Essays on Asset Pricing Anomalies

A thesis presented in fulfilment of the requirement for the degree of

Doctor of Philosophy in Finance at Massey University, Albany, New Zealand.

Kaining Gu

School of Accountancy, Economics and Finance

2026

ABSTRACT

This thesis consists of three essays on stock return anomalies in the Chinese stock market, which is characterised by a high proportion of retail investors and relatively severe information frictions.

The Chinese stock market is dominated by retail investors, whose behavioural biases are likely to give rise to stock anomaly returns. The first essay investigates the presence of such anomalies and examines whether financial information transparency reduces anomaly returns and thus enhances market price efficiency. We find that, among the 11 anomalies identified by Stambaugh et al. (2012) and Chu et al. (2020), most are persistent in the Chinese stock market. These anomalies' returns are negatively related to information availability, proxied by a firm-level opacity measure. Notably, this effect is more pronounced in the long legs of the anomaly portfolio than in the short legs. The observed negative relationship between information opacity and anomaly returns suggests that improving firm-level information disclosure could help reduce anomaly returns and thereby enhance market price efficiency.

The second essay explores the effect of market constraints on stock anomalies. The relaxation of the Chinese short-sale ban started in March 2010, expanded in steps, provides a quasi-experimental setting to investigate the influence of short-sale constraints on stock anomalies. We employ a difference-in-differences (DiD) approach to compare anomaly returns between stocks eligible and stocks ineligible for short sales during each expansion, which refers to a regulatory wave in which additional stocks became eligible for short selling. The results indicate that relaxing the short-sale constraint has heterogeneous and limited effects on anomaly returns. Furthermore, the effect is not uniform across anomalies. To better capture the intensity of short-selling activities, we construct a short-selling depth variable that isolates net shorting pressure from the margin trading effect. We find a negative relationship between short-selling depth and anomaly returns, indicating that anomaly profitability weakens as short-selling intensity increases. This result echoes the argument of Chang et al. (2014), who advocate for the development of a more robust securities lending market to promote market efficiency.

The third essay examines the pricing implications of climate risk disclosure (CRD). We develop a firm-level CRD measure through textual analysis of corporate annual reports and use it to construct a novel climate-risk-disclosure-based anomaly factor. Both stock-level and portfolio-level analyses show that stocks with lower CRD scores consistently earn higher returns, and that the climate-risk-disclosure factor (F_{CRD}) generates significant abnormal returns relative to traditional risk factors and policy uncertainty factors. These findings provide strong evidence that a CRD-characteristic-based anomaly persists in the Chinese stock market, highlighting the return predictive power of climate-related disclosure.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to Professor Xiaoming Li and Dr. Mei Qiu for their long-standing guidance, support, and encouragement throughout my academic training, beginning from my Master's studies and continuing through this PhD. I have benefited greatly from their expertise, and from the time and attention they have devoted to my research. Professor Li's high standards, constructive feedback, and rigorous approach to scholarship shaped my development as a researcher, while his supervision combined intellectual discipline with patience and humour. Dr. Qiu's mentorship, generous encouragement, and support extended beyond academic matters; her advice and perspective were invaluable during the more demanding periods of the PhD, and I am grateful for both the confidence she placed in my work and the personal support she offered throughout this journey.

My thanks extend further, to the faculty and staff of the School of Economics and Finance at Massey University for providing a collegial and supportive environment for doctoral research. Appreciation is also due to the administrative team for their assistance with academic and administrative matters throughout the programme.

My PhD experience was made more meaningful by the companionship and mutual support of my colleagues and office mates. I especially thank my friends Yiming Yang, Yunzhe Zhu, and Xiaoyan Zhang for navigating the ups and downs of the PhD process together. I sincerely wish them a rewarding academic and professional future.

My gratitude also goes to my long-standing friends Xue Yu and Jenny Chao, whose continued friendship and support provided comfort and perspective beyond academic life.

I am also grateful to Professor Christoph Schumacher, whose conversations and insights broadened my horizons and exposed me to perspectives beyond my immediate research field. I appreciate the generosity and openness with which he shared his experiences and knowledge.

I owe special thanks to Dr. Qifang Feng, my partner, for his patience, understanding, and unwavering support throughout this journey. His encouragement and companionship were

essential during the demanding periods of the PhD, and I am deeply grateful for his presence and care.

I am profoundly grateful to my family for their unconditional love, support, and encouragement. Their belief in me made this achievement possible, and I am especially thankful for the sacrifices they made throughout my years of study. I would like to express particular gratitude to my father, Gu Tianjun, and my mother, Wang Dong, for their dedication and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER ONE: INTRODUCTION

This chapter outlines the three essays comprising this thesis. It highlights the specific research gap each essay aims to address, the underlying motivations, and the contributions to the literature. This chapter concludes by describing the structure of the remainder of this thesis.

1.1. Overview of the Thesis

Stock anomalies represent persistent patterns in asset returns that cannot be fully explained by traditional risk-based models. As such, they challenge the foundational assumptions of the efficient market hypothesis and thus the rationality of asset prices. These pricing irregularities, which are often driven by behavioural biases, limits to arbitrage, or information frictions, offer valuable insights into the mechanisms underlying mispricing. By examining stock anomalies, this thesis contributes to the broader asset pricing literature and provides a complementary perspective to existing studies on valuation errors and market inefficiencies.

This thesis is motivated by two prevalent themes in stock anomaly literature. The first theme concerns the importance of the market environment in shaping anomaly patterns. Prior studies in developed markets (Stambaugh et al., 2012; Chu et al., 2020) document the strong return predictive power of stock anomalies. The Chinese stock market, characterised by its distinctive investor composition and trading behaviour (Li & Wang, 2010; Ng & Wu, 2007; Peikun & Jing, 2010; Yu et al., 2019b), is more susceptible to the emergence and persistence of such anomalies. Meanwhile, improved corporate information disclosure enhances investors' decision-making efficiency (Deng & Xu, 2011; Lawrence, 2013), potentially weakening anomaly-driven returns. The first essay of the thesis extends the existing literature by examining stock anomalies and their causal relationship with information transparency within the context of a developing market.

The second essay further explores how market constraints influence stock anomalies. Short-selling constraints prevent pessimistic investors from reflecting negative information in prices, often leading to persistent overvaluation and higher arbitrage cost (Miller, 1977; Shleifer & Vishny, 1997). Stambaugh et al. (2012) find that the anomaly returns are concentrated in stocks with greater short-sale constraints, suggesting that anomaly returns are sustained when arbitrage remains constrained. In the Chinese context, the initial relaxation of short-selling restrictions began in March 2010. Empirical evidence shows that improved access to short sales enhances pricing efficiency (Boehmer & Wu 2013; Chang et al., 2015). However, the development of China's short-selling mechanism has been gradual, with eight distinct phases of expansion and limited market depth. Against this backdrop, the second essay examines how short-sale constraints shape anomaly returns in China, and whether the attenuation of anomaly profitability depends on both eligibility and the intensity of short-selling activity.

The second theme relates to sustainability-related anomalies originating from firms' annual report disclosures. Traditional firm characteristics, such as size and book-to-market ratios in the Fama–French model, or accruals in Hirshleifer et al. (2012), have been shown to predict cross-sectional stock returns, forming the foundation of the characteristic-based anomaly literature. In contrast, climate risk disclosure (CRD), as an emerging non-financial firm characteristic, remains underexplored in the asset pricing literature. Sustainable development in China is policy driven, reaching peak carbon emissions by 2030 and carbon neutrality achieved by 2060. In contrast with the mandatory disclosure in European markets, the voluntary disclosure of ESG-related information gives Chinese firms potential motivations to manage investors' attention towards their firm value. The third essay investigates whether CRD has return predictive power and yields significant abnormal returns as a novel anomaly.

The remainder of this chapter is structured as follows. Sections 1.2–1.4 summarize Essays 1–3 and highlight their contributions to the literature. Section 1.5 summarises the key research outputs. Section 1.6 presents the sequence of the rest of the thesis.

1.2. Essay One

The Chinese stock market is heavily dominated by retail investors, whose behavioural biases are likely to contribute to the emergence of stock return anomalies. The first essay investigates whether stock return anomalies persist, and assesses the extent to which enhanced financial disclosure can reduce these anomalies and promote more efficient pricing.

Stock anomalies often originate from pricing inefficiencies and the heterogeneous beliefs of market participants. These anomalies are particularly difficult to reconcile with standard asset pricing models, including the CAPM and the Fama–French three-factor model (Chen et al., 2011). Stambaugh et al. (2012) document the profits of anomalies in terms of long-short strategies. They find that higher investor sentiment is associated with greater profitability of these anomalies, primarily driven by the short legs of the strategies, while the long positions exhibit no significant relationship with sentiment. Using the same set of anomalies, Chu et al. (2020) shows that the introduction of short-selling reduces the returns of long–short anomaly portfolios. This reduction is driven entirely by the short legs of anomalies. However, due to institutional differences, these measures may not translate well to developing markets, where firm valuation is shaped by distinct market features. The relatively lower efficiency of trading systems and regulatory oversight in such markets may further enhance stock anomalies.

The Chinese stock market provides a unique environment for this study. Individual investors account for over 80% of total trading volume (Li & Wang, 2010; Yu et al., 2019). These investors are highly responsive to attention-grabbing events and prevailing market sentiment, thereby amplifying market volatility (Tian et al., 2018). Notably, retail investors in China exhibit strong herding behaviour (Peikun & Jing, 2010), which often leads to irrational trading activities. Such behaviour contributes to the persistence and profitability of stock anomalies.

Institutional investors can influence share prices through their trading activities, particularly when they possess insider information that is not available to the public. Such opacity increases information risk and leads investors to demand a higher risk premium, thereby hindering the efficient pricing of stocks (Barry & Brown, 1985). In contrast, greater transparency in corporate disclosures can mitigate information risk and enhance the quality of financial reporting, including earnings management practices (Hirst & Hopkins, 1998; Hunton et al., 2006). Overall, enhanced disclosure reduces information asymmetry, improves market liquidity, and lowers firms' cost of equity (Diamond & Verrecchia, 1991).

This essay examines the set of stock anomalies studied by Stambaugh et al. (2012) and Chu (2020) through long–short portfolio strategies. While these anomalies have demonstrated profitability in the U.S. stock market, they remain largely unexplored in the context of the

Chinese market, whose unique institutional features motivate this investigation. We find that most of these stock anomalies persist in the Chinese stock market. Traditional asset pricing factors cannot fully account for the return predictability associated with these anomalies. Importantly, we further explore the influence of information transparency on anomalies. To capture firm-level disclosure quality, we construct an "Opacity" variable that proxies the extent to which firms reveal financial information to the public. Our results show that greater financial transparency mitigates anomaly returns and contributes to more efficient pricing.

This essay makes two primary contributions to the literature on stock anomalies. First, the essay points out the importance of the distinct market environment in shaping the magnitude and persistence of stock anomalies in China. Second, it examines the role of information transparency on anomalies in explaining anomaly returns, which has not been documented in the extant literature.

1.3. Essay Two

Stock anomalies could arise from market inefficiency. For example, market constraints may prevent stock prices from fully incorporating available information. In particular, short-sale constraints restrict pessimistic investors from expressing their negative views, leading to overvaluation of stock prices (Miller, 1977; Diamond & Verrecchia, 1987). Chu et al. (2020) provide empirical evidence from the U.S. stock market, showing that the relaxation of short-sale constraints reduces anomaly-based long–short strategies, with the effect primarily driven by the short legs of the portfolios. This raises an important question: Does the same pattern hold in the context of the Chinese market?

The Chinese government started to gradually relax the short-selling ban in March 2010. Since then, the evolution of China's short-selling framework has progressed gradually, with eight distinct phases of expansion. Chang et al. (2014) provide early evidence that short sellers serve as information conduits in China, with their activity enhancing price efficiency by embedding negative information into stock prices. In contrast, Wan (2024) documents that the imbalanced trading of margin traders and short sellers deteriorates market price efficiency.

Using data from 2010 to 2024, we examine whether relaxing the short sale ban could contribute to reduce stock anomaly returns. Employing a difference-in-differences (DiD) framework, we find heterogeneous effects of short-sale relaxation on anomaly profitability. Nevertheless, Fama–French five-factor-controlled regression results suggest that the return predictive power of these anomalies persists, indicating that mispricing is not substantially corrected. We further construct a short-selling depth variable to capture the intensity of short-selling activities. We find that intensive short-selling activity helps reduce anomaly returns. Overall, the evidence suggests that short-sale eligibility alone does not systematically eliminate anomaly profitability; rather, the intensity of short-selling activity plays a more decisive role in mitigating mispricing. However, this effect is not uniform across all anomalies, highlighting the heterogeneous nature of anomaly performance under different market environments[1]. Moreover, our findings highlight the importance of short-selling depth in shaping the extent to which anomaly-based mispricing persists. This is consistent with the notion that limited short-selling activities, coupled with active margin trading, results in imbalanced market pressures that contribute to persistent mispricing.

This study contributes to the literature by comprehensively investigating the whole eight expansion waves of short sales' impact on anomaly returns over thirteen years in the Chinese market, thereby extending the limits-to-arbitrage literature to an emerging market context. Most literatures focus on only two or three short sale expansions (Chang et al., 2014; Chen & Yang, 2016; Li et al., 2018; Liu et al. 2020; Lv & Wu, 2020; etc). Second, the study introduces a novel transaction-based measure of short-selling depth, distinguishing between the extensive margin of short-sale eligibility and the intensive margin of actual short-selling activity. We argue that short sale depth plays the important role of diminishing anomaly profitability.

[1] Chu et al. (2020) find a uniform effect of short sale activities on the same panel of anomalies in the U.S. stock market, where releasing the short sale constraint reduces long-short anomaly returns.

1.4. Essay Three

Climate risk disclosure (CRD) is increasingly recognised as a vital aspect of firms' sustainability strategies, as climate-related physical and transition risks pose significant challenges to firm valuation (Bansal et al., 2017). Transparent CRD can reduce information asymmetry, allowing investors to better assess firms' long-term prospects. Moreover, CRD is often embedded within broad ESG frameworks, potentially diluting its signal and complicating investor interpretation (Chatterji et al., 2009; Duan et al., 2025). Overconfidence or misreading of disclosures may result in persistent mispricing and return anomalies in CRD-sorted portfolios.

This essay develops a firm-level measure of climate risk disclosure (CRD) based on textual analysis of annual reports from Chinese A-share firms. In this market, retail investors, who contribute over 80% of trading volume (Yu et al., 2019a), often exhibit limited information access and pronounced behavioural biases, which can amplify the pricing impact of disclosure quality. Firms with lower levels of CRD are likely perceived as riskier, prompting investors to demand a higher return premium. At the same time, China's evolving climate policy landscape, including the 2021 launch of a national emissions trading scheme and a sharp rise in voluntary ESG disclosures, creates a unique institutional environment. Unlike the mandatory frameworks in Europe or market-driven mechanisms in the U.S., China's approach is largely policy-guided yet relies on voluntary firm responses. This combination of regulatory structure and investor composition offers a compelling and underexplored setting to investigate the asset pricing implications of CRD in emerging markets.

Using our firm-level CRD measure based on text analysis, we first explore its distribution across industries and firm characteristics such as size, age, and profitability, and observe notable heterogeneity in climate-related disclosures. We then examine the return implications of CRD through stock-level and portfolio-level analyses. Fixed-effect regression analysis show that firms with lower CRD tend to yield higher returns, suggesting that investors demanding a premium for greater information risk. Complementing this result, we construct four zero-cost hedging portfolios (F_{CRD}) based on CRD deciles. Fama–MacBeth regressions confirm that these portfolios earn significantly positive returns, and (F_{CRD}) strongly predicts future returns in the cross-section of 25 size–B/M portfolios. Together, the results provide robust evidence of CRD-characteristic-based return anomalies in China.

This essay contributes to the literature in three aspects. First, this study develops a novel firm-level measure of climate risk disclosure (CRD) based on textual analysis of annual reports, capturing firms' climate transparency over time. Unlike prior ESG-based proxies, our CRD measure isolates climate-specific disclosure behaviour and extends the work of Lin and Wu (2023) by treating CRD as a firm characteristic relevant for asset pricing. Second, we offer the first systematic evidence that CRD has significant explanatory power for cross-sectional variation in stock and portfolio returns in an emerging market context. We document robust return anomalies associated with CRD, an area largely unexplored in the literature, particularly in China. Third, our findings enrich the growing literature on intangible firm characteristics and climate finance by proposing that CRD should be considered a new pricing anomaly. This suggests that the set of characteristic-sorted factor-mimicking portfolios (e.g., size, value, investment, profitability) should now be expanded to include climate-related disclosure dimensions.

1.5. Research Output of the Thesis

Essay Three

The third essay, "Climate Risk Disclosure and Return Anomalies: A Textual Analysis of Chinese A-Shares", was presented at both the 29th New Zealand Finance Colloquium 2025, University of Otago, New Zealand and the Massey Sustainable Finance Conference 2025, Nanjing, China.

1.6. The Structure of the Thesis

The remainder of the thesis is structured as follows. Chapter 2 presents the first essay, investigating the presence of stock return anomalies and examining their underlying causal relationship with firm-level information opacity. Chapter 3 presents the second essay, which focuses on the effect of short sale mechanisms on these anomalies. The third essay, on the return predictive power of climate-related anomalies, is presented in Chapter 4. Finally, Chapter 5 concludes the thesis by summarising the main findings and implications.

CHAPTER TWO: ESSAY ONE

Information Opacity and Asset Pricing Anomalies

This chapter presents the first essay, which examines whether firm-level information opacity conditions the profitability of firm characteristic-based anomalies in the Chinese A-share market. The chapter investigates whether anomalies documented in prior studies remain profitable in this emerging market and whether their returns are systematically related to firms' financial disclosure . It analyses how a firm-level opacity measure is associated with the long-leg, short-leg, and long–short returns of anomaly portfolios after controlling for standard risk factors.

2.1. Introduction

Asset pricing anomalies challenge the view that competitive capital markets fully and instantaneously incorporate all available information into prices. A large body of research shows that portfolios formed on firm characteristics – such as profitability, investment, accruals, issuance, financial distress, and past returns – earn returns that are difficult to reconcile with standard risk-based models. When firm information is transparent and cheaply available, informed traders can more easily arbitrage away such mispricing. When information is opaque, investors face higher information-processing costs and are more likely to disagree about fundamental value, so that mispricing can persist and anomaly-based strategies remain profitable. This essay examines whether firm-level information opacity helps explain the cross-sectional behaviour of characteristic-based return anomalies in China's A-share market.

According to the Efficient Market Hypothesis (EMH), security prices reflect all available market-level and firm-level information, so that stocks trade at fair values and abnormal profits cannot be earned systematically (Fama, 1970). In the classic argument of Friedman (1953), irrational traders who buy high and sell low will eventually lose wealth and exit the market, leaving prices to be set by rational arbitrageurs. The Capital Asset Pricing Model (CAPM) of Sharpe (1964) and Treynor (1965) formalises this view by linking an asset's expected excess return to its market beta, so that cross-sectional differences in mean returns should be explained by exposure to a single systematic risk factor. However, empirical evidence since the 1970s shows that expected returns are also related to non-beta characteristics, including size, book-to-market, dividend yield, and earnings-to-price ratios (Banz, 1981; Basu, 1977; Black & Scholes, 1974; Reinganum, 1981). These findings indicate that the CAPM is incomplete, and motivate multi-factor models and anomaly-based tests of market efficiency.

Fama and French (1993) extend the CAPM to a three-factor model that adds size and book-to-market factors and substantially improves the fit to the cross-section of average returns. Yet momentum strategies, which buy recent winners and sell recent losers, earn sizeable abnormal returns that are not captured by the three-factor model (Jegadeesh & Titman, 1993). Subsequent work documents additional patterns related to liquidity, trading frictions, profitability, investment, and financial distress that also generate apparent abnormal returns (e.g. Brennan et al., 1998; Campbell et al., 2008; Novy-Marx, 2013). In response, Fama and French (2015) introduce a five-factor model that augments the size and value factors with profitability and investment factors. While these factors absorb part of the cross-sectional variation in expected returns, many anomalies remain statistically and economically significant, and their robustness varies across markets and sample periods (Schwert, 2003).

A large strand of the anomaly literature focuses on characteristics constructed from financial statements and other firm-level information. Accrual-based anomalies show that stock prices fail to fully impound the implications of the accrual and cash flow components of earnings (Sloan, 1996; Xie, 2001). Asset-growth anomalies indicate that firms expanding their asset base or pursuing high investment rates tend to earn lower subsequent risk-adjusted returns, consistent with investor underreaction to overinvestment or empire building (Cooper et al., 2008; Fairfield et al., 2003; Fama & French, 2006, 2008; Titman et al., 2004). More profitable firms earn higher future returns even after controlling for traditional factors, and that simple measures such as gross profitability are highly predictive (Ball et al., 2015; Novy-Marx, 2013). Distress-related variables, including the O-score, downside risk measures and other failure-risk indicators, predict returns in ways that are difficult to reconcile with standard risk views because financially distressed stocks often earn low average returns despite being ex ante riskier (Agarwal & Taffler, 2008; Campbell et al., 2008; Lai et al., 2010; Ohlson, 1980). Momentum and reversal effects reinforce the idea that investors underreact to information in fundamentals and overreact to recent price moves.

These anomalies can be interpreted either as compensation for omitted risk factors or as evidence of mispricing driven by behavioural biases and limits to arbitrage. A key input in both explanations is the information environment faced by investors. Grossman and Stiglitz (1980) argue that when information acquisition is costly, prices cannot be perfectly informative in equilibrium because some investors must be rewarded for collecting and processing information. If disclosure is limited or complex, investors may rely more on coarse signals, sentiment, and attention-grabbing events, leading to pricing distortions (Baker & Wurgler, 2006; De Bondt & Thaler, 1985). Studies document that earnings announcements, IPOs, and other corporate events influence stock prices in ways not fully justified by fundamentals (Dontoh et al., 2007; Pastor & Veronesi, 2003), and that stock returns display high-frequency patterns such as monthly, weekly, or even intraday anomalies that correlate with trading frictions and investor mood (Bergsma & Jiang, 2016; Birru, 2018).

The corporate disclosure literature provides a complementary perspective by emphasising how transparency affects information asymmetry, market liquidity, and the cost of capital. Diamond and Verrecchia (1991) show that richer disclosure reduces information asymmetry between informed and uninformed traders and can reduce expected returns by compressing bid–ask spreads and enhancing market depth. Kim and Verrecchia (1994) highlight how disclosure quality influences trading volume and the speed of price discovery. Healy and Palepu (2001) provide evidence that more transparent reporting reduces information risk, improves the information environment for outside investors, and is associated with a lower cost of equity capital. Later work stresses that not only the distribution of information but also its overall accuracy matters for pricing (Lambert et al., 2012). In this view, higher-quality public information can reduce the need for costly private information acquisition, thereby facilitating more efficient risk sharing and narrowing the scope for mispricing.

Corporate transparency is particularly salient in emerging markets. Legal protections for investors are weaker, enforcement is less predictable, and financial intermediaries and information infrastructures are less developed than in mature markets (Ball et al., 2000; Bushman et al., 2004). For such markets, theory and evidence suggest that improving disclosure quality can help mitigate mispricing and promote financial stability by reducing information asymmetry and constraining insiders' ability to exploit less informed investors (Durnev & Kim, 2005). Firth et al. (2015) and related studies on China emphasise that reforms to accounting standards and disclosure regulation have gradually strengthened corporate reporting, yet considerable dispersion remains in disclosure practices and information quality across firms and over time.

China's A-share market offers a useful setting to study the interaction between information opacity and stock return anomalies. The market is dominated by individual investors who rely heavily on public information, media coverage and salient events when forming expectations (Deng & Xu, 2011; Li et al., 2017). Retail participation and attention-based trading amplify the influence of disclosure practices on prices and contribute to pronounced herding and sentiment-driven fluctuations (Lin et al., 2020). Recent evidence shows that a wide range of anomalies documented in developed markets also appear in Chinese equities, though their strength varies across anomaly categories and over time (Han, 2022; Jansen et al., 2021). Anomalies related to trading frictions and limits to arbitrage tend to be particularly strong, and their profitability is closely linked to investor sentiment in this retail-dominated environment (Han, 2022). These features make the A-share market as an ideal environment for analysing whether firm-level information opacity is associated with stronger anomaly profits. In such an environment where retail traders rely predominantly on public disclosure and media signals, pricing errors induced by heterogeneous interpretation of firm disclosures may persist, particularly for stocks with limited transparency. Firm-level opacity may therefore influence the extent to which characteristic-based mispricing persists in China, particularly in comparison to markets with deeper institutional arbitrage.

While the anomaly and disclosure literatures are both extensive, relatively few studies directly connect firm-level information opacity to the cross-sectional variation in anomaly returns. Most anomaly papers take the information environment as given and focus on documenting which characteristics predict returns and how far multi-factor models can go in capturing these patterns (e.g. Chu et al., 2020; Stambaugh et al., 2012). On the accounting side, studies of disclosure and information risk typically examine how disclosure affects average returns, the cost of capital, liquidity, or analyst behaviour, but they rarely ask whether disclosure quality systematically moderates the profitability of characteristic-based strategies. As a result, it remains unclear whether anomalies are stronger among opaque firms because mispricing is harder to arbitrage, or whether anomalies are largely unrelated to firm-specific information environments once risk factors are controlled for.

This essay bridges these strands of literature by combining a broad set of characteristic-based anomalies with a firm-level measure of information opacity for Chinese listed firms. Following Stambaugh et al. (2012) and Chu et al. (2020), we construct eleven anomalies spanning profitability, investment, financing, accruals, asset growth, financial distress, and momentum dimensions. These variables are derived from annual financial statements and monthly stock returns from the CSMAR database and are constructed to match as closely as possible the definitions in the original anomaly papers. We then obtain a firm-level opacity index from CSMAR, which summarises the quality and transparency of financial reporting and disclosure. This allows us to examine whether anomaly profits differ systematically between firms with relatively transparent information environments and those whose disclosures are more opaque.

Our empirical analysis proceeds in three steps. First, we assess whether the eleven characteristic-based anomalies and a composite anomaly portfolio remain profitable in China's A-share market after controlling for the market factor, the Fama–French three factors and the Fama–French five factors. This provides a baseline comparison with the international evidence and establishes whether the anomalies are economically meaningful in our sample. Second, we split firms into high- and low-transparency groups based on the opacity index and compare long-leg, short-leg and long–short anomaly portfolio returns across these groups. If opacity contributes to mispricing, we expect anomaly profits to be more pronounced among firms with more opaque information. Third, we estimate contemporaneous and predictive regressions of anomaly portfolio returns on the opacity measure, controlling for standard risk factors, to test whether transparency has an incremental effect on anomaly profitability beyond conventional factor exposures.

We document three main findings. First, most of the eleven anomalies and the composite anomaly portfolio yield positive and statistically significant profits over our sample period, even after adjusting for the market, three-factor and five-factor benchmarks. This confirms that characteristic-based anomalies are a prominent feature of the Chinese A-share market over our sample window. Second, portfolio sorts by the opacity index show that anomaly returns are systematically related to firms' information environments. In the long-leg portfolios, expected returns are lower for stocks with higher transparency and higher for more opaque firms. Although this pattern is less pronounced on the short-leg side, long–short anomaly profits are generally larger among opaque firms, implying that anomalies are weaker when corporate information is more transparent. Third, regression results using both contemporaneous and lagged opacity measures reinforce this conclusion: higher transparency is associated with lower expected returns from anomaly-based long–short strategies, even after controlling for standard risk factors.

Taken together, these results suggest that firm-level information opacity is an important conditioning variable for understanding anomaly profits in China's A-share market. Anomalies are more difficult to exploit when corporate information is relatively transparent, consistent with the idea that improved disclosure facilitates more accurate valuation and faster incorporation of firm-specific information into prices. Our evidence therefore links the literature on disclosure and information risk with the anomaly literature by showing that the strength of characteristic-based anomalies depends on the quality of firm-level information environments.

This essay contributes to the literature in several ways. First, it extends the evidence on characteristic-based anomalies by documenting the profitability of eleven anomalies and a composite strategy in a large emerging market, using a sample that covers a longer period and a broader cross-section of A-share firms than many prior studies. Second, it provides systematic evidence on how firm-level information opacity relates to anomaly profits, showing that higher transparency is associated with weaker anomaly returns. Third, by focusing on a retail-dominated market that has undergone substantial regulatory and disclosure reforms, the study sheds light on how improvements in disclosure standards may interact with limits to arbitrage in influencing the persistence of return predictability.

The rest of this chapter is organised as follows. Section 2.2 reviews the related literature on asset pricing models, characteristic-based anomalies, the Chinese stock market, and information opacity. Section 2.3 develops the motivation for the study and outlines the conceptual channels through which firm-level opacity may condition anomaly profitability. Section 2.4 describes the data, sample construction, and the measurement of the eleven anomaly characteristics and the firm-level opacity index. Section 2.5 presents the empirical results and discussion. Section 2.6 concludes.

2.2. Literature Review

2.2.1. Asset Pricing in Efficient and Inefficient Markets

The determinants of asset prices have been debated for several decades. A classic perspective can be traced back to the Efficient Market Hypothesis (EMH), which argues that security prices fully reflect all available information so that abnormal risk-adjusted returns are difficult to obtain. Early contributions such as Friedman (1953) emphasise that rational investors quickly exploit and eliminate arbitrage opportunities, while irrational traders who persistently buy high and sell low are eventually driven out of the market through wealth losses. Under this view, observed returns should primarily compensate for systematic risk, and any persistent predictability in the cross-section of returns should be interpreted as risk premia rather than mispricing.

Early empirical work provided some support for this efficiency view using event-study methods. Ball and Brown (1968) examine stock price reactions around earnings announcements and show that prices respond quickly to the information content of reported earnings. Jegadeesh and Titman (1993) document momentum profits over medium horizons, but also find that part of the momentum predictability decays when the portfolio holding periods are narrowed, suggesting that anomalies may be short-lived. These studies helped establish the idea that markets process public information relatively quickly, at least in simple event settings.

However, a large body of literature documents systematic departures from the EMH. One important channel is information asymmetry between informed and uninformed investors. Grossman and Stiglitz (1980) show that when information is costly to acquire, prices cannot be fully revealing in equilibrium; some degree of noise is necessary to compensate informed traders. As a result, equilibrium prices may deviate from fundamental values even in a rational setting.

Institutional frictions further weaken the link between information and prices. Allen and Gorton (1993) highlight how portfolio managers can exploit private information when other investors are less informed, especially when contracting frictions and benchmarking pressures limit the ability of capital providers to monitor risk-taking. Duffie et al. (2002)

develop an asset valuation model in which short sellers must search for securities lenders; they show that short-selling constraints and lending fees can cause constrained assets to trade at higher prices relative to unconstrained ones. Such models imply that deviations from fundamental values can arise from trading frictions and regulatory constraints even without behavioural biases.

Taken together, this strand of research suggests that while market efficiency remains a useful benchmark, institutional features such as information asymmetry and short-selling constraints can generate persistent cross-sectional return patterns. This provides a theoretical foundation for the large anomaly literature that follows, and motivates examination of how market structure and firm-specific information environments influence the strength of anomalies, particularly in emerging markets like China.

2.2.2. Development of Asset Pricing Theories

Modern asset pricing initially develops within the framework of general equilibrium models linking expected returns to systematic risk. The best known is the Capital Asset Pricing Model (CAPM) of Sharpe (1964) and Treynor (1965), which posits a linear relation between the expected excess return on an asset and its beta with respect to the market portfolio. Early empirical tests, such as Black et al. (1972), form portfolios sorted by beta and show a positive association between average returns and market risk, providing initial support for CAPM.

However, subsequent cross-sectional studies identify patterns that are not well explained by market beta. Basu (1977) finds that stocks with high earnings-to-price ratios earn higher average returns than implied by CAPM. Banz (1981) documents that small-capitalisation firms earn higher returns than large firms even after controlling for beta, while Reinganum (1981) and others report similar size and value-related effects. These findings challenge the empirical sufficiency of CAPM and suggest that additional dimensions of risk or mispricing may be relevant.

Similar evidence arises outside the U.S. Fama and French (1992) show that in the U.S. cross-section, firm size and book-to-market equity explain average returns better than beta, and that the simple CAPM using a broad stock index as the market proxy performs poorly. Studies such as Gonzalez (2001) and Iqbal et al. (2010) report analogous inadequacies of CAPM for describing average stock returns in emerging markets.

To relax the single-factor structure of CAPM, Ross (1976) proposes the Arbitrage Pricing Theory (APT), under which expected returns are driven by multiple common factors. APT is less restrictive than CAPM, but it does not specify the identity of the factors. Empirically, factors are typically inferred from macroeconomic variables or from portfolios that proxy for systematic shocks. In either case, the empirical challenge is to identify a parsimonious set of factors that jointly span the cross-section of expected returns.

Fama and French (1993) take an important step by introducing a three-factor model that augments the market factor with size (SMB) and value (HML) factors constructed from portfolios long in small or value stocks, and short in large or growth stocks. They show that this model captures the size and value premiums and substantially improves upon CAPM in explaining the cross-section of average returns. The three-factor model quickly becomes a benchmark specification in empirical asset pricing.

Nevertheless, several anomalies remain unexplained. Jegadeesh and Titman (1993) document a strong momentum effect, whereby past winners continue to outperform past losers over horizons of three to twelve months. Brennan et al. (1998), and later Bauer et al. (2010), show that momentum and liquidity effects are not fully absorbed by the three-factor model. Daniel and Titman (1997) argue that portfolios formed directly on characteristics such as book-to-market perform differently from portfolios formed on factor loadings, suggesting that the three factors may not represent pure risk factors. Petkova (2006) further shows that part of the explanatory power of the Fama–French factors can be traced to underlying macroeconomic risks, but this does not eliminate all anomalies.

To incorporate momentum explicitly, Carhart (1997) proposes a four-factor model that adds a momentum factor to the three-factor specification. This model captures much of the profitability of momentum strategies in U.S. data. In the Chinese context, Narayan and Zheng (2010) find that the Carhart four-factor model augmented with a liquidity factor can explain some of the size, value, and liquidity premia, but that momentum effects are still not fully subsumed.

More recently, Fama and French (2015) introduce a five-factor model that adds profitability (RMW) and investment (CMA) factors to the original three. Motivated by valuation theory and evidence that more profitable firms and firms with lower investment earn higher returns, the five-factor model improves the overall fit to the cross-section. However, they also find that the traditional HML factor becomes largely redundant once profitability and investment factors are included.

Despite these advances, model sufficiency remains an open question. Brennan (1998) uses the principal components approach of Connor and Korajczyk (1988) to examine whether security characteristics retain explanatory power once latent factors are controlled for, and finds that several characteristics still predict risk-adjusted returns. Schwert (2003) emphasises that tests of asset pricing models are joint tests of market efficiency and model specification: rejection of a model may reflect mispricing, misspecification, or both. Avramov and Chordia (2006) similarly raise the question of whether extant factor models are sufficient to subsume return predictability associated with firm characteristics.

These debates motivate the large anomaly literature that examines whether returns associated with specific characteristics can be reconciled with risk-based explanations, or whether they reflect limits to arbitrage and behavioural mispricing. They also motivate examining how the strength of anomalies varies across markets and information environments.

2.2.3. Information Opacity, Corporate Transparency, and Mispricing

Information opacity is a key dimension of the institutional environment that influences both firm financing and asset pricing. Corporate transparency refers to the extent to which firms provide timely, reliable, and sufficiently detailed disclosures that allow outside investors to assess fundamentals and risks. Theoretical and empirical studies show that lower transparency increases information risk, affects liquidity, and can generate risk premia and mispricing.

Barry and Brown (1985) develop a model in which imperfect disclosure leads investors to face non-diversifiable uncertainty about the distribution of future returns. Investors then demand a premium for holding securities with more opaque information environments, so that expected returns vary with disclosure quality. Diamond and Verrecchia (1991) and Kim and

Verrecchia (1994) show that greater voluntary disclosure reduces information asymmetry between informed and uninformed investors, narrows bid–ask spreads, and improves liquidity. By lowering adverse selection costs and reducing the dispersion of beliefs, better disclosure moves prices closer to fundamental values and lowers the cost of equity capital.

Corporate transparency also affects credit markets and investment efficiency. DeBoskey and Gillett (2013) find that higher transparency is associated with better credit ratings and lower cost of capital, consistent with investors requiring smaller risk premia when information quality is high. From an asset pricing perspective, these results imply that firms with poor disclosure may exhibit higher average returns, but that these returns can reflect both compensation for information risk and mispricing due to difficulty in valuing fundamentals.

In emerging markets, where legal enforcement, accounting standards, and media freedom are still developing, disclosure quality is highly heterogeneous. Ball et al. (2000) and Bushman et al. (2004) show that emerging markets generally exhibit weaker investor protection, less conservative accounting, and less informative disclosures than developed markets. In such settings, some firms may deliberately obscure bad news or delay recognition of losses, while others voluntarily provide more detailed disclosures to signal quality and reduce financing costs. Durnev and Kim (2005) argue that firms with strong growth opportunities and good governance have greater incentives to disclose information to distinguish themselves from lower-quality peers.

Information opacity interacts with investor sentiment and limits to arbitrage in determining anomaly strength. When disclosure is poor, valuation is more difficult and beliefs are more dispersed, especially among retail investors. Firth et al. (2015) show that, in the Chinese market, stock prices of firms with lower transparency are more sensitive to sentiment proxies, suggesting that sentiment has a larger effect when fundamentals are hard to observe. In such environments, mispricing associated with characteristics like accruals, investment, and profitability is more likely to arise and persist, particularly if short-selling is constrained.

Empirical work links disclosure quality directly to the magnitude of anomalies. Drake et al. (2009) find that accrual and cash-flow mispricing is more pronounced among firms with lower-quality disclosure. Chiyachantana et al. (2013) show that firms with higher transparency have narrower bid–ask spreads and higher turnover than opaque firms, indicating better liquidity and greater participation by informed investors. Mugaloglu and Erdag (2013) provide evidence that poor transparency increases uncertainty and disrupts efficient price formation. These findings suggest that opacity amplifies the asset pricing effects of accounting-based characteristics.

In the Chinese A-share market, these mechanisms are reinforced by the dominance of individual investors and the variation in firm-level disclosure practices. Retail investors rely heavily on headline financial indicators and may find it difficult to interpret complex notes or risk disclosures. When reporting is opaque or boilerplate, investors are more likely to rely on sentiment and heuristics, increasing the scope for anomalies in characteristics linked to fundamental information. Conversely, firms that provide more granular and credible disclosures can reduce information asymmetry and attract more sophisticated capital, which can dampen anomaly profits.

Against this background, examining how firm-level information opacity is related to the profitability of characteristic-based anomalies in the Chinese A-share market can shed light on the interaction between disclosure practices, market structure, and cross-sectional return predictability. The existing literature suggests that opacity should strengthen mispricing-related anomalies by increasing information risk, widening belief dispersion, and weakening arbitrage, especially in an emerging market dominated by individual investors and subject to institutional constraints.

Economically, the OPAC measure is intended to capture the disclosure-transparency dimension of the firm's information environment. The disclosure literature suggests that higher-quality corporate information can reduce information asymmetry and investor uncertainty and facilitate the incorporation of firm-specific information into market prices (Diamond & Verrecchia, 1991; Healy & Palepu, 2001; Leuz & Verrecchia, 2000). Conversely, a less transparent disclosure environment may increase information frictions and contribute to the persistence of mispricing. This interpretation is also consistent with evidence linking financial reporting opacity to the incorporation of firm-specific information into stock prices (Hutton et al., 2009). Accordingly, OPAC is interpreted as a proxy for disclosure-related information opacity rather than as a direct measure of corporate governance, stock liquidity, or other individual firm characteristics.

2.2.4. Investor Irrationality and Economic Anomalies

In parallel with the development of multifactor risk models, a substantial body of literature examines whether persistent cross-sectional return patterns reflect investor irrationality and limits to arbitrage. Financial market anomalies are typically defined as return patterns associated with firm characteristics or trading strategies that cannot be fully explained by standard asset pricing models under rational expectations (Brav & Heaton, 2002). The existence of such anomalies poses a challenge for the mainstream asset pricing paradigm.

Some researchers argue that anomalies are transient and weaken as the sample period extends or as investors learn. Schwert (2003) reviews a range of anomalies and finds that many become less pronounced after they are documented, suggesting that either market efficiency improves or that earlier findings partially reflect data-snooping. Fama and French (1992) also note that some characteristics can be interpreted as proxies for risk, complicating the distinction between risk and mispricing.

Behavioural finance provides an alternative explanation by relaxing the assumption of fully rational expectations. Daniel et al. (1998) show that overconfidence and biased self-attribution can generate momentum and long-run reversals, as investors overweight private signals and underreact to public information. Shleifer and Vishny (1997) emphasise that even when some investors are rational, arbitrage is limited by capital constraints, career concerns, and model uncertainty; as a result, mispricing caused by irrational traders may persist.

Empirical evidence supports the importance of behavioural mechanisms. Grinblatt and Han (2005) argue that investors' reluctance to realise losses, often termed the disposition effect, leads to underreaction to bad news and creates momentum in subsequent returns, particularly in speculative stocks. Edelen et al. (2016) find that institutional investors, although typically regarded as sophisticated, also exhibit systematic stock-picking preferences that contribute to anomaly returns when their trades are synchronised.

Macroeconomic and environmental conditions can further influence investor mood and sentiment. Economic anomalies arising from holidays, calendar effects, and weather changes are often interpreted through a psychological lens. Bergsma and Jiang (2016) show that positive holiday sentiment and liquidity injections surrounding the Lunar New Year are associated with higher stock prices in China. Birru (2018) documents that returns on speculative stocks are lower on Mondays and higher on Fridays, consistent with mood fluctuations across the week. Dowling and Lucey (2005) present evidence that weather conditions such as precipitation are related to equity returns, although the economic magnitude and persistence of such effects remain to be debated. Using longer samples and broader markets, Athanassakos et al. (2019) find that weather-induced anomalies weaken over time, consistent with the view that they may be exploited and arbitraged away as markets develop.

Overall, this literature suggests that investor sentiment, limited attention, and cognitive biases can generate predictable return patterns, especially in stocks that are hard to arbitrage or difficult to value. Many of these anomalies arise from non-fundamental triggers, such as mood, attention, or trading conventions, rather than from firm fundamentals. This contrasts with the anomalies examined in this essay, which are constructed from accounting variables and firm characteristics, although both sets of anomalies share a common mechanism involving behavioural biases and limits to arbitrage.

2.2.5. Characteristic-Based Anomalies

A large body of empirical literature investigates specific anomalies based on firm-level characteristics. These studies typically sort stocks into portfolios according to accounting or market-based variables and examine whether long–short spreads yield abnormal returns after controlling for standard risk factors.

A major strand focuses on accrual-based anomalies. Sloan (1996) shows that the accrual and cash-flow components of earnings have different persistence, but that investors fail to fully recognise this distinction. Firms with high accruals earn lower subsequent returns than low-accrual firms, consistent with overvaluation of the accrual component. Subsequent work using quarterly data confirms the existence of accrual and cash-flow mispricing and shows that the anomaly is concentrated in abnormal accruals, where managerial discretion is greater. Xie

(2001) documents that the overpricing of total accruals is largely due to abnormal accruals, while Daniel et al. (2003) show that the mispricing is stronger when the proportion of non-institutional investors is high and when arbitrage is more restricted. These findings support the view that limited attention and accounting complexity contribute to mispricing.

Another class of anomalies is based on firm investment and balance-sheet growth. Fairfield et al. (2003) introduce the net operating assets (NOA) anomaly and show that firms with high growth in NOA experience lower subsequent returns. They interpret this as evidence that markets underreact to the negative implications of aggressive investment for future profitability. Titman et al. (2004), Fama and French (2006, 2008), and Cooper et al. (2008) document a broader negative relation between asset growth and risk-adjusted returns: firms that expand their total assets more aggressively tend to earn lower future returns than firms with conservative growth. This investment or asset-growth anomaly is often linked to overinvestment by managers, agency frictions, and investor underreaction.

Profitability-based anomalies further emphasise the role of operating performance in asset pricing. Novy-Marx (2013) demonstrates that gross profitability, sales minus cost of goods sold, scaled by total assets, has strong predictive power for average returns. High-profitability firms earn higher returns than low-profitability firms, even controlling for size and value. Ball et al. (2015) show that gross profitability is a superior predictor than net income because it is less affected by financing and accounting choices and more directly linked to the productivity of assets. Kogan et al. (2019) provide a production-based explanation: if variable production costs are pro-cyclical, they can reduce the systematic risk of low-profitability firms, generating a profitability premium. Behavioural explanations also appear. Bouchaud et al. (2019) argue that analysts' sticky expectations and slow updating contribute to mispricing, while Lam et al. (2020) show that profitability premia are stronger when investor sentiment is high and when ex ante expectations are biased.

Distress and default-risk anomalies challenge the standard risk–return trade-off. Ohlson (1980) develops the O-score as a bankruptcy prediction measure. Campbell et al. (2008) find that stocks with the highest distress risk earn lower, not higher, subsequent returns than less distressed stocks, which is hard to reconcile with rational pricing. Agarwal and Taffler (2008) show that in UK data, momentum profits are largely subsumed by distress risk factors, suggesting that part of momentum reflects mispricing of distressed firms. Lai et al. (2010) provide evidence that a four-factor model including the O-score as a distress factor helps explain returns in U.S. and Japanese markets.

These anomalies are typically implemented as zero-investment hedging strategies, taking long positions in “undervalued” stocks and short positions in “overvalued” stocks according to a given signal. Stambaugh et al. (2012) group a wide range of such characteristics into categories such as valuation, profitability, investment, and distress. They show that many of these anomalies are stronger in periods of high investor sentiment and that their returns are higher among stocks that are difficult to arbitrage, supporting a mispricing interpretation. Fisher et al. (2015) examine combination strategies that jointly use value and momentum signals and find that these can improve performance while reducing turnover and transaction costs.

Finally, trading-based anomalies highlight the role of liquidity. Pastor and Stambaugh (2003) construct a market-wide liquidity factor and show that stocks with higher sensitivity to liquidity innovations command higher expected returns. Lee and Swaminathan (2000) study the relation between turnover and momentum, while Keene and Peterson (2007), Chordia et al. (2002), Jun et al. (2003), and Lee (2011) provide evidence that liquidity and liquidity risk are priced, particularly in emerging markets. Together, these findings indicate that firm characteristics related to accounting, investment, distress, and trading all contain information about expected returns that is not fully captured by standard factor models.

2.2.6. Characteristics of the Chinese Stock Market

The Chinese A-share market has grown rapidly over the past two decades and now represents one of the largest equity markets globally. However, its structural characteristics differ markedly from those of developed markets. The market is dominated by individual investors, has a significant presence of state-owned enterprises, and has historically imposed tight limits on short selling and derivatives trading. These features have important implications for information efficiency and the strength of anomalies.

A defining feature of the A-share market is the high participation of retail investors. Over long sample periods, institutional investors' holdings account for only a small share of the market, while individuals contribute the majority of trading volume. Deng and Xu (2011) find that institutional investors exhibit superior stock selection ability and contribute more to price efficiency, whereas individual investors rely heavily on public information and are more influenced by market sentiment and attention-grabbing events. Ng and Wu (2007) show that institutions tend to trade with momentum on both the buy and sell sides, while individuals more often take contrarian positions after recent price changes. In futures markets, Xu and Wan (2015) report that institutional trading improves price discovery, whereas individual trading can adversely affect information efficiency.

Investor herding behaviour has been identified as another source of instability. Li et al. (2017) document that both institutional and individual investors herd in the Chinese market, but that individual herding is more sensitive to market returns and volatility. Yao et al. (2014) find that herding is more pronounced during market declines and that regulatory reforms have improved information efficiency over time. Zheng et al. (2015) show that herding behaviour has a significant impact on excess returns and can generate long-term reversals, consistent with temporary mispricing followed by correction. When herding is strong and short-selling opportunities are limited, overpricing is especially difficult to arbitrage away, which provides fertile ground for anomalies.

Another characteristic of emerging markets such as China is high stock price synchronicity and low price informativeness. Morck et al. (2000) and later studies show that in markets with weaker investor protection and less developed disclosure systems, individual stock prices co-move more with market and industry indices and incorporate less firm-specific information. Chan and Hameed (2006) and related work document similar patterns in other emerging markets. High synchronicity implies that firm-level news is only partially reflected in prices, so that characteristics related to fundamentals can have stronger predictive power for returns.

Institutional features also matter. The presence of many state-owned enterprises affects corporate objectives, disclosure practices, and responses to policy shocks. State ownership can weaken the link between firm characteristics and expected returns if objectives other than shareholder value maximisation are pursued. At the same time, regulatory reforms, such as the gradual introduction of margin trading and securities lending, the Qualified Foreign Institutional Investor (QFII) scheme, and the development of index futures, have progressively increased the role of institutional investors and hedging instruments. Short-selling, however, remains more restricted than in many developed markets, and the set of stocks eligible for borrowing is limited.

Empirical evidence suggests that value, low-risk, and trading anomalies exist in the A-share market, although their magnitudes and persistence differ from those commonly reported in developed markets. Jansen et al. (2021) examine 32 anomalies in China and show that value, low-risk, and trading strategies generate economically meaningful long-short profits, while size and quality anomalies are comparatively weaker. They further document that anomaly profitability persists across firm-size segments and cannot be fully explained by industry composition or institutional features, suggesting that mispricing mechanisms in China are not driven solely by market segmentation or state ownership. For characteristics-based anomalies constructed from accounting information, these findings imply that the interaction between firm characteristics, investor composition, and institutional frictions is central to understanding observed return patterns in this market.

2.3. Motivation and Conceptual Channels

2.3.1. Motivation

Return anomalies challenge traditional asset pricing by documenting systematic patterns that are difficult to reconcile with risk-based explanations. While many anomalies have been studied extensively in developed markets, their persistence varies across institutional settings. One factor that has received comparatively less empirical attention is the role of the firm-level information environment. When financial reporting and disclosure are opaque, investors face greater difficulty in assessing firm fundamentals, which can impede informed trading and prolong mispricing. Conversely, transparent disclosure may facilitate more accurate valuation, reduce information frictions, and narrow the scope for anomaly-based strategies.

The relevance of transparency is particularly salient in China's A-share market. The market is characterised by a high degree of retail participation, relatively limited institutional arbitrage capital, and heterogeneous investor sophistication. Retail investors rely heavily on public disclosure, media reports, and attention-grabbing signals when forming expectations, and their reactions are sensitive to sentiment and market mood (Lin et al., 2020). Regulatory reforms in accounting and disclosure over the past two decades have materially improved reporting standards (Firth et al., 2015), yet considerable cross-sectional variation remains in the transparency of listed firms. In such an environment, opacity may amplify valuation disagreement and inhibit arbitrage, generating conditions under which characteristic-based anomalies can persist more strongly. Understanding how anomalies interact with firms' information environments is therefore central to assessing whether anomaly profitability in China reflects mispricing that is conditional on disclosure quality rather than unconditional inefficiency.

2.3.2. Conceptual Channels

A firm's information environment can influence anomaly persistence through several reinforcing channels grounded in asset-pricing theory and limits-to-arbitrage arguments.

2.3.2.1. Information-processing and information-risk channel

Grossman and Stiglitz (1980) demonstrate that when information acquisition is costly, prices cannot be perfectly informative in equilibrium because informed traders require compensation for their information processing. Easley and O'Hara (2004) formalise information risk as a priced component of expected returns, while Lambert et al. (2012) highlight that the precision of public information affects risk-sharing and valuations. When disclosure quality is low, investors face higher information-processing costs and greater uncertainty about fundamentals, which can create valuation wedges not immediately eliminated by arbitrage. Characteristic-based anomalies, many of which rely on information extracted from financial statements, may therefore produce greater abnormal returns among firms with more opaque reporting environments.

2.3.2.2. Disagreement and heterogeneous beliefs channel

Opacity can also amplify belief heterogeneity. When disclosures are incomplete or ambiguous, investors may disagree more about a firm's intrinsic value. Miller (1977) shows that in the presence of disagreement and short-sale constraints, prices reflect the views of the most optimistic investors, leading to overpricing. Harris and Raviv (1993) and Hong and Stein (1999) demonstrate related mechanisms in which information dispersion increases trading and price divergence. In China's retail-dominated market, heterogeneous interpretation of corporate announcements and media signals can induce substantial pricing errors, particularly for firms with limited transparency. If anomaly signals capture predictable reversals in such mispricing, anomalies should be more pronounced among firms with greater disclosure-related disagreement.

2.3.2.3. Limits-to-arbitrage channel

Even when informed traders recognise mispricing, anomaly profits may persist if arbitrage is limited or costly (Pontiff, 2006; Shleifer & Vishny, 1997). Opacity can increase these costs by reducing the precision of valuation, weakening collateral value, and heightening uncertainty about payoff horizons. In China, the effects of opacity interact with the market's structural features: institutional arbitrageurs remain a smaller fraction of trading volume relative to developed markets, and retail investors hold substantial influence over price dynamics. These characteristics reduce arbitrage depth, allowing mispricing associated with anomaly characteristics to survive longer when firm transparency is low.

Collectively, these conceptual channels suggest that corporate opacity can condition the profitability of characteristic-based anomalies. If disclosure improvements reduce information-processing costs, diminish disagreement, and deepen arbitrage, then anomaly returns should be weaker among firms with higher transparency. Conversely, when firms disclose less information or disclose it in less interpretable form, anomaly-based mispricing may persist for longer periods, especially in market environments with limited institutional arbitrage capital such as China's A-share market.

2.4. Data and Methodology

2.4.1. Sample and Data

We examine Chinese A-share firms listed on the Shanghai and Shenzhen Stock Exchanges over the period from January 2000 to December 2020, as firm-level opacity measures are available beginning in January 2000, enabling us to maximise the use of available data and ensure sufficient sample coverage for empirical analysis. Monthly stock prices, returns, and market value are obtained from the CSMAR "Monthly Stock Price – Returns" database. Financial statement data are drawn from CSMAR's balance sheet (BS), income statement (IS), and cash flow statement (CF) modules. We restrict the sample to firms with fiscal year-ends on 31 December to ensure alignment between financial reporting dates and portfolio formation. For each anomaly, we require the relevant accounting information to be available in the current fiscal year t and, where necessary, its lagged values, so that anomaly characteristics are well defined at the end of year t.

Firm-level disclosure quality is measured using the Company Opacity (OPAC) index from CSMAR. The measure is based on the annual information disclosure quality assessments conducted by the Shanghai and Shenzhen Stock Exchanges and is reported in CSMAR under the field "CompanyOpacity" ("Transparency of Listed Company"). The ratings provide a comprehensive regulatory assessment of firms' disclosure practices over the fiscal year, taking into account factors such as disclosure compliance, timeliness, accuracy, and completeness. Listed firms are assigned one of four disclosure grades, ranging from A (excellent disclosure) to D (poor disclosure). To facilitate empirical implementation, the alphabetical ratings are converted into a four-level numeric index, assigning 4 to A-grade firms and 1 to D-grade firms, such that higher values indicate greater disclosure transparency and lower values indicate greater information opacity. The exchange-based nature of the measure provides an institutionally grounded assessment of firms' disclosure environments, and recent evidence supports the informativeness of these exchange disclosure-quality ratings in the Chinese market (Yang & Zhao, 2025). OPAC is measured annually based on the corresponding fiscal-year disclosure assessment.

We also obtain monthly Fama–French factor returns for the Chinese A-share market from CSMAR, including the market excess return (MKT), size (SMB), value (HML), profitability (RMW) and investment (CMA) factors. These factor series are constructed by CSMAR as value-weighted factor-mimicking portfolios based on standard Fama–French procedures. The risk-free rate is sourced from the CSMAR benchmark interest-rate module, which reports annualised deposit and government-bond rates together with derived daily and monthly series. We use the monthly risk-free rate when computing excess returns.

In addition, we use the mainland China economic policy uncertainty (EPU) index of Huang and Luk (2020), which is constructed from the frequency of policy-related terms in mainland newspapers, as a macro-level control for policy uncertainty in later regressions.

All firm-level anomaly variables are winsorized at the 1st and 99th percentiles to moderate the influence of outliers. The combined dataset yields an unbalanced firm–month panel, which we use for portfolio formation, factor regressions, and opacity-based analyses.

2.4.2. Construction of Anomaly Characteristics

To examine the impact of information disclosure on anomaly returns, we construct eleven anomaly characteristics following Stambaugh et al. (2012) and Chu et al. (2020). These variables span profitability, investment, accruals, financing activity, financial distress, and market-based mispricing dimensions. Accounting variables are obtained annually from CSMAR financial statements at the December fiscal year-end, and trading-related variables are taken from CSMAR's monthly stock-price files. All anomaly characteristics are updated once per year at the fiscal year-end and are lagged by one year relative to the return evaluation period to avoid look-ahead bias.

The eleven anomaly variables are: gross profitability (GP), asset growth (AG), investment-to-assets (IVA), return on assets (ROA), net operating assets (NOA), accruals (ACC), net equity issuance (ISS), composite equity issuance (CEI), momentum (MOM), O-score (OSC), and downside risk (DSR). A brief description is provided below, and the full construction details and CSMAR item mappings are reported in Appendix A.

1. Gross profitability (GP)

   Novy-Marx (2013) shows that gross profitability scaled by assets is a powerful predictor of cross-sectional returns. We follow Novy-Marx (2013) and construct GP as gross profit scaled by total assets at the fiscal year-end, where gross profit is measured as operating revenue minus operating costs from the income statement. This ratio captures how efficiently the firm's asset base generates gross operating income.

2. Asset growth (AG)

   Cooper et al. (2008) find that firms with high growth in total assets earn lower subsequent returns, consistent with investor overvaluation of aggressive investment. We define AG as the year-over-year percentage change in total assets, computed as the change in total assets between December of year $t-1$ and December of year t, scaled by lagged total assets.

3. Investment-to-assets (IVA)

   Titman et al. (2004) show that higher capital investment is associated with lower subsequent stock returns. Following their approach, we measure IVA as the sum of the annual changes in fixed assets and inventories, scaled by lagged total assets. Fixed assets and inventories are obtained from the balance sheet, and the changes are computed between December of year $t-1$ and December of year t.

4. Return on assets (ROA)

   Chen et al. (2011) and Stambaugh et al. (2012) document that past profitability predicts higher future returns. We measure ROA as net profit divided by total assets at the fiscal year-end, where net profit is taken from the income statement and total assets from the balance sheet.

5. Net operating assets (NOA)

   Hirshleifer et al. (2004) find that firms with higher net operating assets tend to earn lower subsequent returns, suggesting overvaluation of aggressive investment in operating assets. We follow their conceptual definition and construct NOA as operating assets minus operating liabilities, scaled by lagged total assets. Operating assets are defined as total assets minus cash, cash equivalents, and short-term investments, while operating liabilities are obtained by subtracting explicit financing items (short-term loans, long-term loans, minority interests, and total equity) from total assets.

6. Accruals (ACC)

   Sloan (1996) shows that firms with higher accruals earn lower average returns, as investors fail to fully distinguish between the accrual and cash-flow components of earnings. We measure accruals as the annual change in non-cash current assets minus the annual change in non-debt current liabilities, scaled by lagged total assets. Non-cash current assets are defined as current assets minus cash and short-term investments, and non-debt current liabilities are defined as current liabilities minus short-term loans and taxes payable.

7. Net equity issuance (ISS)

   Pontiff and Woodgate (2008) show that net share issuance negatively predicts cross-sectional returns, as firms tend to issue equity when overvalued, and repurchase when undervalued. We follow their approach and measure ISS as the log change in shares outstanding over the previous year. Shares outstanding are inferred at each December portfolio-formation date from total market value and closing price in the monthly stock-price file.

8. Composite equity issuance (CEI)

   Daniel and Titman (2006) introduce composite equity issuance as an alternative financing anomaly, capturing market value of equity growth not explained by stock returns. Following their logic, we construct CEI as the log change in market equity over a five-year window minus the corresponding cumulative stock return over that window. Market equity is taken from December market capitalisation, and the construction uses the market value and return series from CSMAR.

9. Momentum (MOM)

   The momentum effect, first documented by Jegadeesh and Titman (1993), describes the tendency for stocks with high past returns to continue outperforming those with low past returns. We construct MOM as a medium-term return signal based on cumulative returns over the last year excluding the most recent month, in line with Jegadeesh and Titman (1993). Using the monthly stock-price file, we compute cumulative returns from month $t-12$ to $t-2$ as of the December portfolio-formation date.

10. O-score (OSC)

    The O-score of Ohlson (1980) is a financial distress measure constructed from accounting ratios. We obtain OSC from CSMAR's financial-distress module, where the index is computed using the original Ohlson logit model applied to CSMAR accounting data. The O-score is reported at the fiscal year-end; we retain the December observations and merge them to our anomaly panel by firm and year. Higher values indicate higher predicted distress risk.

11. Downside risk (DSR)

    Stambaugh et al. (2012) and Chu et al. (2020) use failure probability as a distress proxy, but the Campbell et al. (2008) failure probability measure is not directly available for Chinese data. As an alternative, we adopt downside risk following Ang et al. (2006), defined as the standard deviation of monthly returns below the firm's mean return over a rolling window. We obtain DSR directly from CSMAR's downside-risk dataset, where downside volatility is computed over rolling 36-month windows and reported at an annual frequency. We retain the December values so that DSR is observed once a year at the portfolio-formation date.

All eleven anomaly characteristics are winsorized at the 1st and 99th percentiles to reduce the influence of extreme observations. They are then merged into the monthly stock-return panel by firm and year and used to rank stocks at each December portfolio-formation date.

Table 2.1 reports summary statistics for the eleven anomaly characteristics used in Essay One. The table provides an overview of the distributional properties, dispersion, and extreme values of each variable.

Table 2. 1. Summary Statistics of Anomaly Characteristics

	No.	Mean	Median	Std	Min	Q1	Q3	Max
GP	14177	0.245	0.188	0.221	-0.022	0.101	0.316	1.271
AG	14177	0.163	0.074	0.449	-0.450	-0.012	0.203	3.333
IVA	14177	0.044	0.015	0.146	-0.289	-0.017	0.072	0.887
ROA	14177	0.016	0.013	0.028	-0.073	0.003	0.028	0.112
NOA	14177	0.679	0.673	0.320	-0.165	0.516	0.812	2.309
ACC	14177	0.024	0.016	0.137	-0.404	-0.042	0.079	0.612
ISS	14177	0.096	0.000	0.229	-0.015	0.000	0.009	1.173
CEI	14177	0.417	0.398	0.930	-1.727	-0.175	1.010	2.880
OSC	13051	0.002	0.000	0.009	0.000	0.000	0.001	0.073
DSR	13051	0.037	0.023	0.046	0.000	0.003	0.055	0.270
MoM	13051	0.003	0.006	0.222	-0.628	-0.137	0.141	0.618

Notes This table reports summary statistics for the eleven anomaly characteristics constructed in Section 2.4.2. All accounting-based variables are measured using annual financial statement information aligned to fiscal year-end observations, while market-based variables are obtained from monthly stock-price data. All anomaly characteristics are updated annually, lagged by one year relative to returns, and winsorized at the 1st and 99th percentiles. Detailed variable definitions are provided in Appendix A.

2.4.3. Anomaly-Based Portfolio Construction

To evaluate the return implications of firm-level anomalies, we construct characteristic-sorted portfolios following standard asset-pricing practice. At the end of each December in year t, we sort stocks into deciles based on each anomaly characteristic using their lagged annual values. For each anomaly, we then form value-weighted long- and short-legs and hold the resulting portfolios from January to December of year $t+1$. Portfolios are rebalanced annually at each December formation date.

The long–short strategies are constructed in line with the theoretical expected return direction of each anomaly. For anomalies that are positively related to expected returns, such as gross profitability (GP), return on assets (ROA), and momentum (MOM), we take long positions in the highest decile (decile 10) and short positions in the lowest decile (decile 1), forming High – Low portfolios. For anomalies associated with overinvestment, financial distortions, or distress, such as AG, IVA, NOA, ACC, ISS, CEI, OSC, and DSR, we take Low – High positions, consistent with the literature that links higher levels of these characteristics to lower future returns. Portfolio weights are based on lagged market capitalisation at the formation date.

For each anomaly i, we compute monthly value-weighted portfolio returns over the holding period, as well as an equal-weighted composite anomaly portfolio that averages the eleven long–short return series. All portfolio returns are expressed as excess returns, calculated as raw portfolio returns minus the corresponding monthly risk-free rate. Time-series averages of these excess returns provide initial evidence on the magnitude and persistence of anomaly profits in China's A-share market.

In addition, we form the combined anomaly portfolio (COMB) which is calculated as the equal-weighted average of the monthly individual anomaly portfolios returns. It represents the average performance of anomaly-based trading strategies and provides a comprehensive measure of anomaly profitability.

2.4.4. Risk-Adjusted Returns and Baseline Regressions

To assess whether anomaly profits remain after controlling for standard risk factors, we estimate time-series regressions of monthly anomaly portfolio excess returns on benchmark asset-pricing models. Specifically, for each anomaly portfolio i, we estimate:

1. CAPM

$$R_{i,t} = \alpha_i + \gamma_{1,i}\mathrm{MKT}_t + \varepsilon_{i,t} \tag{1}$$

2. Fama–French three-factor model (FF3)

$$R_{i,t} = \alpha_i + \gamma_{1,i}\mathrm{MKT}_t + \gamma_{2,i}\mathrm{SMB}_t + \gamma_{3,i}\mathrm{HML}_t + \varepsilon_{i,t} \tag{2}$$

3. Fama–French five-factor model (FF5)

$$R_{i,t} = \alpha_i + \gamma_{1,i}\mathrm{MKT}_t + \gamma_{2,i}\mathrm{SMB}_t + \gamma_{3,i}\mathrm{HML}_t + \gamma_{4,i}\mathrm{RMW}_t + \gamma_{5,i}\mathrm{CMA}_t + \varepsilon_{i,t} \tag{3}$$

where $R_{i,t}$ is the monthly excess return of the long–short anomaly portfolio i in month t, computed as the raw portfolio return minus the risk-free rate for the same month. MKT is the market excess return; SMB is the size factor (small minus big); HML is the value factor (high book-to-market minus low book-to-market); RMW is the profitability factor (robust minus weak); and CMA is the investment factor (conservative minus aggressive).

The intercept α_i from each regression represents the risk-adjusted anomaly return (alpha) under the corresponding benchmark model. Using 21 years of monthly data, we compute the mean excess returns and CAPM-, as well as the FF3- and FF5-adjusted alphas for each of the 11 individual anomalies and for the composite anomaly portfolio. Pearson correlations across the 12 long–short excess return series (11 anomalies plus the composite portfolio) are also computed to assess the degree of overlap among anomaly strategies. The correlation matrix is reported in Panel A of Table 2.2, and the average excess returns and benchmark-adjusted alphas are presented in Panel B of Table 2.2.

2.4.5. Opacity Regressions and Macroeconomic Uncertainty

To examine how firm-level transparency conditions anomaly returns, we augment the baseline factor models with the OPAC measure. The OPAC measure is lagged by twelve months relative to the return evaluation period. This timing structure ensures that the explanatory variable is predetermined with respect to subsequent stock returns, thereby mitigating concerns about reverse causality because future stock returns cannot influence previously observed disclosure quality. Following the standard panel-data framework (Wooldridge, 2010), the lagged OPAC variable is plausibly exogenous with respect to subsequent return innovations under the assumption that the regression disturbances are not serially correlated.

The empirical design in this section is guided by several related expectations. If opacity slows the incorporation of firm-specific information into prices, anomaly strategies should earn higher excess returns among firms with weaker disclosure, particularly in the long legs that buy stocks with apparently favourable characteristics. In contrast, the short legs mainly contain firms that appear overvalued on the anomaly dimension and need not be strongly linked to disclosure quality, so we do not expect opacity to have a symmetric effect on both sides of the trade. Furthermore, if opacity matters beyond standard risk exposures and macroeconomic conditions, its estimated effect should remain present after controlling for Fama–French risk factors and, in an extended specification, the aggregate economic policy uncertainty index. The regression specifications below are chosen to test these expectations in a unified framework.

Specifically, for each anomaly portfolio i, we first estimate the following regression:

$$\begin{aligned} R_{i,t} \quad &= \alpha_i + \beta_i \, \mathrm{OPAC}_{i,t-12} + \gamma_{1,i} \, \mathrm{MKT}_t + \gamma_{2,i} \, \mathrm{SMB}_t + \gamma_{3,i} \, \mathrm{HML}_t \\ &+ \gamma_{4,i} \, \mathrm{RMW}_t + \gamma_{5,i} \, \mathrm{CMA}_t + \varepsilon_{i,t} \end{aligned} \tag{4}$$

where $R_{i,t}$ is the excess return on the long leg, short leg, or long–short portfolio for anomaly i in month t, and $\mathrm{OPAC}_{i,t-12}$ denotes the firm-level transparency rating at the December portfolio-formation date (carried forward into the holding period). The coefficient β_i captures whether anomaly returns are systematically related to firm-level transparency after controlling for standard risk factors. The estimated β_i coefficients for the long, short and long–short portfolios are reported in Table 2.4.

To test whether the effect of transparency on anomaly returns is robust to macro-level policy uncertainty, we further include the mainland China EPU index in the regression:

$$\begin{aligned} R_{i,t} \quad &= \alpha_i + \beta_i \, \mathrm{OPAC}_{i,t-12} + \gamma_{1,i} \, \mathrm{MKT}_t + \gamma_{2,i} \, \mathrm{SMB}_t + \gamma_{3,i} \, \mathrm{HML}_t \\ &+ \gamma_{4,i} \, \mathrm{RMW}_t + \gamma_{5,i} \, \mathrm{CMA}_t + \gamma_{6,i} \, \mathrm{EPU}_t + \varepsilon_{i,t} \end{aligned} \tag{5}$$

where EPU_t is the economic policy uncertainty index in month t. This specification allows us to examine whether the relation between opacity and anomaly returns is stable after controlling for time-varying macro uncertainty that may affect risk premia and investor sentiment. Together with the Fama–French five-factor controls included in Equation (4), these specifications reduce the possibility that the estimated relation between OPAC and anomaly returns is driven by common risk factors or macroeconomic conditions. The estimated coefficients for OPAC in the presence of EPU are reported in Table 2.5.

Together with the baseline factor regressions in Section 2.4.4, these models allow us to assess both the aggregate profitability of anomaly strategies and the extent to which firm-level disclosure quality helps explain variation in their excess returns across different transparency and macroeconomic environments.

2.5. Empirical Results

2.5.1. Baseline Anomalies Performances

Before examining how corporate opacity influences anomaly-based strategies, we first verify whether the eleven anomalies documented in the international literature are present in China's A-share market. For each anomaly, we construct a corresponding long–short portfolio that goes long in stocks with high values of the characteristic and short in stocks with low values, with the direction (High – Low or Low – High) chosen to align with the theoretical prediction of the original study. At each December portfolio-formation date, stocks are sorted into deciles based on the relevant anomaly characteristic, and monthly value-weighted excess returns are calculated as the return difference between the long leg (top or bottom decile) and the short leg (opposite decile) over the subsequent twelve months. In addition to the eleven individual strategies, we construct a composite anomaly portfolio as the equal-weighted average of the eleven long–short excess returns in each month. The detailed construction of the anomaly characteristics and portfolio strategies is described in Section 2.4.

Table 2.2 summarises the properties of these anomaly strategies. Panel A reports the Pearson correlations across the twelve long–short excess return series (eleven individual anomalies plus the composite portfolio). Correlations are generally moderate rather than uniformly high, indicating that the strategies capture partially distinct sources of return variation. Investment-related characteristics (AG, IVA and NOA) tend to be more closely related to each other, whereas profitability-based (GP, ROA), financing-related (ISS, CEI) and distress-related anomalies (OSC, DSR) display more heterogeneous co-movement. The composite portfolio therefore diversifies across different dimensions of firm characteristics rather than simply duplicating any single anomaly.

Panel B reports the time-series averages of the monthly long–short excess returns and the corresponding CAPM-, FF3- and FF5-adjusted alphas. Several anomaly strategies generate

Table 2. 2. Anomalies Return Performances

Panel A	**Pearson correlation for anomaly portfolio returns**											
	GP	AG	IVA	ROA	NOA	ACC	ISS	CEI	MOM	OSC	DSR	COMB
GP	1.00											
AG	0.14	1.00										
IVA	0.17	0.58	1.00									
ROA	0.34	0.62	0.38	1.00								
NOA	0.28	0.01	0.29	-0.08	1.00							
ACC	-0.18	-0.13	-0.35	-0.19	-0.29	1.00						
ISS	0.14	0.37	0.31	0.37	0.18	-0.22	1.00					
CEI	0.18	0.59	0.45	0.51	0.27	-0.26	0.54	1.00				
MOM	0.02	0.10	0.03	0.21	0.01	0.10	0.20	0.06	1.00			
OSC	0.27	0.51	0.25	0.77	-0.24	-0.14	0.19	0.33	0.20	1.00		
DSR	0.14	0.54	0.33	0.56	-0.37	-0.06	0.09	0.18	0.15	0.63	1.00	
COMB	0.39	0.73	0.51	0.78	0.11	0.11	0.53	0.64	0.42	0.66	0.55	1.00

Table 2. 2. (continued)

Panel B	**Raw & benchmark-adjusted anomaly returns**							
	Raw ret α		CAPM α		Fama3 α		Fama5 α	
GP	0.51	*	0.62	**	0.61	**	0.42	
	[1.63]		[2.08]		[1.97]		[1.44]	
AG	1.04	***	1.07	***	1.39	***	1.16	***
	[4.29]		[4.01]		[5.90]		[5.04]	
IVA	0.72	**	0.72	*	1.01	***	0.90	***
	[2.26]		[1.93]		[3.42]		[3.45]	
ROA	1.46	***	1.58	***	2.20	***	1.67	***
	[3.19]		[3.68]		[6.25]		[5.76]	
NOA	0.9	***	0.87	***	0.74	**	0.71	**
	[2.64]		[2.91]		[2.39]		[2.43]	
ACC	0.81		0.68		0.27		0.56	
	[0.99]		[0.74]		[0.26]		[0.68]	
ISS	0.60	*	0.57	**	0.94	***	0.72	**
	[1.77]		[2.12]		[2.94]		[2.30]	
CEI	1.77	***	1.75	***	2.32	***	1.90	***
	[4.11]		[4.9]		[5.83]		[5.32]	
MOM	0.66	*	0.73	***	0.85	**	0.85	***
	[1.67]		[3.21]		[2.51]		[2.74]	
OSC	0.57		0.70	*	1.14	***	0.85	***
	[1.53]		[1.91]		[4.08]		[2.96]	
DSR	-0.15		-0.06		0.26		0.01	
	[-0.45]		[-0.16]		[0.70]		[0.04]	
COMB	0.81	***	0.84	***	1.07	***	0.89	***
	[4.48]		[5.63]		[7.16]		[6.44]	

Notes Panel A of Table 2.2 reports the Pearson correlations among the eleven anomaly-sorted long–short excess return portfolios, gross profitability (GP), asset growth (AG), investment-to-assets (IVA), return on assets (ROA), net operating assets (NOA), accruals (ACC), net equity issuance (ISS), composite equity issuance (CEI), momentum (MOM), O-score (OSC), downside risk (DSR), and the composite anomaly portfolio (COMB), over the sample period January 2000 to December 2020. Panel B reports the time-series averages of the monthly excess returns and the corresponding CAPM-, Fama–French three-factor (FF3)-, and Fama–French five-factor (FF5)-adjusted alphas for each portfolio. Excess returns and alphas are expressed in percentage terms by multiplying the decimal values by 100. Newey–West t-statistics with 12-month lags are reported in parentheses. $*$, $**$, and $***$ indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

statistically significant positive excess returns over the sample period. Profitability-related anomalies (GP and ROA) deliver positive High–Low spreads, whereas investment-related anomalies (AG, IVA and NOA) deliver negative Low–High spreads, consistent with the predictions of the US literature. Distress-based anomalies (OSC and DSR) also yield positive long–short excess returns, indicating that financial distress risk is priced in the Chinese A-share market. The composite portfolio produces an average monthly excess return of X, suggesting that anomaly profits are not confined to a specific characteristic dimension.

After factor adjustments, a subset of anomalies remains statistically significant. Profitability and distress-related anomalies retain significant alphas under the five-factor model, whereas investment-related anomalies attenuate but do not fully disappear. The persistence of positive abnormal returns after controlling for market, size, value, profitability and investment risk factors indicates that standard risk models do not fully span the return premia associated with these characteristics in this market. These results are broadly consistent with Stambaugh et al. (2012) and Chu et al. (2020), but differ in that anomaly profits remain economically meaningful in China during a period in which several US-based anomaly premia have weakened.

Taken together, the evidence confirms the presence of return anomalies in China's A-share market and motivates an examination of whether firm-level transparency conditions their pricing. Given the documented role of heterogeneous beliefs, disclosure frictions and limited arbitrage in shaping asset prices in emerging markets, the interaction between anomaly profitability and firm opacity represents an important mechanism through which mispricing may persist.

2.5.2. Anomalies under High and Low OPAC Classification

Table 2.3. examines whether anomaly-based long–short returns differ across firms with high and low levels of opacity. Firms are classified into two groups using the OPAC measure: High OPAC corresponds to firms rated A or B high disclosure), whereas Low OPAC corresponds to firms rated C or D (low disclosure). For each anomaly, we report the average monthly excess returns of the long–short strategy, as well as the associated long-leg and short-

leg components, together with Newey–West t-statistics using 12-month lags. Excess returns are multiplied by 100 for reporting purposes.

Table 2. 3. Anomaly Returns by Information Opacity (OPAC)

Panel A Long – short strategy

Anomaly	Low OPAC	t-value	High OPAC	t-value	Low-High	t-value
GP	0.37***	[5.29]	0.24***	[7.69]	0.13***	[3.28]
AG	1.42***	[26.07]	1.37***	[59.29]	0.06*	[1.88]
IVA	-1.11***	[-15.64]	-1.20***	[-42.65]	0.10**	[2.51]
ROA	1.53***	[14.70]	1.25***	[28.99]	0.27***	[4.72]
NOA	-0.66***	[-8.91]	-0.76***	[-22.06]	0.11**	[2.46]
ACC	1.14***	[6.07]	0.90***	[12.41]	0.24**	[2.30]
ISS	-0.45***	[-5.78]	-0.52***	[-16.33]	0.08*	[1.81]
CEI	1.78***	[19.52]	1.58***	[41.47]	0.20***	[3.92]
MOM	0.71***	[7.70]	0.62***	[14.81]	0.09*	[1.67]
OSC	0.79***	[9.24]	0.50***	[13.87]	0.28***	[5.92]
DSR	0.33***	[4.80]	0.04	[1.32]	0.29***	[7.39]
COMB	0.93***	[22.26]	0.82***	[46.54]	0.12***	[4.96]

Panel B Long leg

Anomaly	Low OPAC	t-value	High OPAC	t-value	Low-High	t-value
GP	0.44***	[3.22]	0.24***	[4.12]	0.20***	[5.00]
AG	0.75***	[5.26]	0.71***	[11.48]	0.03	[0.79]
IVA	-0.15	[-0.92]	-0.21***	[-2.94]	0.06	[1.13]
ROA	1.00***	[7.70]	0.83***	[15.23]	0.17***	[4.40]
NOA	0.53***	[3.56]	0.50***	[7.94]	0.03	[0.64]
ACC	1.96***	[8.97]	1.53***	[17.79]	0.43***	[7.14]
ISS	0.02	[0.16]	0.09	[1.45]	-0.07	[-1.58]
CEI	0.90***	[6.08]	0.67***	[10.25]	0.24***	[5.22]
MOM	0.70***	[4.90]	0.59***	[9.61]	0.10**	[2.44]
OSC	0.63***	[4.80]	0.41***	[7.36]	0.22***	[5.67]
DSR	0.18	[1.36]	0.11**	[2.04]	0.06*	[1.67]
COMB	0.63***	[4.52]	0.50***	[8.38]	0.13***	[3.23]

Table 2. 3. (continued)

Panel C Short leg

Anomaly	Low OPAC	t-value	High OPAC	t-value	Low-High	t-value
GP	0.07	[0.43]	0.00	[-0.01]	0.07	[0.40]
AG	-0.68***	[-4.39]	-0.65***	[-8.96]	-0.02	[-0.14]
IVA	0.96***	[6.28]	1.00***	[13.48]	-0.04	[-0.26]
ROA	-0.53***	[-3.12]	-0.42***	[-5.26]	-0.11	[-0.59]
NOA	1.19***	[7.79]	1.27***	[16.88]	-0.08	[-0.46]
ACC	0.82***	[5.19]	0.63***	[8.31]	0.19	[1.11]
ISS	0.47***	[3.22]	0.62***	[8.65]	-0.15	[-0.92]
CEI	-0.88***	[-5.69]	-0.91***	[-12.58]	0.03	[0.21]
MOM	-0.01	[-0.04]	-0.02	[-0.30]	0.02	[0.09]
OSC	-0.15	[-0.91]	-0.09	[-1.13]	-0.06	[-0.35]
DSR	-0.15	[-0.98]	0.07	[0.98]	-0.23	[-1.34]
COMB	0.10	[0.67]	0.13*	[1.88]	-0.03	[-0.21]

Notes Table 2.3 examines whether anomaly-based long–short returns differ across firms with high and low levels of opacity. Firms are classified into two groups using the OPAC measure: High OPAC corresponds to firms rated A or B (high disclosure), whereas Low OPAC corresponds to firms rated C or D (low disclosure). For each anomaly, we report the average monthly excess returns of the long–short strategy, as well as the associated long-leg and short-leg components, together with Newey–West t-statistics using 12-month lags. Excess returns are multiplied by 100 for reporting purposes.

Panel A of Table 2.3 shows that the average long–short excess returns are uniformly higher for high-opacity firms than for low-opacity firms. Although several anomalies yield negative long–short averages in levels (e.g., IVA, NOA & ISS), the difference between low OPAC and high OPAC portfolios is positive and statistically significant in all cases. These results support the hypothesis that opacity strengthens anomaly profits, consistent with the view that low information disclosure exacerbates mispricing and slows the incorporation of firm-specific information into prices. The composite anomaly portfolio (COMB) earns 12 basis points per month under High OPAC relative to Low OPAC, with a t-statistic of 4.96. Among the individual anomalies, DSR displays the largest difference of 29 basis points per month (t=7.39), whereas AG shows the smallest difference of 6 basis points (t=1.88). Overall, the evidence in Panel A provides strong support for the first hypothesis that anomaly profits are amplified among more opaque firms.

Several individual anomaly portfolios, particularly IVA, NOA, and ISS, display negative long–short returns within the OPAC-conditioned groups. These results should be interpreted considering the portfolio construction used in Table 2.3. Unlike the unconditional anomaly portfolios reported in Table 2.2, Table 2.3 further partitions firms according to their OPAC classifications before comparing anomaly returns. This additional conditioning changes the composition of the underlying anomaly portfolios and may therefore affect both the magnitude and the realized sign of the corresponding return spreads.

This issue is particularly relevant for IVA, NOA, and ISS. IVA and NOA capture investment-related firm characteristics, while ISS reflects equity-financing activity. Although these measures represent different dimensions of firm behaviour, they are all related to corporate investment, financing, and balance-sheet expansion decisions and may therefore be particularly sensitive to changes in portfolio composition. Accordingly, the negative signs observed for several OPAC-conditioned portfolios are interpreted cautiously and should not be viewed as evidence that the corresponding underlying anomalies are absent.

More importantly, the objective of Table 2.3 is to examine whether anomaly returns differ systematically across firms with different information environments. The relevant evidence is therefore the difference in anomaly returns between the Low- and High-OPAC groups, rather than the unconditional sign of each individual conditional portfolio. This distinction allows the opacity-related result to be evaluated separately from the standalone profitability of the underlying anomaly strategy."

Panels B and C decompose the long–short strategies into their long and short legs. Consistent with the aggregate long–short results, long legs exhibit stronger performance under High OPAC than under Low OPAC for most anomalies. Eight out of twelve long-leg differences are positive and statistically significant. This pattern is consistent with the interpretation that opacity delays the correction of undervaluation: when disclosure is limited, sentiment-driven or inattentive traders may underreact to firm fundamentals, allowing long positions in high-opacity firms to earn higher subsequent returns.

By contrast, short-leg differences are generally small and statistically insignificant. Most OPAC differences in Panel C are negative and insignificant, implying that opacity does not meaningfully alter the payoff to shorting overpriced stocks. This pattern is consistent with the third hypothesis, which predicts that opacity should not materially affect short-leg performance, as short selling tends to require information precision and trading sophistication, and is therefore more sensitive to environments with high disclosure rather than opacity.

Taken together, Table 2.3 shows that opacity primarily affects the long side of anomaly strategies, reflecting stronger returns among undervalued firms when disclosure is limited, while leaving the short side largely unaffected. These results reinforce the role of disclosure in shaping anomaly profitability in the Chinese A-share market.

2.5.3. Impact of OPAC on Anomalies

Tables 2.4 and 2.5 examine how firm-level opacity relates to anomaly returns after controlling for common risk factors. Table 2.4 reports estimates of equation (4), where monthly anomaly portfolio excess returns are regressed on the OPAC index and the Fama–French five factors. Table 2.5 extends this specification by adding the mainland China EPU index, corresponding to equation (5), to assess whether the relation between opacity and anomaly returns remains stable after controlling for macroeconomic policy uncertainty. In both tables, Panel A reports results using the contemporaneous OPAC measure at the portfolio-formation date, whereas Panel B reports robustness checks using OPAC lagged by twelve months.

As with other proxy-based measures of the corporate information environment, OPAC may be correlated with other firm characteristics. For example, firms with different governance structures, market visibility, or liquidity conditions may also differ systematically in their disclosure quality. The empirical results should therefore be interpreted as evidence of a systematic association between disclosure-related information opacity and anomaly profitability rather than as evidence that OPAC is completely independent of these related firm characteristics. The Fama–French five-factor controls and the extended specification incorporating Economic Policy Uncertainty reduce the possibility that the documented relationship reflects common risk exposures or broader macroeconomic conditions. Nevertheless, further analysis using additional firm-level controls and alternative measures of information opacity would provide useful validation of this relationship.

Table 2. 4. Excess Return Regressions on Information Opacity (OPAC)

Panel A OPAC coefficients (contemporaneous)

Anomaly	Long leg		Short leg		Long - short	
	$\hat{\beta}$	t-value	$\hat{\beta}$	t-value	$\hat{\beta}$	t-value
GP	-0.08**	[-2.02]	-0.01	[-0.27]	-0.07***	[-3.13]
AG	0.01	[0.18]	0.04	[1.18]	-0.04**	[-2.22]
IVA	0.02	[0.56]	0.06*	[1.65]	-0.04**	[-2.46]
ROA	-0.08***	[-3.58]	0.10***	[3.18]	-0.18***	[-6.04]
NOA	0.02	[0.49]	0.09**	[2.22]	-0.07***	[-4.95]
ACC	-0.17***	[-3.02]	-0.11***	[-4.94]	-0.06*	[-1.84]
ISS	0.06	[1.54]	0.13***	[3.00]	-0.06***	[-2.97]
CEI	-0.10**	[-2.14]	0.01	[0.35]	-0.12***	[-5.68]
MOM	-0.02	[-0.48]	0.03	[1.02]	-0.05***	[-2.73]
OSC	-0.10***	[-2.65]	0.09***	[3.83]	-0.19***	[-12.28]
DSR	-0.02	[-0.77]	0.17***	[7.71]	-0.19***	[-7.63]
COMB	-0.04	[-1.06]	0.05	[1.09]	-0.07***	[-7.16]

Panel B OPAC coefficients (12-month lag predictive)

Anomaly	Long leg		Short leg		Long - short	
	$\hat{\beta}$	t-value	$\hat{\beta}$	t-value	$\hat{\beta}$	t-value
GP	-0.13***	[-5.69]	-0.05	[-1.34]	-0.08***	[-3.27]
AG	-0.04	[-0.95]	-0.02	[-0.41]	-0.03	[-1.61]
IVA	-0.01	[-0.15]	0.00	[0.03]	-0.01	[-0.65]
ROA	-0.13***	[-5.02]	0.01	[0.33]	-0.14***	[-4.82]
NOA	-0.06	[-1.45]	0.04	[0.73]	-0.10***	[-5.17]
ACC	-0.31***	[-5.22]	-0.22***	[-4.75]	-0.09**	[-2.00]
ISS	0.01	[0.25]	0.09**	[2.23]	-0.08***	[-3.49]
CEI	-0.16***	[-3.25]	-0.01	[-0.11]	-0.16***	[-8.67]
MOM	-0.06	[-1.37]	-0.02	[-0.61]	-0.04**	[-2.06]
OSC	-0.15***	[-4.88]	0.05***	[3.12]	-0.20***	[-9.79]
DSR	-0.08***	[-3.84]	0.13***	[3.08]	-0.21***	[-8.62]
COMB	-0.10***	[-2.68]	0.00	[0.02]	-0.07***	[-6.41]

Notes This table reports coefficient estimates from regressions of monthly anomaly excess returns on OPAC. Panel A uses contemporaneous OPAC values, while Panel B includes OPAC lagged by 12 months to examine its predictive effects. Excess returns are measured relative to the risk-free rate. The regressions do not include additional risk factors. All coefficients are scaled by 100 for ease of interpretation. Newey–West standard errors with 12 lags are used to compute t-statistics. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

In Panel A of Table 2.4, the long–short strategies exhibit a clear and consistent pattern. All twelve OPAC betas on long–short excess returns are negative and statistically significant, with magnitudes ranging from −0.04 to −0.19 after scaling by 100. The composite anomaly portfolio (COMB) has a relatively moderate OPAC beta of −0.07 ($t = -7.16$), indicating that a one-unit increase in the transparency index is associated with a reduction of approximately 7 basis points per month in the composite long–short excess return. Anomalies such as OSC, ROA and DSR display relatively large negative exposures, with long–short OPAC betas close to −0.18 or below in absolute value. Given that OPAC is coded so that higher values correspond to higher disclosure quality, these negative coefficients imply that anomaly profits are larger when firms are more opaque, consistent with the hypothesis that low information disclosure strengthens return predictability.

The decomposition into long and short legs in Panel A further clarifies the mechanism. On the long side, OPAC betas are negative and significant for several anomalies, including ACC, CEI, DSR, GP, OSC and ROA, indicating that higher disclosure quality is associated with lower average returns on the long legs of these strategies. On the short side, OPAC betas are predominantly positive and often significant (e.g., ACC, DSR, ISS, IVA, NOA, OSC and ROA), implying that more transparent firms tend to exhibit higher average returns on the short legs. Taken together, the negative long–short betas arise from a combination of weaker long-leg returns and stronger short-leg returns as disclosure improves, which is consistent with the earlier portfolio evidence that opacity primarily amplifies anomaly profits through undervalued firms.

Panel B of Table 2.4 repeats the analysis using OPAC lagged by twelve months. The results are broadly similar and, if anything, slightly stronger in magnitude. Long–short OPAC betas remain negative and statistically significant for all anomalies, with values ranging between approximately −0.01 and −0.21. The composite portfolio again shows a negative and significant coefficient (−0.07, $t = -6.41$). This robustness to lagging the opacity measure mitigates concerns about mechanical timing overlap between OPAC ratings and return measurement, and supports the interpretation of opacity as a predictor of subsequent anomaly profits rather than a purely contemporaneous correlate.

Table 2. 5. Extended Excess Return Regressions on Information Opacity (OPAC)

Panel A OPAC coefficients (contemporaneous)

Anomaly	Long leg		Short leg		Long - short	
	$\hat{\beta}$	t-value	$\hat{\beta}$	t-value	$\hat{\beta}$	t-value
GP	-0.17***	[-19.22]	-0.10***	[-9.17]	-0.07***	[-5.47]
AG	-0.08***	[-10.77]	-0.09***	[-11.57]	0.01	[0.78]
IVA	-0.13***	[-13.40]	-0.05***	[-5.91]	-0.07***	[-6.14]
ROA	-0.12***	[-12.49]	-0.07***	[-7.74]	-0.05***	[-3.67]
NOA	-0.04***	[-4.44]	-0.04***	[-4.93]	-0.01	[-0.69]
ACC	-0.24***	[-9.83]	-0.14***	[-9.06]	-0.10***	[-3.49]
ISS	-0.06***	[-6.18]	0.05***	[4.36]	-0.10***	[-7.70]
CEI	-0.18***	[-17.07]	-0.12***	[-11.86]	-0.07***	[-4.95]
MOM	-0.13***	[-12.84]	-0.10***	[-8.72]	-0.03*	[-1.81]
OSC	-0.16***	[-16.96]	-0.07***	[-7.84]	-0.09***	[-8.35]
DSR	-0.08***	[-10.14]	0.01	[1.63]	-0.10***	[-10.63]
COMB	-0.13***	[-17.05]	-0.07***	[-10.11]	-0.03***	[-4.80]

Panel B OPAC coefficients (12-month lag predictive)

Anomaly	Long leg		Short leg		Long - short	
	$\hat{\beta}$	t-value	$\hat{\beta}$	t-value	$\hat{\beta}$	t-value
GP	-0.17***	[-19.35]	-0.09***	[-8.12]	-0.08***	[-6.28]
AG	-0.08***	[-10.71]	-0.11***	[-13.75]	0.02***	[2.80]
IVA	-0.12***	[-11.82]	-0.06***	[-6.55]	-0.05***	[-4.49]
ROA	-0.12***	[-12.41]	-0.12***	[-12.16]	-0.01	[-0.61]
NOA	-0.07***	[-7.02]	-0.05***	[-5.64]	-0.03**	[-2.13]
ACC	-0.33***	[-13.22]	-0.19***	[-11.59]	-0.14***	[-4.66]
ISS	-0.07***	[-7.05]	0.06***	[5.29]	-0.12***	[-8.91]
CEI	-0.18***	[-17.38]	-0.09***	[-9.33]	-0.09***	[-6.42]
MOM	-0.13***	[-12.21]	-0.10***	[-8.44]	-0.03	[-1.61]
OSC	-0.16***	[-16.33]	-0.07***	[-7.02]	-0.10***	[-8.32]
DSR	-0.09***	[-10.81]	0.01	[1.62]	-0.11***	[-10.88]
COMB	-0.14***	[-18.25]	-0.07***	[-10.84]	-0.03***	[-4.66]

Notes This table reports estimated OPAC betas from regressions of monthly anomaly excess returns on OPAC after controlling for the Fama–French five factors and Economic Policy Uncertainty (EPU). Panel A uses contemporaneous OPAC; Panel B uses a 12-month lag to assess predictive effects. Excess returns are measured relative to the risk-free rate. Betas are multiplied by 100 for presentation. Newey–West standard errors with 12 lags are used to compute t-statistics. *, ** and *** denote significance at the 10%, 5% and 1% levels, respectively.

Panel A of Table 2.5 introduces the EPU index as an additional control to test whether the OPAC–anomaly relation is driven by time-varying macroeconomic uncertainty. After controlling for both the Fama–French five factors and EPU, OPAC betas on long–short returns remain predominantly negative and statistically significant. In Panel A, ten out of twelve long–short strategies display significantly negative OPAC coefficients, with magnitudes between −0.10 and −0.03. The composite portfolio retains a negative and significant OPAC beta of −0.03 ($t = -4.80$), indicating that the cross-sectional effect of opacity on anomaly returns is not subsumed by common risk factors or macroeconomic policy uncertainty.

On the long side of Table 2.5, OPAC betas are negative and highly significant for all anomalies, confirming that higher disclosure quality is systematically associated with lower subsequent long-leg returns even after accounting for EPU. For the short legs, OPAC betas are mostly negative and significant, with the notable exception of ISS, which exhibits a positive and significant coefficient, and DSR, which is small and statistically insignificant. These patterns indicate that the dominant channel remains a reduction in long-leg returns as disclosure improves, while the impact on short-leg returns is more heterogeneous across anomalies.

Panel B of Table 2.5, which uses OPAC lagged by one year in the presence of EPU, delivers similar conclusions. Long–short OPAC betas remain negative and statistically significant for most anomalies, and the composite portfolio again exhibits a negative and significant coefficient. The persistence of negative long–short betas across specifications, with and without lagging OPAC, and with and without EPU, suggests that opacity is robustly associated with stronger anomaly profits in the Chinese A-share market.

Overall, the evidence in Tables 2.4 and 2.5 is consistent with the hypothesis that lower disclosure quality strengthens anomaly-based return predictability. Higher opacity is associated with larger long–short excess returns, primarily through stronger long-leg performance and, in some cases, weaker short-leg performance. This pattern complements the portfolio-sorting results in Tables 2.2 and 2.3 and reinforces the view that information disclosure is an important state variable for the pricing of anomaly characteristics in this market.

The regression results provide a complementary assessment of the portfolio evidence reported in Table 2.3. Whereas Table 2.3 conditions the anomaly portfolios on OPAC classifications and therefore changes the composition of the underlying portfolios, Tables 2.4 and 2.5 directly estimate the relationship between OPAC and anomaly portfolio returns. The predominantly negative OPAC coefficients on long–short returns indicate that higher disclosure quality is generally associated with weaker anomaly profitability. This pattern is broadly preserved when OPAC is lagged by twelve months and when the Fama–French five factors and Economic Policy Uncertainty are included as controls. The consistency of the regression evidence therefore supports the main opacity-related conclusion while allowing the negative signs observed for several individual OPAC-conditioned portfolios to be interpreted more cautiously.

As an additional assessment of the robustness of the OPAC measure, I examined alternative indicators that capture related aspects of firms' information environments. Analyst-based measures available from CSMAR were considered because analyst evaluations may provide complementary information about firms' disclosure environments. One potentially relevant measure is the analyst ranking variable (Rank), which ranges from 1 (the first) to 5 (the fifth). I conducted preliminary matching and portfolio analyses using this measure to assess its feasibility as an alternative proxy for OPAC.

However, the analyst-based measure is available only for firms that receive analyst coverage and evaluation. This results in substantially narrower firm coverage than the exchange-based OPAC measure and materially reduces the number of firms available for portfolio formation. Because analyst coverage is itself related to firm characteristics and market attention, the resulting subsample may differ systematically from the broader A-share sample used in the baseline analysis. Consequently, portfolios formed using the analyst-based measure are not directly comparable with the baseline OPAC portfolios, and differences in portfolio performance may reflect changes in sample composition rather than differences in the underlying opacity measure. For this reason, I do not report the preliminary estimates as a formal robustness test. Nevertheless, this exercise highlights the limitation of relying on a single primary opacity measure and motivates further validation using alternative disclosure-based or market-based measures when sufficiently comprehensive firm-level data become available.

2.6. Conclusion

Recent evidence suggests that the Chinese stock market has gradually evolved toward greater pricing efficiency, with investors becoming more attentive to firm-specific information (Hilliard & Zhang, 2015). In this setting, the information environment surrounding listed firms becomes relevant for understanding return predictability. This essay examines whether firm-level disclosure quality conditions anomaly returns in the Chinese A-share market. Using 11 anomalies commonly studied in the literature (Chu et al., 2020; Stambaugh et al., 2012), we evaluate the link between firm transparency and anomaly profitability.

Two main findings emerge. First, the 11 anomalies, as well as the combined anomaly, generate statistically significant excess returns, and these profits persist after controlling for the market factor, the Fama–French three-factor model, and the Fama–French five-factor model. Second, both contemporaneous and predictive regressions indicate that firm transparency is negatively related to anomaly returns, implying that anomalies tend to be stronger among firms with lower levels of disclosure. Taken together, the evidence suggests that firm transparency conditions the magnitude of anomaly-based mispricing in this market.

The objective of this essay is not to provide a full explanation for each anomaly, but rather to assess whether the corporate information environment affects the extent of mispricing more broadly. While the analysis incorporates the five Fama–French factors and policy uncertainty, additional firm-level or macroeconomic controls may refine this relationship in future work. The analysis also relies primarily on the exchange-based OPAC measure to capture firm-level disclosure quality. Although alternative analyst-based measures were examined, their substantially narrower firm coverage limits their comparability with the baseline portfolio analysis. Future research could therefore further validate the findings using alternative disclosure-based or market-based opacity measures when sufficiently comprehensive firm-level data become available. The results also show that transparency affects the long and short legs of anomaly strategies in different ways, raising the possibility that institutional features such as short-selling arrangements may interact with disclosure to shape return predictability. Investigating this interaction represents a natural direction for future research.

CHAPTER THREE: ESSAY TWO

Firm Characteristics and Anomaly Returns: Evidence from China's Short-Selling Expansion

This chapter presents the second essay, which examines the impact of the gradual relaxation of the short-sale ban on the profitability of characteristic-based return anomalies in the Chinese A-share market.

3.1. Introduction

Asset pricing anomalies, referring to persistent return patterns unexplained by standard risk models, continue to challenge the Efficient Market Hypothesis (EMH). A large body of empirical work shows that strategies based on firm characteristics, such as profitability, investment, accruals, and issuance, yield abnormal returns (Fama & French, 2008; Hou et al., 2015). The persistence of these patterns contradicts the notion of frictionless and fully rational markets, raising the puzzle of why such return predictability survives despite the presence of arbitrageurs.

Limits to arbitrage provide an important explanation. Short-sale constraints restrict pessimistic investors from expressing negative information through trading, resulting in sustained overpricing (Miller, 1977). Arbitrage is also risky and capital-constrained, making it costly even for rational investors to correct mispricing (Shleifer & Vishny, 1997). Empirical evidence supports this view: Stambaugh et al. (2015) find that U.S. anomalies are stronger among stocks facing higher short-sale impediments, and Chen et al. (2002) document that short-selling costs reduce arbitrage effectiveness. These studies suggest that trading frictions, particularly limits to short-sale constraints are widely recognised as one of the key frictions.

The Chinese A-share market offers a particularly relevant setting for testing this mechanism. As the world's second-largest equity market, it displays many anomalies observed in developed markets (Chen et al., 2010; Hou et al., 2020). However, it has historically been subject to stringent short-sale restrictions. Short selling was prohibited until March 2010, when the China Securities Regulatory Commission (CSRC) introduced the initial pilot programme of margin trading and securities lending (MT/SL) with 90 eligible stocks. Since then, the shortable list has expanded through eight regulatory waves, with each wave introducing additional eligible firms and expanding the universe of stocks permitted for short selling. This phased rollout generated cross-sectional variation in the timing of short-sale eligibility across firms, providing a quasi-experimental setting to study the impact of arbitrage constraints (Chang et al., 2013; Chu et al., 2020).

Despite more than a decade of liberalisation, short selling in China remains modest. The supply of lendable shares is concentrated in large-cap, liquid firms. Borrowing costs are high, and administrative barriers persist (Boone & White, 2015). Consequently, the proportion of shortable securities relative to the full market is small, and trading volumes are limited compared to mature markets. This environment provides a unique opportunity to investigate whether incremental increases in arbitrage opportunities, captured by both eligibility for short selling and variation in short-selling intensity, reduce the profitability of anomaly-based strategies.

While anomalies have been extensively documented in developed markets, much less is known about how they interact with arbitrage constraints in emerging economies. In the U.S., Stambaugh et al. (2015) show that anomalies are stronger where short-sale impediments are binding, and Chen et al. (2002) emphasise that borrowing costs and limited lending supply reduce the ability of arbitrageurs to correct mispricing. Cross-country studies also highlight institutional differences: Beber and Pagano (2013) find that restrictions on short selling are associated with greater mispricing and weaker price efficiency. Yet despite its size and importance, the Chinese equity market has not been systematically studied in this regard.

China provides an especially valuable setting. Trading in A-shares is dominated by retail investors, who account for 70–80% of turnover, compared with less than 20% in the U.S., where institutional investors prevail (Allen et al., 2005; Gao et al., 2018). Behavioural biases among retail traders may therefore amplify anomalies. At the same time, regulatory interventions are frequent, and access to short selling has been tightly controlled since the launch of the margin trading and securities lending (MT/SL) programme in 2010. Participation in MT/SL is restricted to institutions and a limited set of qualified investors, leaving most retail participants unable to establish short positions (Chang et al., 2013). Lending supply also remains concentrated in a subset of large-cap firms, and the aggregate balance of margin and short-selling transactions represents only 2–3% of market capitalisation, far below the 20–30% benchmark observed in the U.S. (Bris et al., 2007; D'Avolio, 2002). Moreover, annualised borrowing costs for A-shares are as high as 8–10%, compared with typically below 1% in the U.S. (Chang et al., 2013). These institutional frictions, dominance of retail investors, limited participation in short selling, concentrated lending supply, and high costs, create uncertainty over whether anomaly returns in China reflect risk premia, behavioural biases, or binding trading constraints.

Against this background, the study addresses two questions: Do anomaly returns systematically decline once firms become eligible for short selling, or is actual trading intensity required for meaningful price correction? Within the subset of eligible firms, does the intensity of short-selling activity further explain variation in anomaly profitability?

The contributions are threefold. First, the paper provides the first systematic evidence on how short-sale constraints shape anomaly returns in China, extending the limits-to-arbitrage literature beyond developed markets. Second, it introduces a transaction-based measure of short-selling depth that reflects the intensity of arbitrage activity, distinguishing between the extensive margin of short-sale eligibility and the intensive margin of actual trading activity. Third, by examining 11 anomalies spanning profitability, investment, accruals, issuance, and distress, the study offers a comprehensive view of how short-sale frictions affect heterogeneous forms of mispricing. Together, these contributions provide new insights into the interaction between market frictions and return predictability in an emerging market context.

This study employs a multi-step empirical strategy that integrates characteristic-sorted portfolio tests, a phased difference-in-differences (DiD) framework, and short-selling depth analysis to evaluate how constraints influence anomaly returns in the Chinese A-share market. This approach draws on the theoretical foundation of limits to arbitrage (Miller, 1977; Shleifer & Vishny, 1997) and recent empirical studies on short-sale frictions (Chu et al., 2020; Stambaugh et al., 2015), while adapting them to the unique regulatory environment of China.

The first part examines the firm-level characteristic based anomaly returns of long-short portfolios. At the end of each fiscal year, firms are ranked according to 11 anomaly characteristics: profitability (Gross Profitability, ROA), investment activity (Asset Growth, Investment-to-Assets), accruals, equity issuance (Net Issuance, Composite Equity Issuance), and mispricing-related or distress signals (Ohlson O-score, Downside Risk, Momentum). Firms are sorted into deciles, and long–short portfolios are formed by taking value-weighted positions in the top and bottom groups. Portfolios are held for 12 months in year t+1, using lagged December market capitalisation to compute anomaly returns. This design ensures comparability with the standard procedures in the anomaly literature (Fama & French, 1993;

Hou et al., 2015) and avoids look-ahead bias.

The second part exploits the phased liberalisation of short selling in China. Since 2010, the CSRC has expanded the list of shortable stocks through eight regulatory waves. For each expansion, newly added firms constitute the treatment group, while those not yet included serve as controls. A DiD specification compares anomaly returns in the twelve months before and after inclusion, effectively isolating the effect of short-sale eligibility from broader market shocks or firm-specific characteristics. By aligning the treatment windows across different waves, the framework leverages quasi-experimental variation to estimate how eligibility alters anomaly profitability, extending the single-pilot designs used in prior studies to a multi-wave setting.

The third part evaluates the role of trading intensity within shortable firms. A monthly measure of short-selling depth (SSD) is constructed as the ratio of the securities-lending balance to market capitalisation. This stock-based measure captures the net shorting pressure applied by arbitrageurs and reflects how aggressively short sellers are positioned in a given stock. Based on this measure, shortable firms are sorted into high- and low-SSD groups, and anomaly returns are compared across these subsamples. In addition, risk-adjusted regressions of anomaly return on an SSD dummy are estimated, controlling for exposures to the Fama–French five factors. This step distinguishes between the extensive margin of arbitrage, defined by eligibility, and the intensive margin, reflected in trading depth.

Together, these three layers of analysis provide complementary perspectives. Portfolio sorts reveal whether anomaly profits differ between shortable and non-shortable stocks. The DiD framework exploits phased policy reforms to identify causal effects of short-sale eligibility. The SSD analysis incorporates transaction-level information to capture heterogeneity in the intensity of short-selling activity. Taken as a whole, this integrated design enables a comprehensive evaluation of how short-sale frictions shape the persistence of anomaly returns in the Chinese A-share market.

The results of this study indicate that short-sale frictions exert a significant influence on the persistence of anomaly-based returns in the Chinese A-share market. Consistent with earlier work, long–short portfolios based on profitability, investment, accruals, and issuance characteristics deliver abnormal profits when formed across the full sample of firms. These excess returns highlight the presence of systematic mispricing beyond the scope of traditional risk models. However, eligibility alone does not systematically attenuate anomaly returns across waves. The difference-in-differences analysis shows heterogeneous effects across expansion waves, suggesting that short-sale liberalisation does not mechanically eliminate anomaly-based return predictability.

The effect is not uniform across all strategies. The decline is sharpest for anomalies related to investment intensity and earnings quality, such as Asset Growth, Investment-to-Assets, and Accruals. These results are consistent with the interpretation that overinvestment and earnings manipulation are harder for uninformed investors to recognise, and thus particularly dependent on arbitrage by sophisticated traders. In contrast, anomalies tied to issuance activity or profitability show weaker or less consistent attenuation, implying that certain pricing irregularities may reflect risk exposures or behavioural biases that are less easily corrected through short selling.

The role of trading intensity provides an additional layer of evidence. Within the subset of shortable stocks, anomaly returns vary systematically with short-selling depth, defined as the net shorting balance relative to market capitalisation. In particular, anomaly profitability is substantially stronger among stocks with low short-selling depth and weakens markedly as short-selling intensity increases. Stocks with higher SSD levels display significantly lower anomaly returns, even after controlling for market, size, value, profitability, and investment factors. This finding underscores that the intensity of arbitrage, not merely the eligibility to short, is an important determinant of anomaly persistence. It also highlights that the effectiveness of short sellers depends on both regulatory access and the scale of actual trading activity.

Overall, the findings suggest that short sellers function as informed arbitrageurs whose participation mitigates mispricing. Yet, the corrective effect is uneven across anomaly categories and incomplete at the market level. Structural frictions, including limited lending supply, high borrowing costs, and administrative restrictions, continue to constrain arbitrage in China. These institutional features help explain why some anomalies remain profitable despite gradual liberalisation. The evidence therefore contributes to a more nuanced understanding of how short-sale frictions interact with asset pricing anomalies in emerging markets, showing both the progress and the remaining limits of efficiency improvements in the Chinese equity market.

The chapter is organised as follows. Section 3.2 reviews the related literature on asset pricing anomalies, limits to arbitrage, and short-sale constraints. Section 3.3 describes the empirical methodology, including the construction of 11 anomaly characteristics, portfolio formation procedures, the difference-in-differences design, and short-selling depth measures. Section 3.4 presents the empirical results, including portfolio-based comparisons, phased policy analyses, and cross-sectional tests on the role of short-selling depth. Section 3.5 concludes the second essay. The reference list for this chapter is included at the end of this thesis.

3.2. Literature Review

3.2.1. Efficient Market Hypothesis and the Puzzle of Anomalies

The Efficient Market Hypothesis (EMH) has long served as a cornerstone of modern finance, positing that security prices fully incorporate all available information, leaving no systematic opportunities for abnormal returns (Fama, 1970). In its strong form, EMH suggests that even private information is impounded into prices through informed trading, while the semi-strong form emphasises the rapid adjustment of prices to all publicly available information. Under this paradigm, predictable return patterns should be arbitraged away, as rational investors act on mispricing until it disappears. The implication is that no simple trading rule based on publicly observable firm characteristics should yield consistent excess returns once risk is properly accounted for.

Yet an extensive body of empirical research documents systematic deviations from the predictions of EMH. These deviations, commonly termed asset pricing anomalies, arise when firm characteristics predict cross-sectional differences in expected returns beyond what can be explained by standard factor models. Early evidence includes the size and value effects, with small firms and high book-to-market firms earning higher average returns than predicted by the Capital Asset Pricing Model (Banz, 1981; Fama & French, 1992, 1993). Later work uncovers momentum, where past winners continue to outperform losers over intermediate horizons (Jegadeesh & Titman, 1993), and reversal patterns at long horizons (De Bondt & Thaler, 1985). More recently, profitability and investment-related anomalies, such as gross profitability, return on assets, and asset growth, have been documented across markets (Fama & French, 2008; Hou et al., 2015).

The persistence of these anomalies represents a puzzle. If markets are competitive and arbitrage is unconstrained, such return patterns should vanish once discovered, as arbitrageurs exploit them until profits are eliminated. However, anomalies have been found not only in U.S. equities but also in international and emerging markets, including China (Chen et al., 2010; Griffin et al., 2003). Furthermore, Fama and French (2008) show that many anomalies remain significant even after adjustments for their three-factor and later five-factor models. This persistence raises the question of whether anomalies reflect unpriced risk factors, behavioural biases, or market frictions that impede arbitrage.

One line of research argues that anomalies may proxy for risk exposures not captured by traditional models. For example, investment-based models (e.g., Hou et al., 2015) interpret profitability and investment patterns as reflections of firms' risk–return trade-offs. Yet the fact that many anomalies yield large and economically meaningful excess returns, particularly in settings with less developed arbitrage infrastructure, suggests that frictions also play an important role. Among these frictions, short-sale constraints are frequently highlighted. When pessimistic investors cannot easily take short positions, prices may disproportionately reflect the views of optimistic traders, sustaining overpricing (Miller, 1977). The persistence of anomaly returns, therefore, may be tied to the inability of markets to incorporate negative information efficiently.

The Chinese A-share market provides an especially relevant setting to study this puzzle. As the world's second-largest equity market, it exhibits anomalies similar to those documented in developed economies (Hou et al., 2020), yet has historically been subject to stringent restrictions on short selling. Short selling was prohibited entirely until 2010, when the China

Securities Regulatory Commission introduced margin trading and securities lending on a limited basis. Since then, liberalisation has progressed gradually but remains incomplete. This combination of well-documented anomalies and significant short-sale frictions offers a natural environment to investigate whether constraints on arbitrage underpin the persistence of predictable return patterns.

3.2.2. Limits to Arbitrage and Short-Sale Constraints

The persistence of anomalies despite widespread awareness among investors has led to the development of the limits-to-arbitrage framework, which emphasises the frictions and risks that prevent rational traders from fully eliminating mispricing. Shleifer and Vishny (1997) argue that arbitrage is constrained not only by transaction costs and institutional restrictions but also by risk and capital constraints. Arbitrageurs typically rely on external capital, and when mispricing persists or worsens in the short-run, they may face redemptions or margin calls, forcing them to liquidate positions prematurely. These dynamic limits their willingness to take large positions against mispriced securities, allowing anomalies to persist even when they are well known.

Among the various frictions identified in this literature, short-sale constraints play an important role. Miller (1977) was among the first to show that when short selling is costly or restricted, stock prices disproportionately reflect the valuations of optimistic investors, since pessimists cannot fully express their views through trading. This leads to systematic overpricing, as negative information is underrepresented in the price formation process. Diamond and Verrecchia (1987) formalise this intuition in a rational expectations model, showing that the relaxation of short-sale constraints improves informational efficiency by facilitating the incorporation of negative signals into stock prices. Importantly, their model also highlights that even the potential for short selling disciplines prices, as the threat of arbitrage reduces the likelihood of overvaluation.

Empirical studies provide substantial evidence consistent with these predictions. Chen et al. (2002) find that U.S. stocks with higher short-sale constraints exhibit stronger momentum effects, suggesting that restrictions allow mispricing driven by investor sentiment to persist.

Jones and Lamont (2002) analyse historical data on borrowing fees and demonstrate that stocks with higher shorting costs are more likely to be overpriced and subsequently underperform. Similarly, Asquith et al. (2005) show that heavily shorted stocks deliver lower subsequent returns, consistent with short interest reflecting informed pessimism. Saffi and Sigurdsson (2011) provide cross-country evidence that stocks with more active lending markets, indicating lower short-sale frictions, display higher price efficiency and less severe anomalies. Collectively, these findings support the view that short-sale constraints are a primary mechanism through which limits to arbitrage sustain predictable return patterns.

The interaction between arbitrage limits and investor behaviour further explains why anomalies persist in practice. When shorting is restricted, optimistic investors dominate price formation, and rational arbitrageurs face both implementation costs and the risk of short-term losses (Shleifer & Vishny, 1997). This combination creates a setting where anomalies, such as overpricing linked to issuance, investment, or accruals, can remain profitable strategies for long-only investors. Chu et al. (2020) provide causal evidence using the U.S. Reg SHO pilot programme, showing that relaxing short-sale constraints reduces the profitability of a broad set of anomaly strategies. Their results demonstrate that the availability of short selling directly mitigates mispricing, highlighting the critical role of short-sale mechanisms in the arbitrage process.

Overall, the limits-to-arbitrage framework underscores that anomalies are not purely behavioural or risk-based phenomena but are sustained by institutional and practical barriers to arbitrage, with short-sale constraints being among the most important. By preventing the incorporation of negative information into stock prices, these constraints reduce market efficiency and allow return predictability to persist. The gradual relaxation of such constraints, therefore, provides a natural experiment to study whether expanding arbitrage capacity reduces anomalies, offering direct insights into the mechanisms through which market efficiency evolves.

3.2.3. Empirical Evidence on Anomalies and Short-Sale Constraints

Empirical studies provide extensive evidence that short-sale frictions influence the behaviour of asset pricing anomalies, though the literature emphasises two distinct perspectives. The first strand of research highlights that constraints allow mispricing to persist, leading to stronger observed anomaly spreads. Jegadeesh and Titman (1993) show that momentum strategies deliver significant profits, difficult to reconcile with standard risk models. Chen et al. (2002) further document that momentum profits are amplified among stocks facing greater short-sale impediments, consistent with Miller's (1977) prediction that pessimistic information is not fully incorporated into prices. Extending this evidence, Stambaugh et al. (2012) examine 11 anomalies, including accruals, net issuance, and profitability, and find that returns are concentrated in stocks with high short-sale constraints, implying that anomalies largely reflect overpricing that arbitrageurs are unable to correct. Diether et al. (2009) similarly show that stocks with higher short-sale costs exhibit larger post-earnings-announcement reversals, consistent with persistent mispricing. Collectively, these studies suggest that when shorting is restricted, anomaly strategies appear more profitable because arbitrage is unable to eliminate overvaluation.

A second strand of evidence emphasises that, while anomaly spreads may appear stronger under constraints, their implement ability is reduced. Jones and Lamont (2002) show that stocks with high borrowing fees underperform subsequently, but the costs of establishing short positions make it difficult to exploit these predictable returns. Asquith et al. (2005) find that heavily shorted stocks earn lower subsequent returns, consistent with short interest reflecting informed pessimism, but they also stress that practical barriers limit the extent to which arbitrageurs can act on this information. Chu et al. (2020) provide causal evidence from the U.S. Reg SHO pilot programme, showing that relaxing short-sale constraints significantly reduces the profitability of a broad set of anomalies, underscoring the corrective role of short sellers. Yet, their findings also imply that many anomaly profits documented under constraints are more reflective of unrealised mispricing than tradable opportunities.

Cross-country evidence reinforces both perspectives. Saffi and Sigurdsson (2011) show that markets with deeper lending supply exhibit smaller anomalies and greater price efficiency, while Beber and Pagano (2013) demonstrate that regulatory bans during the financial crisis amplified mispricing and weakened market quality. At the same time, Bris et al. (2007)

document that stricter short-sale restrictions slow the adjustment of prices to negative news, highlighting the persistence of anomalies in constrained environments.

Overall, the empirical literature suggests that short-sale constraints shape anomalies in two ways: they sustain observed anomaly return spreads by limiting the ability of arbitrageurs to correct overpricing, but they also reduce the practical tradability of these returns by raising the costs and risks of short positions. This dual perspective explains why anomalies remain visible in the data yet are difficult to fully exploit in practice. For markets such as China, where short-selling mechanisms have been gradually liberalised, these insights highlight the importance of examining both the persistence of anomaly patterns and the extent to which expanding arbitrage capacity alters their profitability.

3.2.4. Evidence from Emerging Markets and the Chinese Context

While much of the literature on short-sale constraints and anomalies is based on developed markets, emerging markets provide an equally important setting where institutional frictions are often more binding, and investor bases differ substantially. In these markets, short-sale restrictions tend to be stricter, lending supply is limited, and retail investors dominate trading activity, amplifying behavioural biases and mispricing. Studying such environments offers insights into how the effectiveness of arbitrage varies with market development and regulatory design.

Empirical evidence suggests that short-sale constraints are especially influential in emerging markets. Bris et al. (2007) conduct a cross-country analysis of 46 markets and find that short-sale restrictions are associated with slower price adjustments to negative information and greater overvaluation. They also show that countries with more permissive short-selling regimes exhibit smaller post-event drifts, indicating that the ability to short enhances informational efficiency globally. Similarly, Saffi and Sigurdsson (2011) highlight that lending supply in international markets strongly predicts price efficiency, with effects more pronounced in markets where institutional structures are weaker.

The Chinese A-share market provides a particularly valuable case. Prior to 2010, short selling was prohibited, meaning that pessimistic signals could not be incorporated into stock prices. The launch of the margin trading and securities lending (MT/SL) programme in 2010 marked the first step toward liberalisation, but access was limited to 90 large and liquid stocks. Subsequent expansions between 2010 and 2022 gradually broadened the list of eligible securities, creating eight phases of short-sale liberalisation. This phased rollout provides quasi-experimental variation in arbitrage capacity across firms and time, making China an ideal laboratory for testing limits-to-arbitrage theories (Chang et al., 2013; Chu et al., 2020).

Evidence from the Chinese setting confirms that short-sale access influences pricing efficiency. Boehmer and Wu (2013) show that the introduction of short selling in China improved price informativeness by accelerating the incorporation of earnings information into stock prices. Huang et al. (2016) find that short-sale liberalisation reduced speculative bubbles and enhanced price discovery, while Li et al. (2022) document that the expansion of the eligible list decreased volatility and improved liquidity. These findings demonstrate that short-selling reforms in China improved overall market quality, though the scale of activity remained modest compared to developed markets.

Research has also examined the implications for anomaly returns in China. Chen et al. (2010) document that momentum strategies are profitable in the Chinese market, but their persistence suggests the presence of binding arbitrage frictions. More recent studies suggest that anomalies linked to investment and accruals remain strong in China, consistent with limits-to-arbitrage arguments, but attenuate when stocks become shortable under the MT/SL programme (Feng et al., 2020). Evidence from China's margin trading and securities lending reforms further shows that anomaly profits decline following short-sale liberalisation, consistent with the role of short sellers in mitigating mispricing (Chincarini, 2025). This finding aligns with the U.S. evidence of short-sale frictions sustaining anomaly spreads. These results align with the broader global literature while highlighting the unique institutional constraints in China that sustain return predictability.

Taken together, evidence from emerging markets underscores the critical role of institutional design in shaping arbitrage opportunities. In China, the phased expansion of short-sale eligibility and the limited depth of lending markets imply that efficiency gains from liberalisation are gradual and uneven across anomalies. This context provides an important opportunity to examine not only whether short-sale constraints sustain anomalies, but also whether incremental liberalisation and variation in trading intensity reduce anomaly profitability.

3.3. Methodology and Data Description

3.3.1. Research Design and Institutional Background

In order to examines whether constraints on short selling affect the profitability and persistence of anomaly-based trading strategies in the Chinese A-share market. We implement a multi-step empirical design that incorporates characteristic-based portfolio construction, phased policy liberalisation, and short-sale activity measures to capture both the availability and intensity of short selling. Our identification strategy relies on China's regulatory setting, where the gradual rollout of margin trading and securities lending created exogenous variation in shortable stock access.

Short selling was not permitted in China until March 2010, when the China Securities Regulatory Commission (CSRC) launched the initial pilot programme for margin trading and securities lending. The first batch of 90 eligible stocks marked the beginning of a series of regulatory expansions. From 2010 to 2022, the CSRC and the two main stock exchanges (Shanghai and Shenzhen) announced seven additional expansions of the shortable stock list. Each expansion introduced a new batch of stocks eligible for short selling, creating a phased timeline of accessibility across firms. This institutional feature provides a quasi-natural experiment that enables a difference-in-differences (DiD) approach to estimate the impact of short-sale constraints on anomaly strategy performance.

We define shortable status using the official lists released during each expansion wave. A binary variable, Shortable, equals one starting from the month a firm first appears in any batch and equals zero otherwise. The shortable status is updated monthly and merged with firm-level return and accounting data. This allows us to compare performance between shortable and non-shortable firms over time, as well as to capture the dynamic effects of liberalisation. This institutional classification is consistent with the shortability-based framework of Bai et al. (2017), who examine how short-sale constraints affect asset-pricing model performance in the Hong Kong market by comparing shortable and non-shortable stocks. Their framework provides conceptual support for our empirical design using China's phased margin-trading and securities-lending expansions.

Our empirical design consists of three key components. First, we construct long–short portfolios based on eleven widely studied firm-level anomaly variables, including profitability, investment, asset quality, issuance behaviour, and mispricing-related attributes. These portfolios are formed annually at the end of December and held over a 12-month horizon. Second, we compare returns of these portfolios between shortable and non-shortable groups before and after the firm's short-sale eligibility. Third, we implement a DiD framework that exploits the phased rollout of policy to isolate the causal impact of short-sale access on anomaly returns, controlling for time-varying shocks and firm-level fixed effects.

To complement the baseline DiD analysis, we also explore the role of short-selling activity intensity using two constructed measures. These are derived from transaction-level margin trading and securities lending data and are described in Section 3.2. Together, our empirical design leverages both policy-induced variation in short-selling access and time-varying intensity across firms, enabling a comprehensive assessment of how short-sale frictions influence anomaly return dynamics in an emerging market setting.

3.3.2. Sample Selection and Data Sources

The empirical analysis focuses on Chinese A-share firms listed on the Shanghai and Shenzhen Stock Exchanges over the period from March 2010 to March 2024. This sample window is chosen to align with the introduction and gradual expansion of China's short-selling mechanism, which began in March 2010 and unfolded through multiple regulatory waves. Firms are required to have non-missing data for target financial variables, monthly stock returns, and market capitalisation. Special treatment (ST/*ST) firms and financial sector firms are excluded to avoid distortions arising from regulatory interventions and inconsistent reporting structures. Table 3.1 summarises the regulatory timeline of China's margin-trading and securities-lending (MT/SL) programme from 2010 to 2022, detailing each expansion wave's implementation date, cumulative number of eligible stocks, and newly added securities. The steady increase in shortable stocks, from 96 in the initial pilot to 4,132 by the final expansion, illustrates the gradual relaxation of short-selling constraints that forms the institutional basis of this study.

Firm-level financial data are obtained from the China Stock Market and Accounting Research (CSMAR) database, including annual accounting items from the Balance Sheet (BS), Income Statement (IS), and Cash Flow (CF) modules. These are used to construct the anomaly variables discussed in Section 3.3.3. Monthly stock returns, adjusted for dividends and corporate actions, are sourced from the CSMAR stock returns file. Market capitalisation is computed at the end of each December and used for both return weighting and size-based analyses.

We identify shortable stocks using official eligibility lists for margin trading and securities lending (MT/SL) published by the Shanghai and Shenzhen Stock Exchanges. Each firm is coded as shortable starting from the month it is first included in any expansion batch and remains shortable thereafter. A monthly indicator variable is constructed to capture this status and is merged with the firm-month panel.

We also incorporate measures of short-selling depth constructed from margin-trading and securities-lending transactions, described in Section 3.3.6. This proxy reflects the relative depth of short-selling pressure.

Finally, we incorporate monthly Fama-French five-factor data for the Chinese A-share market, obtained from CSMAR. These factors include the market factor (MKT), size factor (SMB), value factor (HML), profitability factor (RMW), and investment factor (CMA), and are merged with the stock-level dataset to facilitate risk-adjusted performance analysis. MKT is defined as the market excess return from the CSMAR Fama-French five-factor dataset, constructed as the value-weighted excess return on all tradable A-share stocks listed on the Shanghai and Shenzhen Stock Exchanges. The remaining factors (SMB, HML, RMW, and CMA) are obtained from the same CSMAR dataset, in which these factors are constructed following Fama and French (2015). The risk-free rate used to compute excess returns is the one-month risk-free rate provided in the CSMAR factor dataset.

Table 3. 1. Short-Selling Pilot Expansion Timeline (2010–2022)

Expansion Wave	Implementation Date	Total Stocks in Shortable List	Number Added	Remarks
1	31-Mar-10	96	96	Initial pilot programme launched by CSRC
2	5-Dec-11	288	195	First expansion phase
3	31-Jan-13	514	279	Coverage broadened to mid-caps
4	16-Sep-13	715	189	Fourth expansion wave
5	22-Sep-14	922	207	Post-2014 policy broadening
6	12-Dec-16	1,128	222	Minor adjustment
7	19-Aug-19	2,834	1,740	Expansion to main-board and SMEs
8	24-Oct-22	4,132	1,391	Latest expansion before sample end

Notes This table shows the staggered expansion of the short-selling eligibility list in China between 2010 and 2022. Expansion Wave is each round of adjustment to the short-selling pilot program. Implementation Date reports the official effective date of each expansion. Total Stocks in Shortable List denotes the cumulative number of stocks eligible for short selling after the corresponding expansion wave. Number Added captures the incremental number of stocks newly included in the shortable list in that wave. Remarks briefly summarize the regulatory background or key features associated with each expansion.

The full sample of monthly stock returns spans January 2010 to December 2024. Analyses related to short-selling expansions are restricted to 2010–2022, as the final wave of short-selling eligibility occurred in 2022, and we require 152 months of stock data to evaluate the impact of short selling eligibility. Baseline anomaly performance is therefore evaluated using the full return sample, whereas expansion-based difference-in-differences tests are conducted within the expansion window ending in 2022.

3.3.3. Anomaly Variables

We construct a comprehensive set of firm-level variables that capture anomaly signals, short-selling constraints, and standard control characteristics. This study adopts 11 anomaly measures, examined by Stambaugh et al. (2012) and Chu et al. (2020). While the overall selection follows these two studies, each anomaly is defined based on its original source to maintain consistency with the established literature. These anomalies span profitability, investment, financing activity, financial distress, and market-based mispricing dimensions. Table 3.2 summarises the standard definitions and economic interpretations of the 11 anomaly characteristics considered in this study. This table is organised following the structure proposed by Chincarini et al. (2025).

Accounting variables and trading-related data[2] are sourced from the CSMAR database. All anomaly variables are winsorized at the 1st and 99th percentiles and lagged by one year relative to the return evaluation period to mitigate outliers and avoid look-ahead bias. Owing to differences in data availability and variable definitions between Chinese databases and the Compustat/CRSP-based data sources commonly used in the U.S. anomaly literature, the construction of some anomaly variables does not mechanically replicate their original U.S. implementations. In such cases, we follow the economic intuition of the original studies and construct proxy measures using the closest available accounting and market information from CSMAR. This approach is standard in international asset-pricing literature and facilitates cross-market comparability of anomaly signals.

[2] The details of trading-related data, as well as the accounting variables, are shown in Appendix A. Please refer to the sections "Stock prices, returns and market value" and "Accounting data" in Appendix A.

Table 3. 2. Anomaly Characteristics List

Anomaly	Label	Theory Definition	Data Construction (This Study)	Reference
Gross profitability	GP	Gross profit scaled by total assets.	Operating revenue minus operating costs, scaled by total assets at fiscal year-end.	Novy-Marx (2013)
Asset growth	AG	Year-over-year percentage change in total assets.	Change in total assets over previous year, scaled by lagged total assets.	Cooper et al. (2008)
Investment-to-assets	IVA	Capital expenditure scaled by lagged total assets.	Change in fixed assets and inventories over previous year, scaled by lagged total assets.	Titman et al. (2004)
Return on assets	ROA	Net income scaled by total assets.	Net income from income statement scaled by total assets at fiscal year-end.	Chen et al. (2011)
Net operating assets	NOA	Operating assets minus operating liabilities, scaled by assets.	Operating assets minus operating liabilities, scaled by lagged total assets.	Hirshleifer et al. (2004)
Accruals	ACC	Earnings minus operating cash flows, scaled by assets.	Change in non-cash current assets minus change in non-debt current liabilities, scaled by lagged total assets.	Sloan (1996)
Net equity issuance	ISS	Annual change in shares outstanding.	Log change in shares outstanding reconstructed from market equity and price at fiscal year-end.	Pontiff and Woodgate (2008)
Composite equity issuance	CEI	Market equity growth unexplained by returns.	Log change in market equity net of stock returns over a five-year window.	Daniel and Titman (2006)
O-score	OSC	Bankruptcy risk constructed from accounting ratios.	Ohlson O-score obtained from CSMAR, based on logit model using accounting variables.	Ohlson (1980)

Table 3. 2. (continued)

Anomaly	Label	Theory Definition	Data Construction (This Study)	Reference
Downside risk	DSR	Volatility of negative returns below mean.	Standard deviation of monthly returns below mean over rolling 36-month window (CSMAR).	Ang et al. (2006)
Momentum	MOM	Cumulative return over past 12 months excluding most recent month.	Cumulative return from t-12 to t-2 at fiscal year-end.	Jegadeesh and Titman (1993)

Note This table summarizes 11 anomaly characteristics examined in this study. Anomaly reports the name of each asset pricing anomaly. Label provides the abbreviated notation used to refer to each anomaly. Theory Definition describes the standard conceptual definition of each anomaly. Data Construction explains how each anomaly variable is measured using firm-level accounting data in this study. Reference lists the seminal studies that originally document or formally define each anomaly. The details of variable definitions and construction are provided in Appendix A.

Short-selling constraints are captured using both binary and continuous measures. A firm is classified as shortable (Dummy = 1) if it appears in any of the eight regulatory short-sale eligibility batches between 2010 and 2022; otherwise, it is treated as non-shortable (Dummy = 0). This indicator is used in the difference-in-differences design to distinguish treated and control firms.

Our measure of control variables follow standard asset-pricing practice. Size is measured as the logarithm of June-end market capitalisation. Book-to-market ratio (BM) is calculated as the ratio of book value of equity to market capitalisation at fiscal year-end. Prior returns include the past one-month and past twelve-month cumulative returns. Monthly Fama–French five-factor returns (MKT, SMB, HML, RMW, CMA) for the A-share market are obtained from CSMAR and merged at the monthly frequency.

3.3.4. Anomaly-Based Portfolio Construction

To examine the return implications of firm-level anomalies, we construct characteristic-sorted long–short portfolios following standard asset-pricing practice. At the end of each fiscal year t, stocks are sorted into deciles based on each anomaly measure. Portfolios are held for the subsequent twelve months, from January to December of year $t + 1$, and are rebalanced annually each December. The long–short strategy takes value-weighted positions between the top and bottom deciles, going long decile 10 (top 10%) and short decile 1 (bottom 10%) for positively priced anomalies, and vice versa for negatively priced anomalies, consistent with their theoretical expected return directions. Portfolio returns are computed using value weights based on lagged market capitalisation from December of year t. Observations with missing characteristics or return data are excluded, and all returns are matched with lagged predictors to avoid look-ahead bias.

Monthly value-weighted portfolio returns are calculated over the holding horizon to obtain the long–short return series. Monthly portfolio returns are expressed as excess returns, computed as raw returns minus the corresponding monthly risk-free rate. Following the asset pricing literature (Fama & French, 1993; Fama & French, 2015; Stambaugh et al., 2012), we evaluate both raw long–short returns and alphas obtained from time-series regressions that control for standard risk factors, including the CAPM, Fama–French three-factor, and Fama–French five-factor models. All factor returns are sourced from CSMAR.

For clarity, each anomaly portfolio is constructed following its theoretically expected return direction, consistent with prior asset pricing studies. Specifically, we take long–short positions for anomalies that are positively related to expected returns, such as gross profitability (GP), return on assets (ROA), and momentum (MoM), as higher values of these characteristics indicate stronger fundamentals or return continuation (Jegadeesh & Titman, 1993; Novy-Marx, 2013). Conversely, we take Low – High positions for anomalies that are negatively related to future returns, including asset growth (AG), investment-to-assets (IVA), net operating assets (NOA), accruals (ACC), equity issuance (ISS), composite equity issuance (CEI), O-score (OSC), and downside risk (DSR). These characteristics capture overinvestment, accounting distortions, or distress risk, which prior studies show to predict lower subsequent returns (Ang et al., 2006; Daniel & Titman, 2006; Ohlson, 1980; Pontiff & Woodgate, 2008; Richardson et al., 2005; Sloan, 1996). The signs of the long–short construction are thus consistent with the theoretical return implications of each anomaly.

3.3.5. Difference-in-Differences (DiD) Analysis

To identify the impact of short-selling liberalisation on the profitability of anomaly-based trading strategies, we employ a difference-in-differences (DiD) framework that leverages China's phased short-sale liberalisation as a quasi-natural experiment. DiD model allows us to isolate the incremental impact of the short-sales relaxation by differencing out common time trends and group-specific fixed effects and hence avoiding reducing concerns associated with differences between treated and control stocks[3].This analysis follows the difference-in-differences (DiD) framework employed by Chu, Hirshleifer, and Ma (2020) while adapting to the phased nature of China's short-sale expansions, which differ from the one-time pilot program in the U.S. context. The phased expansion of the shortable stock list across eight major waves provides a useful setting to compare anomaly portfolio returns for stocks newly eligible

[3] In the Chinese stock market, the stocks which are allowed to sold short are qualified. Such as large market value, good liquidity, and etc. For more details, please refers to relevant content in Section 3.4.1.

for short selling (treatment group) against those that remain non-shortable (control group) within the same months.

In our DiD design, the treatment group consists of stocks that become newly eligible for short selling during a specific expansion wave, while the control group includes stocks that remain non-shortable during the same calendar months. In implementation, we construct a panel of monthly anomaly long–short portfolio returns for treated and control groups and define the post indicator relative to each expansion event. Specifically, the post period covers the 12 months following the effective date of a given expansion, while the pre period includes the 12 months preceding the expansion. Based on this event-time window, the DiD estimator is defined as follows:

$$\text{Effect} = (\text{Post\_T} - \text{Pre\_T}) - (\text{Post\_C} - \text{Pre\_C}) \quad (6)$$

Pre_T denotes the average value of the anomaly portfolio constructed by stocks that will become eligible for short selling in a given expansion wave, measured over the 12-month window prior to the effective date of the expansion. Post_T denotes the average value of the anomaly portfolio constructed by the same group of stocks during the 12-month window following the expansion, when short-selling eligibility becomes effective. Pre_C denotes the average value of the anomaly portfolio constructed by stocks that remain non-shortable during the same expansion wave, measured over the identical 12-month pre-expansion window. Post_C denotes the average value of the anomaly portfolio constructed by the control stocks during the 12-month window after the expansion, when they still remain ineligible for short selling.

The difference-in-differences estimator compares changes in outcomes for anomaly portfolios formed by newly shortable stocks before and after each expansion with contemporaneous changes for stocks that remain non-shortable. Specifically, $(\text{Post}_T - \text{Pre}_T)$captures the total change in the outcome for treated portfolios formed by stocks following the expansion, which reflects both the effect of short-selling eligibility and any common time trends. In contrast, $(\text{Post}_C - \text{Pre}_C)$measures the change for control portfolios formed by stocks over the same period, isolating the evolution driven solely by market-wide or time-specific factors. By differencing these two changes, the DiD estimator accounts for common shocks and captures the differential change in anomaly returns associated with becoming eligible for short selling within each expansion wave.

This comparison is implemented separately for each expansion wave by aligning observations relative to the effective date of the corresponding expansion. As a result, the treatment and control groups are redefined in each wave, allowing the analysis to exploit the staggered introduction of short-selling eligibility across stocks and over time. Short-selling eligibility in China is not randomly assigned, and stocks admitted to the short-selling list may differ systematically from non-shortable stocks in observable characteristics such as firm size and liquidity. This creates a potential selection concern when constructing the treatment and control groups. Propensity Score Matching (PSM) provides a complementary approach to improving pre-treatment comparability by matching newly eligible stocks with non-shortable stocks that exhibit similar observable pre-treatment characteristics (Rosenbaum & Rubin, 1983; Heckman et al., 1998). In the present analysis, the principal identification strategy remains the wave-specific DiD design, while the comparability of the resulting treatment and control portfolios is further assessed through formal pre-treatment trend tests.

In addition, we include the Fama–French five factors as time-series controls to absorb common risk movements in monthly returns. Specifically, the regressions control for the Fama–French five factors, including the market risk premium (MKT), size (SMB), value (HML), profitability (RMW), and investment (CMA). This specification isolates the incremental change in anomaly portfolio returns associated with short-selling eligibility beyond contemporaneous market-wide risk factor realizations.

Accordingly, the baseline DiD specification is estimated using the following model:

$$Ret_{p,t} = \alpha + \beta_1 Post_t + \beta_2 Treated_p + \beta_3 \left(Post_t * Treated_p\right) + \gamma' F_t + \epsilon_{p,t}, \tag{7}$$

where $Ret_{p,t}$s are the monthly returns of the anomaly-based long–short portfolios (by using shortable stocks and non-shortable stocks, respectively.) p in month t; $Post_t$ is the post-expansion indicator, defined as one for months falling within the 12-month window after the effective date of a given short-selling expansion and zero for months in the corresponding 12-month pre-expansion window. $Treated_p$ equals one for portfolios formed from stocks newly eligible for short selling in a given expansion wave and zero for portfolios formed from stocks that remain non-shortable during that wave. $Post_t$ equals one for months in the corresponding post-expansion window and zero for months in the pre-expansion window. The interaction term $Post_t * Treated_p$ captures the differential change in anomaly portfolio returns for newly eligible stocks following the expansion relative to the contemporaneous change in the control portfolios. The coefficient β_3 captures the DiD effect, measuring whether anomaly portfolio returns for newly shortable stocks change differentially after the expansion relative to non-shortable stocks. F_t denotes the vector of the Fama–French five factors (MKT, SMB, HML, RMW, and CMA).

The results are reported in Table 3.5, which presents DiD estimates based on the interaction between the wave-specific treatment indicator (Pilot) and the post-expansion indicator. These results provide direct evidence on whether the phased introduction of short-selling eligibility attenuates pricing inefficiencies and reduces the abnormal returns associated with anomaly portfolios.

3.3.6. Short-Selling Depth Regression Analysis

For the subsample of stocks that are eligible for short selling, we further examine whether variation in short-selling activity helps explain the cross-sectional differences in anomaly returns. To further examine how the intensity of short-selling activity influences anomaly returns, we conduct a cross-sectional analysis based on short-selling depth. While the DiD framework captures the binary effect of short-sale eligibility, this section explores the heterogeneous impact of short-selling intensity across firms.

This construction follows the short-selling activity framework used in the Chinese margin-trading and securities-lending literature, such as Chang et al. (2013), who proxy short-selling intensity using trading-based measures (e.g., short-selling turnover). Consistent with this approach, we employ a stock-based measure of short-selling activity using securities-lending balances. Specifically, we define short-selling depth (SSD) as:

$$SSD_{i,t} = \frac{ShortBalance_{i,t}}{MarketValue_{i.t}}, \tag{8}$$

where ShortBalance is the outstanding value of borrowed shares sold short and not yet repurchased, and MarketValue is the firm's total market capitalisation. This ratio captures the intensity of short-selling participation scaled by firm size.

The SSD measure parallels the spirit of short-selling activity indicators in Chang et al. (2013) but uses stock-based balances rather than trading-flow data. SSD is winsorized at the 1st and 99th percentiles to mitigate the influence of extreme values.

We apply two SSD-based classifications in the empirical analysis. First, in Table 3.7, firms are sorted at the beginning of each month (t + 1) into ten deciles of SSD based on their values observed in the previous month (t), with Decile 1 representing the lowest SSD and Decile 10 the highest. This design uses lagged SSD to avoid look-ahead bias and to capture the prior month's short-selling intensity when examining how firm characteristics vary across the deciles of short-selling intensity. Second, in Table 3.8, SSD is collapsed into a high–low indicator based on the monthly cross-sectional median, where the indicator equals 1 for firms in the high-SSD group (top 50 percent of the distribution) and 0 for firms in the low-SSD group (bottom 50 percent). Within each SSD group, we form long–short portfolios for each of the 11 anomaly signals, allowing us to assess whether anomaly-based profits differ systematically under varying levels of short-selling activity.

The underlying hypothesis is that more active short-selling facilitates the correction of mispricing, resulting in weaker anomaly returns, whereas limited short-selling activity allows anomalies to persist. Tables 3.8 and 3.9 provide descriptive and portfolio-level evidence consistent with this mechanism.

To formally evaluate whether short-selling activity attenuates anomaly-based returns, we estimate panel regressions that link anomaly portfolio performance to the SSD-based high–low classification defined in Section 3.3.5. For each anomaly, monthly returns are computed separately for high-SSD and low-SSD stocks and aggregated into two value-weighted portfolio series. The regression specification parallels the regime-interaction design of Bai (2021), who investigates whether firm-characteristic pricing differs between opposite short-sales regimes using panel regressions. In our context, the SSD-group dummy captures the intensive margin of short-selling activity within the liberalised regime.

We estimate the following regression model:

$$Y_{i,t} = \alpha + \beta Dummy_{i,t} + \delta_{MKT}MKT_t + \delta_{SMB}SMB_t + \delta_{HML}HML_t + \delta_{RMW}RMW_t + \delta_{CMA}CMA_t + \varepsilon_{i,t}, \tag{9}$$

where $Y_{i,t}$ denotes the excess return of the SSD-sorted anomaly long–short portfolio for anomaly i in month t. Each anomaly portfolio is formed separately within the high-SSD and low-SSD groups. The variable $Dummy_{i,t}$ is an SSD-group indicator. With the SSD-group indicator equal to one for portfolios formed from firms in the bottom 50 percent of the SSD distribution (low-SSD group) and zero for portfolios formed from firms in the top 50 percent (high-SSD group). MKT_t, SMB_t, HML_t, RMW_t, CMA_t denote the Fama–French five factors. The coefficient β captures the return differential between high- and low-SSD portfolios after controlling for systematic risk exposures. A negative and statistically significant β would indicate that deeper short-selling activity is consistent with lower anomaly profitability, consistent with the view that short sellers mitigate mispricing.

The regressions are estimated using ordinary least squares (OLS), excluding observations with missing factor returns, SSD group assignments, or portfolio returns. Table 3.9 reports the estimated coefficients on the SSD Dummy variable and the corresponding Fama–French factor loadings.

3.4. Empirical Results

This section investigates whether short-selling constraints influence the persistence of anomaly-based returns in the Chinese A-share market. We employ a comprehensive empirical framework combining portfolio-level analyses, difference-in-differences (DiD) designs, and cross-sectional regressions to explore how market liberalization affects anomaly profitability.

Our analysis focuses on eleven well-documented accounting- and return-related anomalies covering profitability, investment, asset composition, equity issuance, and mispricing-related signals. By exploiting China's unique phased short-selling policy expansions, this setting allows for an identification strategy that contrasts newly shortable stocks with those that remain non-shortable at each stage of regulatory reform.

To examine how short-sale constraints shape anomaly returns, we adopt a multi-step empirical framework. First, we compare the profitability of anomaly portfolios between shortable and non-shortable firms around short-selling policy expansions using a difference-in-differences approach. Second, we investigate how anomaly performance evolves following the introduction of short-selling eligibility, capturing the dynamic effects of liberalization. Third, we study cross-sectional variation in anomaly returns through high-minus-low portfolio strategies that reflect differences in short-selling activity. Finally, we examine heterogeneity in anomaly profitability across firm characteristics such as size and book-to-market ratio using double-sorted portfolios.

From a theoretical and empirical perspective, many anomaly-based return patterns are intricately linked to limits to arbitrage and short-sale constraints. Prior studies provide evidence that, in markets where short selling is restricted, mispricing tends to persist, and anomalies related to investment, issuance, and accruals often generate positive long–short profits. In particular, asset growth (AG), investment-to-assets (IVA), equity issuance (ISS), and accruals (ACC) are commonly associated with overvaluation that is difficult to correct in the absence of short selling. In contrast, distress-related signals such as O-score (OSC) and downside risk (DSR) frequently predict negative long–short returns under short-sale constraints, consistent with settings in which pessimistic beliefs are unable to be fully incorporated into prices (Stambaugh, Yu, and Yuan, 2012). Momentum (MOM), which reflects return continuation, is expected to remain profitable unless arbitrage activity becomes sufficiently active to eliminate it. These empirical regularities, grounded in theories of limits to arbitrage, provide a useful benchmark for interpreting anomaly returns in the Chinese A-share market.

Our empirical findings broadly align with these regularities and with prior evidence from markets characterized by short-sale constraints. The positive returns associated with AG and IVA in our sample are consistent with the investment-based asset pricing framework of Chen, Novy-Marx, and Zhang (2011), which links firm investment decisions to expected returns and predicts the persistence of overvaluation when arbitrage is limited. Similarly, the negative returns associated with OSC and DSR conform to the findings of Stambaugh, Yu, and Yuan (2012), who document that distress-related mispricing is more pronounced under short-sale restrictions. However, some anomalies exhibit behaviour that differs from patterns documented in U.S. markets. For example, the profitability of net operating assets (NOA) and accrual-based strategies is weaker in our sample, potentially reflecting China-specific institutional features such as higher retail investor participation, state ownership, and differences in accounting practices. These comparisons suggest that while the overall structure of anomaly returns in China resembles that observed in other constrained markets, the magnitude and persistence of individual anomalies are shaped by the unique institutional frictions of the A-share market, particularly the availability and intensity of short selling.

3.4.1. Descriptive Statistics and Preliminary Analyses

3.4.1.1. Summary Statistics of Anomaly Variables

Table 3.3 reports anomaly performance for eleven firm-characteristic based long–short strategies. These anomalies span five broad economic dimensions: profitability (Gross Profitability, ROA), investment activity (Asset Growth, Investment-to-Assets, CEI), asset composition (Net Operating Assets, Accruals), equity issuance (Net Issuance), and mispricing-related signals (Ohlson O-Score, Downside Risk, Momentum). All strategies are constructed under a unified portfolio formation procedure to ensure consistency across characteristics.

Table 3.3 shows that most anomalies deliver positive and statistically significant excess returns, indicating that anomaly-based strategies remain profitable in the Chinese stock market. Although the raw return of DSR is statistically insignificant, it retains a positive sign. Overall, these findings are broadly consistent with Chu et al. (2020).

In addition to mean returns, Table 3.3 reports the standard deviation and Sharpe ratio of each anomaly portfolio. Several anomalies exhibit economically meaningful Sharpe ratios, suggesting that their profitability is not solely driven by elevated return volatility. In particular, profitability- and investment-related anomalies demonstrate relatively strong risk-adjusted

Table 3. 3. Performance of Anomaly Portfolios

	GP	AG	IVA	ROA	NOA	ACC	ISS	CEI	OSC	DSR	MoM
Anomaly returns											
mean	0.009***	0.016***	0.010***	0.017***	0.006***	0.009***	0.012***	0.021***	0.008**	0.003	0.026***
t-statistics	[2.80]	[5.08]	[4.16]	[4.87]	[2.75]	[3.98]	[4.70]	[5.23]	[2.01]	[1.01]	[5.14]
Std. dev.	0.003	0.003	0.002	0.004	0.002	0.002	0.003	0.004	0.004	0.003	0.005
Sharpe ratio	0.16	0.30	0.25	0.29	0.16	0.23	0.28	0.31	0.12	0.06	0.30
Obs.	288										
Benchmark-adjusted returns											
mean	0.012***	0.017***	0.012***	0.022***	0.006***	0.007***	0.013***	0.021***	0.013***	0.006***	0.024***
t-statistics	[4.98]	[7.01]	[5.72]	[9.56]	[2.73]	[3.46]	[6.07]	[6.54]	[5.62]	[2.84]	[4.95]
Obs.	288										

Note Table 3.3 reports monthly high–minus–low (H–L) anomaly portfolio returns and benchmark-adjusted returns for eleven firm-characteristic based anomalies spanning profitability (Gross Profitability, ROA), investment activity (Asset Growth, Investment-to-Assets, CEI), asset composition (Net Operating Assets, Accruals), equity issuance (Net Issuance), and mispricing-related signals (Ohlson O-Score, Downside Risk, Momentum). Anomaly returns are reported as raw anomaly portfolio returns in excess of the risk-free rate. Standard deviations are computed from monthly returns, and Sharpe ratios are calculated as the mean return divided by the standard deviation. Benchmark-adjusted returns are obtained as intercepts from time-series regressions of raw anomaly returns on the Fama–French five factors; t-statistics are reported in brackets. Portfolios are rebalanced annually on each fiscal year end. Portfolio returns are calculated as the value-weighted return spread between the top and bottom deciles.

performance, implying that their abnormal returns extend beyond simple compensation for systematic risk exposure.

The benchmark-adjusted results further confirm this pattern. For a subset of anomalies, the alphas from the Fama–French five-factor regressions remain statistically significant, indicating that traditional asset pricing factors do not fully absorb their return predictability.

Return magnitudes vary across anomaly signals, reflecting heterogeneity in their economic mechanisms. The eleven anomalies span six broad dimensions—profitability, investment activity, asset composition, equity issuance, downside risk, and market behaviour—each capturing distinct sources of mispricing. This cross-anomaly variation provides a suitable setting to examine how short-selling constraints affect different types of anomalies returns.
It is noteworthy that the anomaly return magnitudes reported in Table 3.3 differ from those documented in Table 2.2 (Panel B) of Essay 1. In particular, several anomalies exhibit weaker statistical significance in the post-2010 subsample. This variation highlights the time-sensitive nature of anomaly profitability in the Chinese market.

While Essay 1 spans the 2000–2020 period, capturing episodes such as the 2005 split-share reform and the 2008 global financial crisis, Essay 2 focuses on the post-2010 era marked by the gradual liberalization of short-selling mechanisms. The shift in institutional environment and trading constraints appears to coincide with changes in anomaly performance. This temporal variation provides further motivation to examine how market frictions, particularly short-selling constraints, shape the persistence and magnitude of anomaly returns.

3.4.1.2. Summary of Anomaly Factors

Baker and Wurgler (2006) suggest that the valuation of stocks with similar characteristics tends to be subjective and more likely to be biased. Moreover, the pilot stocks which are eligible to be sold short are “blue chip” ones with large market capitalization, low volatility, and stable earnings performance. In the process of stock expansion, eligible firms should also meet the requirements of completing the share split reform and having shareholders numbering more than 4,000. Hence, we compare stock characteristics of shortable and non-shortable stocks on each pilot expansion, before testing the short sale constraints on anomaly

Table 3. 4. The Mean Value of Anomaly Variables (Shortable vs Non-shortable)

Panel A Initial List (Shortable obs. = 63, Non-shortable obs. = 1317)

	GP	AG	IVA	ROA	NOA	ACC	ISS	CEI	OSC	DSR	MOM
Shortable mean	0.586***	0.378***	0.138***	0.062***	0.581***	-0.049***	0.123***	1.874***	-8.868***	0.028***	0.022
Non-shortable mean	0.538***	0.176***	0.043***	0.020***	0.627***	0.003	0.069***	1.046***	-7.397***	0.062***	-0.035***
Difference mean	0.048	0.202***	0.095***	0.042***	-0.046	-0.052***	0.055**	0.828***	-1.470***	-0.034**	0.057***
T-value	[0.70]	[4.22]	[4.12]	[5.50]	[-0.85]	[-2.80]	[2.05]	[9.10]	[-4.91]	[-2.30]	[2.65]

Panel B Expand 1 (Shortable obs. = 210, Non-shortable obs. = 1169)

	GP	AG	IVA	ROA	NOA	ACC	ISS	CEI	OSC	DSR	MOM
Shortable mean	0.594***	0.195***	0.068***	0.034***	0.626***	0.032***	0.075***	1.423***	-7.464***	0.064***	-0.049
Non-shortable mean	0.594***	0.195***	0.068***	0.034***	0.626***	0.032***	0.075***	1.423***	-7.464***	0.064***	-0.049***
Difference mean	0.002	0.075***	0.048***	0.032***	-0.010	0.025***	0.078***	0.511***	-1.216***	-0.022***	0.046***
T-value	[0.04]	[3.11]	[4.07]	[7.82]	[-0.41]	[2.75]	[4.28]	[10.41]	[-7.53]	[-2.76]	[3.99]

Panel C Expand 2&3 (Shortable obs. = 544, Non-shortable obs. = 1009)

	GP	AG	IVA	ROA	NOA	ACC	ISS	CEI	OSC	DSR	MOM
Shortable mean	0.529***	0.190***	0.063***	0.045***	0.602***	0.022***	0.118***	0.064**	-8.294***	0.059***	0.292***
Non-shortable mean	0.597***	0.133***	0.045***	0.018***	0.615***	0.017***	0.068***	-0.244***	-7.309***	0.066***	0.308***
Difference mean	-0.068***	0.057***	0.018**	0.027***	-0.013	0.005	0.050***	0.308***	-0.985***	-0.007	-0.016**
T-value	[-2.71]	[3.25]	[2.53]	[8.19]	[-0.82]	[0.75]	[4.36]	[7.17]	[-8.79]	[-1.17]	[-2.37]

Table 3. 4. (continued)

Panel C Expand 4 (Shortable obs. = 689, Non-shortable obs. = 961)

	GP	AG	IVA	ROA	NOA	ACC	ISS	CEI	OSC	DSR	MOM
Shortable mean	0.533***	0.170***	0.055***	0.043***	0.610***	0.033***	0.089***	0.993***	-8.373***	0.053***	-0.194***
Non-shortable mean	0.584***	0.162***	0.051***	0.019***	0.633***	0.036***	0.071***	0.805***	-7.343***	0.062***	-0.166***
Difference mean	-0.051**	0.008	0.004	0.024***	-0.023	-0.003	0.018*	0.188***	-1.030***	-0.009	-0.028***
T-value	[-2.21]	[0.49]	[0.62]	[8.26]	[-1.59]	[-0.42]	[1.82]	[5.13]	[-9.90]	[-1.48]	[-4.05]

Panel E Expand 5 (Shortable obs. = 868, Non-shortable obs. = 1054)

	GP	AG	IVA	ROA	NOA	ACC	ISS	CEI	OSC	DSR	MOM
Shortable mean	0.461***	0.213***	0.046***	0.037***	0.613***	0.018***	0.160***	0.606***	-8.577***	0.035***	-0.004
Non-shortable mean	0.453***	0.215***	0.043***	0.010***	0.676***	0.031***	0.158***	0.641***	-7.934***	0.058***	0.108***
Difference mean	0.008	-0.002	0.002	0.027***	-0.064***	-0.014**	0.002	-0.035	-0.643***	-0.023***	-0.113***
T-value	[0.43]	[-0.09]	[0.35]	[9.03]	[-4.19]	[-2.19]	[0.13]	[-1.02]	[-6.56]	[-9.98]	[-12.27]

Panel F Expand 6 (Shortable obs. = 1377, Non-shortable obs. = 1147)

	GP	AG	IVA	ROA	NOA	ACC	ISS	CEI	OSC	DSR	MOM
Shortable mean	0.488***	0.112***	0.036***	0.034***	0.599***	0.020***	0.060***	0.433***	-8.676***	0.035***	-0.038***
Non-shortable mean	0.459***	0.030***	0.012***	-0.021***	0.574***	0.007*	0.056***	0.245***	-7.474***	0.052***	-0.026***
Difference mean	0.029	0.082***	0.024***	0.056***	0.025**	0.013***	0.004	0.188***	-1.202***	-0.018***	-0.011**
T-value	[1.62]	[7.61]	[5.70]	[12.72]	[2.40]	[2.87]	[0.66]	[6.90]	[-12.38]	[-9.58]	[-2.41]

Table 3. 4. (continued)

Panel G Expand 7	(Shortable obs. = 1868, Non-shortable obs. = 1134)										
	GP	AG	IVA	ROA	NOA	ACC	ISS	CEI	OSC	DSR	MOM
Shortable mean	0.504***	0.131***	0.039***	0.034***	0.582***	0.016***	0.044***	0.053***	-8.743***	0.027***	0.046***
Non-shortable mean	0.463***	0.061***	0.019***	-0.013***	0.537***	0.008**	0.033***	-0.675***	-7.556***	0.063***	0.114***
Difference mean	0.041**	0.069***	0.019***	0.046***	0.045***	0.007	0.011***	0.728***	-1.186***	-0.036***	-0.068***
T-value	[2.46]	[7.01]	[5.18]	[12.11]	[4.55]	[1.55]	[2.61]	[29.55]	[-13.07]	[-12.15]	[-10.43]

Notes This table compares anomaly variables/firm characteristics of shortable and non-shortable firms across the eight expansion waves of China's short-selling program (Panels A–G). Within each panel, "Shortable mean" and "Non-shortable mean" are the time-series averages of the corresponding anomaly variables. "Difference" represents the value difference between anomaly variables on shortable stocks and the ones on non-shortable stocks (There are eleven anomalies including eight positive return predictors, two negative return predictors (OSC, DSR) and one market momentum performance (MOM)). Abbreviations: GP (gross profitability), AG (asset growth), IVA (investment-to-assets), ROA (return on assets), NOA (net operating assets), ACC (accruals), ISS (net issuance), CEI (composite equity issuance), OSC (Ohlson O-score), DSR (downside risk), MOM (momentum). Obs. denotes the number of observations in each group. T-value reports the associated t-statistics. Statistical significance at the 1%, 5%, and 10% levels is denoted by ***, **, and *, respectively.

factors. We cluster stock into shortable and non-shortable groups on each pilot expansion and calculate the average stock characteristics in Table 3.4.

In Table 3.4, we find heterogeneous performance of different types of anomaly variables on pilot expansions. The difference means of anomalies on investment activity (AG & IVA) and equity issuance (ISS & CEI) themes show positive and significant signs. This positive signal indicates that stocks which are allowed to be sold short have stronger investment ability and willingness to raise capital by issuing shares. In contrast, The difference means of anomalies on tail risk theme (OSC & DSR) shows negative and significant signs across pilot expansions, suggesting that shortable firms exhibit lower financial distress risk and downside risk compared to non-shortable firms. This finding is not surprising, as shortable stocks are "blue chips", with solid financial performance. These shortable stocks are less likely to experience a bankruptcy or financial distress than non-shortable stocks. However, anomalies on asset composition theme (NOA & ACC) do not show consistent significant signs across pilot expansions. These two firms' characteristics seems not sensitive to short selling activities.

Momentum of firms shows positive and significant signs on the first two rounds of expansions then flips to negative in the following expansions. The reason is that the selected stocks to be sold have better performance than non-shortable stocks in the first two rounds. During the early period, the number of shortable stocks is limited (Chang et al, 2014). Investors also face comparative higher thresholds to sold stocks in the Chinese stock market. These limit investors to express their negative opinions on eligible stock prices, thus contributing to a good performance on momentum. While, increasing stocks are allowed to be sold short across expansions. Securities lending refinancing[4] releases more spaces to investors to sold short on stocks, leading to down prices around expansions. Hence, the sign on momentum flips after the second expansion.

In sum, the finding of significant positive/negative signs on several anomaly variables is consistent with the fact that Chinese stocks which are eligible to sold short need to meet

[4] Securities lending refinancing refers to the provision of securities by China Securities Finance Corporation to securities firms for the purpose of facilitating securities lending activities.
http://www.csrc.gov.cn/csrc/c106256/c1653885/1653885/files/a7650b7dc10740829f071881bdf13bc9.pdf

certain requirements on firm characteristics. Even if not each anomaly variable shows significant signs. While the sufficient anomalies with heterogenous performance in Table 3.4 indicate that the anomalies used in this study is comprehensive and thus facilitate us to investigate the short sale constraints on anomalies.

3.4.2. Short-Sale Constraint Expansions and Anomaly Returns

The previous results have shown that certain anomaly variables/firm characteristics of shortable firms are significantly different from that of non-shortable firms. An intuitive thinking is that does anomaly portfolios formed on these firm characteristics have the same pattern. In this section, we explore whether there is significant difference between anomalies formed on shortable stocks and those on non-shortable stocks. In other words, we investigate whether these anomalies are sustained by limits to arbitrage rather than by compensation for fundamental risk.

Based on equation (7), we estimate a series of difference-in-differences (DiD) regressions, as reported in Table 3.5. The setting features eight short-selling expansion waves. For each wave, we define the treatment group as stocks that become newly shortable in that wave and construct the corresponding anomaly long–short portfolio returns. The key treatment indicator, Pilot, equals one for treated (newly shortable) portfolios and zero for control portfolios. The post indicator, Dummy, is defined in event time relative to the effective month of each expansion. Specifically, for wave w, we compute EventTime as the number of months between calendar month (Yr, Month) and that wave's expansion effective date, and define pre as EventTime $\in [-12, -1]$ (Dummy = 0) and post as EventTime $\in [0, 11]$ (Dummy = 1). Observations outside the $[-12, 11]$ window are excluded. The DiD effect is captured by the interaction term Pilot × Dummy, which measures whether anomaly portfolio returns for newly shortable portfolios change differentially from pre to post, relative to the control portfolios, within the same event-time window for each wave.

Before examining the DiD estimates, we assess the parallel-trends assumption by conducting pre-treatment trend tests separately for each of the eight short-selling expansion waves and eleven anomaly portfolios. For each expansion wave, the test is estimated using the

12 months preceding the corresponding expansion and examines whether the pre-treatment time trend differs significantly between portfolios formed from stocks that subsequently become eligible for short selling and portfolios formed from stocks that remain non-shortable. The regressions additionally control for the Fama–French five factors. Detailed results are reported in Appendix Table B.1.

The pre-treatment trend tests provide broad support for the parallel-trends assumption. Among the 88 anomaly-wave tests, 78 (88.64%) show no statistically significant differential pre-treatment trend at the 5% level, while 83 (94.32%) are insignificant at the 1% level. At the 10% level, 74 of the 88 tests (84.09%) remain insignificant. Although significant pre-trends are observed for a limited number of anomaly-wave combinations, they are not systematically concentrated in the same direction or across the same expansion waves. Overall, the treated and control portfolios generally exhibit comparable pre-treatment trends, while some heterogeneity remains across individual anomalies and expansion events.

Table 3.5 shows that the effects of relaxing short-sale constraints on anomaly returns remain limited, with most differences being economically small and statistically insignificant across expansion waves, even after controlling for common risk factors using the Fama–French five-factor model. This pattern is consistent with recent evidence suggesting that many anomaly profits remain economically meaningful even when short-sale costs and risk factor exposures are taken into account (Muravyev et al., 2025).

Among the eighty-eight estimated interaction effects, only a small subset reaches conventional levels of statistical significance, and these significant estimates are scattered across anomalies and expansion waves. For example, gross profitability and composite equity issuance exhibit statistically detectable responses in isolated expansion phases, with coefficients that are negative in sign, indicating partial attenuation of anomaly-based returns among newly shortable stocks. However, these effects do not persist across subsequent waves. Several anomalies, including net operating assets, accruals, net equity issuance, and return on assets, display statistically significant DiD estimates in specific expansion waves. However, these responses are episodic and not uniform in sign across waves. In contrast, downside risk, momentum, and the Ohlson O-score show only weak and sporadic significance, with the

Table 3. 5. Difference-in-Differences Estimates of Anomaly Portfolio Returns Around Short-Selling Expansions

	Initial issuance	Expand 1	Expand 2	Expand 3	Expand 4	Expand 5	Expand 6	Expand 7
GP	-0.006	0.000	-0.027**	0.000	-0.024	0.004	0.019	-0.021
	[-0.27]	[-0.04]	[-2.22]	[-0.01]	[-1.43]	[0.20]	[1.13]	[-1.47]
AG	0.007	-0.008	-0.002	-0.013	-0.004	0.022	0.002	-0.032
	[0.35]	[-0.58]	[-0.12]	[-0.80]	[-0.17]	[1.47]	[0.10]	[-1.44]
IVA	-0.023	-0.009	-0.012	0.006	0.005	0.002	0.008	-0.029
	[-1.05]	[-0.41]	[-0.74]	[0.44]	[0.27]	[0.18]	[0.44]	[-1.09]
ROA	0.004	0.000	0.013	-0.008	-0.027	0.034*	0.017	0.007
	[0.19]	[0.01]	[0.74]	[-0.43]	[-1.23]	[1.90]	[1.02]	[0.23]
NOA	0.009	0.030**	-0.005	-0.001	-0.029	0.004	0.014	-0.013
	[0.40]	[2.28]	[-0.37]	[-0.06]	[-1.00]	[0.28]	[0.83]	[-0.51]
ACC	-0.010	-0.021	0.030**	0.004	0.016	0.023**	0.011	0.002
	[-0.46]	[-1.34]	[2.22]	[0.24]	[0.89]	[2.03]	[0.91]	[0.17]
ISS	0.017	-0.016	0.039***	0.011	-0.046*	0.007	0.002	0.002
	[0.81]	[-0.94]	[2.91]	[0.87]	[-1.85]	[0.61]	[0.17]	[0.08]
CEI	0.005	-0.029*	0.019	-0.027	0.023	0.024	-0.012	-0.018
	[0.19]	[-1.79]	[1.05]	[-1.26]	[0.73]	[1.51]	[-0.36]	[-0.74]
OSC	0.016	-0.004	0.009	-0.002	-0.001	0.033**	0.000	-0.010
	[0.72]	[-0.18]	[0.47]	[-0.12]	[-0.03]	[2.36]	[-0.01]	[-0.29]
DSR	0.023	-0.033	0.026	0.030*	-0.021	-0.010	0.015	0.007
	[0.72]	[-1.48]	[1.46]	[1.84]	[-1.04]	[-0.77]	[0.92]	[0.36]
MOM	0.016	-0.027	0.047*	-0.016	-0.025	-0.022	0.021	-0.004
	[0.41]	[-0.86]	[1.81]	[-0.48]	[-0.60]	[-0.68]	[0.60]	[-0.04]

Table 3. 5. (continued)

Note This table reports the difference-in-differences (DiD) results, as specified in equation (7), around the introduction and subsequent expansions of China's short-selling pilot program, covering eight expansion waves in total. For each expansion wave, we estimate the DiD specification separately using a GENMOD regression model. For each wave, the sample is restricted to a symmetric event window spanning 12 months before and 12 months after the effective expansion date. The treatment group consists of portfolios formed from stocks that become newly shortable in the corresponding expansion wave (Pilot = 1), while the control group includes portfolios formed from stocks that remain non-shortable during that wave (Pilot = 0). The post-period indicator (Dummy) equals one for months in the post-expansion period and zero for months in the pre-expansion period. The interaction term Pilot × Dummy captures the differential change in average monthly anomaly portfolio returns for newly shortable portfolios relative to control portfolios following the expansion. All regressions control for the Fama–French five factors, including the market risk premium (MKT), size (SMB), value (HML), profitability (RMW), and investment (CMA). Eleven anomaly strategies are examined: accruals (ACC), asset growth (AG), composite equity issuance (CEI), gross profitability (GP), net issuance (ISS), investment-to-assets (IVA), net operating assets (NOA), return on assets (ROA), downside risk (DSR), momentum (MOM), and Ohlson O-score (OSC). Corresponding z-statistics are reported in parentheses beneath each coefficient. Returns are expressed in monthly decimal form. Statistical significance at the 1%, 5%, and 10% levels is denoted by ***, **, and *, respectively.

Majority of estimates remaining statistically indistinguishable from zero across expansion phases. No single anomaly exhibits a stable and systematic response across multiple expansion waves, indicating that changes in short-selling eligibility at the extensive margin alone are insufficient to generate broad-based corrections in anomaly-based return patterns.

In many cases, newly shortable stocks do not exhibit materially different return dynamics relative to non-shortable stocks, implying that short-selling eligibility alone is insufficient to induce systematic pricing adjustments.[5] This observation highlights an important distinction between the extensive margin of short-selling eligibility and the intensive margin of actual short-selling activity.

These findings contrast with early evidence from the Chinese market. Using an event-study design, Chang et al. (2014) document improvements in price efficiency following the first two short-selling pilot expansions, but also show that such improvements are primarily driven by margin trading rather than short selling itself. Moreover, their analysis focuses on short event windows surrounding policy implementation. By contrast, our study exploits eight expansion waves and adopts a difference-in-differences framework that compares anomaly returns over a longer post-expansion horizon, reducing concerns that informed traders front-run policy announcements within narrow event windows.

One explanation for the results of Table 3.5 lies in the institutional features of China's trading mechanism, under which margin trading and short selling are introduced jointly. While short selling facilitates the incorporation of negative information into prices, margin trading simultaneously amplifies positive demand pressure. In practice, margin-buying activity dominates short-selling activity in both scale and frequency, reflecting higher participation thresholds and limited securities supply on the short-selling side (Wan et al., 2024). As a result, relaxing short-selling eligibility alone may have a muted net effect on price correction when actual short-selling activity remains limited. To illustrate this imbalance in trading activity, Table 3.6 reports annual summary statistics on margin-buying and short-selling volumes in the

[5] We also examine how short-sale liberalization affects anomaly-related mispricing, and the results are presented and discussed in Appendix B.2.

Chinese A-share market, showing that margin-buying activity consistently exceeds short-selling activity by a wide margin throughout the sample period.

Table 3. 6. Margin Trading and Short-Selling Activity in China

Year	Obs.	MT (%)	SS (%)	MT-SS (%)
2010	797	0.0268	0.0000	0.0268
2011	1261	0.0535	0.0013	0.0522
2012	3322	0.0833	0.0017	0.0816
2013	6465	0.2905	0.0002	0.2903
2014	8949	0.4172	0.0006	0.4166
2015	10113	0.1233	-0.0006	0.1239
2016	10252	-0.1437	0.0001	-0.1438
2017	10999	-0.0385	0.0000	-0.0385
2018	11109	-0.1926	0.0000	-0.1927
2019	14887	0.0561	-0.0006	0.0567
2020	21036	0.0584	0.0006	0.0578
2021	24699	0.0107	-0.0091	0.0197
2022	30339	-0.0435	-0.0095	-0.0339
2023	39889	0.0246	-0.0042	0.0288
2024	39587	0.0218	-0.0085	0.0303

Notes This table reports annual average ratios of margin trading and short-selling balances over market value in the Chinese A-share market. Margin trading balance (MT) refers to the outstanding financing balance (denominated in Chinese Yuan), while short-selling balance (SS) refers to the outstanding securities-lending balance. MT − SS captures the net imbalance ratio between long and short positions.

In this sense, the largely insignificant results in Table 3.5 should not be interpreted as evidence against the role of short selling in improving price efficiency. Rather, they suggest that short-selling eligibility at the extensive margin is insufficient to ensure effective arbitrage in the absence of meaningful short-selling activity. Accordingly, the next section examines whether variation in short-selling depth provides a more informative channel for understanding how short selling affects anomaly-based return patterns.

3.4.3. Anomaly-Based Strategies and Short-Selling Depth

3.4.3.1. Long–Short Portfolio Performance Across SSD Groups

In the previous sections, we have discussed that whether the short-sale relaxation influence stock anomalies. However, Chinese investor face relatively high barriers[6] to short selling and constraints on securities lending supply. These frictions limit the extent to which an eligible stock can actually be sold short. Hence, this essay introduces the short-selling depth (SSD) to measure short-selling intensity of a stock over a given period. Table 3.7 examines how firm-level anomaly characteristics vary with short-selling depth by sorting shortable stocks into deciles. The short-selling depth (SSD). is measured as the ratio of short-selling volume to total trading volume following Chang et al. (2013).

We find that the high-minus-low spread for most anomaly-related firm characteristics is positive and statistically significant, except for OSC and DSR. These characteristics are generally associated with positive return predictability. These firm characteristics are associated with positive return predictability. In contrast, negative return-predictive signals, such as OSC and DSR, exhibit negative high-minus-low spreads, although the DSR spread is statistically insignificant. While, negative return-predictive signals, such as OSC and DSR, exhibit negative high-minus-low spreads, even if DSR spread is insignificant.

These findings indicate that stocks with greater short-selling intensity tend to be concentrated in segments characterized by stronger anomaly signals or more extreme firm characteristics. The largely monotonic patterns across SSD deciles further suggest that short sellers do not randomly target stocks. Instead, higher short-selling activity is systematically associated with firms exhibiting stronger investment and issuance signals, such as AG, IVA, and CEI.

This pattern implies that short-selling activity is more prevalent in segments where characteristic-based mispricing is likely to be more pronounced. In contrast, the insignificant spread in DSR suggests that risk-related signals may not attract the same degree of arbitrage attention. Overall, the evidence supports the view that short sellers selectively engage with

[6] Only qualified investors are allowed to sold short in the Chinese stock market. Qualified investors need to have investment knowledge and historical trading records, with a maintenance capital of 500,000 Chinese Yuan minimum in their trading accounts, etc. Moreover, Chinese stock markets charge the interest rate for short sellers varies from 8% to 10.5%. The changes of the interest cost are evaluated based on the qualifications of investors and financial agents. The interest fee may vary across different brokerage firms.

Table 3. 7. Average Firm Anomaly Characteristics Across Short-selling Depth Deciles

	Low SSD									High SSD		
	1	2	3	4	5	6	7	8	9	10	H-L	t-value
GP	0.470	0.487	0.485	0.472	0.465	0.477	0.479	0.489	0.494	0.474	0.004	0.935
AG	0.120	0.115	0.108	0.104	0.100	0.101	0.113	0.118	0.124	0.133	0.013***	5.854
IVA	0.032	0.033	0.032	0.030	0.030	0.029	0.032	0.034	0.035	0.035	0.003***	3.972
ROA	0.033	0.034	0.034	0.027	0.021	0.024	0.030	0.034	0.037	0.036	0.003***	4.539
NOA	0.553	0.557	0.565	0.571	0.578	0.581	0.577	0.571	0.568	0.561	0.007***	2.939
ACC	0.015	0.014	0.015	0.014	0.013	0.013	0.015	0.015	0.016	0.016	0.002*	1.790
ISS	0.061	0.055	0.053	0.055	0.058	0.059	0.060	0.057	0.057	0.065	0.004**	2.546
CEI	0.388	0.330	0.279	0.233	0.213	0.229	0.291	0.322	0.385	0.453	0.065***	9.016
OSC	-8.794	-8.813	-8.786	-8.573	-8.345	-8.428	-8.601	-8.743	-8.845	-8.879	-0.085***	-4.255
DSR	0.027	0.023	0.024	0.028	0.034	0.034	0.028	0.025	0.024	0.026	0.000	-0.970
MoM	-0.015	-0.016	-0.009	-0.005	-0.003	0.001	-0.004	0.000	-0.003	-0.002	0.013***	8.004

Note This table reports the time-series mean value of firm-level characteristics across deciles. For each month in the sample, shortable firms are ranked into ten SSD deciles (1-10), and the time series means of eleven anomaly-related firm characteristics are computed within each decile. Decile 1 contains stocks with the lowest SSD. Decile 10 contains stocks with highest SSD. The column "H-L" reports the difference between the top and bottom deciles, and the final column presents the corresponding t-statistics. Characteristics include Gross Profitability (GP), Asset Growth (AG), Investment-to-Assets (IVA), Return on Assets (ROA), Net Operating Assets (NOA), Accruals (ACC), Net Issuance (ISS), Composite Equity Issuance (CEI), Ohlson O-Score (OSC), Downside Risk (DSR), and 12-month Momentum (MoM). ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

firms displaying stronger anomaly-related characteristics rather than distributing their activity uniformly across the cross-section.

3.4.3.3. Cross-Sectional Effects of Short-Selling Depth on Anomalies

Table 3.8 examines how anomaly-based portfolio returns differ between high and low short-selling depth groups within shortable stocks. We cluster all shortable stocks into high and low groups each month based on their short-selling depth (SSD), defined as short-selling volume over total trading volume. Stocks in the top five SSD deciles form the high SSD group, while those in the bottom five deciles form the low SSD group, capturing differences in investor attention and arbitrage pressure across firms with varying degrees of short-selling activity. Table 3.8 reports the average monthly returns for each anomaly-based long–short portfolio within the high and low SSD groups, alongside the difference between the two groups.
Notably, Asset Growth (AG), Investment-to-Assets (IVA), Return on Assets (ROA), Accruals (ACC), and Composite Equity Issuance (CEI) exhibit positive and significant differences between the high and low SSD groups. The positive spread is primarily driven by the substantially lower and often statistically insignificant anomaly returns in the high-SSD group. For example, AG yields 2.7% per month in the low-SSD group compared with 1.0% in the high-SSD group, resulting in a significant difference of 1.6% (t = 3.32). Similarly, ROA declines from 3.5% in the low-SSD group to –0.5% in the high-SSD group, producing a spread of 4.0% (t = 10.35). These results indicate that anomaly profitability weakens as short-selling depth increases.

For negative-return predictive signals, such as OSC and DSR, the sign of the differential reverses. OSC yields –2.0% in the low-SSD group but becomes statistically insignificant in the high-SSD group, resulting in a spread of –2.2% (t = –5.72). DSR displays an even stronger contrast, with –2.2% in the low-SSD group versus +1.6% in the high-SSD group. Momentum (MoM) also shows a significant reduction in magnitude across SSD groups. These results imply that short sellers are particularly active in stocks characterized by distress or downside risk, accelerating price adjustment in those segments.

Table 3. 8. Anomaly Portfolio Returns by Short-Selling Depth (High vs. Low SSD Groups)

	Low	High	Difference (Low – High)
GP	0.001	0.000	0.001
	[0.36]	[0.06]	[0.24]
AG	0.027***	0.010**	0.016***
	[7.16]	[2.28]	[3.32]
IVA	0.016***	0.000	0.016***
	[4.65]	[-0.09]	[3.93]
ROA	0.035***	-0.005	0.040***
	[9.92]	[-1.00]	[10.35]
NOA	0.005	-0.006	0.011***
	[1.50]	[-1.44]	[3.04]
ACC	0.010***	-0.001	0.010*
	[2.83]	[-0.25]	[1.88]
ISS	0.005	0.005	0.000
	[1.58]	[1.54]	[0.08]
CEI	0.024***	0.008	0.016***
	[4.73]	[1.59]	[3.51]
OSC	-0.020***	0.003	-0.022***
	[-6.03]	[0.80]	[-5.72]
DSR	-0.022***	0.016***	-0.038***
	[-5.16]	[4.09]	[-7.14]
MOM	-0.028***	-0.014**	-0.014***
	[-4.80]	[-2.51]	[-2.96]

Note This table reports the average long–short returns of anomaly portfolios among shortable stocks for low and high groups, classified by short-selling depth (SSD). Each month, firms are split into High-SSD and Low-SSD groups. SSD is defined as the ratio of short-selling volume to total trading volume. The securities-lending balance is defined as the net trading value (denominated in Chinese Yuan) of short sales on a stock within a month. Within each SSD group, stocks are sorted by each anomaly characteristic and value-weighted long–short portfolios are constructed. The column "Diff" reports the difference in long–short returns between High-SSD and Low-SSD groups. T-statistics (in brackets) are based on robust standard errors clustered at the firm level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

In contrast, Gross Profitability (GP) and Net Issuance (ISS) show economically small and statistically insignificant differences across SSD groups, indicating that the influence of short-selling intensity is heterogeneous across anomaly dimensions.

Taken together, Table 3.8 provides direct cross-sectional evidence that anomaly profitability is systematically stronger when short-selling depth is limited and weakens as short-selling intensity increases. This pattern supports the view that the intensity of arbitrage activity, rather than short-selling eligibility alone, is an important determinant of whether characteristic-based mispricing persists in the Chinese A-share market.

3.4.4 Risk-Adjusted Evidence on Short-Selling Depth

Previous results have shown that stocks with higher short-selling intensity are more likely to have higher values of anomaly firm characteristics which may act as trading signal to attract investors. The significant differences between the high and the low short selling intensity groups reported in Table 3.8 indicates that the intensive short selling activities may help reducing the profitability of anomaly portfolios. To formally verify this conjecture, we further conduct regressions of anomaly returns on short-selling depth (SSD) while adjusting for the Fama–French five factors in Table 3.9. This regression formula is shown in equation (9) in Section 3.3.6. The anomaly portfolios in Table 3.9 are constructed in an SSD-sorted way. At each month, the shortable stock sample is partitioned into high and low short-selling depth groups, defined by the top and bottom five SSD deciles, respectively. Then, long–short portfolios are constructed within these groups, and monthly value-weighted returns are computed. Portfolio returns are then regressed on a dummy variable indicating low SSD exposure, alongside the Fama–French five factors, including the market excess return, size, value, profitability, and investment.

The results in table 3.9 reveals that short-selling depth significantly affects the profitability of anomalies by reducing anomaly portfolio returns. Among the eleven anomalies examined, six display significantly positive coefficients on the SSD dummy: Asset Growth (AG), Investment-to-Assets (IVA), Return on Assets (ROA), Net Operating Assets (NOA), Accruals (ACC), and Composite Equity Issuance (CEI). For example, the dummy coefficient for ROA is 0.040 ($t = 7.85$), and for AG it is 0.016 ($t = 2.86$), indicating that these anomalies yield higher returns in stocks with lower short-selling activity. These results suggest that limited short-selling depth sustains mispricing associated with these firm characteristics, consistent

Table 3. 9. Effects of SSD on Anomaly Returns

	GP	AG	IVA	ROA	NOA	ACC	ISS	CEI	OSC	DSR	MOM
Intercept	0.000	0.010**	0.000	-0.004	-0.005	-0.002	0.004	0.008*	0.002	0.016***	-0.014***
	[-0.15]	[2.55]	[-0.05]	[-1.07]	[-1.49]	[-0.46]	[1.37]	[1.68]	[0.53]	[4.27]	[-2.64]
SSD Dummy	0.001	0.016***	0.016***	0.040***	0.011**	0.011**	0.000	0.016**	-0.024***	-0.038***	-0.014*
	[0.21]	[2.86]	[3.43]	[7.85]	[2.49]	[2.17]	[0.07]	[2.49]	[-5.26]	[-7.36]	[-1.89]
RiskPremium2	-0.133***	-0.026	-0.064	-0.167***	0.165***	-0.164***	-0.048	0.075	0.198***	0.040	-0.293***
	[-3.40]	[-0.52]	[-1.51]	[-3.70]	[4.21]	[-3.64]	[-1.37]	[1.31]	[4.85]	[0.88]	[-4.31]
SMB2	-0.020	0.016	-0.024	-0.043	-0.056	0.135	0.067	-0.103	0.096	-0.063	0.185
	[-0.23]	[0.15]	[-0.26]	[-0.44]	[-0.66]	[1.39]	[0.89]	[-0.84]	[1.10]	[-0.63]	[1.26]
HML2	-0.908***	-0.487***	-0.529***	-0.564***	0.896***	-0.414***	-0.430***	-1.318***	-0.037	-0.089	0.022
	[-7.42]	[-3.09]	[-4.01]	[-3.99]	[7.33]	[-2.94]	[-3.93]	[-7.45]	[-0.29]	[-0.62]	[0.11]
RMW2	0.377**	0.746***	0.461**	1.071***	0.192	0.358	-0.060	0.881***	-0.518***	-1.279***	-0.786**
	[2.00]	[3.08]	[2.27]	[4.93]	[1.02]	[1.64]	[-0.36]	[3.23]	[-2.66]	[-5.68]	[-2.41]
CMA2	0.876***	-0.009	-0.308	0.128	-0.450**	-0.237	-0.307*	0.494*	0.434**	-0.236	-0.551*
	[4.83]	[-0.04]	[-1.58]	[0.61]	[-2.49]	[-1.12]	[-1.89]	[1.89]	[2.26]	[-1.11]	[-1.76]

Note This table reports panel regression results of anomaly-portfolio returns regressed on short-selling depth (SSD), as specified in equation (9). The dependent variable is the monthly value-weighted long–short return for each anomaly. The key independent variable is SSD Dummy variable. SSD Dummy equals 1 for portfolios formed from firms with the bottom 50% SSD and 0 for those from firms with the top 50% SSD. SSD (short-selling depth) is measured as the ratio of the outstanding securities-lending balance (shares borrowed and not yet returned) to the firm's market capitalization. All regressions include the Fama–French five-factor controls: Market excess return (RiskPremium), Size (SMB2), Value (HML2), Profitability (RMW2), and Investment (CMA2). T-statistics (in brackets) are based on Newey–West standard errors with 12 lags. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

with the limits-to-arbitrage hypothesis. This interpretation aligns with Stambaugh, Yu, and Yuan (2012), who show that mispricing-related return predictability persists even after controlling for standard risk factors when arbitrage is constrained. In contrast, Downside Risk (DSR), Momentum (MOM) and Ohlson O-Score (OSC) display negative SSD coefficients, although the effect is statistically significant only for DSR, suggesting that short-sellers selectively exploit certain anomaly-based mispricing to accelerate price correction. The remaining anomalies, including Gross Profitability (GP) and Net Issuance (ISS), display economically small and statistically insignificant SSD coefficients, suggesting that short-selling activity plays a limited role in shaping their profitability. To sum, the short-selling depth reducing portfolio returns for anomalies, based on positive predictors, and increases returns for anomalies based on negative predictors.

The mixed responses of anomalies to SSD highlight the heterogeneous role of short sellers in price formation. For anomalies linked to investment intensity and earnings manipulation, such as AG, IVA, ACC, and CEI, reduced SSD appears to allow prices to deviate further from fundamentals, sustaining the profitability of long–short strategies. This pattern aligns with the idea that lacking arbitrage will causing systematic mispricing. Anomalies tied to risk-based characteristics, such as DSR and OSC, show evidence that active short-sellers trade against mispricing, thereby eroding anomaly returns where SSD is higher. Taken together, these findings suggest that short-sellers play a corrective role in pricing dimensions.

The results present a nuanced picture of how short-selling constraints shape anomaly returns. Notably, the SSD dummy remains economically large and statistically significant after controlling for the Fama–French five factors, indicating that the documented effects are not driven by systematic risk exposures. This provides incremental evidence that short-selling intensity explains cross-sectional variation in anomaly returns beyond conventional factor models. Table 3.5. documents that the relaxation of short-sale restrictions only partially attenuates anomaly premiums, with significant effects concentrated in asset growth and profitability-related factors, consistent with the limits-to-arbitrage hypothesis. Table 3.6. corroborates these findings by showing that stocks newly eligible for short-selling exhibit smaller mispricing signals, but the impact is neither uniform across anomalies nor persistent over all phases. Table 3.7. further confirms the cross-sectional predictive power of several anomalies using value-weighted long–short portfolios, while Table 3.8. highlights that anomaly profitability varies systematically across firm size and book-to-market dimensions, indicating heterogeneous effects of arbitrage frictions.

Table 3.9. delivers the most rigorous evidence by jointly controlling for Fama-French five-factor exposures and short-selling depth. The findings reveal that anomalies such as asset growth, investment-to-assets, return on assets, and accruals remain significantly more profitable among stocks with limited short-selling activity, underscoring the persistent role of arbitrage constraints even after accounting for systematic risk factors. Overall, the empirical results demonstrate that while short-selling liberalization enhances market efficiency, its effects are neither uniform across anomalies nor complete, suggesting that frictions continue to sustain predictable return patterns in China's emerging market.

3.5. Conclusion

This essay examines how short-selling constraints affect anomaly-based trading strategies in the Chinese A-share market. Using phased expansions of the short-selling programme and stock-level short-selling depth (SSD) constructed from securities-lending activity, we evaluate whether variations in arbitrage capacity, captured by both short-sale eligibility and short-selling depth, are associated with the profitability of characteristic-based anomalies. Eleven anomalies spanning profitability, investment, accruals, issuance, distress, and trading frictions are analysed through portfolio sorts, DiD estimation, SSD-based comparisons, and factor-adjusted panel regressions.

First, a broad set of characteristics continues to deliver economically meaningful and, in many cases, statistically significant long–short returns in A-shares. Many of these premiums remain after adjusting for Fama–French five-factor exposures, indicating that the anomalies are only partially absorbed by standard risk factors.

Second, short-selling eligibility has heterogeneous effects across characteristics. For several anomalies associated with overvaluation (e.g., asset growth), long–short returns are lower among stocks eligible for short selling. For other anomalies, including accruals and composite issuance, differences are small or not consistently negative. Difference-in-differences estimates around pilot expansions deliver limited and scattered evidence: while individual waves and signals show reductions in returns for newly shortable stocks, the majority of coefficients are economically small and statistically insignificant.

Short-selling depth plays a more systematic role. Sorting shortable stocks by SSD shows that anomaly returns are generally higher when short-selling depth is low, and substantially weaker when depth is high. SSD effects are strongest for anomalies linked to investment intensity, earnings quality, and issuance activity. By contrast, distress-related characteristics (O-score and downside risk) exhibit more negative long–short returns in the high-SSD group, consistent with short sellers targeting firms with elevated default or downside risk. Panel regressions controlling for Fama–French factors reinforce these patterns: SSD is negatively associated with anomaly returns for several profitability-, investment-, and issuance-based signals, indicating that anomaly profitability weakens as short-selling intensity increases.

Overall, the findings indicate that short-selling liberalisation and the subsequent development of short-selling depth have contributed to a partial reduction in anomaly profitability, but that many premiums persist in segments of the market where short-selling depth remains low. For regulators, the results highlight that expanding eligibility is only one component of improving price efficiency; the functioning and depth of the lending market are also important. For investors, the evidence suggests that characteristic-based strategies continue to yield positive returns in environments where short-selling constraints remain binding.

CHAPTER FOUR: ESSAY THREE

Climate Risk Disclosure and Return Anomalies: A Textual Analysis of Chinese A-Shares

This chapter presents the third essay, which examines whether firm-level climate risk disclosure gives rise to return anomalies in China's A-share market and whether a climate-risk-disclosure factor commands a return premium beyond conventional risk and policy uncertainty factors.

4.1. Introduction

We explore the return-anomaly outcome of firms' climate risk disclosure (CRD) via constructing CRD scores and applying them to China's A-share firms. Specifically, aided by the scores, we investigate whether firm-level CRD is relevant for generating some return anomalies on top of historically observed characteristic-based anomalies. This question has not yet been touched in the existing literature but deserves examination for the reasons detailed below.

Our study views CRD as a new firm characteristic, and coins the term "CRD characteristic". This aspect goes beyond conventional firm characteristics such as firm size, book-to-market, profitability, and so on. It is well documented that some characteristic-based factors can predict cross-sectional variation in expected returns. Fama and French (1993) find that factors related to firm size and book-to-market, i.e., SMB and HML, have return predictive power beyond a market factor (MKT), and conclude that size and book-to-market equity proxy for sensitivity to common risk factors in stock returns. Hirshleifer et al. (2012) document that an accrual-based factor has the ability to predict returns after controlling for other common factors, suggesting that investors misprice the accrual characteristic, causing the accrual anomaly to persist. Alti and Titman (2019) take the profitability and growth rates of firms as behavioural elements in their proposed dynamic model. The authors outline that, due to investors' overconfidence in their ability to assess the disruption climate, firms with different exposures to disruption – e.g., value versus growth firms – exhibit predictable return differences. While such a literature thread is far too large to fully discuss here, the above findings and explanations regarding characteristic-based anomalies suffice to motivate us to delve into the CRD-characteristic-based anomalies, [7] whether grounded on a risk narrative or on a misvaluation narrative.

Note that one would not take into account the CRD-characteristic-based anomalies from time periods when the detrimental effects of climate change, and hence climate risk management, had not yet been brought to the attention of governments and researchers world-

[7] The predictable patterns are dubbed "anomalies" as they are inconsistent with maintained theories of rational asset pricing behaviour, and they indicate either market inefficiency or inadequacies in the underlying asset pricing model (Schwert, 2003).

wide. It is now widely recognised that climate change as a risk can be categorised into physical risks and transition risks. These risks pose significant challenges to firm performance and valuation (Bansal et al., 2017). Physical risks, such as extreme weather events, directly impact firms' operations and supply chains. Transition risks, stemming from regulatory changes and evolving consumer preferences, affect firms' strategic planning. These risks are not only material but also systemic, influencing both industries and financial markets. The disclosures of climate risks could play a pivotal role in mitigating information asymmetry between firms and investors. More transparent disclosures enable investors to better evaluate a firm's long-term viability and hence reduce the risk exposures of their investments in the firm or reduce the degree of stock misvaluation. This is particularly so in uncertain environments (Chatterji et al., 2009).

Physical and transition climate risks may affect the cross-section of stock returns through distinct economic channels. Physical climate risk primarily affects firms through direct exposure to acute events, such as floods, storms, and extreme temperatures, and chronic changes in climatic conditions. For example, following a severe flood, a manufacturing firm with production facilities concentrated in the affected region may experience plant closures, asset damage, and supply-chain disruptions, whereas a firm with geographically diversified facilities may experience little direct impact. Even when the two firms have otherwise similar financial characteristics, this difference in physical exposure can generate different revisions to expected cash flows and, consequently, different stock-return responses. Because these exposures vary across firms and industries, investors may require different expected returns as compensation for bearing physical climate risk.

Transition climate risk, in contrast, arises from the adjustment toward a low-carbon economy. Changes in environmental regulation, carbon pricing, technology, energy systems, and consumer preferences can alter firms' compliance costs, investment requirements, competitive positions, and the value of carbon-intensive assets. For example, the introduction or tightening of carbon-emission regulation may impose substantially greater compliance and investment costs on a carbon-intensive coal, steel, or power producer than on a low-emission firm. At the same time, firms specialising in renewable energy or energy-efficient technologies may benefit from the same transition through increased demand and improved growth opportunities. These heterogeneous exposures can therefore produce different valuation and return effects across firms even when they face the same regulatory change. Firms that are more exposed to these transition pressures may therefore face different expected cash flows and required returns from firms that are better positioned to adapt (TCFD, 2017; Bolton et al., 2020; Sautner et al., 2023).

These differences also provide potential channels through which climate risk can contribute to return anomalies. If investors do not fully or immediately incorporate heterogeneous climate exposures into prices, firms may become systematically misvalued relative to their underlying climate-related risks. For instance, two firms with similar book-to-market ratios may appear comparable under a conventional value strategy, but a firm with substantial physical exposure to flood-prone locations may face materially different future cash-flow risk from a firm with limited physical exposure. Similarly, among firms classified as high investment or high asset growth, a carbon-intensive firm undertaking large expenditures to comply with transition policies may have a different return implication from a firm whose investment reflects ordinary business expansion. Climate exposure can therefore introduce economically meaningful heterogeneity within portfolios formed on conventional anomaly characteristics. Physical risks may be difficult to price because the timing and severity of climate events are uncertain, whereas transition risks depend on evolving regulation, technological change, and market expectations. In both cases, incomplete information or differences in the interpretation of climate-related information can generate cross-sectional return predictability. Climate-risk disclosure is therefore economically relevant because it can reduce information asymmetry regarding these exposures and facilitate their incorporation into stock prices.

Meanwhile, climate risk disclosures are often embedded in broad Environmental, Social, and Governance (ESG) frameworks, which may dilute the focus on climate-specific factors (Chatterji et al., 2009; Duan et al. 2025). Such dilution likely hampers the precise assessment of how climate risk disclosures impact stock returns. If investors are overconfident about their ability to evaluate the messages conveyed by the disclosures, the CRD-characteristic-sorted portfolios may exhibit persistent mispricing, leading to return anomalies.

All in all, there emerges an urgent need to bring to light the relevant cause-effect nexus (Ilhan et al., 2021) between CRD and anomalies. To achieve this, we propose a firm-level CRD measure using textual data from annual reports of firms traded in the Chinese A-share market, in which retail investors' trades provide 82% of the total market liquidity (Yu et al., 2019a). Retail investors' behavioural biases and limited access to relevant information amplify market inefficiencies (Li, 2017; Pan et al., 2016; Gwilym et al., 2016). The lack of transparency may lead irrational or less sophisticated investors to misestimate, and overreact to, the risks associated with lower-disclosure firms due to incomplete or ambiguous information. As a result, they would demand an equity premium to compensate for perceived uncertainty and risk, which is then reflected by expected returns on lower-disclosure stocks being higher than those on higher-disclosure stocks.

Meanwhile, the significance of climate governance has been increasing in China. In 2021, the Chinese government launched a national emissions trading scheme to encourage firms to enhance their disclosure practices, aligning with its interim goal of reaching peak carbon emissions by 2030 and its long-term objective of achieving carbon neutrality by 2060. Although ESG development in China began later than in Europe and the U.S., its progress has been rapid, with the number of social responsibility reports increasing significantly from 32 in 2006 to over 2,000 by 2019 (Albuquerque et al., 2019). However, unlike Europe where mandatory ESG disclosure policies are prevalent, China relies on a voluntary approach, emphasising firms' proactive engagement in sustainability initiatives. By contrast, the U.S. emphasises market-driven mechanisms and shareholder activism to influence corporate disclosure practices. These differences in regulatory frameworks highlight the varying roles of policy and market forces in shaping climate risk transparency across the East and the West. Together with the behavioural biases of retail investors, China's policy-driven approach also provides a unique context in which to explore how climate risk disclosure influences asset pricing and return anomalies.

Quantifying firms' climate risk disclosures is challenging but conducive to our research. This study follows Lin and Wu (2023) and constructs a robust CRD measure to leverage textual analysis of corporate annual reports of Chinese A-share firms from 2002 to 2022. Unlike traditional financial data analysis, textual analysis captures nuanced contextual and linguistic patterns in corporate disclosures, enabling the identification of implicit signals about firms' climate strategies and risks (Ilhan et al., 2021; Luo et al., 2015). Our CRD measure is based on a lexicon of 155 climate-related keywords derived from authoritative sources, including Chinese government work reports and international policy documents. We calculate CRD as the ratio of the total count of climate-related terms to the total word count in annual reports, to generate a standardised and scalable metric for assessing firms' climate risk communication.

With the CRD measure constructed, we first examine its distribution across industries and firm characteristics, to identify sector-specific patterns in CRD and in its connections with other firm attributes, such as size, age, and profitability. These patterns demonstrate the heterogeneity in corporate climate communication practices. Subsequently, we investigate the impact of CRD on stock returns through both stock-level and portfolio-level analyses. The stock-level analysis relies on fixed-effect panel regressions, and the portfolio-level analysis utilises Fama-Macbeth regressions. The stock-level results reveal that stocks with lower levels of CRD consistently exhibit higher returns, suggesting that investors demand a risk premium for holding stocks with lower climate risk transparency. In the portfolio-level analysis, we form four zero-investment hedging portfolios (F_{CRD}) based on CRD deciles to robustify the result that F_{CRD} commands a significantly positive premium. In other words, we find strong evidence that F_{CRD} does possess return predictive power for future 25 size-B/M portfolio returns in the cross-section. Thus, we can conclude that both the stock-level and the portfolio-level analyses point to the presence of CRD-characteristic-based return anomalies in China.

The implications of our results are threefold, pertaining respectively to investors, policymakers, and corporate managers. For investors, the results suggest that incorporating CRD into investment strategies can identify mispriced assets, enhance portfolio performance while address sustainability goals (Bolton & Kacperczyk, 2021). For policymakers, the results underscore the need for a more standardised, transparent, and even mandatory climate risk disclosure framework to reduce CRD discrepancies and information asymmetries to improve market efficiency. Initiatives such as the Task Force on Climate-related Financial Disclosures

(TCFD, 2017) provide a useful foundation but require further adaptation for emerging markets like China (Xie et al., 2021a). For corporate managers, the results imply the tangible benefits of improving climate transparency, including attracting long-term investors, aligning with regulatory expectations and reducing the cost of capital, as a lower equity premium generally implies a lower cost of capital for a company.

Our study contributes to the literature in three main aspects. First, we construct a new firm-level measure of CRD using textual analysis techniques following Lin and Wu (2023). However, we utilise the measure as an additional firm characteristic – i.e., CRD characteristic, to capture the extent of firms' willingness to voluntarily disclose and manage climate risks over the full period that witnesses the evolution of China’s climate policy regulations and corporate CRD practices. Second, we establish novel evidence about how the CRD measure explains/predicts stock/portfolio returns, by isolating the role of the F_{CRD} factor from a broad ESG framework in asset pricing. Third, our China-related findings imply that the set of characteristic-sorted factor-mimicking portfolios should now be updated to include a CRD-characteristic-sorted one. This update is propelled by the recent emergence and development of climate governance around the world.

The rest of this chapter is organised as follows. Section 4.2 reviews the relevant literature, providing theoretical and empirical foundations for the present study. Section 4.3 is concerned with the methodologies employed in this study. Section 4.4 describes the data and provides summary statistics. Section 4.5 carries out the empirical analyses, and reports the characteristics-related results, and the stock-level and portfolio-level findings. In section 4.6, we offer concluding remarks by summarising our results and briefly discussing their policy implications.

4.2. Literature Review

This section reviews selected articles that are strongly relevant to our study, covering three areas: (1) characteristic-based anomalies; (2) China’s A-share market; and (3) measuring climate risk disclosure.

4.2.1. Characteristic-based anomalies

An anomaly return is commonly defined as a security's actual return that deviates significantly from what is expected based on established financial models or theories. The unexpected returns can arise from various factors, including, among other things, firm characteristics. Existing studies have found that the following firm characteristics matter for stock returns: market capitalisation (Banz, 1981), book-to-market ratio (Basu, 1977), prior returns (Jegadeesh & Titman, 1993), asset growth measures (Titman et al., 2004), investment activities (Lyandres et al., 2008), financing activities (Bradshaw et al., 2006), profitability measures (Novy-Marx, 2013), financial distress (Avramov et al., 2013), accrual (Hirshleifer et al., 2012), and stock styles (Barberis et al., 1998) among others. While dissecting seven characteristic-based anomalies, Fama and French (2008) outline the underlying rationale to be either risk-related characteristics of the expected return or the results of behavioural biases. Alti and Titman (2019) treat the profitability and growth rates of firms as behavioural elements in their dynamic behavioural model, and refer to the resultant return predictability as characteristic-based anomalies. All these traditional firm characteristics are financial in nature.

Recently, researchers have turned to exploring whether some non-financial firm characteristics, such as climate-change-related ones, also affect investors and stock returns. This gradual shift is driven by the fact that climate change has been increasingly considered to be a danger to the economic/financial system, and so governments throughout the world have paid increasing attention to climate risk. As a result, regulators now require firms to disclose climate risk in some form, thereby catalysing research that incorporates climate change issues in general and climate risk disclosures in particular into asset pricing models.

We deem some representative studies especially relevant to our current research. Bolton and Kacperczyk (2021) find that firms with higher carbon emissions earn a carbon premium, particularly in climate-sensitive industries. Therefore, targeting climate risk facilitates precise risk assessment and portfolio construction, aligning with investor preferences for sustainable investment strategies. Ilhan et al. (2023) probe the relation between climate risk disclosures and institutional investors. Their work documents that institutional investors value and demand climate risk disclosures, so better firm-level climate risk disclosures encourage institutional investors to increase their holdings of the firms' stocks. The authors rely on survey data and firm-level climate risk disclosure data readily available from the Climate Disclosure Project. Lin and Wu (2023) focus on China in their examination of the association between CRD and stock price crash risk. Using a self-constructed CRD measure, the authors show that climate risk disclosures play a critical role in reducing stock price crash risk by improving transparency and reducing information asymmetry.

Though we have exhausted the literature, reviewing the above selected representative studies already enables us to draw the following conclusions. Traditional firm characteristics, or some of them, have been found to explain/predict stock return anomalies. Since climate change poses a risk to the economic/financial system, the requirements for firms to disclose information about climate risk have led to the emergence of a new firm characteristic. We refer to this non-financial characteristic as the "CRD characteristic" to differentiate it from traditional firm characteristics, as inspired by Alti and Titman (2019). Our endeavour in this study is to reveal its implications for return anomalies. Filling the void in the literature as to whether and how the CRD characteristic can explain/predict stock returns is an important topic to pursue.

4.2.2. China's A-share market

China's A-share market is characterised by some unique features that distinguish it from developed markets, notably the dominance of retail investors and the significant information asymmetry in the market. Retail investors account for the majority of total trading volumes, often exhibiting speculative behaviour and high sensitivity to public announcements (Tian et al., 2018a; 2018b). This investor composition exacerbates market volatility and heightens the role of corporate disclosures in shaping investor sentiment. Additionally, the limited presence of institutional investors reduces the market's capacity to efficiently process and reflect fundamental information in asset prices (Yu et al., 2019b).

In recent years, China has made significant strides in environmental and climate policies, aligning its domestic initiatives with global frameworks. Early legislation, such as the

Cleaner Production Promotion Law and the Renewable Energy Law (2005), established a legal foundation for sustainable development. The 2007 National Climate Change Program outlined strategies for reducing greenhouse gas emissions and adapting to climate change. In 2009, the government pledged to cut carbon intensity by 40%-45% by 2020 compared to 2005 levels. On the international stage, China joined the Paris Agreement in 2016, demonstrating its alignment with global climate governance. Building on this foundation, the 2020 announcement of China's carbon neutrality target by 2060 marked a pivotal moment, leading to stringent environmental regulations and incentives for corporate environmental responsibility. The introduction of China's national emissions trading scheme (ETS) in 2021 further underscored the government's commitment to integrating climate considerations into economic policies (Xie et al., 2021b). These policy developments have prompted firms to enhance their climate risk disclosures, as investors and regulators increasingly demand transparency regarding environmental risks and opportunities.

Given the afore-mentioned market and regulatory conditions, climate risk disclosures hold increasing significance in China. In a market dominated by speculative trading and short-termism, climate risk disclosures provide a standardised metric that captures firms' climate risk management and communication. By addressing information asymmetry and reducing speculative distortions, the disclosures help enhance the market's ability to evaluate firm-level resilience to climate risks. This aligns with prior research highlighting the importance of tailored climate risk disclosure in improving market efficiency, particularly in emerging economies (Luo et al., 2015). Thus, we consider China an ideal research laboratory to probe into the CRD-characteristic-based anomalies.

4.2.3. Measuring climate risk disclosure

One prominent method of measuring climate risk disclosures involves a textual analysis of corporate documents, such as annual reports and sustainability disclosures. Early studies, including Michelon et al. (2015), used keyword frequency analysis to quantify the extent of climate-related information in corporate communications. By identifying terms such as "carbon", "emissions" and "renewable", and standardising their occurrence relative to document length, these studies provide approximate measures of climate risk communication. However, such approaches often lack contextual nuance, which is critical for understanding the depth and intent of disclosures.

More advanced approaches employ Natural Language Processing (NLP) techniques to capture thematic and tonal content in climate risk disclosures. For instance, Jiang (2019) used Latent Dirichlet Allocation (LDA) to categorise climate-related topics within corporate reports, offering a nuanced understanding of how firms address specific climate risks. Machine learning models, such as those based on Bidirectional Encoder Representations from Transformers (BERT) or Generative Pre-trained Transformer (GPT), further enhance precision by identifying specific disclosure elements, including regulatory compliance and adaptation strategies (Huang & Li, 2020). However, these sophisticated methods require substantial computational resources and large, annotated datasets, posing challenges for scalability.

In the context of the Chinese market, Lin and Wu (2023) propose a CRD measure tailored to the unique regulatory and market environment. Their methodology involves a textual analysis of corporate annual reports, drawing on climate-related keywords selected from Chinese government work reports. This approach leverages official policy documents as authoritative sources for climate-related terminology, ensuring that the lexicon reflects the priorities and language of China's regulatory framework. Government work reports provide a consistent and policy-related basis for understanding how firms align their disclosures with national climate goals. This alignment is particularly relevant given China's strong policy-driven market dynamics.

While Lin and Wu's (2023) CRD ratio offers a useful reference point, we construct our own CRD measure. There are three notable differences between our measure and Lin and Wu's. The first is a longer policy horizon, which aligns with national planning terminology. We expand the coverage of government work reports, from 2016–2020 in Lin and Wu (2023), to 2002–2022, to allow the keyword set to reflect a broader and more stable set of policy signals. This enables us to capture climate-related terms associated with long-term national development strategies, such as "Scientific Outlook on Development", "West-East Gas Transmission", "West-East Power Transmission", and "Western Development Strategy". These terms reflect institutional emphasis on climate governance through infrastructure and regional planning, which are less observable in shorter time windows.

The second is manual synonym extension over vector-based expansion. Lin and Wu (2023) rely on a Word2Vec-based similarity model for expanding the keyword list. While this technique has an automation advantage, it may introduce ambiguity and reduce transparency, particularly in the context of Chinese-language analysis where semantic precision is more difficult to guarantee. In contrast, we adopt a manual approach to synonym extension, which ensures that all added terms are directly relevant, context-specific, and semantically consistent with core policy vocabulary.

The third is the exclusive use of annual reports for disclosure consistency. Due to their sample period being only approximately five years, Lin and Wu (2023) supplement annual reports with quarterly announcements. In our case, the longer historical window (stretching over approximately eleven years) provides sufficient observations from annual reports alone, which are standardised, regulator-reviewed, generally more comprehensive, and relatively inerratic. Conversely, as they cover a shorter period, quarterly reports tend to be more operational in focus, introducing considerable variation in disclosure quality and structure.

4.3. Methodology

4.3.1. Measuring CRD

To construct the CRD measure, we first download the annual government work reports from the official website of the Chinese government.[8] The period of the reports is from 2002 to 2022. Then, we analyse the frequency of words related to environment, energy, and climate in these reports, and record the frequency of each keyword across different years, following the method of Luo et al. (2015). Through this analysis, we identify a set of the top 110 most frequently appearing keywords as a seed word collection. Next, we manually add common synonyms for some of these keywords to minimise matching omissions, which results in a final set of 155 keywords. These steps help us understand the Chinese government's attention to

[8] The official website of the Chinese government is www.gov.cn

climate risk-related issues.

The use of Government Work Reports as the starting point for dictionary construction is intended to ensure that the keyword set reflects climate-, environmental-, and energy-related terminology that is relevant to China’s institutional and policy setting over the sample period. Rather than relying exclusively on a climate dictionary developed in another market or language, the procedure identifies terminology repeatedly used in official Chinese policy documents. The initial frequency-based screening retains the 110 most frequently appearing climate-related terms as the seed-word collection. Common synonyms and closely related expressions are subsequently added to reduce potential matching omissions arising from variation in the terminology used by listed firms in their annual reports. After screening and removing duplicative or unsuitable expressions, the resulting dictionary contains 155 climate-related keywords. This dictionary-based approach is consistent with established financial textual-analysis methods that construct domain-specific lexicons to extract economically meaningful information from corporate texts (Loughran & McDonald, 2011), as well as the growing literature using textual information to measure firm-level climate-related exposure and disclosure (Sautner et al., 2023; Bingler et al., 2022).

To improve the transparency and economic interpretation of the dictionary, the final 155 keywords are organised using a structured climate-risk taxonomy based primarily on the TCFD framework. The keywords are first classified into two overarching types, Risk and Opportunity. Risk is further divided into Physical Risk and Transition Risk: Physical Risk comprises Acute and Chronic Physical Risk, while Transition Risk comprises Policy and Legal, Technology, and Market Risk. Climate-related Opportunity comprises Resource Efficiency, Energy Source, Products and Services, Markets, and Resilience. The classification is additionally informed by terminology relevant to China’s climate-policy and institutional setting. AI-assisted semantic classification is used to support the allocation of the final keywords across these categories, followed by consistency checks against the TCFD framework and the original meaning of the Chinese terms. The classification is used to improve the interpretability of the dictionary and does not alter the baseline frequency-based construction of the CRD measure. The complete classification and English translations of all

155 keywords are reported in Appendix C.

Although the keyword taxonomy distinguishes physical risk from transition risk for classification and interpretation purposes, the baseline CRD measure aggregates climate-related disclosure across the complete dictionary. It should therefore be interpreted as a measure of firms' overall climate-related disclosure rather than as a separate empirical measure of physical or transition climate risk. The distinction between these dimensions provides an economic framework for understanding the heterogeneous channels through which climate-related information may affect firm valuation and cross-sectional stock returns.

Subsequent to obtaining this climate risk seed word library, we use Python's network request and JSON parsing functions to download the basic information of Chinese A-share stocks listed on the Shanghai Stock Exchange (SSE) and the Shenzhen Stock Exchange (SZSE) from the China Information Network (CNINFO). The information includes key fields such as company code and name, which are processed using pandas for further analysis. By matching company codes and years, we retrieve the PDF links for each company's annual report. Before the batch download of these reports, we clean the data, removing irrelevant summaries, cancelled announcements, and English titles to ensure data accuracy and consistency. This produces a file containing company codes, years, and PDF links of firms' annual reports. Guided by this file, we use Python to automate the batch download of annual report PDFs and convert them into text format to enable direct text analysis.

The frequency of keywords appearing in the annual reports are obtained using Python's text processing (jieba) function, and weighted calculations are applied based on the character count of each keyword to determine the total keyword frequency for each company's report. Finally, the CRD measure is calculated as follows:

$$CRD = \frac{n}{N} \times 100 \tag{10}$$

where n represents the total count of climate-related keywords appearing in a company's annual report in a given year, and N is the total word count of the report.

Although equation (10)'s notation appears the same as in Lin and Wu (2023, page 24), there are three main differences, as already articulated above (see subsection 4.2.3.). These differences mean our CRD measure contains richer, more inerratic, and more comprehensive information than that of Lin and Wu (2023).

The CRD measure is frequency-based and should therefore be interpreted primarily as a measure of the relative intensity of climate-related discussion in firms' annual reports. It does not directly distinguish between substantive and symbolic disclosure, positive and negative climate-related language, or boilerplate and firm-specific discussion. Moreover, although the expanded keyword dictionary improves the coverage of climate-related terminology, dictionary-based measures inevitably involve some judgement in keyword selection and may be affected by changes in climate-related terminology over time. Accordingly, CRD should not be interpreted as a direct measure of firms' underlying climate exposure, environmental performance, or the quality of their climate-related actions. Rather, it provides a systematic and comparable proxy for the extent to which climate-related issues are discussed in corporate reporting.

4.3.2. Stock-level analysis: Fixed-effects panel regression

To investigate the direct impact of CRD on stock returns at the firm level, we employ a fixed-effects panel regression approach. This method helps control for other characteristics that are unobservable and time-invariant that might be correlated with independent variables, allowing for a more accurate estimation of the relationship between CRD and stock returns. The baseline model is specified as follows:

$$\begin{aligned} RET_{i,t} = \alpha_0 + \alpha_1 CRD_{i,t-1} + \alpha_2 LOGME_{i,t} + \alpha_3 LOGBM_{i,t} + \alpha_4 RET_{i,t-1} \\ + \alpha_5 RET_{i,t-2,t-12} + \alpha_6 ISS_{i,t} + \alpha_7 IVOL_{i,t} + \alpha_8 ACC_{i,t} + \alpha_9 AG_{i,t} \\ + \alpha_{10} IVA_{i,t} + \alpha_{11} DE_{i,t} + \varepsilon_{i,t} \end{aligned} \tag{11}$$

where i indexes a firm and t a month. In equation (11), RET_{it} is the monthly stock return in excess of the risk-free rate; $CRD_{iy\text{-}1}$ is the climate risk disclosure score lagged by one year, calculated using equation (10); $LOGME_{it}$ is the logarithm of market capitalisation (ME);

$LOGBM_{it}$ is the logarithm of the book-to-market ratio (BM); $RET_{it\text{-}1}$ is the stock return from the previous month; $RET_{it\text{-}2,\ t\text{-}12}$ is the cumulative return over the past 12 to 2 months; ISS_{it} represents share issuance, measured as the logarithmic change in shares outstanding over a 12-month period; $IVOL_{it}$ captures idiosyncratic volatility, calculated as the standard deviation of residuals from regressing daily stock returns on market returns; ACC_{it} represents accruals/total assets (Sloan, 1996); AG_{it} denotes asset growth (Cooper et al., 2008); IVA_{it} is the investment-to-asset ratio (Lyandres et al., 2008); and DE_{it} represents leverage, calculated as the ratio of total liabilities to the market value of equity. α_0 through to α_{11} are coefficients to be estimated, and ε_{it} is the error term. To address the problems of heteroskedasticity and autocorrelation in the error term, we cluster standard errors by firms.

From equation (11), one can see that we add $CRD_{iy\text{-}1}$ as a non-traditional firm characteristic, i.e., CRD-characteristic, while considering ten traditional firm characteristics as control variables to allow for their effects on stock returns. The data frequency of $CRD_{iy\text{-}1}$ is annual because firms' reports for climate-related disclosures are released at the end of each year. Using $CRD_{iy\text{-}1}$ for the monthly regression model (11) amounts to assuming that investors have the previous year's $CRD_{iy\text{-}1}$ available in each month t of current year y available to them when making investment decisions about purchasing, holding, or selling stock i, while also utilising data on the ten conventional firm characteristics as described above. This practice broadly follows Bofinger et al.'s (2022).

The use of lagged CRD also helps mitigate potential endogeneity arising from reverse causality, since climate risk disclosure measured in the previous year is used to predict subsequent monthly stock returns. In addition, the firm fixed-effects specification absorbs time-invariant unobserved firm characteristics that may jointly affect climate-related disclosure and stock returns, while the control variables account for a range of observable firm characteristics associated with expected returns. These features therefore help reduce, although do not fully eliminate, potential endogeneity concerns, particularly those arising from reverse causality and omitted firm characteristics.

4.3.3. Portfolio-level analysis: Fama-MacBeth two-stage regression

4.3.3.1. F_{CRD} construction

In addition to stock-level analysis employing fixed-effects panel regressions, we further carry out portfolio-level examinations to determine whether CRD-characteristic-based anomalies are present or, equivalently, whether the F_{CRD} factor commands a positive or negative premium carried by the loadings on the factor. This requires us to apply a Fama-MacBeth (Fama & MacBeth, 1973) two-stage regression, a different methodology than fixed-effects panel regressions. To this end, we construct F_{CRD} in the following way.

First, we sort all stocks into deciles based on their CRD scores at the end of each June. The decile M1 includes firms with the lowest 10% CRD scores; the decile M2 contains firms with CRD scores higher than the firms' in M1 but lower than the firms' in M3; the decile M3 contains firms with CRD scores higher than the firms' in M2 but lower than the firms' in M4; …; and the decile M10 embraces firms with the highest 10% CRD scores. Then, we form four zero-investment hedging portfolios (denoted as F_{CRD}) by going long on the stocks with the bottom 10% CRD scores (in M1) and going short on the stocks with the top 10% CRD scores (in M10); and by going long on the stocks with the bottom 40% CRD scores (in M1+M2+M3+M4) and going short on the stocks with the top 40% CRD scores (in M7+M8+M9+M10). Each of the long-leg and each of the short-leg portfolio returns is respectively calculated as value-weighted on a monthly basis, and each of the four zero-investment hedging portfolios is held for the next 12 months from July of year t to June of year t+1. That is, each F_{CRD} portfolio is rebalanced annually to ensure alignment with the most recent climate disclosure data.

4.3.3.2. Policy uncertainty factor-mimicking portfolios

Our portfolio-level analysis also allows for two policy uncertainty factor-mimicking portfolios as control variables. One is related to China's Economic Policy Uncertainty (EPU), utilising the EPU index developed by Huang and Luk (2020). The other concerns China's Climate Policy Uncertainty (CPU), and utilises the CPU index developed by Ma et al. (2023). We obtain EPU and CPU factor-mimicking portfolios, F_{EPU} and F_{CPU}, as described below.

Following Li (2017) but employing 60-month rolling windows, we regress monthly changes in the EPU and CPU indices respectively on the excess returns of 25 size-B/M portfolios:

$$\Delta \log EPU_t = \theta_{EPU} + \kappa_{EPU}^{'} R_t + \delta_t^{EPU} \Rightarrow F_{EPU} = \kappa_{EPU}^{'} R_t \tag{12}$$

$$\Delta \log CPU_t = \theta_{CPU} + \kappa_{CPU}^{'} R_t + \delta_t^{CPU} \Rightarrow F_{CPU} = \kappa_{CPU}^{'} R_t \tag{13}$$

where $\Delta\log EPU_t$ and $\Delta\log CPU_t$ are the log differences of, respectively, the EPU and CPU indices; θ_{EPU} (θ_{CPU}) denotes intercept; $\kappa_{EPU}^{'}$ ($\kappa_{CPU}^{'}$) is a row vector of 25 coefficients; R_t is a column vector of 25 size-B/M portfolio excess returns; and δ_t^{EPU} (δ_t^{CPU}) denotes the error term.

4.3.3.3. Fama-Macbeth regression analysis

Constructing the F_{CRD} factor aims at investigating its return predictive power for the cross-section of 25 size-B/M portfolio returns. In other words, we attempt to determine whether the factor commands an equity premium, positive or negative, controlling for conventional risk factors and policy uncertainty factors. To this end, we employ the Fama-MacBeth regression methodology.

The Fama-MacBeth regression involves two stages: a time-series regression and a cross-sectional regression. In the first stage, we perform time-series regressions to estimate the betas of F_{CRD} and other factors for each of the 25 size-B/M portfolios, with a rolling window of 60 months. These estimated betas reflect the time-varying loadings of each size-B/M portfolio on F_{CRD} and other factors over the entire sample period. The first stage regression model is:

$$PRET_{jt} = \alpha_j + \beta_{j,F_{CRD}} F_{CRD,t} + \beta_{j,MKT} MKT_t + \beta_{j,SMB} SMB_t + \beta_{j,HML} HML_t \\ + \beta_{j,RMW} RMW_t + \beta_{j,CMA} CMA_t + \beta_{j,EPU} F_{EPU,t} + \beta_{j,CPU} F_{CPU,t} + e_{j,t} \tag{14}$$

where j (= 1, 2, …, 25) denotes a size-B/M portfolio and t a month. In equation (14), $PRET_{jt}$ is the monthly return on a size-B/M portfolio j in excess of the risk-free rate; $F_{CRD,t}$ is the return on the climate-risk-disclosure factor portfolio; MKT_t is the return on the market factor; SMB_t is the return on the size factor (Small Minus Big); HML_t is the return on the book-to-market factor (High Minus Low); RMW_t is the return on the profitability factor (Robust Minus Weak); CMA_t is the return on the investment factor (Conservative Minus Aggressive); $F_{EPU,t}$ is the return on the EPU factor-mimicking portfolio; $F_{CPU,t}$ is the return on the CPU factor-mimicking portfolio; and $e_{i,t}$ is the error term. Equation (14) as the Stage 1 regression uses data at t within each window to estimate alpha and all eight betas (factor loadings) for that window. As the window is rolled forward one month and the Stage 1 regression is repeated, we obtain a series of each of eight coefficient estimates corresponding to a series of rolling windows.

In Stage 2, we use the estimated 25 betas (loadings) as independent variables in cross-sectional regressions to determine their effects on the cross-section of 25 size-B/M portfolio returns each averaged over the future 12 months. This allows us to obtain the estimated premiums associated with F_{CRD} and the other factors. The estimation procedure is restated as follows. Over the first window of 60 months (month 1 to month 60), the Stage 1 regression yields a set of beta estimates which are used in the Stage 2 regression; the window is then rolled forward one month (month 2 to month 61) to repeat the Stage 1 regression and then the Stage 2 regression; and so on, until the end of the sample period. In general, the Stage 2 cross-sectional regression model can be expressed as:

$$PRET_{j,avg} = \gamma_0 + \gamma_{FCRD} \beta_{j,F_{CRD}} + \gamma_{MKT} \beta_{j,MKT} + \gamma_{SMB} \beta_{j,SMB} + \gamma_{HML} \beta_{j,HML} \\ + \gamma_{RMW} \beta_{j,RMW} + \gamma_{CMA} \beta_{j,CMA} + \gamma_{FEPU} \beta_{j,F_{EPU}} + \gamma_{FCPU} \beta_{j,F_{CPU}} + \mu_j \tag{15}$$

where $PRET_{j,\text{avg}}$ is the average return on a size-B/M portfolio j averaged over the next 12 months subsequent to the previous window of 60 months; $\beta_{j,FCRD}$, $\beta_{j,MKT}$, $\beta_{j,SMB}$, $\beta_{j,HML}$, $\beta_{j,RMW}$,

$\beta_{j,CMA}$, $\beta_{j,EPU}$ and $\beta_{j,CPU}$ are a size-B/M portfolio j's loadings on the eight respective factors, estimated using equation (14) over the previous 60-month window which moves forward each time by one month till the end of our sample period; γ_{FCRD}, γ_{MKT}, γ_{SMB}, γ_{HML}, γ_{RMW}, γ_{CMA}, γ_{FEPU} and γ_{FCPU} measure the premiums associated with the eight respective betas (loadings); $\gamma_{i,0}$ is the intercept; and μ_j is the error term. We exclude stocks with a survival time shorter than 36 months to maintain consistent time-varying beta estimates over time. By using the Fama-MacBeth regression, we can assess whether F_{CRD} offers significant predictive power for future size-B/M portfolio returns in the cross-section and serves as an independent risk factor in the Chinese A-share market. In other words, we can investigate the CRD-characteristic-based return anomalies in China through the lens of the F_{CRD} factor.

4.4. Data Description and Summary Statistics

4.4.1. Data and variable definitions

We obtain monthly data on stock returns and the risk-free rate, along with quarterly accounting data, from the Chinese Stock Market & Accounting Research (CSMAR) database. We classify all A-share stocks into 11 industry sectors, based on the Bloomberg industry classification system, including Communications, Consumer Discretionary, Consumer Staples, Energy, Financials, Real Estate, Health Care, Industrials, Materials, Technology, and Utilities. Additionally, monthly excess return data on the market factor (MKT), the size factor (SMB), the book-to-market factor (HML), and the momentum factor (MOM) are also downloaded from CSMAR. These factors are used for portfolio-level analysis (equations (14) and (15)).

In the stock-level analysis (equation (11)), we use a set of traditional firm characteristics as control variables including the following; ME is market capitalisation, BM is the book-to-market ratio, RET is monthly stock returns, AGE is the number of years since a stock first appears on the CSMAR database, NI is the net income, DE is the leverage ratio (Ferguson and Shockley, 2003), IVA is the investment-to-asset ratio (Lyandres et al. 2008), AG represents the asset growth rate (Cooper et al. 2008), ACC is the operating accruals/total assets (Sloan, 1996), and ISS is the share issuance defined as the logarithm change of outstanding shares over a 1-year period, as suggested by Pontiff and Woodgate (2008). The data on these firm characteristics are also sourced from CSMAR.

4.4.2. Summary statistics and preliminary analysis

4.4.2.1. CRD in each year

The CRD measure captures the extent of a company's climate risk disclosure and intent on climate risk management. To provide an overview of the measure, Table 4.1 presents the summary statistics of the CRD scores (defined in equation (10)) over the period from 2002 to 2022, including the number of observations (Obs), mean (Mean), standard deviation (Std Dev), minimum (Min), and maximum (Max) for each year. The mean increased from 0.045 in 2002 to 0.251 in 2022, showing the increasing importance companies place on climate risk disclosure. It is notable that the maximum value of CRD scores increased sharply in 2015, more than doubling the maximum value in preceding years and reaching their highest value of 2.820 in 2018. Thereafter, the maximum value stayed at a high level albeit slightly lower than 2.820 due largely to the Covid-19 pandemic. Such sharp increases in the maximum value in turn contributed to the gradual increases in Mean, likely because increasing numbers of Chinese companies started to pay attention to CRD. These observations could be driven by several government policies and regulations being put in place. First, the impact of the Paris Agreement adopted in 2016 began to take effect. Second, Guidelines for Establishing the Green Financial System was published in 2016, which prompted financial institutions and companies to enhance their climate risk disclosures. Third, the implementation of the Environmental Protection Tax Law in 2018 further pressured companies to disclose more information on climate and environmental risks. Due to these international and domestic regulation/policy pressures, climate risk disclosures in Chinese companies' annual reports became broader and deeper.

The time-series evolution of CRD is also broadly consistent with the increasing prominence of climate-related policy in China. Over the sample period, climate and environmental issues received progressively greater attention through policies related to environmental protection, green development, energy transition, carbon reduction, and climate-related corporate disclosure. The corresponding increase in annual CRD therefore provides policy-consistency evidence that the textual measure responds to the changing climate-information environment faced by Chinese firms. This comparison is interpreted as validation of the broad time-series pattern rather than as evidence that any individual policy action mechanically causes changes in CRD.

As an additional external validation exercise, the annual CRD series is compared with independently constructed Environmental Scores from CNRDS and Bloomberg. Firm-level observations for each measure are first averaged within each year, and Pearson and Spearman correlations are calculated over the corresponding overlapping sample periods. CRD exhibits strong positive associations with both external measures. The Pearson correlations between CRD and the CNRDS and Bloomberg Environmental Scores are 0.803 and 0.934, respectively, while the corresponding Spearman correlations are 0.971 and 0.990. These strong positive associations provide external validation that CRD captures a climate- and environment-related information dimension that is also reflected in independently constructed third-party environmental indicators, while recognising that these external scores capture broader environmental dimensions than climate-risk disclosure alone.

Figure 4. 1. Trend in CRD scores (mean) over time

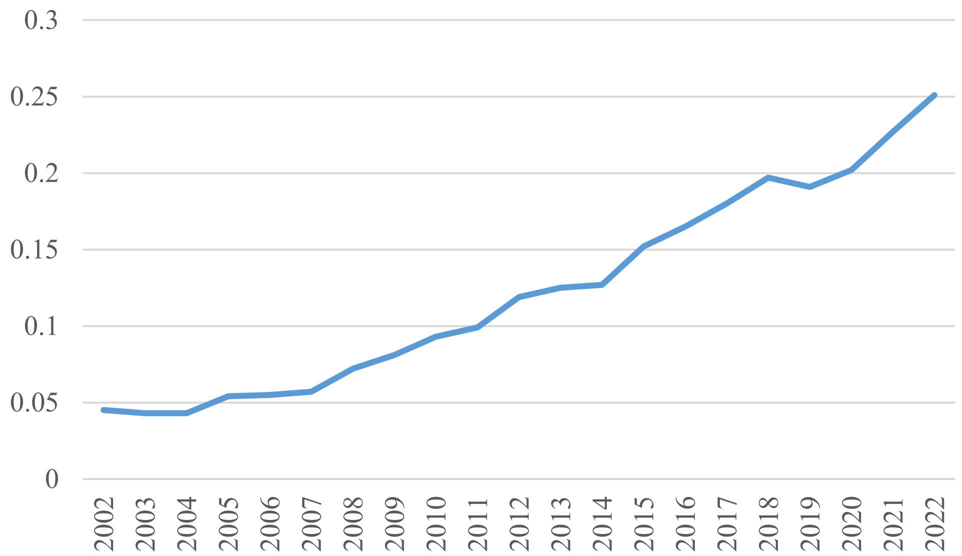


Figure 4.1. graphically illustrates the upward trend in China’s CRD scores, depicted using the mean value. The trend is largely increasing monotonically, indicating that, over time, increasing numbers of companies were disclosing climate-related risks in their annual reports.

Table 4. 1. Summary Statistics of Climate Risk Disclosure Scores by Year

Yr	Obs	Mean (%)	Std Dev (%)	Min (%)	Max (%)
2002	895	0.045	0.049	0.003	0.595
2003	926	0.043	0.046	0.002	0.569
2004	1,015	0.043	0.050	0.002	0.528
2005	635	0.054	0.059	0.003	0.583
2006	1,175	0.055	0.061	0.004	0.634
2007	1,264	0.057	0.067	0.002	1.261
2008	1,371	0.072	0.073	0.001	0.869
2009	1,517	0.081	0.084	0.001	1.254
2010	1,840	0.093	0.099	0.002	1.349
2011	2,121	0.099	0.105	0.002	1.440
2012	2,242	0.119	0.115	0.002	1.391
2013	2,304	0.125	0.116	0.002	1.359
2014	2,426	0.127	0.112	0.002	1.198
2015	2,612	0.152	0.139	0.004	2.611
2016	2,890	0.165	0.151	0.006	2.458
2017	3,266	0.180	0.153	0.002	2.489
2018	3,359	0.197	0.152	0.006	2.820
2019	3,551	0.191	0.145	0.004	2.786
2020	3,935	0.202	0.146	0.014	2.554
2021	4,024	0.227	0.156	0.012	2.439
2022	4,038	0.251	0.163	0.001	2.251

Note This table reports the summary statistics of CRD scores over the period from 2002 to 2022, including observation, mean, standard deviation, minimum and maximum.

4.4.2.2. CRD in each industry

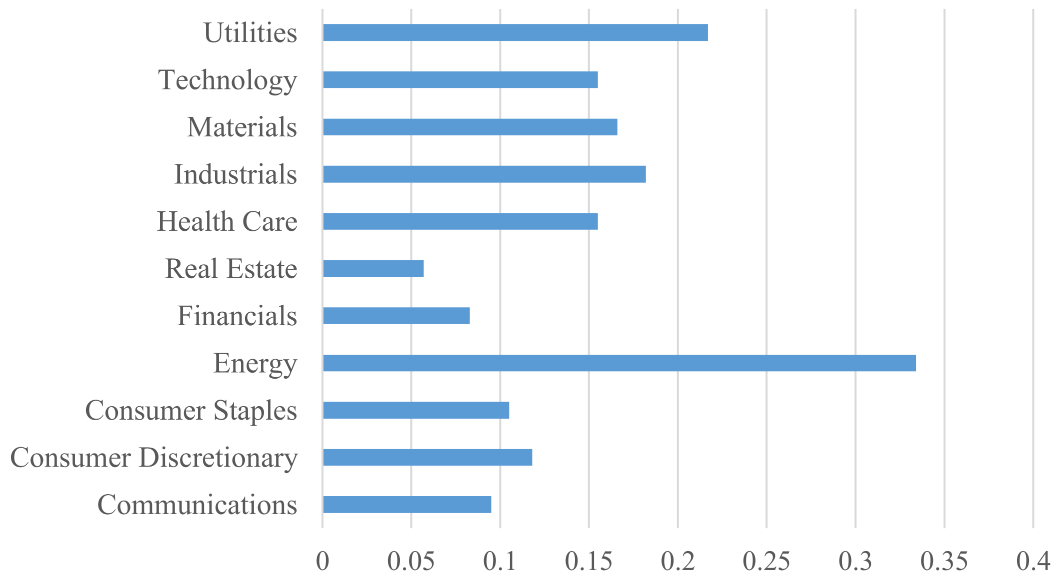
Table 4.2 categorises the CRD-related summary statistics into 11 industries, using Bloomberg industry classification. The statistics reveal that the energy sector has the highest mean value (0.334%) of CRD, indicating that it has the largest number of companies reporting climate risk in their annual reports. This finding is consistent with the fact that the energy sector tends to have high carbon emissions and strict environmental regulations which mandatorily require energy firms to disclose such information. In contrast, the real estate sector shows the lowest mean value (0.057%) of CRD, reflecting itself to be the sector with either the fewest climate disclosures, or the weakest connection to climate-related risks. The largest Max value (2.820%) is found in Utilities. This is because the utilities industry typically involves energy production and distribution, which are associated with carbon emissions and environmental impacts, leading to the strictest and highest regulatory requirements for climate-related disclosures. The Financials sector exhibits the lowest mean CRD (0.083%) due to its nature as an office-based industry with almost negligible environmental impacts.

Table 4. 2. Summary Statistics of Climate Risk Disclosure by Industry

Code	Industry	Obs	Mean (%)	Std Dev (%)	Min (%)	Max (%)
1	Communications	1734	0.095	0.07	0.003	0.638
2	Consumer Discretionary	7081	0.118	0.104	0.002	1.315
3	Consumer Staples	2822	0.105	0.099	0.002	1.005
4	Energy	1314	0.334	0.324	0.002	1.686
5	Financials	1202	0.083	0.056	0.001	0.463
6	Real Estate	2032	0.057	0.056	0.003	0.904
7	Health Care	4120	0.155	0.101	0.002	0.751
8	Industrials	10926	0.182	0.154	0.001	1.503
9	Materials	8436	0.166	0.131	0.003	1.435
10	Technology	6024	0.155	0.099	0.002	1.345
11	Utilities	1715	0.217	0.26	0.002	2.820

Note This table reports the summary statistics of CRD scores across industries over the period from 2002 to 2022, including observation, mean, standard deviation, minimum and maximum.

Figure 4. 2. The bar chart of CRD scores (mean) across industries



Note This figure shows the mean CRD scores across eleven industries.

As alluded to above, the differences in CRD scores across industries can be attributed to differences in China's economic structure and regulatory requirements. For example, Energy and Utilities industries are heavily regulated and directly affected by climate policies, such as emissions reduction and environmental compliance requirements. These sectors also face more critical reviews from investors and regulators. Conversely, sectors like Real Estate and Communications face fewer immediate regulatory pressures and have less direct exposure to environmental impacts, which may explain their relatively lower levels of CRD.

4.4.2.3. Firm characteristics

Our data sample includes all China's A-share stocks listed on the Shanghai Stock Exchange (SSE) and the Shenzhen Stock Exchange (SZSE), covering the period from December 2002 to December 2022. We use the firm characteristics of these stocks as control variables in our stock-level regression analysis (equation (11)). Table 4.3 presents their summary statistics, including the mean (Mean), median (Median), quartiles (Q1 and Q3), and standard deviation (Std Dev).

The table reveals the fundamental attributes of the sample firms. The CRD score, in percentage terms, captures the extent of climate-related reporting as the ratio of climate-related keywords to the total word count in a firm's annual report. ME is the market capitalization, expressed in millions. The book-to-market ratio (BM) is defined as the total equity minus preference shares and minority interests, divided by the market value of tradable shares. Monthly stock return (RET) represents each firm's return for the month. Firm age (AGE) measures the number of years since the firm's first listing in the Chinese A-share market. Net income (NI), in millions, represents the net profit. The leverage ratio (DE) follows the definition of Ferguson and Shockley (2003), calculated as the sum of long-term and short-term loans over the market value of equity. The investment-to-asset ratio (IVA) reflects the annual changes in net fixed assets plus inventories, scaled by lagged total assets, as per Lyandres et al. (2008). Asset growth (AG) is calculated, following Cooper et al. (2008), as the change in total assets over the previous year, divided by the lagged total assets. Operating accruals (ACC), following Sloan (1996), are computed by taking the change in current assets (excluding cash equivalents) minus the change in current liabilities (excluding short-term loans and taxes payable), scaled by total assets. Idiosyncratic volatility (IVOL), expressed as a percentage, measures the standard deviation of the residuals from regressing daily stock returns on the Fama–French three factors, as per Ang et al. (2006). Lastly, we follow Pontiff and Woodgate (2008) and define share issuance (ISS) as the logarithmic change in adjusted shares outstanding over a 12-month period. These variables provide a comprehensive overview of the financial characteristics of firms used to explore the relationship between these characteristics and climate risk disclosures.

Table 4. 3. Summary Statistics of Firm Characteristics

		Mean	Median	Q1	Q3	Std Dev
ME	(in millions)	10,543	3,146	1,412	7,073	49,840
BM		0.773	0.569	0.325	0.971	1.094
RET		0.010	-0.004	-0.073	0.073	0.162
AGE		16	16	10	24	9
CRD	(in per cent)	0.153	0.119	0.061	0.195	0.145
NI	(in millions)	598	50	9	179	6177
DE		0.282	0.124	0.035	0.323	0.503
IVA		0.022	0.003	-0.007	0.020	1.261
AG		0.071	0.016	-0.014	0.053	4.366
ACC		0.030	0.004	-0.024	0.032	22.364
IVOL	(in per cent)	0.017	0.016	0.011	0.022	0.008
ISS		0.024	0.000	0.000	0.000	0.113

Note This table presents the summary statistics of key financial and firm characteristics, including the mean, median, lower quartile (Q1), upper quartile (Q3), and standard deviation. ME represents market capitalisation (in millions). BM is the book-to-market ratio. RET refers to monthly stock returns. AGE is the number of years since a firm's listing. CRD (in percent) represents the climate risk disclosure score. NI (in millions) denotes net income. DE is the leverage ratio, defined as the book value of total liabilities over the market value of equity following Ferguson and Shockley (2003). IVA is the investment-to-asset ratio, calculated as the annual change in gross property, plant, and equipment plus the annual change in inventories, scaled by lagged total assets following Lyandres et al. (2008). AG refers to asset growth following Cooper et al. (2008). ACC measures operating accruals following Sloan (1996). IVOL captures monthly idiosyncratic volatility of stock returns, measured as the standard deviation of residuals from regressing daily stock returns on the Fama–French three factors. ISS refers to the share issuance measure, defined as the change in the logarithm of shares outstanding over a 12-month period following Pontiff and Woodgate (2008).

4.5. Empirical Results

With the CRD measure developed, this section investigates its relationship with return anomalies. First, we examine Table 4.4 for stock characteristics under different CRD deciles. Second, we analyse Table 4.5 for Alpha as a measure of abnormal returns (also known as return anomalies). Note, in the interest of space, Panel B of Table 4.5 only reports the results associated with size quintiles and BM quintiles.

Third, we report the results of the fixed-effects panel regression in Table 4.6, to determine how the non-traditional CRD-characteristic impacts stock returns. Fourth, we compare the performance (mean return) of the F_{CRD} factor with those of conventional risk factors, based on Table 4.7. Fifth, we explore the relation between the F_{CRD} factor and other risk factors in Table 4.8. These explorations aim to determine whether the F_{CRD} factor offers abnormally high returns relative to the conventional risk factors. Finally, we probe the predictive power of the F_{CRD} factor in the cross-section of 25 size-B/M portfolio returns.

4.5.1. Stock characteristics across CRD deciles

Table 4.4 explores the association between CRD deciles and selected stock characteristics. In each year, firms are sorted into deciles according to their CRD measure, from the lowest (M1) to the highest (M10). Figure 4.3. displays the average values of the considered stock characteristics for each decile, providing insights into how these characteristics vary with climate risk disclosure levels.

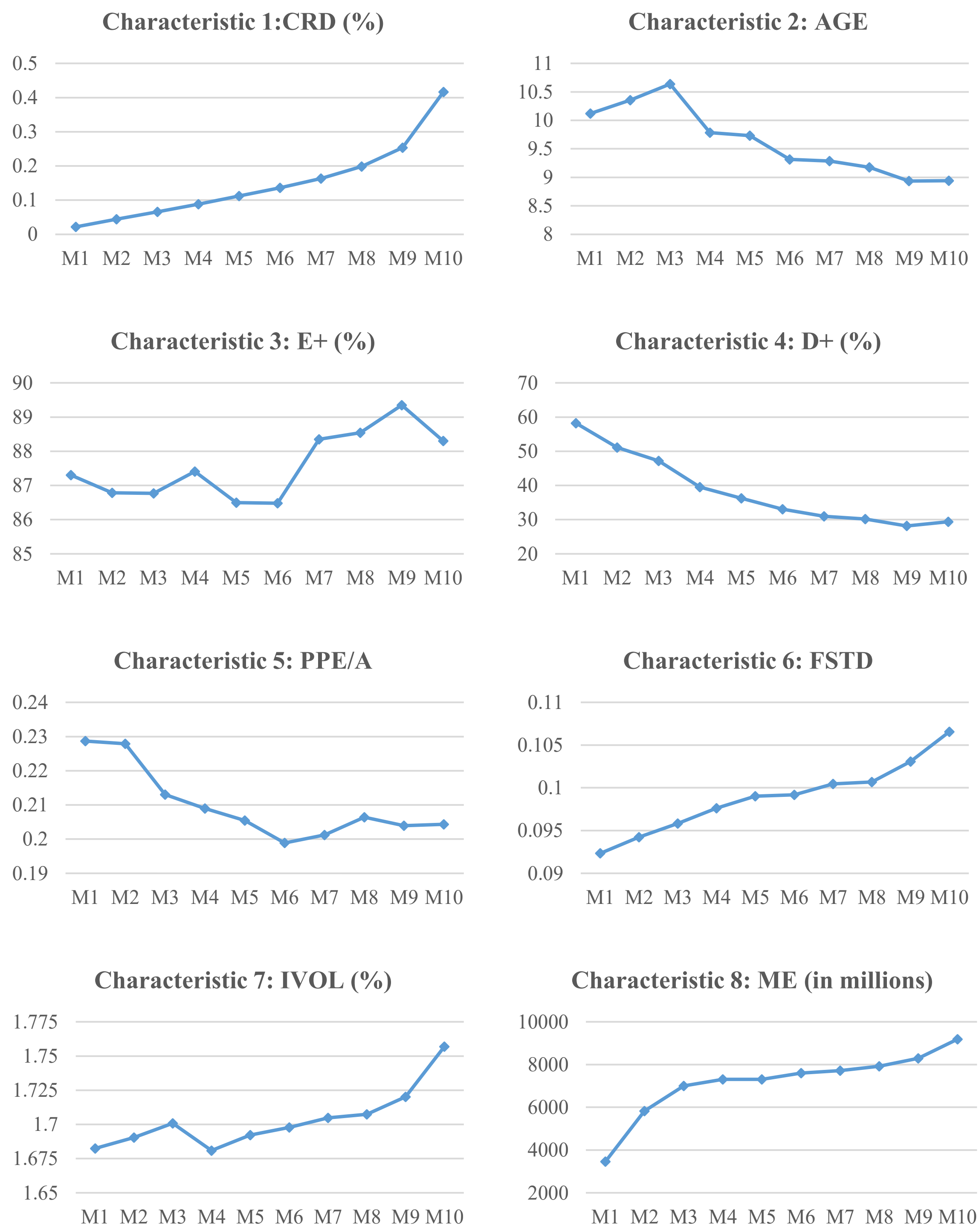
The CRD value consistently increases from 0.022 in M1 to 0.416 in M10, meaning that companies in higher deciles report more on climate risk. AGE (years since IPO) shows that firms in the lower CRD deciles tend to be older, with a gradual decrease in age as CRD scores rise. This suggests that younger companies have more willingness to disclose climate risk, because young companies are more likely to meet environmental expectations, whereas companies with long history are less active in this area. E+ (percentage of firms with positive earnings) shows a slight decrease, from 87.3% in M1 to 80.3% in M10, implying that firms with high CRD scores may have slightly lower earnings, since they have higher costs associated with climate risk management. D+ (percentage of firms paying dividends) decreases as CRD

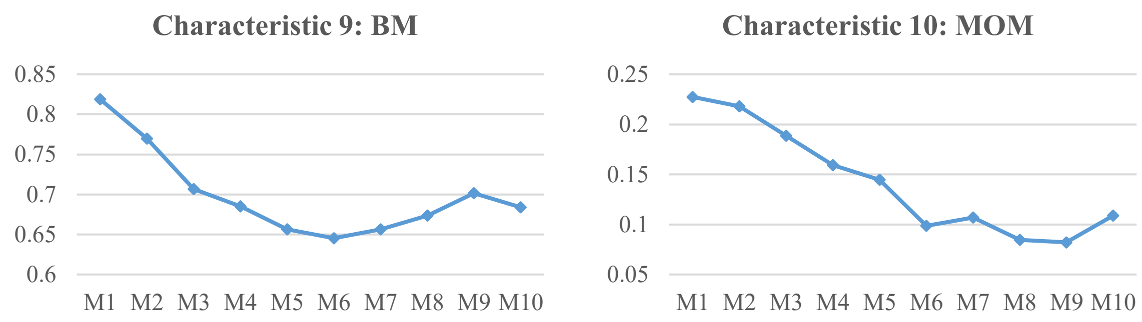
Table 4. 4. Stock characteristics versus CRD deciles

		M1(L)	M2	M3	M4	M5	M6	M7	M8	M9	M10(H)	M10-M1		t-stat.	M10-M5		t-stat.
CRD		0.022	0.044	0.066	0.089	0.112	0.136	0.163	0.198	0.254	0.416	0.394	***	[836.06]	0.304	***	[644.84]
AGE		10	10	11	10	10	9	9	9	9	9	-1	***	[-28.17]	-1	***	[-17.35]
E+	(in per cent)	87.301	86.784	86.766	87.404	86.501	86.480	88.347	88.540	89.348	88.304	1.004	***	[4.84]	1.803	***	[8.58]
D+	(in per cent)	58.241	51.118	47.203	39.548	36.255	33.083	30.990	30.219	28.152	29.385	-28.855	***	[-95.87]	-6.869	***	[-23.15]
PPE/A		0.229	0.228	0.213	0.209	0.205	0.199	0.201	0.206	0.204	0.204	-0.024	***	[-23.77]	-0.001		[-1.17]
FSTD		0.092	0.094	0.096	0.098	0.099	0.099	0.100	0.101	0.103	0.107	0.014	***	[44.27]	0.008	***	[22.62]
IVOL	(in per cent)	0.017	0.017	0.017	0.017	0.017	0.017	0.017	0.017	0.017	0.018	0.001	***	[12.88]	0.001	***	[10.5]
ME	(in millions)	3461	5831	6996	7306	7310	7598	7722	7926	8298	9186	5725	***	[79.18]	1876	***	[22.28]
BM		0.819	0.770	0.707	0.685	0.656	0.645	0.656	0.673	0.701	0.684	-0.135	***	[-38.03]	0.028	***	[8.89]
MOM		0.227	0.218	0.189	0.159	0.145	0.098	0.107	0.084	0.082	0.109	-0.119	***	[-29.44]	-0.036	***	[-10.10]

Note This table reports stock characteristics across CRD deciles ranging from M1 (low CRD) to M10 (high CRD). M1 includes the firms with the lowest CRD scores (most undervalued), while M10 contains those with the highest CRD scores (most overvalued). The differences in stock characteristics between M10 and M5, as well as M10 and M1, and M10 and M5 are also reported. CRD is the firm's climate risk disclosure score. AGE refers to the number of years since the firm's initial public offering. E+ represents the percentage of firms with positive earnings, and D+ is the percentage of firms paying dividends. PPE/A is the ratio of fixed assets to total assets. FSTD is forecast standard deviation. IVOL captures monthly idiosyncratic volatility of stock returns, measured as the standard deviation of the residuals from regressing daily stock returns on the Fama–French three factors. ME is the market capitalisation, reported in millions. BM is the book-to-market ratio. MOM represents the stock returns over the past 12 months. * Significance at the 10% level. ** Significance at the 5% level. *** Significance at the 1% level.

Figure 4. 3 Characteristics of portfolios formed based on CRD measure





Note This figure shows the stock characteristics across deciles based on the climate risk disclosure (CRD) levels. In each year, firms are sorted into deciles according to their CRD measure, from the lowest (M1) to the highest (M10). For each decile, the average values of firm age (AGE), percentage of firms with positive earnings (E+), percentage of dividend payers (D+), the ratio of fixed assets to total assets (PPE/A), standard deviation of analysts' forecasts (FSTD), idiosyncratic volatility (IVOL), market capitalisation (ME), book-to-market ratio (BM), and stock returns over the past 12 months (MOM) are calculated.

deciles increase, from 58.241% in M1 to 29.385% in M10, with a significant difference of -28.855%. This suggests that firms with higher CRD are less likely to pay dividends, as they may reinvest earnings in sustainable projects or climate risk management.

PPE/A (fixed assets to total assets) is highest in the lowest CRD deciles and decreases slightly as CRD scores increase, suggesting that firms with higher climate risk disclosure may rely less on fixed assets. FSTD (forecast standard deviation) shows a slight decrease as CRD deciles increase, going from 0.093 in M1 to 0.070 in M10, with a difference of -0.023. This suggests that firms with higher CRD scores tend to have lower forecast uncertainty among analysts. This is because more detailed climate risk disclosures provide analysts with additional information to utilise, which helps reduce the uncertainty in a company's future performance. IVOL (idiosyncratic volatility) is the most stable across the different CRD deciles compared to the other firm characteristics.

ME (market capitalisation) declines from 8,361 million in M1 to 575 million in M10, implying that firms with lower CRD scores tend to be larger sized. Smaller companies may focus more on climate risk disclosure to attract environmentally conscious investors. BM (book-to-market ratio) also decreases with higher CRD scores, implying that companies with high CRD scores have relatively low book-to-market ratios and are typically growth companies.

MOM (momentum) declines from 0.227 in M1 to 0.109 in M10, suggesting that low CRD companies have relatively better past performance and may be favoured by traditional investors.

In summary, Table 4.4 provides a comprehensive view of how the traditional firm characteristics vary across deciles sorted by climate risk disclosure levels. Firms with high CRD scores are generally younger and smaller. They are less likely to pay dividends and have worse past performance. These firms tend to be valued as growth stocks, with lower book-to-market ratios and more emphasis on climate risk disclosure. Additionally, they may focus more on growth and attracting climate-conscious investors.

4.5.2. Performance of CRD in terms of returns

4.5.2.1. Predictive power pre-test on sort of CRD

Table 4.5 presents the raw returns and abnormal returns (alphas) for portfolios sorted using CRD scores, as a preliminary assessment of the potential return predictive power of CRD. Overall, the results show how climate risk disclosure impacts returns within different Size and BM groups. Significant t-statistics indicate that climate risk disclosure may have a meaningful impact on returns, with larger or higher-BM firms having different sensitivities than smaller or lower-BM firms.

Specifically, in Panel A of Table 4.5, stocks are sorted into deciles based on their CRD scores, from the lowest decile (M1, representing firms with minimal climate risk exposure) to the highest decile (M10, representing firms with maximum climate risk exposure). The raw returns (Raw RET), and the alphas adjusted separately by three asset-pricing models, are reported for each decile. The return differences between the bottom decile (M1) and top decile (M10), L-H, are also presented.

From Panel A, the raw returns (Raw RET) decrease from M1 (1.685%) to M10 (1.123%), and the return difference between M1 and M10 is 0.562% (with $t = 5.19$) in the L-H column. The significant positive value of L-H suggests that firms with the highest CRD tend to yield lower raw returns than those with the lowest CRD, indicating a strong return predictive ability of CRD. This finding is consistent with the rationale that high-risk investment should offer great returns to compensate investors who take the risk that firms minimally disclose

Table 4. 5. Raw return and Alphas based on sort of CRD

Panel A Raw returns and alphas of portfolios sorted on CRD

	M1(L)	M2	M3	M4	M5	M6	M7	M8	M9	M10(H)	L-H		t-stat.
Raw RET	1.760	1.722	1.281	1.468	1.399	1.178	0.952	0.994	0.772	1.002	0.758	***	[6.82]
CAPM α	1.555	1.515	1.080	1.277	1.225	1.021	0.802	0.850	0.629	0.856	0.698	***	[14.80]
FF3 α	1.555	1.516	1.081	1.276	1.226	1.020	0.798	0.848	0.628	0.854	0.702	***	[13.05]
FF3 + MOM α	1.555	1.517	1.081	1.277	1.226	1.020	0.798	0.848	0.628	0.854	0.702	***	[13.02]

Panel B The FF3+MOM alphas of portfolios sorted first by Size or BM and then by CRD

Size Quintiles	1(S)		2		3		4		5(B)	
L-H Raw RET	1.209	***	1.604	***	1.066	***	1.038	***	0.267	**
t-stat.	[9.32]		[10.78]		[9.08]		[7.97]		[2.13]	
BM Quintiles	1(LBM)		2		3		4		5(HBM)	
L-H Raw RET	0.043		0.912	***	1.164	***	1.048	***	0.675	***
t-stat.	[0.26]		[7.01]		[8.48]		[9.49]		[6.25]	

Note This table presents the results of sorting stocks based on Climate Risk Disclosure (CRD) and analysing returns across different groups. In Panel A, stocks are first sorted into deciles based on their CRD scores, with the lowest decile representing firms with the lowest CRD score (indicating minimal climate risk disclosure) and the highest decile representing firms with the highest CRD score (indicating the most extensive disclosure). For each decile, Panel A reports the average raw returns (Raw RET) and the abnormal returns (alphas) after controlling for the market factor (CAPM α), the Fama-French three factors (FF3 α), and the Fama-French three factors plus the momentum factor (FF3 + MOM α). The raw returns and alphas are reported in percentages. Additionally, Panel A provides the return difference (L-H) between the bottom decile (lowest CRD score, L) and the top decile (highest CRD score, H), along with the corresponding t-statistics (based on Newey-West standard errors) to test the significance of the differences. In Panel B, stocks are first sorted into quintiles based on their market capitalisation (Size) or book-to-market ratio (BM) at the end of June each year.

LBM denotes the lowest BM, and HBM denotes the highest BM. Within each Size or BM quintile, stocks are further sorted into three groups based on their CRD score; the top 30% (H), the middle 40% (M), and the bottom 30% (L); and Panel B reports the FF3+MOM alphas for the resultant L-H return difference in each quintile, along with the corresponding t-statistics. The results show how climate risk disclosure impacts stock returns within different Size and BM groups. Significant t-statistics indicate that climate risk disclosure may have a meaningful impact on stock returns, with larger or higher-BM firms showing different sensitivities compared to smaller or lower-BM firms.

climate information to the public, and vice versa. The CAPM alpha also shows a decreasing trend across CRD deciles, from 1.478% in M1 to 0.977% in M10, with a significant difference of 0.501% between M1 and M10. This finding shows that the predictive power of CRD on stock returns are not solely explained by market exposure. When controlling for the Fama-French three factors, the alpha decreases from 1.478% in M1 to 0.975% in M10, showing a significant L-H difference of 0.503%. After adjusting for the Fama-French three factors and the momentum factor, the alpha decreases from 1.552% in M1 to 0.963% in M10, with a significant L-H difference of 0.589%. These results reinforce each other in pointing to the presence of CRD-characteristic-based anomalies.

Panel B of Table 4.5 examines whether the return predictive power of CRD is influenced by Size and Book-to-Market (BM) characteristics. Stocks are first sorted into quintiles based on either market capitalisation (Size) or book-to-market ratio (BM) at the end of June each year. Within each Size or BM quintile, stocks are further divided into three groups by CRD scores: the top 30% (H), middle 40% (M), and bottom 30% (L). L-H Raw RET is the raw return difference between the bottom 30% CRD and top 30% CRD within each Size or BM quintile.

In the Size quintiles, the L-H return difference is significantly positive across all quintiles, with the value ranging from 0.267% in quintile 5(B) (the largest quintile) to 1.604% in quintile 2 (the second smallest quintile). This suggests that the relationship between CRD and returns is not due to firm size. Similarly, throughout the five BM quintiles, the L-H return difference is also significantly positive, ranging from 0.043% in quintile 5(HBL) (the highest quintile) to 1.164% in quintile (3) (the middle quintile). This suggests that the predictive relevance of CRD for returns cannot be subsumed by the book-to-market effect.

Overall, Table 4.5 suggests that CRD *per se* may have return predictive ability. Panel A indicates that portfolios with low CRD scores generally outperform those with high CRD scores, and this observation holds after controlling various conventional risk factors. Panel B further confirms that this observation remains robust across different firm sizes and book-to-market ratios, indicating that the association between CRD and returns is not driven by Size or BM effects. These findings suggest that the CRD characteristic may serve as a meaningful metric in stock-picking analysis, potentially enabling investors to identify high abnormal returns by focusing on firms with lower CRD levels.

4.5.2.2. Fixed-effects panel regression results

To further conduct stock-level analysis of the relation between stock returns (RET) and end-of-last-year's Climate Risk Disclosure (lag_CRD), Table 4.6 presents the results of the fixed-effects panel regressions, with various control variables added. Lagged CRD (lag_CRD) is utilised because the current year's CRD is not available to investors in each month prior to the year-end. Using contemporaneous control variables follows the existing practice (see Bofinger et al., 2022).

Model 1 regresses stock returns on lag_CRD alone. Models 2 to 9 then adds the control variables firm size (LOGME), book-to-market ratio (LOGBM), past returns (RET_{t-1} and $\mathrm{RET}_{t-12,t-2}$), stock issuance (ISS), idiosyncratic volatility (IVOL), accruals (ACC), asset growth (AG), investment-to-assets (IVA), and leverage ratio (DE) one by one in sequence. This horse race approach allows us to observe whether the incremental effect of each control variable will negate the relation between CRD and returns.

A notable observation is that, across all the nine models, the coefficient on lag_CRD is consistently negative and highly significant. That is, adding more-and-more control variables to the regression does not weaken the significant, negative impact of CRD on stock returns in a qualitative analysis. In particular, Model 9 has all the control variables included, but the estimated coefficient on lag_CRD is -0.022 with a t-value of -3.64. Some of the control variables are highly significant at the 1% level, so their contributions to the variations in stock returns are properly accounted for. In other words, the potential omitted-variable bias of the results should be reduced to a minimum.

The possible economic interpretations of the detected negative CRD-RET relationship are as follows. Investors might view high levels of climate risk disclosure as a signal that companies are facing potential challenges related to environmental risks, regulatory compliance, or resource allocation toward sustainability initiatives. These factors could lead to

Table 4. 6. CRD and Stock Returns: Fixed-Effects Panel Regressions

	1	2	3	4	5	6	7	8	9
Intercept	0.090 ***	00.736 ***	0.696 ***	0.368 ***	0.801 ***	0.753 ***	0.750 ***	0.803 ***	0.315 ***
	[112.2]	[25.00]	[24.8]	[13.95]	[26.72]	[23.66]	[24.87]	[26.60]	[11.91]
lag_CRD	-0.026 ***	-0.043 ***	-0.039 ***	-0.032 ***	-0.043 ***	-0.050 ***	-0.044 ***	-0.033 ***	-0.022 ***
	[-8.01]	[-5.01]	[-4.85]	[-5.35]	[-4.72]	[-5.26]	[-4.96]	[-3.64]	[-3.64]
LOGME		-0.001 ***	-0.000	0.000	-0.001	-0.001	-0.001	0.000	0.000
		[-1.01]	[-0.62]	[0.88]	[-1.06]	[-1.05]	[-0.95]	[-0.70]	[0.83]
LOGBM		-0.719 ***	-0.684 ***	-0.443 ***	-0.735 ***	-0.731 ***	-0.733 ***	-0.720 ***	-0.435 ***
		[-34.20]	[-34.10]	[-23.90]	[-34.39]	[-32.53]	[-34.28]	[-32.86]	[-24.10]
RET_{t-1}		-0.001	-0.003	-0.047 ***	0.001	-0.001	-0.001	0.002	-0.047 ***
		[-0.44]	[-1.44]	[-10.52]	[0.57]	[-0.47]	[-0.53]	[0.92]	[-8.19]
$RET_{t-12,t-2}$		-0.008 ***	-0.007 ***	-0.015 ***	-0.008 ***	-0.008 ***	-0.008 ***	-0.008 ***	-0.016 ***
		[-12.47]	[-13.07]	[-15.32]	[-13.45]	[-13.03]	[-14.23]	[-12.76]	[-17.35]
ISS			0.024 ***						-0.003
			[11.20]						[-1.10]
IVOL				5.302 ***					5.341 ***
				[94.35]					[93.76]
ACC					0.021 ***				0.020 ***
					[4.15]				[3.21]
AG						0.000			0.000
						[1.02]			[0.74]
IVA							0.000		0.000
							[0.11]		[-0.55]
DE								-0.016 ***	-0.012 ***
								[-11.43]	[-10.75]

Table 4. 6. (Continued)

	1	2	3	4	5	6	7	8	9
R^2 (%)	0.55	2.24	2.21	9.20	2.39	2.06	2.32	2.62	9.47
Obs.	554,041	520,216	514,636	391,074	457,497	518,136	509,112	395,695	307,141

Note This table reports the results of fixed-effects panel regressions of the monthly stock returns (RET) on Climate Risk Disclosure (CRD) and other variables as the control variables. Model 1 shows the regression of RET on lag_CRD alone. Models 2 through 9 incrementally add control variables, including firm size (LOGME), book-to-market ratio (LOGBM), past returns (RET_{t-1} and $RET_{t-12,t-2}$), stock issuance (ISS), idiosyncratic volatility (IVOL), accruals (ACC), asset growth (AG), investment-to-assets (IVA), and leverage ratio (DE). All models account for heteroskedasticity and clustering. The reported R^2 and the number of observations are presented at the bottom of each column. * Significance at the 10% level. ** Significance at the 5% level. *** Significance at the 1% level.

increased costs or reduced short-term profitability, making high-CRD firms less attractive from an investment perspective. This finding is consistent with the result of Table 4.5, where we observe significantly negative raw and abnormal returns for the H-L CRD differences. In Table 4.6, this negative relationship remains across Models 2 through 9, with the CRD's coefficients ranging from -0.050 to -0.022. Thus, the result that climate risk disclosure negatively impacts stock returns is robust.

4.5.3. Return predictability of loadings on F_{CRD}

This subsection turns to portfolio-level analysis to further explore the anomaly implications of climate risk disclosures. We construct four CRD factor portfolios, F_{CRD}, in the way described in section 4.3.1. Recall that, in essence, F_{CRD} represents a zero-investment strategy that goes long on firms with low disclosures and short on firms with extensive disclosures. It captures the pricing effects of climate risk disclosure discrepancies and provides a systematic perspective in which to evaluate the anomaly relevance of CRD. By isolating the returns associated with different climate risk disclosure levels, F_{CRD} enriches the understanding of stock anomalies and offers novel insights into how climate risk related information shapes asset pricing.

4.5.3.1. Preliminary analysis of factor performance

Table 4.7 sets out the summary statistics of the climate risk disclosure factor F_{CRD}, the market factor (MKT), the size factor (SMB), the book-to-market factor (HML), the profitability factor (RMW), the investment factor (CMA), the economic policy uncertainty factor (F_{EPC}), and the climate policy uncertainty factor (F_{CPU}). Note, for the sake of space, Table 4.7 only selects the F_{CRD} factor that is formed as bottom 10% - top 10% in terms of CRD scores for reporting. Each factor's returns are evaluated in terms of the mean excess return (Mean), t-statistic (t-stat), Sharpe ratio (SR), and the percentage of months with positive returns (Positive %).

Table 4. 7. Summary statistics of factor returns (%)

	Mean		t-stat	SR	Positive %
F_{CRD} (bottom 10% - top 10%)	0.609	**	[2.15]	0.132	55.30
MKT	0.427		[0.93]	0.058	53.03
SMB	0.596	**	[2.05]	0.126	51.52
HML	0.044		[0.22]	0.013	50.76
RMW	-0.039		[-0.20]	-0.012	48.86
CMA	0.071		[0.49]	0.030	51.89
F_{EPU}	0.507		[0.40]	0.024	49.62
F_{CPU}	3.193	***	[2.68]	0.165	59.09

Note This table shows the performances of CRD-based factors and traditional risk factors, including market factor (MKT), size factor (SMB), book-to-market factor (HML), profitability factor (RMW), and investment factor (CMA). The selected F_{CRD} represents the overall return spread between firms with the lowest, and firms with the highest, level of climate risk disclosure. F_{EPU} is a factor-mimicking portfolio for EPU. F_{CPU} is a factor-mimicking portfolio for CPU. Mean ret shows the average monthly excess return for each factor, calculated relative to the risk-free rate. t-stat is the t-statistic associated with the mean return, indicating the statistical significance of each factor's return. Sharpe ratio measures the risk-adjusted return by dividing the mean return by its standard deviation. Positive % reports the percentage of months where the factor generated a positive return. * Significance at the 10% level. ** Significance at the 5% level. *** Significance at the 1% level.

The mean return of F_{CRD} is 0.619%, with a t-statistic of 2.15. This mean return is the second highest among the eight factors, only below that of F_{CPU}. In terms of the Sharpe ratio of 0.130, F_{CRD} outperforms all other factors except F_{CPU}. The observation that F_{CRD} outperforms traditional risk factors implies that the factor may carry certain return predictive power for 25 size-B/M portfolios in the cross-section, thereby reinforcing the stock-level findings of Tables 4.5 and 4.6.

4.5.3.2. Relationship between F_{CRD} and other factors

Going further from the preliminary analysis of eight individual factors in Table 4.7, Table 4.8 examines the relationship between them. Again, in the interest of space, the table continues to select the F_{CRD} factor formed as the bottom 10% - top 10% in terms of CRD scores. Panel A of the table presents the Pearson correlation coefficients, and Panel B reports the results of regressing F_{CRD} on the other seven factors.

Table 4. 8. Relationship between F_{CRD} (B10-T10) and other factors

Panel A Pearson correlation among factors

	F_{CRD}	MKT	SMB	HML	RMW	CMA	F_{EPU}
MKT	0.251						
SMB	0.275	0.123					
HML	0.010	-0.156	-0.560				
RMW	-0.406	-0.283	-0.743	0.294			
CMA	0.404	0.103	0.380	0.110	-0.681		
F_{EPU}	0.001	-0.138	-0.099	0.148	0.081	0.056	
F_{CPU}	0.062	0.071	0.068	-0.054	-0.024	0.068	0.007

Panel B Regression of F_{CRD} on other factors

	(1)		(2)		(3)		(4)	
Intercept	0.032	***	0.026	***	0.021	***	0.021	**
	[41.94]		[10.81]		[7.14]		[2.26]	
MKT	0.111	**	0.108	**	0.078		0.078	
	[2.06]		[2.14]		[1.59]		[1.58]	
SMB			0.414	***	0.150		0.151	
			[3.43]		[1.25]		[1.19]	
HML			0.397	***	0.212	*	0.211	
			[2.75]		[1.72]		[1.72]	
RMW					-0.248		-0.248	
					[-1.52]		[-1.48]	
CMA					0.402	**	0.401	*
					[2.35]		[2.32]	
F_{EPU}							0.001	
							[0.07]	
F_{CPU}							-0.001	
							[-0.05]	
R2 (%)	0.1889		0.304		0.3587		0.3588	
Obs.	264		264		264		264	

Note This table presents the relationship between the Climate Risk Disclosure factor (F_{CRD}) and other market and policy factors, including the market factor (MKT), size factor (SMB), book-to-market factor (HML), profitability factor (RMW), investment factor (CMA), factor-mimicking portfolio for EPU (F_{EPU}), and factor-mimicking portfolio for CPU (F_{CPU}). F_{CRD} (B10-T10) represents the return spread between firms with the bottom 10% and firms with the top 10% of climate risk disclosure. Panel A presents Pearson correlation coefficients between F_{CRD} and other factors. Panel B reports the results of OLS regressions of F_{CRD} on other risk factors. Model 1 uses the CAPM framework. Model 2 extends it to the Fama-French 3-factor model. Model 3 further expands to the Fama-French 5-factor model. Model 8 incorporates policy variables into the 5-factor model to assess the impact of economic and climate policy uncertainties on returns. * Significance at the 10% level. ** Significance at the 5% level. *** Significance at the 1% level.

In Panel A, the correlations between F_{CRD} and other factors are generally low. This result suggests that F_{CRD} may capture information distinct from conventional market risk factors and policy uncertainty risk factors, to become a unique factor.

In Panel B, the regression results further reveal F_{CRD}'s independence from other factors. Across all models, the intercept remains positive and statistically significant, suggesting that F_{CRD} contains extra information not accounted for by the other factors. In Model 2, only MKT and SMB show significant positive coefficients. In Models 3 and 4, the only significant factor is CMA. When the policy uncertainty risk factors F_{EPU} and F_{CPU} are added in Model 4, their coefficients are insignificant both statistically and economically. Thus, F_{EPU} and F_{CPU} do not contribute to the variations in F_{CRD}.

4.5.3.3. Fama-Macbeth regression results

A widely applied econometric method to estimate return premiums in the cross-section of portfolios is the two-stage Fama-MacBeth (FM) regression. In stage 1, each of the 25 size-BM portfolios is regressed on N selected factors over a rolling window (set at 60 months). These time-series regressions yield the loading of a portfolio on each of the N selected factors, giving N betas. The 25 size-BM portfolios result in 25×N betas, depending on the value of N. Then, we repeat the above time-series regressions by rolling the window one month forward, until the end of the sample period is reached. In stage 2, the cross-section of 25 future 12 months' average portfolio returns are regressed against the betas obtained from stage 1. In our case, N = 8, so there will be eight estimated betas (such as β_{FCRD}, β_{MKT}, …, β_{FCPU}) in model (14) of Table 4.9.

Table 4.9 presents the results of the FM regressions to gauge the return predictive power of F_{CRD} either independently or in combination with other risk factors. This time, we incorporate all the four versions of F_{CRD} into the analysis, as the results should be more important and of more practical relevance than those reported by Tables 4.7 and 4.8. Model (10) through Model (13) respectively incorporate the four F_{CRD} factors in the two-stage FM regressions. The first observation to make is that the coefficient (γ_{FCRD}) on β_{FCRD} is always positive and significant at a higher than 5% level, with two at a higher than 1% level, across

Table 4. 9. Predictive power of F_{CRD} for 25 size/BM portfolio returns in the cross-section

	(1) B10-T10		(2) B20-T20		(3) B30-T30		(4) B40-T40	
Intercept	0.021	***	0.018	***	0.004	*	0.015	***
	[4.55]		[6.52]		[1.87]		[2.99]	
β_{FCRD}	0.082	***	0.029	**	0.029	**	0.029	**
	[3.20]		[2.16]		[2.05]		[1.99]	
β_{MKT}	-0.020		-0.046	***	-0.077	***	-0.076	***
	[-0.85]		[-3.34]		[-10.32]		[-5.05]	
β_{SMB}	-0.024	***	-0.021	***	0.000		-0.023	***
	[-7.43]		[-3.23]		[-0.05]		[-5.86]	
β_{HML}	0.011	***	0.017	***	0.006	*	0.020	***
	[2.86]		[5.43]		[1.95]		[7.13]	
β_{RMW}	-0.017		-0.010		-0.010		-0.007	
	[-1.08]		[-0.84]		[-1.07]		[-0.66]	
β_{CMA}	0.007		0.003		-0.008		-0.008	
	[0.53]		[0.27]		[-0.84]		[-0.80]	
β_{FEPU}	-0.386		0.361		0.663	**	0.464	
	[-1.54]		[1.05]		[2.00]		[1.35]	
β_{FCPU}	0.044	***	0.050	***	0.056	***	0.047	***
	[5.96]		[5.32]		[7.21]		[4.40]	
Adj. R^2	38.32		38.13		38.69		39.33	
No. obs.	4422		4422		4422		4422	

Note This table reports the results of the Fama-MacBeth two-stage regression of monthly excess returns on 25 size/BM portfolios against various factors including F_{CRD}, MKT, SMB, HML, RMW, CMA, F_{EPU} and F_{CPU}. In Models (1), (2), (3), and (4), F_{CRD} is formed as, respectively, the bottom 10% minus the top 10%, the bottom 20% minus the top 20%, the bottom 30% minus the top 30%, and the bottom 40% minus the top 40% in terms of firms' CRD scores. Reported are the time-series averaged coefficients on factor loadings (betas). The Newey-West standard errors with 12-month lags are used to adjust for heteroscedasticity and autocorrelation in the error terms. * Significance at the 10% level. ** Significance at the 5% level. *** Significance at the 1% level.

Models (1) to (4). In addition, the adjusted R^2 (Adj. R^2) stays at a reasonably high level between 38.13% and 39.33%. It should be stressed that the four models all allow for the full-set of seven standard risk factors as control variables. The control variables include the Fama-French three factors (MKT, SMB, and HML), RMW (profitability factor), CMA (investment factor), F_{EPU} (economic policy uncertainty factor), and F_{CPU} (climate policy uncertainty factor). These results consistently suggest that the loadings of the 25 size-B/M portfolios on the four differently-formed F_{CRD} portfolios positively predict their future returns in the cross-section, and add incrementally to the predictive power of their loadings on the seven other factors. Thus, we can conclude that the climate risk disclosure factor F_{CRD} commands a positive equity premium – which is the focus of this study. This is despite the fact that it is formed as the bottom 10% - top 10%, bottom 20% - top 20%, …, or bottom 40% - top 40%, in terms of firms' CRD scores. This is also the case despite the market factor (MKT), the size factor (SMB), the value factor (HML) and the climate policy uncertainty factor (F_{CPU}) being found to command either a negative or positive equity premium. Our portfolio-level results further corroborate the stock-level findings: CRD-characteristic-based anomalies are present in China, as evidenced by F_{CRD}'s return predictive power for the cross-section of 25 size-B/M portfolios.

4.6. Conclusion

This study explores the return-anomaly implications of climate risk disclosure in the Chinese A-share market. First, we construct a robust firm-level CRD measure using textual analysis of corporate annual reports. Through quantifying climate-related disclosures using a lexicon of 155 keywords derived from authoritative sources, this measure discloses the information on firms' climate-related risk. We treat this measure as a new firm characteristic and coin the term "CRD characteristic", which can be examined against and alongside conventional firm characteristics. Our descriptive analysis reveals significant variations in CRD across industries and conventional firm characteristics. The findings reflect the variability in climate-risk-disclosure practices among companies in the Chinese stock market.

Second, we investigate the relationship between CRD and stock returns at both stock-level and portfolio-level analyses. Our results indicate that firms with lower CRD scores tend to deliver higher subsequent returns, suggesting a potential equity premium demanded by investors for holding stocks with low climate risk disclosure. This relationship is further validated through the construction of a climate-risk-disclosure factor (F_{CRD}), as we find that F_{CRD} yields significant abnormal returns relative to traditional risk factors and policy uncertainty factors. All these findings imply the presence of return anomalies based on the CRD-characteristic.

Thirdly, Fama-MacBeth regressions demonstrate that loading on the F_{CRD} factor can positively and significantly predict the future cross-section of 25 size-B/M portfolio returns. The results offer further evidence in support of CRD-characteristic-based anomalies. The distinction between physical and transition climate risk also provides additional economic interpretation of the documented CRD-characteristic-based return predictability. Physical exposures can affect firms through direct operating and asset-related losses, whereas transition exposures can operate through regulatory, technological, and market-adjustment costs. The aggregate CRD measure does not separately identify the return effects attributable to these two channels; accordingly, examining physical- and transition-risk disclosure separately represents a useful direction for future research.

Although the external validation results support the economic relevance of the CRD measure, third-party environmental and ESG indicators are not treated as direct substitutes for CRD because they capture broader constructs and have different sample coverage. Direct substitution would therefore alter both the underlying construct and the empirical sample rather than provide a directly comparable robustness test. Future research could further validate the findings using alternative climate-disclosure dictionaries or context-sensitive textual measures when sufficiently comprehensive firm-level data become available.

A related limitation concerns potential endogeneity in the CRD–return relationship. Although the use of lagged CRD, firm fixed effects, and firm-level controls helps mitigate reverse causality and omitted-variable concerns, these procedures do not establish strict causal identification. The findings should therefore be interpreted primarily as evidence that prior climate risk disclosure predicts subsequent stock returns, while future research could employ suitable instrumental-variable or quasi-natural-experimental designs to provide stronger causal identification.

Two facts are well-known in China: Climate risk disclosures are implemented via a voluntary approach; and the majority of investors are retail investors characterised by irrational behaviour. In view of these two facts, we can derive two policy implications from our results. The Chinese government should: (1) make climate risk disclosures mandatory to narrow the difference in CRD between the bottom and top deciles; and (2) should encourage the rapid growth of institutional investors like pension funds, mutual funds, and insurance companies, while reducing the proportion of retail investors in financial markets. These two policy measures would help shrink the CRD-characteristic-based anomalies and increase the efficiency of the equity market.

CHAPTER FIVE: CONCLUSION

This chapter concludes the thesis by summarising the main findings and implications of the three essays, and by outlining limitations and directions for future research.

5.1. Thesis Summary and Overall Contributions

This thesis consists of three empirical essays that examine asset return anomalies in the Chinese A-share market from the perspective of informational and institutional frictions. Although each essay focuses on a distinct source of market friction, the three studies jointly develop an integrated framework for understanding how asset pricing anomalies arise, persist, and evolve in the Chinese stock market. The first essay examines financial disclosure quality as an information friction, showing how limited transparency contributes to stock anomalies, which impedes investors' ability to correctly price firms' fundamentals. The second essay studies short-selling constraints as a trading friction, examining how limits to arbitrage influence the persistence of anomaly returns and distinguishing between the effects of short-selling eligibility and the intensity of short-selling activity. The third essay extends the analysis to non-traditional firm characteristics, using climate risk exposure to reveal how new dimensions of information generate novel forms of mispricing that are not captured by traditional financial variables. Taken together, the three essays portray anomalies not as isolated empirical regularities, but as the outcome of a market environment. By focusing on China, a market where retail investors dominate trading and regulatory reforms continue to reshape information flows and market mechanisms, the thesis provides context-specific evidence on the pricing implications of non-standard characteristics and institutional frictions.

The three essays contribute to the literature on characteristic-based stock pricing anomalies by expanding the scope of firm characteristics beyond traditional financial variables, and highlighting the role of market design and disclosure practices in influencing mispricing. Importantly, the three essays develop new firm-level proxies, including opacity metrics, short-selling depth measures, and climate risk disclosure scores, that enable the empirical investigation of anomalies in settings where comparable data are scarce or absent.

Beyond academic relevance, the thesis speaks to broader debates on the functioning of emerging markets, the limits to arbitrage in the presence of institutional constraints, and the implications of evolving corporate disclosure requirements. By documenting how informational and trading frictions affect anomaly profitability, the thesis suggests that market reforms, whether through improvements in disclosure quality, the liberalisation of short-selling mechanisms, or the standardisation of climate-related reporting, may influence both pricing efficiency and the opportunity set available to investors.

5.2. Summary of Main Findings by Essay

5.2.1 Essay One: Information Opacity and Anomaly Returns in China

Essay One examines whether firm-level information opacity contributes to the persistence of anomaly returns in the Chinese A-share market. Using eleven well-established anomaly variables from Stambaugh et al. (2012) and Chu et al. (2020), the analysis first confirms that most anomaly-based long–short strategies generate positive profits during the sample period, consistent with mispricing under limits to arbitrage. The study then evaluates whether these premiums are conditional on firm-level disclosure quality.

The results show that anomaly returns are significantly higher among high-opacity firms than among low-opacity firms, indicating that weak information disclosure amplifies mispricing. Long-leg returns are generally stronger under high opacity, while short-leg returns exhibit limited sensitivity to opacity, implying that pessimistic valuations are less effectively expressed when disclosure is weak. Regression tests further demonstrate that opacity remains negatively associated with anomaly returns after controlling for Fama-French factors and macro-related uncertainty, suggesting its effect is incremental to common risk components.

Overall, Essay One provides evidence that limited information disclosure is a material friction contributing to anomaly persistence in China. These findings highlight the role of corporate transparency in shaping market efficiency and suggest that policies enhancing disclosure quality may facilitate price correction in markets with high retail participation.

5.2.2 Essay Two: Short-Selling Constraints and Anomaly Returns

Essay Two examines whether short-sale constraints contribute to the persistence of return anomalies in the Chinese A-share market, using phased liberalisation of the margin-trading and securities-lending mechanism as a quasi-natural experiment. The empirical framework integrates characteristic-sorted portfolios, phased difference-in-differences estimation, and short-selling depth measures to evaluate both the extensive (eligibility) and intensive (trading activity) margins of arbitrage capacity.

The main findings indicate that short-sale eligibility has limited and heterogeneous effects on anomaly profitability, rather than producing a uniform attenuation across anomalies. For certain profitability-, investment-, and issuance-related characteristics, long–short spreads decline following short-sale eligibility, though the magnitude and significance vary across expansion waves, suggesting that short-sale access may facilitate the correction of overpricing embedded in aggressive investment and issuance behaviour. The DiD tests further show that the effect of short-sale eligibility is concentrated in a subset of anomalies and varies across expansion waves, reflecting the incremental nature of China's regulatory rollout rather than an immediate, market-wide adjustment.

Beyond eligibility, variation in short-selling depth provides additional evidence on the role of informed arbitrage. Within the shortable sample, anomalies are consistently weaker among firms with deeper short-selling activity, even after adjusting for Fama–French five-factor exposures. This pattern suggests that the profitability of anomaly-based strategies relies on limits to arbitrage, and that expanding short-selling capacity improves price efficiency along dimensions related to investment intensity, equity issuance, and earnings quality.

Taken together, the evidence supports the view that structural short-sale frictions contribute to the persistence of anomaly returns in the Chinese market. However, attenuation remains incomplete, as institutional constraints on lending supply, borrowing costs, investor access, and retail market dominance limit the ability of arbitrageurs to fully eliminate mispricing. Consequently, efficiency improvements associated with short-selling liberalisation appear gradual and selective, with anomaly profitability declining but not disappearing during the sample period.

5.2.3 Essay Three: Climate Risk Disclosure and Return Anomalies

Essay Three investigates whether firms' climate risk disclosure constitutes a non-financial firm characteristic that generates return anomalies in the Chinese A-share market. To this end, a textual analysis approach is used to construct firm-level climate risk disclosure (CRD) scores based on annual reports, and the CRD measure is treated as a firm characteristic analogous to profitability, investment, or financing characteristics. The analysis examines both stock-level and portfolio-level implications, with particular attention to whether climate risk disclosure embeds information relevant for cross-sectional return predictability.

The empirical results consistently show that lower climate risk disclosure is associated with higher subsequent returns. Fixed-effects regressions confirm that the CRD characteristic negatively and significantly predicts excess returns. Portfolios formed on CRD deciles display significantly positive long–short spreads even after adjusting for conventional factor models, indicating the presence of a CRD-characteristic-based premium. These findings imply that investors demand compensation for holding firms with limited disclosure of climate-related risks, consistent with either a misvaluation-based story or a compensation-for-transparency-premium interpretation.

Portfolio-level analysis further examines whether CRD-related return patterns can be summarised into a factor structure. A climate-risk-disclosure factor (FCRD) is constructed by taking long positions in low-disclosure firms and short positions in high-disclosure firms. The FCRD factor exhibits positive mean returns, favourable Sharpe ratios relative to conventional factors, and low correlation with market, size, value, profitability, investment, and policy uncertainty factors. In two-stage Fama–MacBeth regressions, factor loadings on FCRD positively and significantly predict the cross-section of 25 size–book-to-market portfolio returns, controlling for the seven conventional factors. These results collectively indicate that climate risk disclosure contains information not subsumed by existing factor structures and that the CRD characteristic contributes to explaining anomaly returns in the Chinese market.

Taken together, the study documents that climate risk disclosure forms a non-financial firm characteristic that carries return predictive power and commands a positive premium. In the Chinese setting, voluntary disclosure regimes combined with a retail-investor-dominated market appear to create conditions under which climate-related transparency becomes a priced attribute. This suggests that environmental disclosure practices have asset pricing relevance beyond conventional ESG frameworks and contributes to the broader literature on characteristic-based anomalies.

5.3 Implications, Limitations and Future Research

Collectively, the three essays highlight the asset pricing relevance of informational and institutional frictions in the Chinese A-share market. The findings have several implications. For academic research, the results suggest that anomalies need not arise solely from traditional financial characteristics. Non-financial disclosure attributes also shape the cross-section of returns and expand the empirical scope of characteristic-based pricing studies. For investors, the evidence implies that anomaly-based strategies remain profitable where arbitrage capacity is limited and disclosure is imperfect, although profitability declines as mechanisms facilitating short selling develop. For regulators and policymakers, improvements in disclosure practices and the gradual liberalisation of short-selling channels may enhance market efficiency, but institutional constraints can delay price adjustment.

The thesis faces several limitations. The analysis focuses on Chinese A-share firms, and the external validity of the findings remains an open question, particularly in markets with different investor compositions or regulatory structures. In addition, the empirical frameworks emphasise reduced-form pricing performance rather than structural modelling.

These limitations generate opportunities for future research. Cross-country comparisons may reveal how disclosure standards, climate reporting regimes, or short-selling institutions differentially shape anomaly persistence. Further work could explore whether derivatives, ETF participation, or institutional investor development alters the pricing of disclosure-related characteristics over time. Finally, integrating textual machine learning approaches may allow for more nuanced measures of climate and non-financial disclosure, enabling sharper tests of their asset pricing implications.

REFERENCES

Allen, F., & Gorton, G. (1993). Churning bubbles. *The Review of Economic Studies*, *60*(4), 813-836.

Allen, F., Qian, J., & Qian, M. (2005). Law, finance, and economic growth in China. *Journal of Financial Economics, 77*(1), 57-116.

Alti, A., & Titman, S. (2019). A dynamic model of characteristic-based return predictability. *The Journal of Finance, 74*(6), 3187-3216.

Andrikopoulos, A., Babalos, V., Balios, D., & Filis, G. (2019). Is there still a weather anomaly? An investigation of stock and foreign exchange markets. *International Review of Financial Analysis*, *61,* 38-51.

Andrikopoulos, A., Wang, C., & Zheng, M. (2019). Is there still a weather anomaly? An investigation of stock and foreign exchange markets. *Finance Research Letters*, *30,* 51-59.

Ang, A., Chen, J., & Xing, Y. (2006). Downside risk. *The Review of Financial Studies*, *19*(4), 1191-1239.

Asquith, P., Pathak, P. A., & Ritter, J. R. (2005). Short interest, institutional ownership, and stock returns. *Journal of Financial Economics, 78*(2), 243-276. https://doi.org/10.1016/j.jfineco.2005.01.001

Avramov S., Chordia, T., Jostova, G., & Philipov, A. (2013). Anomalies and financial distress. *Journal of Financial Economics, 108*, 139-159.

Avramov, D., & Chordia, T. (2006). Asset pricing models and financial market anomalies. *The Review of Financial* Studies, *19*(3), 1001-1040.

Bai, M. (2021). Are firm characteristics priced differently between opposite short-sales regimes? *International Finance*, *24*(1), 95-118.

Bai, M., Li, X. M., & Qin, Y. (2017). Shortability and asset pricing model: Evidence from the

Hong Kong stock market. *Journal of Banking & Finance*, *85*, 15-29.

Ball, A., Owen, D. L., & Gray, R. (2000). External transparency or internal capture? The role of third-party statements in adding value to corporate environmental reports1. *Business Strategy and the Environment*, *9*(1), 1-23.

Ball, R., & Brown, P. (1968). An empirical evaluation of accounting income numbers. *Journal of Accounting Research*, *6*(2), 159-178.

Ball, R., Gerakos, J., Linnainmaa, J. T., & Nikolaev, V. V. (2015). Deflating profitability. *Journal of Financial Economics*, *117*(2), 225-248.

Bansal, R., Ochoa, M., & Kiku, D. (2017). Climate risk and the cross-section of stock returns. *The Journal of Finance, 74*(6), 2493-2530.

Banz, R. (1981). The relationship between return and market value of common stock. *Journal of Financial Economics 9*, 3-18.

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics, 9*(1), 3-18. https://doi.org/10.1016/0304-405X(81)90018-0

Barberis, N., & Thaler, R. (2003). A survey of behavioral finance. *Handbook of the Economics of Finance*, *1,* 1053-1128.

Barberis, N., Huang, M., & Santos, T. (2001). Prospect theory and asset prices. *Quarterly Journal of Economics*, *116*(1), 1-53. https://doi.org/10.1162/003355301556310

Barberis, N., Shleifer, A., & Vishny, R. (1998). A model of investor sentiment. *Journal of Financial Economics, 49*(3), 307-343.

Barry, C. B., & Brown, S. J. (1985). Differential information and security market equilibrium. *Journal of Financial and Quantitative Analysis*, *20*(4), 407-422.

Basu, S. (1977). Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis. *The Journal of Finance*, *32*(3), 663-682.

Bauer, R., Cosemans, M., & Schotman, P. C. (2010). Conditional asset pricing and stock market anomalies in Europe. *European Financial Management*, *16*(2), 165-190.

Beber, A., & Pagano, M. (2013). Short-selling bans around the world: Evidence from the 2007–09 crisis. *Journal of Finance, 68*(1), 343-381. https://doi.org/10.1111/j.1540-6261.2012.01802.x

Bergsma, K., & Jiang, D. (2016). Cultural New Year holidays and stock returns around the world. *Financial Management*, *45*(1), 3-35.

Bingler, J. A., Kraus, M., Leippold, M., & Webersinke, N. (2022). Cheap talk and cherry-picking: What ClimateBert has to say on corporate climate risk disclosures. *Finance Research Letters*, *47*, 102776.

Birru, J. (2018). Day of the week and the cross-section of returns. *Journal of Financial Economics*, *130*(1), 182-214.

Black, F., Jensen, M. C., & Scholes, M. (1972). The capital asset pricing model: Some empirical tests. Available from *SSRN*: https://ssrn.com/abstract=908569

Board, F. S. (2017). Task force on climate-related financial disclosures. *Final Report: Recommendations of the Task Force on Climate-Related Financial Disclosures*.

Boehmer, E., & Wu, J. (2013). Short selling and the price discovery process. *Review of Financial Studies, 26*(2), 287-322. https://doi.org/10.1093/rfs/hhs097

Bofinger, Y., Heyden, K. J., & Rock, B. (2022). Corporate social responsibility and market efficiency: Evidence from ESG and misvaluation measures. *Journal of Banking and Finance, 134*, 106322.

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

Bolton, P., Despres, M., Da Silva, L. A. P., Samama, F., & Svartzman, R. (2020). The green swan. *BIS Books*.

Bouchaud, J. P., Krueger, P., Landier, A., & Thesmar, D. (2019). Sticky expectations and the profitability anomaly. *The Journal of Finance*, *74*(2), 639-674.

Bouchaud, J.-P., Krueger, P., Landier, A., & Thesmar, D. (2019). Sticky expectations and the profitability anomaly. *Journal of Finance*, *74*(2), 639-674.

Bradshaw, M., Richardson, S., & Sloan, R. (2006). The relation between corporate financing activities, analysts' forecasts, and stock returns. *Journal of Accounting and Economics, 42*, 53-85.

Brav, A., & Heaton, J. B. (2002). Competing theories of financial anomalies. *The Review of Financial Studies*, *15*(2), 575-606.

Brennan, M. J., Chordia, T., & Subrahmanyam, A. (1998). Alternative factor specifications, security characteristics, and the cross-section of expected stock returns. *Journal of Financial Economics*, *49*(3), 345-373.

Bris, A., Goetzmann, W. N., & Zhu, N. (2007). Efficiency and the bear: Short sales and markets around the world. *Journal of Finance, 62*(3), 1029-1079.

Bushman, R. M., Piotroski, J. D., & Smith, A. J. (2004). What determines corporate transparency? *Journal of Accounting Research*, *42*(2), 207-252.

Campbell, J. Y., Hilscher, J., & Szilagyi, J. (2008). In search of distress risk. *The Journal of Finance*, *63*(6), 2899-2939.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, *52*(1), 57-82.

Chae, S. J., Nakano, M., & Fujitani, R. (2020). Financial reporting opacity, audit quality and crash risk: Evidence from Japan. *The Journal of Asian Finance, Economics and Business, 7*(1), 9-17.

Chang, E. C., Luo, Y., & Ren, J. (2013). Short-selling, margin-trading, and price efficiency: Evidence from the Chinese market. *Journal of Banking & Finance, 37*(9), 3419-3436. https://doi.org/10.1016/j.jbankfin.2013.05.007

Chang, E. C., Luo, Y., & Ren, J. (2014). Short-selling, margin-trading, and price efficiency: Evidence from the Chinese market. *Journal of Banking & Finance*, *48*, 411-424. doi.org/10.1016/j.jbankfin.2013.10.002

Chatterji, A. K., Levine, D. I., & Toffel, M. W. (2009). How well do social ratings actually measure corporate social responsibility? *Journal of Economics & Management Strategy, 18*(1), 125-169.

Chen, J., Hong, H., & Stein, J. C. (2002). Breadth of ownership and stock returns. J*ournal of Financial Economics, 66*(2-3), 171-205. https://doi.org/10.1016/S0304-405X(02)00223-4

Chen, J., Kadapakkam, P. R., & Yang, T. (2016). Short selling, margin trading, and the incorporation of new information into prices. *International Review of Financial Analysis*, *44*, 1-17.

Chen, L., Kim, H., Yao, T., & Yu, T. (2010). On the predictability of Chinese stock returns. *Journal of Empirical Finance, 17*(4), 595-610. https://doi.org/10.1016/j.jempfin.2010.04.001

Chen, L., Novy-Marx, R., & Zhang, L. (2011). An alternative three-factor model. *Available at SSRN 1418117*.

Chincarini, L. B., Lazo-Paz, R., & Moneta, F. (2025). Crowded spaces and anomalies. *Journal of Banking & Finance*, 182, 107579. https://doi.org/10.1016/j.jbankfin.2025.107579

Chiyachantana, C. N., Nuengwang, N., Taechapiroontong, N., & Thanarung, P. (2013). The effect of information disclosure on information asymmetry. *Investment Management and Financial Innovations*, *10*(1), 225-234.

Chu, Y., Hirshleifer, D., & Ma, L. (2020). The causal effect of limits to arbitrage on asset pricing anomalies. *The Journal of Finance*, *75*(5), 2631-2672.

Connor, G., & Korajczyk, R. A. (1988). Risk and return in an equilibrium APT: Application of a new test methodology. *Journal of Financial Economics*, *21*(2), 255-289.

Cooper, M. J., Gulen, H., & Schill, M. J. (2008). Asset growth and the cross-section of stock returns. *The Journal of Finance*, *63*(4), 1609-1651.

D'Avolio, G. (2002). The market for borrowing stock. *Journal of Financial Economics, 66*(2-3), 271-306.

Daniel, K. D., Hirshleifer, D., & Subrahmanyam, A. (2001). Overconfidence, arbitrage, and equilibrium asset pricing. *The Journal of Finance*, *56*(3), 921-965.

Daniel, K. D., Hirshleifer, D., & Subrahmanyam, A. (2005). Investor psychology and tests of factor pricing models. *SSRN Working Paper 854024*.

Daniel, K., & Titman, S. (2006). Market reactions to tangible and intangible information. *Journal of Finance, 61*, 1605-1643

De Bondt, W. F. M., & Thaler, R. (1985). Does the stock market overreact? *Journal of Finance, 40*(3), 793-805. https://doi.org/10.1111/j.1540-6261.1985.tb05004.x

DeBoskey, D. G., & Gillett, P. R. (2013). The impact of multi-dimensional corporate transparency on us firms' credit ratings and cost of capital. *Review of Quantitative Finance and Accounting*, *40*(1), 101-134.

Deng, Y., & Xu, Y. (2011). Do institutional investors have superior stock selection ability in China? *China Journal of Accounting Research*, *4*(3), 107-119.

Diamond, D. W., & Verrecchia, R. E. (1987). Constraints on short-selling and asset price adjustment to private information. *Journal of Financial Economics, 18*(2), 277-311. https://doi.org/10.1016/0304-405X(87)90042-0

Diamond, D. W., & Verrecchia, R. E. (1991). Disclosure, liquidity, and the cost of capital. *The Journal of Finance*, *46*(4), 1325-1359.

Dichev, I. D. (1998). Is the risk of bankruptcy a systematic risk? *The Journal of Finance*, *53*(3), 1131-1147.

Diether, K. B., Lee, K. H., & Werner, I. M. (2009). Short-sale strategies and return predictability.

Review of Financial Studies, 22(2), 575-607. https://doi.org/10.1093/rfs/hhn036

Dontoh, A., Ronen, J., & Sarath, B. (2003). On the rationality of the post-announcement drift. *Review of Accounting Studies*, *8*(1), 69-104.

Dowling, M., & Lucey, B. M. (2005). Weather, biorhythms, beliefs and stock returns—some preliminary Irish evidence. *International Review of Financial Analysis*, *14*(3), 337-355.

Drake, M. S., Myers, J. N., & Myers, L. A. (2009). Disclosure quality and the mispricing of accruals and cash flow. *Journal of Accounting, Auditing & Finance*, *24*(3), 357-384.

Duan, D., Wei, R., Wang, C., & Xia, B. (2025). Opportunity or obstacle? Climate risk disclosure and corporate ESG performance. *International Review of Economics & Finance, 100*, 1-14.

Duffie, D., Garleanu, N., & Pedersen, L. H. (2002). Securities lending, shorting, and pricing. *Journal of Financial Economics*, *66*(2-3), 307-339.

Durnev, A., Errunza, V., & Molchanov, A. (2009). Property rights protection, corporate transparency, and growth. *Journal of International Business Studies*, *40*(9), 1533-1562.

Easley, D., Hvidkjaer, S., & O'Hara, M. (2002). Is information risk a determinant of asset returns? *The Journal of Finance, 57*(5), 2185-2221.

Edelen, R. M., Ince, O. S., & Kadlec, G. B. (2016). Institutional investors and stock return anomalies. *Journal of Financial Economics*, *119*(3), 472-488.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance, 25*(2), 383-417. https://doi.org/10.2307/2325486

Fama, E. F., & French, K. R. (1992). The cross‐section of expected stock returns. *Journal of Finance, 47*(2), 427-465. https://doi.org/10.1111/j.1540-6261.1992.tb04398.x

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 33*(1), 3-56. https://doi.org/10.1016/0304-405X(93)90023-5

Fama, E. F., & French, K. R. (2006). Profitability, investment and average returns. *Journal of Financial Economics*, *82*(3), 491-518.

Fama, E. F., & French, K. R. (2008). Dissecting anomalies. *Journal of Finance, 63*(4), 1653-1678. https://doi.org/10.1111/j.1540-6261.2008.01371.x

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, *116*(1), 1-22.

Feng, G., Giglio, S., & Xiu, D. (2020). Taming the factor zoo: A test of new factors. *Journal of Finance, 75*(3), 1327-1370. https://doi.org/10.1111/jofi.12883

Firth, M., Wang, K., & Wong, S. M. (2015). Corporate transparency and the impact of investor sentiment on stock prices. *Management Science*, *61*(7), 1630-1647.

Fisher, G. S., Shah, R., & Titman, S. (2015). Combining value and momentum. *Journal of Investment Management*, Forthcoming.

Friedman, M. (1953). The case for flexible exchange rates. *Essays in Positive Economics*, *157*(203), 33.

Gao, X., Lee, C. M. C., & Murphy, D. (2018). Financing and valuation in China's A-share market. *Journal of Financial Economics, 127*(1), 18-42.

Gonzalez, M. F. (2001). CAPM performance in the Caracas Stock Exchange from 1992-1998. *International Review of Financial Analysis*, *10*(3), 333-341.

Griffin, J. M., Ji, X., & Martin, J. S. (2003). Momentum investing and business cycle risk: Evidence from pole to pole. *Journal of Finance, 58*(6), 2515-2547. https://doi.org/10.1046/j.1540-6261.2003.00613.x

Grinblatt, M., & Han, B. (2005). Prospect theory, mental accounting, and momentum. *Journal of Financial Economics*, *78*(2), 311-339.

Grossman, S. J., & Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *The American Economic Review*, *70*(3), 393-408.

Gwilym, O. A., Hasan, I., Wang, Q., & Xie, R. (2016). In search of concepts: the effects of speculative demand on stock returns. *European Financial Management 22*(3), 427-449.

Harris, M., & Raviv, A. (1993). Differences of opinion make a horse race. *The Review of Financial Studies, 6*(3), 473-506.

Healy, P. M., & Palepu, K. G. (2001). Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. Journal of accounting and economics, 31(1-3), 405-440.

Heckman, J. J., Ichimura, H., & Todd, P. (1998). Matching as an econometric evaluation estimator. *The review of economic studies*, *65*(2), 261-294.

Hilliard, J., & Zhang, H. (2015). Size and price-to-book effects: Evidence from the Chinese stock markets. *Pacific-Basin Finance Journal*, *32*, 40-55.

Hirshleifer, D., Hou, K., & Teoh, S. H. (2012). The accrual anomaly: Risk or mispricing. *Management Science, 58*(2), 320-335.

Hirshleifer, D., Hou, K., Teoh, S. H., & Zhang, Y. (2004). Do investors overvalue firms with bloated balance sheets? *Journal of Accounting and Economics, 38*, 297-331.

Hirshleifer, D., Hou, K., Teoh, S. H., & Zhang, Y. (2004). Do investors overvalue firms with bloated balance sheets? *Journal of Accounting and Economics*, *38*, 297-331.

Hirst, D. E., & Hopkins, P. E. (1998). Comprehensive income reporting and analysts' valuation judgments. *Journal of Accounting Research*, *36*, 47-75.

Hong, H., & Stein, J. C. (1999). A unified theory of underreaction, momentum trading, and overreaction in asset markets. *The Journal of Finance, 54*(6), 2143-2184.

Hou, K., Xue, C., & Zhang, L. (2015). Digesting anomalies: An investment approach. *Review of Financial Studies, 28*(3), 650-705. https://doi.org/10.1093/rfs/hhu068

Hou, K., Xue, C., & Zhang, L. (2020). Replicating anomalies. *Review of Financial Studies, 33*(5), 2019-2133. https://doi.org/10.1093/rfs/hhz101

Huang, Y., & Li, J. (2020). Machine learning-based analysis of climate risk disclosures: A comprehensive approach. *Journal of Financial Data Science, 2*(3), 31-47.

Huang, Y., & Luk, P. (2020). Measuring economic uncertainty in China. *China Economic Review, 59*, 101367.

Huang, Y., Li, D., & Zhou, H. (2016). Short selling and market efficiency: Evidence from the Chinese market. *China Economic Review, 39,* 1-14. https://doi.org/10.1016/j.chieco.2016.03.008

Hunton, J. E., Libby, R., & Mazza, C. L. (2006). Financial reporting transparency and earnings management (retracted). *The Accounting Review*, *81*(1), 135-157.

Hutton, A. P., Marcus, A. J., & Tehranian, H. (2009). Opaque financial reports, R2, and crash risk. *Journal of financial Economics*, *94*(1), 67-86.

Ilhan, E., Krueger, P., & Sautner, Z. (2021). Carbon emissions and the cost of capital. *Management Science, 67*(9), 5812-5835.

Ilhan, E., Krueger, P., Sautner, Z., & Starks, L. T. (2023). Climate risk disclosure and institutional investors. *The Review of Financial Studies, 36*, 2617-2650.

Iqbal, J., Brooks, R., & Galagedera, D. U. (2010). Testing conditional asset pricing models: An emerging market perspective. *Journal of International Money and Finance*, *29*(5), 897-918.

Jagannathan, R., & Wang, Z. (1996). The conditional CAPM and the cross-section of expected returns. *The Journal of Finance*, *51*(1), 3-53.

Jansen, M., Swinkels, L., & Zhou, W. (2021). Anomalies in the China A-share market. *Pacific-Basin Finance Journal, 68,* 101607. https://doi.org/10.1016/j.pacfin.2021.101607

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance, 48*(1), 65-91. https://doi.org/10.1111/j.1540-6261.1993.tb04702.x

Jiang, W. (2019). Leveraging big data and natural language processing in corporate climate risk disclosures. *Journal of Environmental Economics and Management, 94,* 174-187.

Jones, C. M., & Lamont, O. A. (2002). Short-sale constraints and stock returns. *Journal of Financial Economics, 66*(2-3), 207-239. https://doi.org/10.1016/S0304-405X(02)00224-6

Kim, O., & Verrecchia, R. E. (1994). Market liquidity and volume around earnings announcements. *Journal of Accounting and Economics*, *17*(1-2), 41-67.

Kogan, L., Li, J., & Zhang, H. H. (2021). Operating hedge and gross profitability premium. *Available at SSRN 3947165*.

Lam, F. Y., Wang, S., & Wei, K. C. (2015). The profitability premium: Macroeconomic risks or expectation errors? In *Financial Management Association (FMA)* 2014 conference. Available at SSRN https://ssrn.com/abstract=2726027

Lawrence, A. (2013). Individual investors and financial disclosure. *Journal of Accounting and Economics*, *56*(1), 130-147.

Lettau, M., & Ludvigson, S. (2001). Resurrecting the (C) CAPM: A cross-sectional test when risk premia are time-varying. *Journal of Political Economy*, *109*(6), 1238-1287.

Leuz, C., & Verrecchia, R. E. (2000). The economic consequences of increased disclosure. *Journal of accounting research*, 91-124.

Li, D., & Zhang, L. (2010). Does q-theory with investment frictions explain anomalies in the cross section of returns? *Journal of Financial Economics*, *98*(2), 297-314.

Li, R., Li, N., Li, J., & Wu, C. (2018). Short selling, margin buying and stock return in China market. *Accounting & Finance*, *58*(2), 477-501.

Li, W., Rhee, G., & Wang, S. S. (2017). Differences in herding: Individual vs. institutional investors. *Pacific-Basin Finance Journal*, *45*, 174-185.

Li, W., Shen, D., & Wang, J. (2022). Short selling, volatility, and liquidity: Evidence from China's margin trading and securities lending program. *International Review of*

Financial Analysis, 79, 101975. https://doi.org/10.1016/j.irfa.2021.101975

Li, X-M. (2017). New evidence on economic policy uncertainty and equity premium. *Pacific-Basin Finance Journal, 46,* 41-56.

Lin, B., & Wu, N. (2023). Climate risk disclosure and stock price crash risk: The case of China. *International Review of Economics and Finance, 83*, 21-34.

Lin, X., & Zhang, L. (2013). The investment manifesto. *Journal of Monetary Economics*, *60*(3), 351-366.

Lin, Y. E., Chu, C. C., Omura, A., Li, B., & Roca, E. (2020). Arbitrage risk and the cross-section of stock returns: Evidence from China. *Emerging Markets Review*, *43*, 100609.

Lintner, J. (1965). Security prices, risk, and maximal gains from diversification. *The Journal of Finance*, *20*(4), 587-615.

Liu, L. X., Whited, T. M., & Zhang, L. (2009). Investment-based expected stock returns. *Journal of Political Economy*, *117*(6), 1105-1139.

Liu, L., Luo, D., & Zhao, N. (2020). Short-selling activity and return predictability: Evidence from the Chinese stock market. *Emerging Markets Finance and Trade*, *56*(14), 3445-3467.

Lochstore, L. A., & Tetlock, P. C. (2020). What drives anomaly returns. *Journal of Finance. 75*(3), 1417-1455.

Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of finance*, *66*(1), 35-65.

Luo, L., Tang, Q., & Lan, Y.-C. (2015). Comparison of propensity for carbon disclosure between developing and developed countries: A resource constraint perspective. *Accounting Research Journal, 28*(1), 42-65.

Lv, D., & Wu, W. (2020). Margin trading and price efficiency: information content or price-adjustment speed?. *Accounting & Finance*, *60*(3), 2889-2918.

Lyandres, E., Sun, L., & Zhang L. (2008). The new issues puzzle: Testing the investment-based explanation. *The Review of Financial Studies, 21*, 2825-2855.

Ma, Y. R., Liu, Z., Ma, D., Zhai, P., Guo, K., Zhang, D., & Ji, Q. (2023). A news-based climate policy uncertainty index for China. *Scientific Data, 10,* 881.

Menezes, C., Geiss, C., & Tressler, J. (1980). Increasing downside risk. *The American Economic Review*, *70*(5), 921-932.

Michelon, G., Pilonato, S., & Ricceri, F. (2015). CSR reporting practices and the quality of disclosure: An empirical analysis. *Critical Perspectives on Accounting, 33*, 59-78.

Miller, E. M. (1977). Risk, uncertainty, and divergence of opinion. *Journal of Finance, 32*(4), 1151-1168. https://doi.org/10.1111/j.1540-6261.1977.tb03317.x

Mugaloglu, Y. I., & Erdag, E. (2013). Corporate governance, transparency and stock return volatility: empirical evidence from the Istanbul Stock Exchange. *Journal of Applied Economics & Business Research, 3*(4), 207-221.

Muravyev, D., Pearson, N. D., & Pollet, J. M. (2025). Anomalies and their short-sale costs. *The Journal of Finance, 80*(6), 3639-3694.

Narayan, P. K., & Zheng, X. (2010). Market liquidity risk factor and financial market anomalies: Evidence from the Chinese stock market. *Pacific-Basin Finance Journal*, *18*(5), 509-520.

Ng, L., & Wu, F. (2007). The trading behavior of institutions and individuals in Chinese equity markets. *Journal of Banking & Finance*, *31*(9), 2695-2710.

Novy-Marx, R. (2013). The other side of value: The gross profitability premium. *Journal of Financial Economics*, *108*(1), 1-28.

Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research, 18*(1), 109-131.

Pan, L., Tang, Y., & Xu, J. (2016). Speculative trading and stock returns. *Review of Finance, 20*, 1835-1865.

Pastor, L., & Veronesi, P. (2003). Stock prices and IPO waves. *NBER Working Paper 9858.* Available at https://www.nber.org/system/files/working_papers/w9858/w9858.pdf

Paule-Vianez, J., Lobão, J., Gómez-Martínez, R., & Prado-Román, C. (2021). Momentum strategies in times of economic policy uncertainty. *Journal of Financial Economic Policy*, *13*(3), 285-300.

Petkova, R. (2006). Do the Fama–French factors proxy for innovations in predictive variables? *The Journal of Finance*, *61*(2), 581-612.

Pontiff, J. (2006). Costly arbitrage: Evidence from closed-end funds. *The Quarterly Journal of Economics, 121*(4), 1361-1395.

Pontiff, J., & Woodgate, A. (2008). Share issuance and cross-sectional returns. *Journal of Finance, 63*, 921-945.

Reinganum, M. R. (1981). Misspecification of capital asset pricing: Empirical anomalies based on earnings' yields and market values. *Journal of Financial Economics*, *9*(1), 19-46.

Renewable Energy Law of the People's Republic of China. (2005). Available at: http://www.npc.gov.cn

Richardson, S. A., Sloan, R. G., Soliman, M. T., & Tuna, I. (2005). Accrual reliability, earnings persistence and stock prices. *Journal of Accounting and Economics*, *39*(3), 437-485.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41-55.

Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory, 13*(3), 341-360.

Saffi, P. A. C., & Sigurdsson, K. (2011). Price efficiency and short selling. *Review of Financial Studies, 24*(3), 821-852. https://doi.org/10.1093/rfs/hhp095

Sautner, Z., Van Lent, L., Vilkov, G., & Zhang, R. (2023). Firm-level climate change exposure. *The Journal of Finance*, *78*(3), 1449-1498.

Schwert, G. W. (2003). Anomalies and market efficiency. *Handbook of the Economics of Finance*, *1*, 939-974.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, *19*(3), 425-442.

Shenzhen Stock Exchange. (2020). *Measures for the assessment of information disclosure work of listed companies of Shenzhen Stock Exchange.*

Shleifer, A., & Vishny, R. W. (1997). The limits of arbitrage. *Journal of Finance, 52*(1), 35-55. https://doi.org/10.1111/j.1540-6261.1997.tb03807.x

Sloan, R. G. (1996). Do stock prices fully reflect information in accruals and cash flows about future earnings? *Accounting Review*, *71*(3), 289-315.

Stambaugh, R. F., Yu, J., & Yuan, Y. (2012). The short of it: Investor sentiment and anomalies. *Journal of Financial Economics, 104*(2), 288-302. https://doi.org/10.1016/j.jfineco.2011.12.001

Task Force on Climate-related Financial Disclosures (TCFD). (2017). *Final Report: Recommendations of the Task Force on Climate-related Financial Disclosures.* Available at: https://www.fsb-tcfd.org/publications/

Tian, G. G., Wan, G., & Guo, M. (2018b). Investor sentiment and stock returns: Evidence from China's A-share market. *Emerging Markets Review, 34,* 55-69.

Tian, G., Yu, J., & Guo, J. (2018a). Investor sentiment and stock returns: Evidence from China. *Pacific-Basin Finance Journal, 51*, 68-87.

Titman, S., Wei, K. J., & Xie, F. (2004). Capital investments and stock returns. *Journal of Financial and Quantitative Analysis*, *39*(4), 677-700.

Titman, S., Wei, K. J., & Xie, F. (2013). Market development and the asset growth effect: International evidence. *Journal of Financial and Quantitative Analysis*, *48*(5), 1405-1432.

Titman, S., Wei, K.C.J., & Xie, F. (2004). Capital investments and stock returns. *Journal of Financial and Quantitative Analysis, 39,* 677-700.

Wan, X., 2024. Margin-buying, short-selling, and stock valuation: Why is the effect reversed over time in China. *Journal of Empirical Finance*, 76, 101476.

doi.org/10.1016/j.jempfin.2024.101476

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data* (2nd ed.). MIT Press.

Xie, L., Zhao, L., & Yao, Y. (2021a). The effect of carbon disclosure on stock liquidity: Evidence from China's emission trading scheme. *Journal of Corporate Finance, 66,* 101881.

Xie, Z., Wang, Q., & Wu, Z. (2021b). The impact of China's national carbon emissions trading scheme on corporate environmental performance. *Environmental Science and Pollution Research, 28*(32), 43657-43671.

Xing, Y. (2008). Interpreting the value effect through the Q-theory: An empirical investigation. *The Review of Financial Studies*, *21*(4), 1767-1795.

Xu, F., & Wan, D. (2015). The impacts of institutional and individual investors on the price discovery in stock index futures market: Evidence from China. *Finance Research Letters*, *15*, 221-231.

Yang, S., & Zhao, J. (2025). Disclosure quality ratings by the stock exchanges: evidence from China. *China Accounting and Finance Review*, *27*(3), 340-363.

Yao, J., Ma, C., & He, W. P. (2014). Investor herding behaviour of Chinese stock market. *International Review of Economics & Finance, 29*, 12-29.

Ying, Q., Kong, D., & Luo, D. (2015). Investor attention, institutional ownership, and stock return: Empirical evidence from China. *Emerging Markets Finance and Trade*, *51*(3), 672-685.

Yonezawa, Y., & Hin, T. K. (1992). An empirical test of the CAPM on the stocks listed on the Tokyo Stock Exchange. *Japan and the World Economy*, *4*(2), 145-161.

Yu, J., Li, D., & Wang, S. (2019a). Retail investors and market anomalies: Evidence from the Chinese stock market. *Emerging Markets Review, 39*, 1-15.

Yu, J., Sun, S., & Zheng, B. (2019b). The impact of retail and institutional investors on stock prices: Evidence from China's A-share market. *China Economic Review, 55,* 1-16.

Zhang, L. (2005). The value premium. *The Journal of Finance*, *60*(1), 67-103.

Zheng, D., Li, H., & Zhu, X. (2015). Herding behavior in institutional investors: Evidence from China's stock market. *Journal of Multinational Financial Management*, *32*, 59-76.

LIST OF ABBREVIATIONS

Abbreviation 1

Abbreviation	**Full term / definition**
ACC	Accruals (accounting accruals measure)
AG	Asset growth
AGE	Firm age (years since IPO)
APT	Arbitrage Pricing Theory
BERT	Bidirectional Encoder Representations from Transformers
BM	Book-to-market ratio
BS	Balance Sheet
CAPM	Capital Asset Pricing Model
CEI	Composite equity issuance
CF	Cash Flow Statement
CMA	Conservative-minus-Aggressive factor (investment factor)
CNINFO	China Information Network (CSRC-designated disclosure platform "cninfo")
COMB	Combination anomaly (composite anomaly signal)
CPU	Climate Policy Uncertainty
CRD	Climate Risk Disclosure
CRSP	Centre for Research in Security Prices
CSMAR	China Stock Market and Accounting Research database
CSR	Corporate Social Responsibility
CSRC	China Securities Regulatory Commission
DE	Leverage ratio (debt-to-equity or similar leverage measure used in the thesis)
DiD	Difference-in-Differences
DSR	Downside Risk (downside risk measure)
EMH	Efficient Market Hypothesis

EPU	Economic Policy Uncertainty
ESG	Environmental, Social, and Governance
ETF	Exchange-Traded Fund
ETS	Emissions Trading Scheme
FCPU	Climate Policy Uncertainty factor (factor return based on CPU)
FCRD	Climate Risk Disclosure factor (factor return based on CRD)
FEPU	Economic Policy Uncertainty factor (factor return based on EPU)
FM	Fama–MacBeth two-stage regressions
FSTD	Forecast standard deviation (analyst forecast dispersion)
GP	Gross profitability
GPT	Generative Pre-trained Transformer
HBL	High-minus-Low book-to-market portfolio (value portfolio difference)
HML	High-minus-Low book-to-market factor (value factor)
IPO	Initial Public Offering
IS	Income Statement
ISS	Net equity issuance
ISSUE	Equity issuance (share issuance measure)
IVA	Investment-to-Assets ratio
IVOL	Idiosyncratic volatility
JSON	JavaScript Object Notation
LBM	Low book-to-market portfolio
LDA	Latent Dirichlet Allocation
LIQ	Liquidity (stock liquidity measure)
LOGBM	Log book-to-market ratio
LOGME	Log market equity (log market capitalization)
ME	Market equity (market capitalization)
MKT	Market factor (excess market return)
MOM	Momentum (stock return momentum)
NI	Net income

NLP	Natural Language Processing
NOA	Net Operating Assets
OLS	Ordinary Least Squares
OPAC	Opacity measure constructed from financial statements (OPAC index)
OSC	O-Score (Ohlson’s distress score)
PPE	Property, Plant, and Equipment
PPT	PowerPoint
RET	Stock return (monthly stock return)
RMB	Renminbi (Chinese yuan)
RMW	Robust-minus-Weak profitability factor
ROA	Return on Assets
SAS	Statistical Analysis System (SAS software)
SMB	Small-minus-Big size factor
SME	Small and Medium-sized Enterprise
SR	Sharpe ratio
SSD	Short-Selling Depth (short-selling intensity measure)
SSE	Shanghai Stock Exchange (Shanghai A-share market)
SSRN	Social Science Research Network
SZSE	Shenzhen Stock Exchange
TCFD	Task Force on Climate-related Financial Disclosures
UK	United Kingdom
US	United States
VAT	Value-Added Tax

APPENDIX

Appendix A. Construction of Anomaly Characteristics

Appendix A. 1. 1. General data description

We construct anomaly characteristics by using annual accounting variables, which are obtained from CSMAR financial statement databases for Chinese A-share firms. Balance sheet (BS), income statement (IS), and cash flow (CF) information are sourced from financial statements modules, which treat December as the end of fiscal year in the Chinese accounting report framework. For each anomaly, we require the relevant accounting information to be available in the current year t and, where needed, its lagged values in earlier years, so that anomaly variables are well constructed at the end of current fiscal year t. Trading information, such as monthly stock closing prices and market value of shares, are also drawn from CSMAR. The monthly trading data allows us to form monthly anomaly/factor returns[9] via annually updated firm characteristics/anomaly variables, which are used to rank stocks.

Appendix A. 1. 2. Accounting data

Firm-level accounting variables are obtained from CSMAR's balance sheet (BS), income statement (IS), and cash-flow statement (CF) databases. All three statements report either consolidated or parent financials identified by the statement-type code Typrep; we restrict the sample to A-share reports (Typrep = "A"). The gathered accounting items are used to construct gross profitability, asset growth, investment-to-assets, return on assets, net operating assets, and accruals, the construction process of which is shown in detail in Appendix A.2.1.-A.2.6.

From the **balance sheet (BS)**, we use the following items (CSMAR English labels in brackets, with brief definitions based on the CSMAR documentation):

- **Total assets** (Total Assets, item A001000000): the sum of all asset items.

[9] For the details of the factor/portfolio construction progress, please refers to Section 4.3.3.1.

- **Cash and cash equivalents** (Cash and Cash Equivalents, A001101000): cash on hand, bank deposits, and other immediately available monetary funds.
- **Short-term investments** (Net Short-term Investments, A001109000): short-term marketable securities net of impairment, intended to be held for less than one year.
- **Net inventories** (Net Inventories, A001123000): inventories net of inventory depreciation reserves, including goods for sale and materials for production.
- **Total current assets** (Total Current Assets, A001100000): the sum of all current asset items.
- **Total current liabilities** (Total Current Liabilities, A002100000): the sum of all current liability items.
- **Short-term loans** (Short-term Loans, A002101000): interest-bearing loans with original maturity of one year or less.
- **Taxes payable** (Taxes Payable, A002113000): various taxes due but not yet paid, such as VAT, income tax, and local taxes.
- **Total liabilities** (Total Liabilities, A002000000): the sum of all liability items.
- **Preference shares** (Including: Preference Share, A003112101): equity securities with preferential rights over ordinary shares in profit and residual-asset distribution.
- **Minority interests (equity)** (Minority Interests, A003200000): the portion of equity attributable to minority shareholders of subsidiaries.
- **Total shareholders' equity** (Total Shareholders' Equity, A003000000): the sum of all equity items attributable to shareholders.
- **Net fixed assets** (Net Fixed Assets, A001212000): fixed assets net of accumulated depreciation and impairment, including plant, equipment, and related facilities.

From the **income statement (IS)**, we use the following items defined in the CSMAR income statement module (CSMAR English label in brackets, with item code in parentheses):

- **Operating revenue** (Operating Revenue, B001101000): revenue from the company's main business operations.
- **Operating costs** (Operating Costs, B001201000): cost of goods and services corresponding to operating revenue.
- **Operating profit** (Operating Profit, B001300000): profit from continuing operations before non-operating items.
- **Net profit** (Net Profit, B002000000): profit after taxes and all expenses attributable to shareholders.

From the **cash-flow statement**, we use:

- **Depreciation and amortisation** (Depreciation & Amortization, F061201B): depreciation of fixed, oil and gas, and bearer biological assets plus amortisation of intangible assets and long-term deferred expenses.

Appendix A. 1. 3. Stock prices, returns and market value

Stock price and trading information are taken from CSMAR's **"Monthly Stock Price – Returns"** file for A-share listings on the Shanghai and Shenzhen Stock Exchanges. These files provide one data observation for each stock on per calendar month[10]. Trading date and closing price are abbreviated as "Trdmnt" and "Clsdt", respectively. We also obtain annual trading-relevant variables. For the construction of Net Equity Issuance, Composite Equity Issuance, and Momentum, we extract year-end (December) observations from the monthly stock file and treat them as annual inputs, as described in Appendix A.2.7, A.2.8, and A.2.11. We use the following variables from this file:

- **Monthly closing price** (Mclsprc): the stock's closing price on the last trading day of the month.

[10] Please note that we obtain two versions of trading relevant variables. Month trading data, for example, uses the stock price to construct anomaly portfolios. Yearly trading data facilitates the construction of such anomalies as Net Equity Issuance, Composite Equity Issuance, and Momentum.

- **Total market value** (Msmvttl): total number of shares multiplied by the monthly closing price, reported in thousand RMB for A shares. We multiply Msmvttl by 1,000 to obtain market value in RMB.
- **Number of shares traded in the month** (Mnshrtrd): total monthly trading volume in shares, used for sample diagnostics but not directly in anomaly construction.
- **Market type** (Markettype): exchange and board identifier (e.g., SSE A share, SZSE A share, ChiNext, STAR Market), used to restrict the sample to A-share listings.

Appendix A. 1. 4. Risk-free rate and distress measures

The **risk-free rate** is taken from the module of benchmark interest-rate in CSMAR, which reports annualised deposit and government-bond rates together with derived daily and monthly rates. We use the **monthly risk-free rate** (Nrrmtdt), defined by CSMAR as the monthly rate converted from the annualised benchmark rate under compound interest, and match this series to our monthly return data when computing excess returns. These datasets are used to construct the O-score and downside-risk anomalies described in Appendix A.2.9 and A.2.10. Two distress-related variables are obtained from dedicated CSMAR financial-distress datasets:

- **O-score (OSC)**: CSMAR's O-score file provides the Ohlson (1980) bankruptcy-risk index for each firm at the fiscal year-end. The index is constructed from a set of accounting ratios (size, leverage, profitability, liquidity, and performance trends) using the original logit model coefficients. Higher values of OSC indicate a greater predicted probability of financial distress, and vice versa. We take the reported O-score for each firm–year and align it with our anomaly panel by stock code and year.
- **Downside risk (DSR)**: CSMAR's downside-risk file reports a volatility measure based on return distributions. Following Ang et al. (2006), downside risk is defined as the standard deviation of monthly returns that are below the firm's mean return over a rolling 36-month window. It captures firm-level exposure to adverse return outcomes

and is commonly interpreted as a distress-related risk indicator. A higher DSR therefore implies a greater likelihood that a firm experiences unfavourable or downturn market conditions. The dataset provides the resulting downside-volatility series at the fiscal year-end; we take the December value for each firm–year and merge it into the anomaly panel.

Appendix A. 2. 1. Gross profitability (GP)

Gross profitability (GP) follows Novy-Marx (2013) and is measured as gross profit scaled by total assets. Gross profit is obtained from CSMAR's income statement extension as operating revenue minus operating costs, reflecting the firm's core operating profitability before financing, tax, and non-operating components. For each A-share firm with a December fiscal year-end, we compute annual gross profit at the fiscal year-end and match it with total assets from the balance sheet for the same date. Gross profitability is then defined as

$$GP_{i,t} = \frac{\mathrm{GrossProfit}_{i,t}}{TA_{i,t}},$$

where $TA_{i,t}$ denotes total assets at the end of year t. This measure captures how efficiently the firm's asset base generates gross operating income in the cross-section.

Appendix A. 2. 2. Asset growth (AG)

Asset growth (AG) follows Cooper et al. (2008) and is defined as the year-over-year percentage change in total assets. Using the balance sheet data, we take total assets for each firm in December of year t and construct its lagged value from December of year $t-1$. Ordering observations by firm and year ensures that lagged quantities are calculated within the same firm. Asset growth is calculated as:

$$AG_{i,t} = \frac{TA_{i,t} - TA_{i,t-1}}{TA_{i,t-1}},$$

and is only defined when both current and lagged total assets are positive and observed. This variable captures broad expansion in the firm's balance sheet, including growth driven by investment, acquisitions, and asset revaluations, and is measured at an annual frequency using year-end balance sheet data.

Appendix A. 2. 3. Investment-to-assets (IVA)

The investment-to-assets ratio (IVA) follows Titman et al. (2004) and measures tangible investment intensity relative to firm size. In line with their definition, we focus on changes in gross property, plant and equipment, and inventories, which correspond to long-lived capital investment and working capital tied up in production. From the balance sheet, we extract net fixed assets and net inventories for each firm at the December fiscal year-end and compute their firm-specific annual changes. For each firm–year, we sum the change in fixed assets and the change in inventory, and scale this by lagged total assets:

$$IVA_{i,t} = \frac{\Delta\text{FixedAssets}_{i,t} + \Delta\text{Inventory}_{i,t}}{TA_{i,t-1}}.$$

This construction closely matches the investment-to-assets measure in Titman et al. (2004), while implementing it using the fixed-asset and inventory fields available in CSMAR. It yields an annual indicator of how aggressively the firm expands its tangible asset base relative to its size.

Appendix A. 2. 4. Return on assets (ROA)

Return on assets (ROA) follows Chen et al. (2011) and is measured as net profit scaled by total assets. Using the income-statement and balance-sheet data described above, we take net profit (Net Profit) from the income statement and total assets (Total Assets) from the balance sheet at the December fiscal year-end and align them by firm and year. ROA is defined as:

$$ROA_{i,t} = \frac{\text{NetProfit}_{i,t}}{\text{TA}_{i,t}}$$

where $\text{TA}_{i,t}$ denotes total assets at the end of year t. This measure captures overall profitability after operating, financing and non-operating items, normalised by the firm's asset base.

Appendix A. 2. 5. Net operating assets (NOA)

Net operating assets (NOA) follow Hirshleifer et al. (2004) and measure the portion of operating assets financed by non-equity, non-interest-bearing liabilities. To adapt this construction to CSMAR's Chinese reporting format, we classify operating and financial components using the balance-sheet fields defined earlier.

Operating assets are defined as total assets minus financial assets, where financial assets consist of cash and cash equivalents, and short-term investments:

$$OPAssets_{i,t} = TA_{i,t} - Cash_{i,t} - ShortTermInvestment_{i,t}$$

Operating liabilities are defined by removing explicit financing items from total assets. Specifically, we subtract short-term loans, long-term loans, minority interests, and total shareholders' equity:

$$OPLiab_{i,t} = TA_{i,t} - ShortTermLoan_{i,t} - LongTermLoan_{i,t} - MinorityInterests_{i,t} - TE_{i,t}$$

Net operating assets are then obtained as operating assets minus operating liabilities, scaled by lagged total assets:

$$NOA_{i,t} = \frac{OPAssets_{i,t} - OPLiab_{i,t}}{TA_{i,t-1}}$$

This construction preserves the economic interpretation in Hirshleifer et al. (2004) while mapping each component directly to CSMAR accounting fields, ensuring consistency with the Chinese balance-sheet structure.

Appendix A. 2. 6. Accruals (ACC)

Accruals (ACC) follow Sloan (1996) and are constructed using a balance-sheet-based measure of working-capital accruals scaled by lagged assets. From the balance sheet, we obtain current assets, cash, short-term investments, current liabilities, short-term loans, tax payable, and total assets at each December fiscal year-end. Non-cash current assets are defined as current assets minus cash and short-term investments, and non-debt current liabilities are defined as current liabilities minus short-term loans and tax payable. For each firm–year, we compute the annual change in non-cash current assets, $\Delta NCA_{i,t}$, and the annual change in non-debt current liabilities, $\Delta NDCL_{i,t}$, between December of year $t-1$ and December of year t. Accruals are then defined as:

$$ACC_{i,t} = \frac{\Delta NCA_{i,t} - \Delta NDCL_{i,t}}{TA_{i,t-1}}.$$

This measure focuses on the extent to which earnings are driven by non-cash adjustments rather than cash flows. In line with the data structure in CSMAR, we rely on balance sheet changes and lagged total assets rather than subtracting depreciation explicitly from cash flows.

Appendix A. 2. 7. Net equity issuance (ISS)

Net equity issuance (ISS) follows Pontiff and Woodgate (2008) and is based on changes in shares outstanding inferred from market value. Using the same monthly stock price file as described above, we retain, for each firm–year, the December observation from the "Monthly Stock Price – Returns" file. We take total market value (Msmvttl) as market capitalisation, multiply it by 1,000 to convert from thousands of RMB to RMB, and infer the implied number of shares outstanding as market capitalisation divided by the corresponding December closing price. Ordering observations by firm and year, we take the lagged share count from December of year $t-1$ and define net equity issuance as:

$$\text{ISS}_{i,t} = \log(\text{Shares}_{i,t}) - \log(\text{Shares}_{i,t-1}).$$

This quantity is positive when the firm experiences net share issuance, and negative when repurchases dominate. Consistent with the implementation in Stambaugh et al. (2012), the measure is based on changes in shares outstanding as reflected in the market data and does not separately adjust for stock splits or similar corporate actions.

Appendix A. 2. 8. Composite equity issuance (CEI)

Composite equity issuance (CEI) follows Daniel and Titman (2006) and captures the component of market equity growth that is not explained by stock returns. Using the same December stock observations as for ISS, we obtain a time series of market capitalisation and associated stock returns from CSMAR's "Monthly Stock Price – Returns" file. For each firm–year, we retain the December market value and the corresponding simple return, and construct a multi-year market equity growth term together with a return term over a fixed window.

For each firm i, let $\mathrm{MV}_{i,t}$ denote the market value in December of year t, and let $R_{i,t}$ denote the corresponding simple stock return. Composite equity issuance over a five-year window is defined as:

$$\mathrm{CEI}_{i,t} = [\log(\frac{\mathrm{MV}_{i,t}}{\mathrm{MV}_{i,t-5}})] - (R_{i,t} - R_{i,t-5}),$$

where the five-year horizon ($k = 5$) follows Daniel and Titman (2006) and matches the implementation used in Essay Two. The first term measures cumulative growth in market equity over the window, while the second term removes the part attributable to stock returns. The residual component reflects net equity issuance (seasoned issues, share-based acquisitions, repurchases, and related actions) and provides a market-based proxy for financing activity that is consistent with the CEI anomaly in Daniel and Titman (2006), adapted to the return and market value series available in CSMAR.

Appendix A. 2. 9. O-score (OSC)

The O-score (OSC) is a bankruptcy risk measure based on Ohlson's (1980) logit model. In the original formulation, the O-score is constructed from nine accounting ratios summarising firm size, leverage, profitability, liquidity, and performance trends, such as the log of total assets, total liabilities to total assets, working capital to total assets, and indicators for recent losses. We obtain an O-score index for Chinese A-share firms from CSMAR's risk and financial distress modules, which implement the Ohlson model using CSMAR financial statement data. The resulting O-score is reported at the fiscal year-end for each firm; we retain the December observations and merge them with the anomaly panel by firm and year. Higher values of OSC indicate a greater predicted probability of financial distress. This implementation preserves the economic interpretation of the O-score anomaly used in Ohlson's (1980), while relying on the accounting inputs and estimation provided in the Chinese data environment.

Appendix A. 2. 10. Downside risk (DSR)

Downside risk (DSR) is used in our setting as a distress-related anomaly in place of the failure-probability measure employed in Stambaugh et al. (2012). In the U.S. literature, one component of the distress dimension is captured by the failure-probability measure of Campbell et al. (2008), which combines market-based and accounting information to estimate the likelihood of corporate default. The Campbell failure-probability series is not available in CSMAR for Chinese A-share firms. To preserve the conceptual role of a distress proxy, we adopt a downside volatility measure following Ang et al. (2006), who define downside risk as the volatility of return realisations below the conditional mean.

The downside risk indicator is obtained directly from the CSMAR financial markets database, where it is computed from firm-level stock returns over a rolling 36-month window. For each firm and window, only return observations below the corresponding mean return are included, and the volatility of these negative observations is computed as:

$$DSR_{i,t} = \sqrt{\frac{1}{n_{i,t}^{-} - 1} \sum_{j \in N_{i,t}^{-}} (R_{i,j} - \bar{R}_{i,t})^2}\,,$$

where $R_{i,j}$ denotes the return for firm i in month j; $\bar{R}_{i,t}$ is the mean return over the 36-month window; $N_{i,t}^{-}$ is the set of months in the window with returns below the mean; and $n_{i,t}^{-}$ is the cardinality of this set. The resulting downside risk series is reported at an annual frequency, and we retain the December values so that DSR is observed once per year at the portfolio formation date.

This approach maintains consistency with the distress anomaly in the U.S. setting, where distress-related characteristics capture exposure to adverse tail outcomes rather than unconditional volatility, while remaining feasible for the Chinese market where Campbell-style failure probabilities are unavailable.

Appendix A. 2. 11. Momentum (MOM)

Momentum (MOM) follows Jegadeesh and Titman (1993) and captures medium-term return continuation using the return history available in CSMAR. Using the monthly stock price file described above, we compute simple monthly returns for each firm from end-of-month closing prices. At the December portfolio-formation date of year t, we define the momentum characteristic as the cumulative return over the last year, excluding the most recent month:

$$\mathrm{MOM}_{i,t} = \prod_{j=t-12}^{t-2} (1 + R_{i,j}) - 1,$$

where $R_{i,j}$ denotes the monthly return on stock i in month j. This construction matches the implementation in Essay Two and Table 3.2, and is consistent with the standard momentum definition in the anomaly literature.

Appendix B. 1. Additional Results

Table B.1. Pre-Treatment Parallel-Trends Tests across Short-Selling Expansion Waves

Anomaly	No. of Waves	Significant at 10%	Significant at 5%	Significant at 1%	Mean p-value
AG	8	3	1	0	0.3418
Accruals	8	3	2	1	0.3007
CEI	8	0	0	0	0.5045
DSR	8	1	1	1	0.3565
Gprofit	8	1	1	1	0.5159
IVA	8	1	1	1	0.5247
Issuance	8	2	2	0	0.3144
MoM	8	0	0	0	0.5104
NOA	8	2	1	1	0.3738
Oscore	8	0	0	0	0.6059
ROA	8	1	1	0	0.3883

Note This table summarises the pre-treatment parallel-trends tests for the eight short-selling expansion waves and eleven anomaly portfolios. For each anomaly and expansion wave, the test is estimated over the 12 months preceding the corresponding expansion. The coefficient of interest is the interaction between the treatment indicator and the event-time trend. A statistically insignificant interaction coefficient indicates no evidence of a differential linear pre-treatment trend between the treated and control portfolios. All regressions control for the Fama–French five factors.

Appendix B. 2. Additional Results

Table B. 2. Anomaly Returns by Short-Selling Expansion Wave: Shortable vs Non-Shortable Stocks (Newly Eligible)

Panel A: Initial issuance (Shortable obs. = 89, Non-shortable obs. = 2121)

	GP	AG	IVA	ROA	NOA	ACC	ISS	CEI	OSC	DSR	MOM
Shortable mean	0.012	0.011	-	-0.070	0.003	0.029	0.008	-0.059	0.050	-0.010	-0.012
Non-shortable mean	0.020	-0.031	-	-0.022	0.012	-0.001	-0.013	-0.006	0.027	-0.008	0.003
Difference mean	-0.008	0.041	-	-0.048	-0.009	0.030	0.021	-0.052*	0.023	-0.001	-0.014
t-value	[-0.37]	[0.56]	-	[-0.73]	[-0.38]	[1.48]	[0.59]	[-1.71]	[1.24]	[-0.05]	[-0.47]

Panel B: Expand 1 (Shortable obs. = 184, Non-shortable obs. = 2379)

	GP	AG	IVA	ROA	NOA	ACC	ISS	CEI	OSC	DSR	MOM
Shortable mean	-0.016	-0.011	0.008	-0.015	-0.005	-0.002	-0.001	-0.003	-0.001	-0.014	-0.003
Non-shortable mean	-0.008	-0.018	0.001	-0.002	-0.009	0.002	-0.015	-0.002	0.015	-0.010	0.004
Difference mean	-0.008	0.007	0.007	-0.013	0.003	-0.003	0.014	-0.001	-0.016	-0.004	-0.007
t-value	[-0.33]	[0.24]	[0.58]	[-0.71]	[0.12]	[-0.23]	[0.92]	[-0.10]	[-0.55]	[-0.21]	[-0.54]

Panel C: Expand 2 (Shortable obs. = 274, Non-shortable obs. = 2347)

	GP	AG	IVA	ROA	NOA	ACC	ISS	CEI	OSC	DSR	MOM
Shortable mean	-0.008	0.014	-0.011	0.010	0.011	0.018	-0.031	0.001	-0.017	0.037	0.001
Non-shortable mean	0.001	-0.019	-0.002	-0.016	0.010	-0.006	-0.018	0.011	0.014	0.024	0.013
Difference mean	-0.009	0.033***	-0.008	0.026**	0.001	0.023	-0.013	-0.010	-0.031***	0.013	-0.012
t-value	[-0.36]	[2.68]	[-0.68]	[2.11]	[0.05]	[1.21]	[-0.78]	[-0.91]	[-2.60]	[1.09]	[-0.81]

Table B.2. (continued)

Panel D: Expand 3	(Shortable obs. = 186, Non-shortable obs. = 2468)										
	GP	AG	IVA	ROA	NOA	ACC	ISS	CEI	OSC	DSR	MOM
Shortable mean	-0.010	-0.015	-0.022	0.003	-0.024	0.008	-0.009	-0.007	-0.025	0.005	-0.009
Non-shortable mean	0.011	-0.024	-0.018	-0.030	0.011	-0.003	-0.017	-0.009	0.038	0.018	-0.012
Difference mean	-0.021	0.009	-0.004	0.033	-0.035**	0.011	0.008	0.002	-0.063**	-0.013	0.003
t-value	[-1.17]	[0.32]	[-0.16]	[1.34]	[-2.03]	[0.43]	[0.44]	[0.08]	[-2.31]	[-0.82]	[0.16]

Panel E: Expand 4	(Shortable obs. = 200, Non-shortable obs. = 2686)										
	GP	AG	IVA	ROA	NOA	ACC	ISS	CEI	OSC	DSR	MOM
Shortable mean	-0.018	-0.028	-0.036	-0.046	0.011	-0.012	0.017	-0.044	0.032	0.002	-0.034
Non-shortable mean	-0.015	-0.021	-0.013	-0.008	-0.020	0.023	-0.008	-0.018	0.012	0.005	-0.036
Difference mean	-0.004	-0.007	-0.023	-0.038	0.031	-0.035	0.026	-0.026	0.021	-0.003	0.002
t-value	[-0.11]	[-0.18]	[-0.42]	[-1.17]	[0.66]	[-1.25]	[0.93]	[-1.24]	[0.65]	[-0.16]	[0.07]

Panel F: Expand 5	(Shortable obs. = 143, Non-shortable obs. = 3421)										
	GP	AG	IVA	ROA	NOA	ACC	ISS	CEI	OSC	DSR	MOM
Shortable mean	0.015	0.019	0.002	0.012	0.022	-0.007	0.012	-0.007	0.019	0.026	0.032
Non-shortable mean	-0.007	0.002	0.015	0.026	-0.027	-0.019	-0.008	-0.026	-0.031	-0.023	-0.002
Difference mean	0.023*	0.017	-0.013	-0.014*	0.049**	0.011	0.020*	0.019	0.050**	0.050*	0.034*
t-value	[1.84]	[1.36]	[-0.80]	[-1.68]	[2.18]	[0.44]	[1.93]	[1.29]	[2.13]	[1.84]	[1.67]

Table B.2. (continued)

Panel G: Expand 6	(Shortable obs. = 842, Non-shortable obs. = 3275)										
	GP	AG	IVA	ROA	NOA	ACC	ISS	CEI	OSC	DSR	MOM
Shortable mean	0.022	0.026	-0.002	0.036	0.003	-0.016	-0.018	0.012	-0.030	-0.016	0.001
Non-shortable mean	0.015	0.031	0.029	0.028	0.019	0.012	-0.010	0.004	-0.023	-0.012	0.002
Difference mean	0.007	-0.005	-0.031***	0.008	-0.016	-0.028**	-0.007	0.009	-0.008	-0.004	-0.001
t-value	[0.56]	[-0.43]	[-2.87]	[0.54]	[-0.63]	[-2.17]	[-0.82]	[0.47]	[-0.49]	[-0.31]	[-0.08]

Panel H: Expand 7	(Shortable obs. =1104, Non-shortable obs. = 4324)										
	GP	AG	IVA	ROA	NOA	ACC	ISS	CEI	OSC	DSR	MOM
Shortable mean	-0.034	-0.053	-0.007	-0.019	-0.016	-0.017	0.001	-0.037	0.004	0.002	-0.016
Non-shortable mean	-0.005	-0.022	-0.015	-0.002	-0.019	-0.010	-0.019	-0.015	-0.004	0.001	-0.004
Difference mean	-0.028	-0.032	0.008	-0.017	0.003	-0.007	0.021	-0.022*	0.009	0.001	-0.012
t-value	[-1.13]	[-1.17]	[0.95]	[-1.08]	[0.29]	[-0.30]	[1.43]	[-1.69]	[0.49]	[0.06]	[-0.99]

Note This table reports average monthly returns of eleven anomaly-based long–short portfolio strategies around short-selling expansions in the Chinese A-share market. The sample is organized by short-selling eligibility at each expansion wave. For each wave, the shortable group consists of stocks that become eligible for short selling for the first time in that specific expansion, while the non-shortable group includes stocks that remain ineligible for short selling during the same period. Anomaly portfolios are formed by sorting stocks into high and low groups based on each anomaly signal at the formation date. Monthly value-weighted returns are then computed for each portfolio over the 12-month period following the implementation of the short-selling policy. Long–short anomaly returns are calculated as the difference between the high- and low-ranked portfolios. The table reports average monthly long–short returns for shortable and non-shortable stocks, as well as their differences (Difference = Shortable − Non-shortable). T-statistics are based on the time-series variation of monthly portfolio returns within each expansion wave. The anomalies include gross profitability (GP), asset growth (AG), investment-to-assets (IVA), return on assets (ROA), net operating assets (NOA), accruals (ACC), net issuance (ISS), composite equity issuance (CEI), Ohlson O-score (OSC), downside risk (DSR), and momentum (MOM). Returns are expressed in monthly decimal form. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

In Table B.2, we examine how short-sale liberalization affects anomaly-related mispricing by comparing anomaly portfolio returns between shortable and non-shortable stocks across China's sequential short-selling expansions, following the framework of Chu, Hirshleifer, and Ma (2020). Newly eligible stocks are identified based on the first appearance of short-selling activity and classified as the shortable group in each expansion wave, while firms that remain ineligible during the same period serve as the non-shortable control group, with classification conducted separately for each wave. This table shows that differences in anomaly-based long–short returns between newly shortable and non-shortable stocks are generally small, statistically insignificant, and heterogeneous across expansion waves, indicating that short-sale liberalization does not produce a consistent or immediate reduction in anomaly profitability in expansion periods. This finding is consistent with Table 3.5.

Appendix C. Climate Risk Disclosure Dictionary

Appendix C.1. TCFD-Based Taxonomy of Climate-Related Keywords

Type	Dimension	Keyword Group	Word Counts	Example Keywords
Risk	Physical Risk	Acute Physical Risk	14	洪水（Flood)，台风（Typhoon)，地震（Earthquake)
		Chronic Physical Risk	16	气候变化（Climate Change)，干旱（Drought)，降水（Precipitation)
	Transition Risk	Policy and Legal Risk	32	碳中和（Carbon Neutrality)，碳排放（Carbon Emissions)，环境保护（Environmental Protection)
		Technology Risk	3	科学与技术（Science and Technology)，技术（Technology)，科技 (Technology Innovation)
		Market Risk	10	化石能源（Fossil Fuels)，能源（Energy)，煤炭（Coal)
Opportunity	—	Energy Source	14	风能（Wind Energy)，光伏（Photovoltaics)，氢能（Hydrogen Energy)
		Products and Services	7	电动汽车（Electric Vehicles)，新材料（New Materials)，绿色 (Green Development)
		Markets	5	碳市场（Carbon Market)，碳交易（Carbon Trading)，绿色金融 (Green Finance)
		Resource Efficiency	14	节能（Energy Conservation)，能源效率（Energy Efficiency)，回收利用（Recycling)
		Resilience	40	森林（Forest)，生态（Ecosystem)，防洪（Flood Prevention)

Notes The keyword dictionary is organised according to the hierarchical taxonomy of the Task Force on Climate-related Financial Disclosures (TCFD). Climate-related keywords are first classified into two overarching types (Risk and Opportunity). Risk is further divided into Physical Risk and Transition Risk, while Opportunity comprises five keyword groups reflecting climate-related opportunities. The number of keywords reported corresponds to the final dictionary used in constructing the CRD measure. Selected example keywords are provided for illustration, and the complete classification of all 155 keywords is presented in Appendix C.2.

Appendix C.2. Full TCFD-Based Climate-Related Keyword Dictionary

This appendix presents the complete climate-related keyword dictionary used to construct the Climate Risk Disclosure (CRD) measure. The keyword taxonomy follows the hierarchical framework of the Task Force on Climate-related Financial Disclosures (TCFD) and is adapted to the Chinese institutional context. Chinese keywords are presented together with their corresponding English translations. Official English translations are adopted for Chinese government policies, national programmes, and institutional terminology where available. For general climate- and environment-related terms, English translations follow the terminology adopted by the Intergovernmental Panel on Climate Change (IPCC) and other internationally recognised climate-related frameworks where applicable.

C.2.1. Risk

C.2.1.1 Physical Risk

C.2.1.1.1 Acute Physical Risk (14 Keywords)

Chinese Keyword	Official English Translation
洪水	Flood
内涝	Urban Waterlogging
自然灾害	Natural Disaster
低温	Low Temperature
冰冻	Freezing Conditions
台风	Typhoon
地质灾害	Geological Disaster
地震	Earthquake
山洪	Flash Flood
泥石流	Debris Flow
洪涝	Flooding
灾害	Disaster
雨雪	Rain and Snow
洪水和内涝	Flood and Waterlogging

C.2.1.1.2 Chronic Physical Risk (16 Keywords)

Chinese Keyword	Official English Translation
气温	Temperature
气候变化	Climate Change
气候	Climate
降水	Precipitation

干旱	Drought
气象	Meteorological Conditions
水土流失	Soil Erosion
自然	Natural Environment
海洋	Ocean
人畜饮水	Drinking Water for Humans and Livestock
土地	Land
生物多样性	Biodiversity
石漠化	Rocky Desertification
荒漠化	Desertification
水资源	Water Resources
大江大河大湖	Major Rivers and Lakes

C.2.1.2 Transition Risk

C.2.1.2.1 Policy and Legal Risk (32 Keywords)

Chinese Keyword	Official English Translation
雾霾	Haze
颗粒物	Particulate Matter
能耗	Energy Consumption
耗能	Energy Consumption
碳中和	Carbon Neutrality
碳峰值	Carbon Peak
减排	Emission Reduction
二氧化碳	Carbon Dioxide
碳减排	Carbon Emission Reduction
碳排放	Carbon Emissions
碳达峰	Carbon Peaking
温室气体	Greenhouse Gases
降碳	Carbon Reduction
污染预防	Pollution Prevention
废气	Waste Gas Emissions
废水	Wastewater
环保	Environmental Protection
环境保护	Environmental Protection
PM2.5	PM2.5
二氧化硫	Sulfur Dioxide
尾气	Vehicle Exhaust Emissions
排放	Emissions
排污	Pollutant Discharge

污染	Pollution
污水	Sewage
垃圾	Solid Waste
扬尘	Fugitive Dust
废弃物	Waste
臭氧	Ozone
减污	Pollution Reduction
环境	Environment
物耗	Material Consumption

C.2.1.2.2 Technology Risk (3 Keywords)

Chinese Keyword	Official English Translation
科学与技术	Science and Technology
技术	Technology
科技	Technological Innovation

C.2.1.2.3 Market Risk (10 Keywords)

Chinese Keyword	Official English Translation
化石能源	Fossil Fuels
天然气	Natural Gas
煤炭	Coal
石油	Petroleum
能源	Energy
西气东输	West-to-East Gas Transmission Project
西电东送	West-to-East Electricity Transmission Project
产业结构	Industrial Structure
能源资源结构	Energy Resource Structure
产能过剩	Industrial Overcapacity

C.2.2. Opportunity

C.2.2.1 Energy Source (14 Keywords)

Chinese Keyword	Official English Translation
新能源	New Energy
风能	Wind Energy
水能	Hydropower
核能	Nuclear Energy
太阳能	Solar Energy

光伏	Photovoltaics
氢能	Hydrogen Energy
能源储存	Energy Storage
可再生	Renewable Energy
核电	Nuclear Power
水电	Hydroelectric Power
生物质能	Biomass Energy
风电	Wind Power
电池	Battery Technology

C.2.2.2 Products and Services (7 Keywords)

Chinese Keyword	Official English Translation
电动汽车	Electric Vehicles
低碳	Low-Carbon Development
创新	Innovation
新型	Emerging Technologies
绿色	Green Development
研发	Research and Development (R&D)
新材料	Advanced Materials

C.2.2.3 Markets (5 Keywords)

Chinese Keyword	Official English Translation
碳市场	Carbon Market
绿色金融	Green Finance
绿色债务	Green Debt
碳交易	Carbon Trading
西部大开发	Western Development Strategy

C.2.2.4 Resource Efficiency (14 Keywords)

Chinese Keyword	Official English Translation
能效	Energy Efficiency
能源效率	Energy Efficiency
降耗	Reduction in Energy Consumption
清洁	Cleaner Production
节能	Energy Conservation
旱作节水	Water-Saving Dry Farming
高效	High Efficiency

优化	Optimization
回收利用	Recycling
循环	Circular Economy
资源	Resource Utilization
节材	Material Conservation
节水	Water Conservation
节约	Resource Conservation

C.2.2.5 Resilience (40 Keywords)

Chinese Keyword	Official English Translation
可持续发展	Sustainable Development
科学发展观	Scientific Outlook on Development
碳汇	Carbon Sink
森林	Forest
生态	Ecosystem
净土	Clean Land
休牧还草	Grazing Withdrawal and Grassland Restoration
天然林	Natural Forest
治沙	Desertification Control
碧水	Clean Water
绿水青山	Lucid Waters and Lush Mountains
绿化	Afforestation
草原	Grassland
蓝天	Blue Sky
还林	Reforestation
还湖	Lake Restoration
还草	Grassland Restoration
退耕	Returning Farmland to Nature
植树	Tree Planting
湿地	Wetlands
三峡工程	Three Gorges Project
南水北调	South-to-North Water Diversion Project
水利	Water Conservancy
水库	Reservoir
引水	Water Diversion
清淤	Sediment Removal
疏浚	Waterway Dredging
堤防	Flood Embankment
城市绿化	Urban Greening
调水	Water Transfer

减灾	Disaster Mitigation
抗旱	Drought Resistance
抗灾	Disaster Response
救灾	Disaster Relief
防汛	Flood Control
防沙	Sand Control
防洪	Flood Prevention
防灾	Disaster Prevention
基础设施	Infrastructure
和谐	Social Harmony