.0011
Panel F: Dependent Variable - Redemption										
Cap Gain	2.6956***	1.5077***	1.2219***	0.8732***	0.2357***	0.8490***	0.6065***	0.5635***	0.4309***	0.2793***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
R^2	0.0504	0.0231	0.0163	0.0096	0.0012	0.0216	0.0130	0.0115	0.0067	0.0026
Observations	481764	499993	502310	504233	507888	403537	412397	414180	415585	418461
Time FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Entity FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

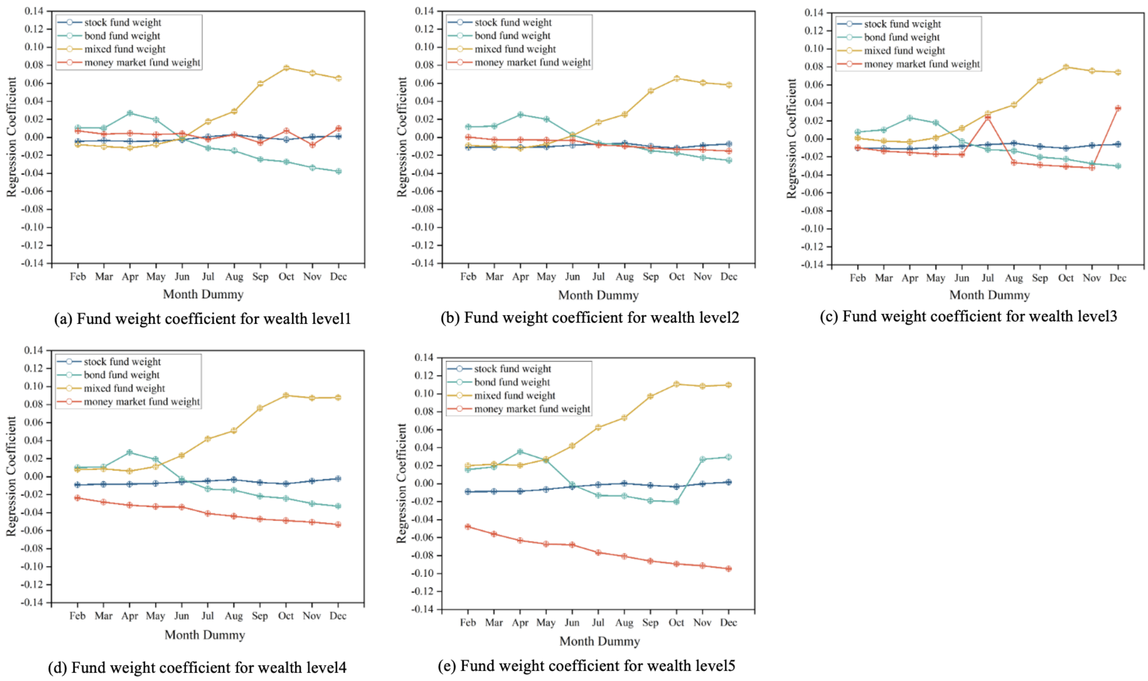
3.4 Portfolio Rebalancing and Timing Test

While metrics like weight may capture common time series variations in a fund's portfolio weight over the previous month, commony used activeness measure include active weight, which correspond to the senerio where fund managers respond to market changes by actively adjusting their asset allocation.

The regression equation for the analysis is as follows:

$$\begin{aligned} y_{i,t} = \alpha_i + \gamma_t &+ \beta_1 \text{Feb\_Dummy}_{i,t} + \beta_2 \text{Mar\_Dummy}_{i,t} + \beta_3 \text{Apr\_Dummy}_{i,t} \\ &+ \beta_4 \text{May\_Dummy}_{i,t} + \beta_5 \text{Jun\_Dummy}_{i,t} + \beta_6 \text{Jul\_Dummy}_{i,t} \\ &+ \beta_7 \text{Aug\_Dummy}_{i,t} + \beta_8 \text{Sep\_Dummy}_{i,t} + \beta_9 \text{Oct\_Dummy}_{i,t} \\ &+ \beta_{10} \text{Nov\_Dummy}_{i,t} + \beta_{11} \text{Dec\_Dummy}_{i,t} + e_{i,t} \end{aligned} \tag{8}$$

Dependent variables y include weight and active weight, respectively for stock fund, bond fund, gold, mixed fund, money market fund, index fund, and DQII fund. I also conducted other metrics like fund amount, weight of fund amount, fund unit netflow, netflow but due to space constraint, I've decided not to report. In the following figures, the regression coefficient is reported for the four fund types that have most prominent transition trend.



(a) Fund weight coefficient for wealth level1
(b) Fund weight coefficient for wealth level2
(c) Fund weight coefficient for wealth level3
(d) Fund weight coefficient for wealth level4
(e) Fund weight coefficient for wealth level5

Figure 1: Timing test of fund weight regressed on monthly dummy for different wealth levels

(a) Active fund weight coefficient for wealth level1

(b) Active fund weight coefficient for wealth level2

(c) Active fund weight coefficient for wealth level3

(d) Active fund weight coefficient for wealth level4

(e) Active fund weight coefficient for wealth level5

Figure 2: **Timing test of active fund weight regressed on monthly dummy for different wealth levels** These are timing tests with each month being dummy variables starting from February 2020 till December 2020. These tests revealed what I call the October effects, where active weights in asset allocation demonstrated a significant surge in the rebalancing towards mixed funds in October.

3.5 The CAPM Model and Alpha Risk

In this dedicated section, I will delve into the intricacies of the Capital Asset Pricing Model (CAPM), thoroughly examining its components such as alpha and beta risk and address any noteworthy shifts brought about by the pandemic. To start, let's look at the basic formula:

$$\begin{aligned} \mathrm{E}(r_i) &= \mathrm{rf} + \beta_m \times [E(r_m) - \mathrm{rf}] \\ &= \mathrm{Riskfree\_Rate} + \beta_m \times (\mathrm{Market\_Return} - \mathrm{Riskfree\_Rate}) \\ &= \mathrm{Riskfree\_Rate} + \beta \times \mathrm{Market\_Risk\_Premium} \end{aligned} \tag{9}$$

$$\begin{aligned} \alpha &= \mathrm{Rp} - \mathrm{rf} - \beta_m \times [E(R_m) - \mathrm{rf}] \\ &= \mathrm{Observed\_Return} - \mathrm{Riskfree\_Rate} - \beta \times \mathrm{Market\_Risk\_Premium} \end{aligned} \tag{10}$$

where Rp is the portfolio observed return, rf is the risk free rate, Rm is the expected market return. Usually, I use the interest rate of Short-Term National Debt as risk free rate. Beta Risk is denoted by β, it is the systematic risk of a combination of assets. Usually, the Beta Risk could decrease approximately to 0 through portfolio diversification in large corporations or investment banking. Alpha Risk is the return unrelated to the fluctuation of the market shock, denoted by α in equation. It could be calculated using observed real return minus market-fluctuation-related return. Alpha could imply the investment ability of investors.

The regression equation for the analysis is as follows:

$$\begin{aligned}\text{Investor\_Return}_{i,t} = \alpha_i + \beta_1 \text{Market\_Rf}_{i,t} + \beta_2 \text{Pandemic\_Dummy}_{i,t} \\ + \beta_3 \text{Market\_Rf}_{i,t} \times \text{Pandemic\_Dummy}_{i,t} + e_{i,t}\end{aligned} \tag{11}$$

For the specified models, only entity fixed effect is included. Each dependent variable represents investor i in month t. α_i, d represents the entity fixed effect.

Table 3: CAPM related alpha and beta risk regressions

	Wealth Level				
	Level 1	**Level 2**	**Level 3**	**Level 4**	**Level 5**
	(1)	(2)	(3)	(4)	(5)
Panel A: Dependent variable - Turnover					
Market Rf	0.1923^{***}	0.226^{***}	0.2197^{***}	0.2092^{***}	0.1848^{***}
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Pandemic Dummy	-0.0002^{***}	0.0001^{***}	7.68E-5***	9.78E-5***	0.0006^{***}
	(0.0102)	(0.0333)	(0.2219)	(0.1110)	(0.0000)
Market Rf * Pandemic Dummy	0.5351^{***}	0.5663^{***}	0.5597^{***}	0.5502^{***}	0.4711^{***}
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
R^2	0.3172	0.3928	0.3935	0.3920	0.3617
Observations	854888	882446	886912	888957	893888
Time Fixed Effect	NO	NO	NO	NO	NO
Entity Fixed Effect	YES	YES	YES	YES	YES

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

3.6 Pandemic Severity and Mutual Funds

With the aim of delving into the spatio-temporal influence, this research also introduced a measurement of pandemic severity at the city level, which I incorporated into the platform. The severity measure is calculated by taking the logarithm of confirmed COVID-19 cases reported by the end of 2020, divided by the total population data obtained from the China Statistical Yearbook for China and its various provinces.

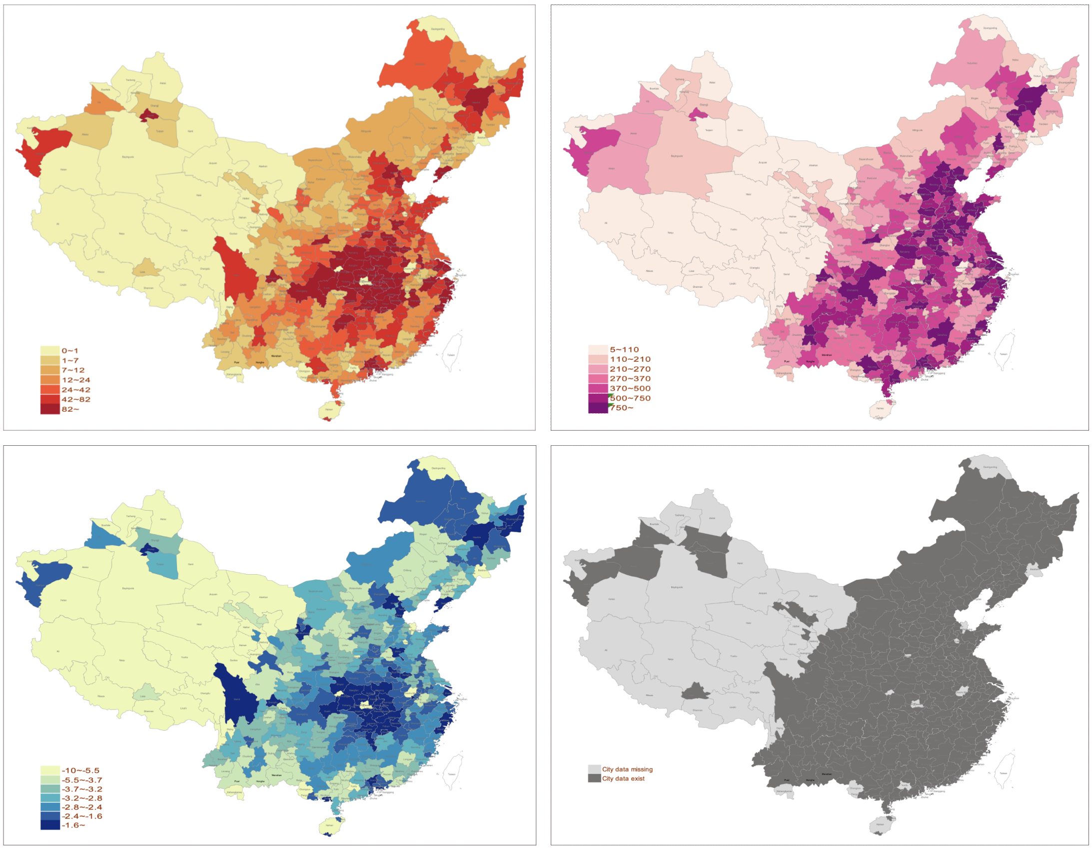


Figure 3: **visualization of the proposed severity measurement in my research** upper left corner is the graph of city level confirmed cases; upper right graph is city level population; lower left is the severity measure that I calculated and incorporated into my research, lower right cornor is whether the city data existence, if missing then may due to 0 in either confirmed cases or population data for that specific city.

Here the dependent variables are respectively turnover, unit net flow, and net flow, weights of different fund type, and active weights of different fund type. The calculation of weight and active weights are on user, month, fund level.

$$\begin{aligned} y_{i,t} = \alpha_i + \gamma_t + \beta_1 \text{Pandemic\_Dummy\_amt}_{i,t} + \beta_2 \text{Pandemic\_Severity}_{i,t} \\ + \beta_3 \text{Pandemic\_Dummy\_amt}_{i,t} \times \text{Pandemic\_Severity}_{i,t} + e_{i,t} \end{aligned} \tag{12}$$

Dependent variables y include turnover, unit net flow, net flow, weight, and active weight. For the specifications above, both time-fixed effect and entity-fixed effect are added. In terms of weight and active weight specifically, each dependent variable represents investor i in month t belonging to fund type x. α_i, x represents the entity fixed effect, and γ_t, x represents the time fixed effect where x belongs to which fund type the weight and active weights are calculating, like stock, bond, gold, mixed fund, money market fund, index fund, DQII fund.

Table 4: Pandemic severity related regressions

Panel A: Dependent variable being turnover, unit net flow and net flow			
	Turnover	Unit net flow	Net flow
Pand x Severity	0.0133***	-0.01 ***	9020.1 ***
	(0.0000)	(0.0000)	(0.0003)
R^2	0.000006351	0.000005941	0.000003163
Observations	4540610	4540610	4540610
Time fixed effects	YES	YES	YES
Entity fixed effects	YES	YES	YES

Panel B: Dependent variable being weights of different fund type							
	Stock	Bond	Gold	Mixed	Money	Index	DQII
Pand*Sev	0.0002*	0.0005***	-0.000033	0.0017***	-0.0016***	0.000037	-0.0008***
	(0.0271)	(0.0000)	(0.6422)	(0.0000)	(0.0000)	(0.7581)	(0.0000)
R^2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
N	4540370	4540370	4540370	4540370	4540370	4540370	4540370

Panel C: Dependent variable being active weights of different fund type							
	Stock	Bond	Gold	Mixed	Money	Index	DQII
Pand*Sev	0.0002	0.0002	0.000021	0.0006**	-0.0005***	-0.0002	-0.0008***
	(0.1984)	(0.2045)	(0.8135)	(0.0029)	(0.0000)	(0.3427)	(0.0000)
R^2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
N	4489509	4489509	4489509	4489509	4489509	4489509	4489509
TF	YES	YES	YES	YES	YES	YES	YES
EF	YES	YES	YES	YES	YES	YES	YES

3.7 Triple Interaction and Consumption Response

As I delve into the overall consumption response, my model incorporates a triple interaction, outlined in Equation (13), to capture relationship between consumption response and the triple interplay of pandemic severity, capital gain, and the pandemic dummy variable.

$$\begin{aligned} Consumption_{i,t} =& \alpha_i + \gamma_t + \beta_1 \text{Pandemic\_severity}_{i,t} \\ & + \beta_2 \text{Pandemic\_Severity}_{i,t} \times \text{Pandemic\_dummy}_{i,t} \\ & + \beta_3 \text{Capital\_gain}_{i,t} \times \text{Pandemic\_dummy}_{i,t} + \\ & + \beta_4 \text{Capital\_gain}_{i,t} \times \text{Pandemic\_severity}_{i,t} \\ & + \beta_5 \text{Pandemic\_Dummy}_{i,t} \times \text{Pandemic\_Severity}_{i,t} \times \text{Capital\_gain}_{i,t} \\ & + e_{i,t} \end{aligned} \tag{13}$$

For this specified models, both time and entity fixed effects are included. Each dependent variable represents investor i in month t. α_i represents the entity fixed effect, and γ_t represents the time fixed effect.

Table 5: Consumption, severity, capital gain, pandemic dummy, and interaction term

Dependent Variable	**Consumption**
Interaction term: severity x pandemic dummy x capital gain	0.0069
	(0.0000)
Interaction term: severity x capital gain	-0.0025
	(0.0027)
Interaction term: pandemic dummy x capital gain	0.0280
	(0.0000)
Interaction term: pandemic dummy x severity	-13430
	(0.0000)
R^2	0.0002
Observations	4564251
Time fixed effects	**YES**
Entity fixed effects	**YES**

4 Results and Discussions

4.1 Results

Turnover and Flow Patterns: Wealthier individuals have increased their turnover following the pandemic, resulting in a decrease in unit net flow. This decline is driven by a rise in both unit inflow and outflow, with outflow increasing more significantly, suggesting a trend of cashing out and reduced investments. However, despite the reduction in unit net flow, absolute inflow and outflow amounts have grown, indicating higher overall investment levels among wealthier individuals after the pandemic.

Disposition Effect, Capital Gains and Sensitivity Shift: The study also used panel regressions of flow patterns on capital gains for different wealth levels. The disposal effect means the higher the gain, the easier it is to sell. Yet, my research around redemption signifies the disposition effect reduced after the pandemic. I discovered a shift in sensitivity to market conditions among different wealth groups. Less wealthy groups tended to be momentum investors before the pandemic, buying when gains were up, while wealthier groups acted as contrarian investors, buying when gains were down. Interestingly, both tendencies were downgraded after the pandemic.

Portfolio Rebalancing and Timing Test: My research revealed that wealthier individuals shifted their portfolio weight from money market funds to mixed funds during the pandemic. I undertook timing tests with each month being dummy variables starting from February 2020 till December 2020. These tests revealed what I call the October effects, where active weights in asset allocation demonstrated a significant surge in the rebalancing towards mixed funds in October, indicating that fund managers may respond to market changes by actively adjusting their asset allocation.

The CAPM and Alpha Risk: In the context of the Capital Asset Pricing Model, I examined changes in alpha represented by the pandemic dummy and the beta with the market risk premium before and after the pandemic. The pandemic dummy reflects the changes in alpha post-pandemic, so the alpha for wealthier individuals increased.

Pandemic Severity and Mutual Funds: The research calculated a severity measure at the city level based on confirmed COVID-19 cases per population, and the confirmed cases taken by the end of 2020, which corresponds with the sample period. It found that turnover and netflow in mutual funds increased with pandemic severity, while unit netflow decreased. The severity of the pandemic had a significant impact on Chinese mutual fund investors, with many rebalancing different types of mutual funds. Stock and mixed funds appeared to be favored during severe pandemic conditions, while money market funds were less favored.

Triple Interaction and Consumption Response: For severity to increase by 1%, consumption's response to capital gain sensitivity increases by 0.0069 * 1%. This implies that in places where the pandemic was more severe, the response of consumption to capital gain became stronger.

4.2 Discussions

4.2.1 Challenges

This research addresses novel challenges not explored in previous studies, likely due to various reasons. Firstly, the critical data acquired from Alipay is essential for precise identification and analysis, posing potential challenges for replication by other researchers due to data constraints. Secondly, the research platform is elusive and time-consuming, with persistent issues like limited memory space and computational power. Lastly, the empirical nature of financial research requires substantial effort. To attain the results presented in this paper, I explored numerous specifications; some fell short of our expectations and are therefore not included here.

4.2.2 Limitation

This research has limitations and caveats. Firstly, despite a large sample size, there may be concerns about generalizability, as the data is sourced from Alipay, potentially not capturing all Chinese mutual fund investors. Secondly, the time frame is short-term, covering 2019 to 2020, and lacks data beyond that period. The brevity of the time interval means that in other research or from a more long-term view, limited context or significance can be drawn when attempting to make comparisons or establish connections with different points in time. Thirdly, the limitations extend to the scope of data coverage in the consumption subcategory, confined solely to Taobao. This restriction hinders the meaningfulness of category-wise analysis, as the exclusive focus on a single platform may not capture the entirety of consumer behavior and trends. Therefore, the categorical consumption requires more data to fully explore deeper implications and potential insights.

4.2.3 Future Work

So far, I have analyzed whether investors' consumption and reinvestment responses differ before and after the pandemic, including flow patterns, turnover, and rebalancing. I also examined whether these responses varied among investors with different wealth levels before and after the pandemic outbreak. The remaining work will explore more potential heterogeneous patterns across investor demographics, such as age and gender. Further work on consumption category would be also feasible as long as I secure another source of data entailing detailed consumption categories. I plan to continue addressing these tasks in the following months' service as a research assistant and continue my exploration of the topic.

5 Conclusion

The study reveals a significant shift in investment patterns among wealthier and less wealthy individuals following the pandemic. Wealthier individuals demonstrated a trend of cashing out and reducing called investments yet maintained higher overall investment levels. Research around redemption signifies that the disposition effect reduced after the pandemic. Less wealthy groups tended to be momentum investors before the pandemic, buying when gains were up, while wealthier groups acted as contrarian investors, buying when gains were down. Interestingly, both tendencies were downgraded after the pandemic. Rebalancing behaviors varied, with wealthier individuals shifting portfolio weight from the money market to mixed funds, and the shift was observed in October 2020. This indicates significant rebalancing towards mixed funds, suggesting active asset allocation adjustments by fund managers in response to market changes. The study extended the analysis to the CAPM model, unveiling an increased alpha risk for the return of the wealthier group post-pandemic. Heterogeneous findings highlight that city-level severity impacts mutual funds turnover and net flow, favoring stock and mixed funds in more severe pandemic conditions. Additionally, consumption response to capital gain was stronger in the face of a more severe pandemic. These findings contribute to a comprehensive understanding of the dynamic changes in investment behaviors and risk factors in the post-pandemic landscape.

In conclusion, the Data Science Capstone project provides a comprehensive analysis of the behavior of Chinese mutual fund investors during the COVID-19 pandemic. Other future directions can be explored, such as delving into the long-term impact once there is data beyond 2020. Additionally, more future work can be done once I obtain more accurate and reasonable categorical data in the consumption realm to make up the missing piece. Future research may even incorporate a third phase contrasting the period when restrictions of COVID-19 are lifted, shedding more light on more complicated responses of investors based on income levels, the impact of pandemic severity, and shifts in portfolio strategies.

References

[1] "Who coronavirus (covid-19) dashboard." World Health Organization, 2020. [Online]. Available: https://covid19.who.int/

[2] W. B. Blogs, "Covid-19 will hit the poor hardest. here's what we can do about it," https://blogs.worldbank.org/voices/covid-19-will-hit-poor-hardest-heres-what-we-can-do-about-it, 2020.

[3] C. Borio, "The covid-19 economic crisis: dangerously unique," *Business Economics*, p. 181–190, 2020. [Online]. Available: https://doi.org/10.1057/s11369-020-00184-2

[4] L. Zhang, F. Sun, and X. Chu, "China's policy experience in responding to covid-19 shock," *United Nations Conference on Trade and Development (UNCTAD)*, no. 4, 2022. [Online]. Available: https://unctad.org/system/files/official-document/BRI-Project_RP24_en.pdf

[5] P. Antràs, S. J. Redding, and E. R.-H. Antràs, "Globalization and pandemics." *American Economic Review*, vol. 113, no. 4, pp. 939–981, 2023. [Online]. Available: https://www.aeaweb.org/articles?id=10.1257/aer.20201479

[6] S. Lai, N. W. Ruktanonchai, L. Zhou, O. Prosper, W. Luo, J. R. Floyd, A. Wesolowski, M. Santillana, C. Zhang, X. Du, H. Yu, and A. J. Tatem, "Effect of non-pharmaceutical interventions to contain covid-19 in china." *Nature*, vol. 585, no. 7825, p. 410–413, 2020. [Online]. Available: https://doi.org/10.1038/s41586-020-2293-x

[7] C. Databytes, "China's retail sales of consumer goods decreased by 3.9

[8] M. S. Eichenbaum, S. Rebelo, and M. Trabandt, "The macroeconomics of epidemics." *The Review of Financial Studies*, vol. 34, p. 5149–5187, 2021. [Online]. Available: https://doi.org/10.1093/rfs/hhab040

[9] S. Galiani, "Pandemic economics." *Journal of Economic Behavior Organization*, vol. 193, p. 269–275, 2022. [Online]. Available: https://doi.org/10.1016/j.jebo.2021.10.016

[10] C. Insights, "The chinese bond market amid the impact of the covid-19 pandemic," 2021. [Online]. Available: https://info.ceicdata.com/zh-cn/china-bond-market-2020-12-cn

[11] A. M. Al-Awadhi, K. Alsaifi, A. Al-Awadhi, and S. Alhammadic, "Death and contagious infectious diseases: Impact of the covid-19 virus on stock market returns." *Journal of Behavioral and Experimental Finance*, 2020. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7144859/

[12] O. Haroon and S. A. R. Rizvi, "Covid-19: Media coverage and financial markets behavior—a sectoral inquiry," *Journal of Behavioral and Experimental Finance*, vol. 27, p. 100343, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2214635020301386

[13] T. Piketty, *Capital in the Twenty-First Century.* Belknap Press, 2017.

[14] J. Patel, F. Nielsen, A. Badiani, A. S., V. Unadkat, R. Ravindrane, and H. Wardle, "Poverty, inequality and covid-19: the forgotten vulnerable." *Public Health*, vol. 183, 2020. [Online]. Available: https://doi.org/10.31389/lseppr.26

[15] B. Decerf, F. H. G. Ferreira, D. G. Mahler, and O. Sterck, "Lives and livelihoods: Estimates of the global mortality and poverty effects of the covid-19 pandemic," Institute of Labor Economics (IZA), Bonn, IZA Discussion Paper 13549, 2020. [Online]. Available: https://covid-19.iza.org/publications/dp13549/

[16] A. Deaton, "Covid-19 and global income inequality." *LSE public policy review*, vol. 1, 2021. [Online]. Available: https://doi.org/10.31389/lseppr.26

[17] U. Malmendier and L. Shen, "Scarred consumption," National Bureau of Economic Research, Working Paper 24696, 2018. [Online]. Available: http://www.nber.org/papers/w24696

[18] J. Chevalier and G. Ellison, "Risk taking by mutual funds as a response to incentives," *Journal of Political Economy*, vol. 105, no. 6, pp. 1167–1200, 1997. [Online]. Available: http://www.jstor.org/stable/10.1086/516389

[19] A. Ataullah, H. Li, and G. Wood, "Common risk factors in the returns on stocks and bonds," *British Journal of Management*, vol. 33, pp. 634–656, 2022. [Online]. Available: https://doi.org/10.1111/1467-8551.12601

[20] A. N. Wahyuni and Y. P. Astuti, "The influence of demographic factors on investment decisions in indonesia in the pandemic time covid-19," in *Proceedings of the 3rd International Conference on Banking, Accounting, Management and Economics (ICOBAME 2020).* Atlantis Press, 2021, pp. 350–355. [Online]. Available: https://doi.org/10.2991/aebmr.k.210311.070

[21] S. J. Grossman and J. E. Stiglitz, "On the impossibility of informationally efficient markets," *The American Economic Review*, vol. 70, no. 3, pp. 393–408, 1980. [Online]. Available: https://www.jstor.org/stable/1805228

[22] T. Odean, "Volume, volatility, price, and profit when all traders are above average," *Journal of Finance*, vol. 53, no. 6, pp. 1887–1934, 1988. [Online]. Available: https://doi.org/10.1111/0022-1082.00078

[23] S. Gervais and T. Odean, "Learning to be overconfident," *The Review of Financial Studies*, vol. 14, no. 1, pp. 1–27, 2001. [Online]. Available: http://www.jstor.org/stable/2696755.

[24] B. M. Barber and T. Odean, "Trading is hazardous to your wealth: The common stock investment performance of individual investors," *Journal of Finance*, vol. 116, pp. 261–292, 2000. [Online]. Available: https://doi.org/10.1111/0022-1082.00226

[25] S. Andersen, T. Hanspal, and K. M. Nielsen, "Once bitten, twice shy: The power of personal experiences in risk taking." *Journal of Financial Economics*, vol. 3, pp. 97–117, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0304405X18303088