ESSAYS IN THE INFLUENCE OF TECHNOLOGY ON ECONOMICS AND FINANCE

By

Linxian Huang

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of Philosophy in Economics

Middle Tennessee State University

August, 2023

Dissertation Committees:

Keith Gamble, Chair

Anne Anderson

Wisarut Suwanprasert

Contents

1 Relationship Between Technology-Based Lending and Banking System - Beneficial or Detrimental?

Abstract

This paper investigates the impact technology-based lending on the profitability of traditional banks, employing the database from Cornelli et al. (2020). I confirm the detrimental effect of technology-based lending growth on bank profitability, with a 5.2% decrease in the ROAA ratio for each 1% increase in technology-based lending. These results are consistent after accounting for the instrumental variables regression to address endogeneity and alternative profitability indicators. Notably, BigTech lending has more substantial negative impact on bank profitability than FinTech lending, predominantly in developing countries. Technology-based lending negatively impacts bank profitability due to country characteristics rather than bank characteristics, rather than the characteristics of the bank itself. The additional negative shocks to banks are stronger when the banking sector is relatively competitive or credit accessibility is lower in the country where the bank is located. By further exploring the alternative channels, I observe that technology-based lending reduces banks' profitability by issuing more unsecured loans to maintain their original market share of lending services, reducing lending efficiency.

Keywords: Bank profitability, technology-based lending

JEL classification numbers: G21, G23

1.1 Introduction

Technology-based lending entered a rapid growth period after the 2007-2008 financial crisis, further promoting the efficiency and diversification of financial products. This literature explores the challenges the traditional banking industry faced during this wave. Theoretically and empirically, the relationship between technology-based lending and the traditional banking industry is multi-faceted. On the one hand, technology-based lending can take banks' market share of banks' lending services and become a competitive relationship, thus affecting banks' profitability (Buchak et al., 2018; Jagtiani and Lemieux, 2018; Thakor, 2020). On the other hand, technology-based lending can help banks expand their business and serve the no- or low-credit customers whom banks' services are uncovered, becoming a complementary relationship with traditional banks (Campanella et al., 2017). This paper considers the factors influencing the relationship between emerging technology-based lending and the traditional banking industry. More specifically, the relationship between the two can vary depending on the bank's specific characteristics or the country or financial market environment in which the bank operates.

My analysis is based on the database of Cornelli et al. (2020) and utilizes a cross-country perspective to analyze how technology-based lending growth affects banks' profitability worldwide. This database shows the total volume change of technology-based lending (sum of FinTech lending and BigTech lending) in 99 countries from 2013 to 2019. This database also addressed the previous technology-based lending data limitation. Most studies considering the relationship between technology-based lending and banks are limited to a particular country or region. Meanwhile, Cornelli et al. (2020) distinguish and discuss two main categories of technology-based lending: FinTech lending and BigTech lending, confirming the potential differences in the impact of these two types of lending on traditional banks.

In the context of the digital revolution in lending services, the literature explores how banks' performance has been affected by the explosive growth of technology-based lending, suggesting that it is gradually becoming a substantial challenge for banks' future business development (Bejar et al., 2022; Hughes et al., 2022; Murinde et al., 2022; Zhao et al., 2022). Based on previous studies, I estimate the impact of technology-based lending on banks' profitability, investigating the consistency with and without instrumental variable regression consideration. I also explore bank-specific and country-specific characteristics that determine the influence of technology-based lending on banks and apply robustness checks to prove the main results' consistency. After assessing the impact of technology-based lending on bank profitability, I investigate the underlying channels that induce such effects, such as decreasing the market share of bank lending services or increasing the default risk of existing loans. Meanwhile, based on the definitions of these two types of lending and banks, I discuss the differences in the relationship between FinTech lending and BigTech lending.

Using data from 4,200 banks in 92 countries, the results from the panel regression with both bank-specific and macroeconomic factors suggest a detrimental effect of technology-based lending growth on a bank's profitability. This effect is statistically significant and economically meaningful, with a 1% increase in technology-based lending leading to a 5.623 % decrease in the return on average asset (ROAA) ratio. This equates to an average $2,135 billion loss in the bank's net income. Bank profitability is more responsive to the total volume of technology-based lending rather than the growth rate. BigTech lending has dominated the impact of technology-based lending on bank profitability. This result is robust to a two-stage least square regression, applying the country's information technology infrastructure penetration and degree of financial services sophistication indicators as external shocks. As expected, the estimated relationship still returns a statistical and negative significant sign, demonstrating the consistency of the results. The detrimental effect is also consistent when considering other banks' profitability indicators as the dependent variable.

I then investigate the factors influencing the relationship between technology-based lending and traditional bank profitability. The findings indicate that the country-specific characteristics of where banks are located make them more vulnerable to additional negative shocks to bank profitability from the growth of technology-based lending than their operating scale characteristics. Banks in developing countries saw significantly negative results compared to developed countries with insignificant influence. I also find that banks in developing countries with a relatively competitive banking sector or a low credit accessibility environment are more sensitive to the growth of technology-based lending, receiving more negative shocks to their profitability. Regarding bank-specific characteristics, banks with fewer branches experience a four times greater negative shock from technology-based lending than banks with fewer branches. The expansion of technology-based lending does not affect small operating-scale banks.

After performing a robustness check to prove the consistency of the baseline model, I further consider the channel in how technology-based lending influences a bank's profitability. The growth of technology-based lending significantly affects the market share of a bank's lending services. In contrast, the increase in a bank's non-performing loans is the primary channel through which technology-based lending leads to decreased profitability. As a result of the impact of technology-based lending growth on bank profitability, banks in developed countries experienced roughly six times more severe negative shocks in the market share of lending services than banks in developing countries. In contrast, the growth of technology-based lending significantly affects the lending efficiency of banks in developing countries. At the same time, there is no significant effect on the lending efficiency of banks in developed countries. Such phenomenon is consistent when comparing the banks with or without small operating scale characteristics, or low credit accessibility environment and competitive market in which the bank operates.

I also further investigate the impact of FinTech and BigTech lending on bank profitability under different bank- and country-specific characteristics as interaction terms. I confirm that consistent with the results of the previous baseline model, BigTech lending still dominates the negative impact of technology-based lending on bank profitability after accounting for these interactions. However, in countries with low credit accessibility, FinTech lending has a higher negative effect on bank profitability than BigTech lending.

Overall, my empirical findings are consistent with previous studies that the growth of technology-based lending can significantly lower banks' profitability (Buchak et al., 2018; Jagtiani and Lemieux, 2018; Thakor, 2020), with two implications. First, for traditional banks, the growth of BigTech lending is more threatening to their business, while FinTech lending has less impact on banks because banks can replicate most of the FinTech lending services. Second, the environment in which financial services are offered and the development level of the countries where banks are located determines the relationship between technology-based lending and traditional banks. When a bank is in a developed country, a less competitive banking market, or a country with a higher degree of credit accessibility, the competition between banks and technology-based lending is less significant and becomes a complementary relationship. Bank profitability is unaffected by maintaining the quality of their original loans and controlling risk. However, when banks are in a developing country, the banking sector is competitive, or in a country where credit accessibility is low, the competition between banks and technology-based lending is vital. Banks will secure their market share by issuing additional loans of lower quality, thus making their lending efficiency lower and thus affecting their profitability.

This paper contributes to several strands of literature. First, it adds to the literature documenting the significant detrimental effect of technology-based lending growth on bank profitability. Current research holds different arguments about the potential impact of technology-based lending on banks. On the one hand, the emergence of financial innovations can hurt bank performance, as online lending and investment platforms cut their profitability (Buchak et al., 2018; Jagtiani and Lemieux, 2018; Thakor, 2020) . Also, financial innovation partially fills the financial service gap of traditional banks due to regulatory constraints (Buchak et al., 2018). Such potential competition could lead to a decrease in the credit market share of financial intermediaries, disrupting the primary activity of the banking system (Greenbaum et al., 2019). On the other hand, technology-based lending helps banks further expand their business (Campanella et al., 2017). It complements the banking system business rather than displaces it because of the limited liquidity technology-based lending lenders provided (Navaretti et al., 2018). While the literature focuses on identifying the relationship between the two, this paper examines the potential factors that contribute to the relationship, exploring the factors that trigger banks to reap additional shocks on the one hand and the factors in technology-based lending that cause banks' profitability to change on the other.

Second, this paper investigates the sensitivity of the relationship between technology-based lending and bank performance changes when considering the different countries and bank-specific characteristics. Previous literature considers credit volume in different countries and the disparity in the impact caused by different types of credit Cornelli et al. (2020); however, technology-based lending solutions can dominate the choices for people accessing financial services in developing countries, due to the lack of a traditional financial infrastructure or the scarcity of financial services due to the geographical characteristics of a specific country Thakor (2020). In this manner, technology-based lending can play different roles in countries with different development levels. Using cross-country data, focusing on the interaction of the bank's own operations scale and the characteristics of the country in which it is located, and employing panel regression, I use exogenous variables for regression to address the potential endogeneity and provide more convincing evidence for the relationship found, indicating that the degree of development of the country in which the bank is located has a determining relationship between the bank and technology-based lending.

Third, by distinguishing the difference of FinTech and BigTech lending, I extend the literature on the different impact of these two on bank performance. Prior research has primarily focused on FinTech lending, whereas its growth is not as sharp as it once was. The new format of technology-based lending, such as BigTech lending, is currently in the spotlight. For example, the BNPL service in various e-commerce platforms provides consumers with alternative lending options. Unlike the traditional banking industry that relies on depositors for lending, BNPL is based on balance sheet lending, with the funds coming from the e-commerce company's funding. Therefore, the profits earned go to the e-commerce firm. The existing literature has begun to assess the impact of FinTech lending on the banking system, but there are a few papers that categorize the different FinTech lending services and the types of companies that provide the services.

The remainder of this paper proceeds as follows. Section 1.2 provides the literature review and concludes the empirical hypothesis. Section 1.3 presents a detailed description of the data and research methodology. Section 1.4 reflect the baseline model result with potential endogeneity consideration, Section 1.5 discuss the factors that may influence the relationship between technology-based lending and bank's profitability. Section 1.6 provide a robustness check for baseline model, and Section 1.7.1 discuss the channel that determine the relationship between technology-based lending and bank profitability, and the different impact of FinTech and BigTech lending on banks. Section 1.8 is the conclusion of this paper.

1.2 Literature Review

1.2.1 Development of technology-based lending

In the aftermath of the 2007–2009 financial crisis, the traditional banking industry faced unprecedented challenges due to its damaged reputation, increasing regulations, and stagnant credit growth. During the same period, information technology has proliferated, and smartphones have become an integral part of people's daily lives, opening up a whole new way of thinking about the future of financial services. In this context, financial technology has emerged, intending to use new technologies to reduce operating costs and provide financial services that differ from those provided by banks. Data, computing, and interfaces are major elements of FinTech. The ideal way to see this is that many FinTech firms have consumer-friendly products that can be used on a mobile phone. In its 2018 U.S. FinTech Market Report, S&P Global categorizes FinTech activity into six categories: payments, digital lending, digital banking, digital investment, personal finance, and blockchain. For further discussion, this paper only focuses on digital lending based on information technology development, named technology-based lending.

Traditional financial intermediaries collect and reallocate funds, including banks and venture capitalists. Individual account holders cannot control how bank deposits are allocated to loans and other financing activities. At the same time, market outcomes are subject to information asymmetries (Stiglitz, 1975; Stiglitz and Weiss, 1981), agency problems (Coase, 1995; Jensen and Meckling, 1976; Fama, 1980), and the distribution of contracts through the of residual control (Grossman and Hart, 1986; Hart and Moore, 1990), which make it harder for businesses to obtain financing (La Porta et al., 1997). The rapid development of cellular technology primarily drives the growth of technology-based lending, which addresses the relative inefficiency of traditional financing channels and financial services. Traditional financial intermediaries improve investment efficiency by carefully screening and allocating returns efficiently while emphasizing their screening and detection strengths through superior information-gathering and processing capabilities (Diamond, 1984; Berger et al., 2005; Boot and Thakor, 2000). While technology-based lending generates credit scores for small businesses and consumers using more sophisticated algorithms, the original process is sped up, and the cost of loan evaluation is reduced. FinTech firms bring new information technologies and innovative approaches to the marketplace, opening up new opportunities for firms to enter the market and attract investors who were not previously involved, and there is significant potential for these market participants to shift to these more productive and efficient means of intermediation. In such a context, many studies have begun to focus on the relationship between technology-based lending and traditional bank lending, as well as the potential impact on bank performance.

1.2.2 Impact of technology-based lending growth

So far, the potential impact of technology-based lending on the banking industry is multifaceted. On one hand, the emergence of financial innovations can have a negative impact on bank performance, as online lending and investment platforms cut into their profitability (Buchak et al., 2018; Jagtiani and Lemieux, 2018; Thakor, 2020). Also, financial innovation partially fills the financial service gap of traditional banks due to regulatory constraints (Buchak et al., 2018). The potential competition could lead to a decrease in the credit market share of financial intermediaries, disrupting the primary activity of the banking system (Greenbaum et al., 2019). On the other hand, technology-based lending helps banks further expand their business (Campanella et al., 2017), and complements the banking system business rather than displace it because of the limited liquidity provided by FinTech or BigTech lenders (Navaretti et al., 2018). Navaretti et al. (2018) also found that FinTech firms are like "full-reserve or narrow banks" in that they can pool collected funds and customers can have access funds when needed, but they cannot use these funds to provide illiquid loans or obtain assets with lower liquidity.

Prior research finds that the development of technology-based lending has a significant positive effect on the banking industry. Lee et al. (2021) argued that FinTech innovations not only improve the cost efficiency of banks, but also enhance the technology used by banks. Iyer et al. (2016) and Jagtiani et al. (2021) found that by analyzing non-traditional personal information that technology-based lending rely on more, such as e-commerce transactions and payment histories, social media postings and viewings, banks can develop a better understanding of potential user risk, and this can enhance the traditional credit information used in credit scoring. Buchak et al. (2018) and Fuster et al. (2019) illustrated that the presence of FinTech in the US mortgage lending market has helped reduce inefficiencies, such as lengthy loan processing.

However there are also studies that find that the development of technology-based lending will further reduce the viability and market share for the traditional banking industry. Hodula (2022) investigated banking sector reactions to FinTech credit and shows mixed results depending on the banking sector concentration. Kowalewski and Pisany (2022a) found the relationship between FinTech credit providers and banks was similar and competitive in developed markets, and BigTech companies compete even more with banks and push some banks out of the market. Compliance with progressively more encumbering regulations is a burden for banks, which often allocate significant capital to relax these requirements and pay occasional penalties when they are unable to comply. However, the vast proportion of FinTech and BigTech firms do not have a full banking business and therefore do not have tight regulations to restrict their business activities. In this context, consumers may choose to have easier access to financial services when they need them, thus reducing the customer share of traditional banks. Thus, I strive to contribute to this paper by empirically verifying the following hypothesis:

Hypothesis I: The growth of technology-based lending have a detrimental effect on bank profitability.

Meanwhile, the bank's characteristics or the features of the country in which it is located may also affect the relationship between technology-based lending and the bank. On the country-specific factors side, Cornelli et al. (2020) considers credit volume in different countries and the disparity in the impact caused by different types of credit. For example, in developing countries, the lack of a traditional financial infrastructure or the scarcity of financial services due to the geographical characteristics of particular areas make FinTech solutions the dominant choice for people accessing financial services (Thakor, 2020). Wang et al. (2023) also confirmed that FinTech credit is positively related to economic and technological development, with the relationship being more pronounced in countries with lower inflation rates. In such a context, technology-based lending can serve different roles in different countries based on their specific macroeconomic factors and considerations, impacting a bank's profitability differently. On the bank-specific factors side, Lee et al. (2023) found that different types of commercial banks are affected differently by FinTech lending, with urban commercial and rural commercial banks being the most affected by FinTech and joint-stock banks being the least affected.

Hypothesis II: Influence of technology-based lending growth in bank profitability varies based on bank-specific and country-specific perspectives.

1.2.3 Difference between FinTech and BigTech lending

FinTech lending is the first category to emerge in technology-based lending. As previously stated, the core of FinTech lending is using technology to provide new and improved financial services. According to Philippon (2015), the unit cost of financial intermediation in the United States has remained around 2% for the past 130 years. In such a context, FinTech lending products endeavor to introduce less costly methods to overcome financial, contractual frictions and reduce the cost of financial services, thereby improving consumer welfare. According to Berg et al. (2021), the volume of FinTech lending in the US exceeded 23.6 trillion in 2020. Even though FinTech lending still represents a small share of the total lending market in the United States, its growth rate outpaces that of the overall lending market. The most well-known FinTech lending company in United State is LendingClub, which was founded in 2006, is now publicly listed, and had a market capitalization of $1.15 billion in 2018. LendingClub is a peer-to-peer lender, while it also offers traditional loans through a bank subsidiary, and its peer-to-peer loans are typically less than $40,000.

In addition to FinTech firms devoting resources to the development of new lending products, BigTech firms are dipping their toes into the financial markets to create unique lending products based on their lucrative customer base and e-commerce platforms. BigTech lending is a late bloomer compared to FinTech lending. However, it is already rapidly expanding and offering financial services to multiple types of borrowers, especially in China, Japan, Korea, Southeast Asia, and some countries in Africa and Latin America. China's BigTech lending market is the largest in the world, with offerings that can serve the needs of many categories of borrowers. Ant Financial and Tencent's WeBank, which provide retail and small business consumers with financial services (Xie et al., 2016), were able to leverage the enormous payment data from their mobile payment services to price credit (Frost et al., 2019).

The most apparent difference between FinTech and BigTech lending is the main business of the firm providing the lending service and whether the firm providing the lending service has an original customer base. The core business of FinTech credit platforms remains financial services. Furthermore, banks can replicate what most FinTech firms can do. FinTech firms benefit from an uneven playing field because they are less regulated than banks. Because of the uneven playing field, non-bank FinTech firms can compete with banks for specific products whose success has nothing to do with what makes them unique but cannot replace banks. At the same time, many FinTech companies do not have a customer base at the start of their business. They can only grow their customer base by improving the efficiency of the screening and monitoring processes, as well as serving low- and no-credit customers. Figure 1.1 shows the relationship between FinTech lending and customers and traditional banks, further demonstrating that FinTech lending still needs to rely on banks for its development. FinTech firms analyze the customer's information to assess his or her creditworthiness and provide FinTech lending services based on that information. Because FinTech firms do not have full banking capabilities, customers who receive FinTech lending will still need to go through a bank and use some of their deposits to repay the lending service if they need to make payments.

On the other hand, BigTech companies have many different business lines, of which lending is only a small part. Their main business activities are usually not related to financial activities. These companies already have a user base, making the onboarding process easier for borrowers. They can reduce information asymmetry by leveraging large-scale micro-level data about users obtained from non-financial activities. BigTech firms have distinct advantages that banks cannot easily replicate, and they pose a more significant challenge to established banks in consumer finance and small business lending. BigTech firms acquire enough customers to run their platform banks. A platform bank constructed by BigTech firm competes with traditional banks across the board, from deposits to payments and wealth management. Figure 1.2 illustrates the relationship between BigTech lending and customers. In contrast to Figure 1.1, the lending relationship between BigTech lending and the customer does not necessarily need to go through a bank. BigTech firms also analyze their customers' information to assess their credit and provide lending services based on that information. However, when BigTech firms initially provide lending services, customers can only purchase the company's existing online business or products. In this case, BigTech firms can complete the lending service without going through the bank, just by the balance sheet, which means recording BigTech lending as accounts payable instead of processing by financial intermediaries. Such a process sometimes does not require the intervention of banks, showing that BigTech lending is a more significant challenge for traditional banks.

In conclusion, banks currently provide many of the services provided by FinTech companies, representing a replicatability of FinTech lending. A large technology company with a platform bank is not required to rely on an existing bank. It may establish a subsidiary bank to provide customers with deposit accounts, credit cards, and digital cash. It can also provide its customers with various third-party financial services. It can assist them in selecting between these services. With the technical knowledge and cutting-edge systems that come with the same scale as the big banks, BigTech companies may have a significant advantage over banks and FinTech companies. They do not have the legacy issues that banks do, nor do they have the organizational issues. Thus, I propose the following hypothesis:

Hypothesis III: BigTech and FinTech lending have a different impact on bank profitability,

1.3 Empirical strategy: Data and methodology

This section briefly describes the data sources used for the paper's analysis. I then highlight key trends and distribution of technology-based lending growth, as well as the bank-specific and country-specific characteristics included in the estimation, and discuss the estimation methodology.

1.3.1 Data sources

I compile data from the following sources: (a) the bank-level data retrieved from the BankFocus database provided by Bureau van Dijk; (b) the country-specific technology-based lending credit volume data from the new database constructed by Cornelli et al. (2020); (c) the country-level macroeconomics data are compiled from the World Bank World Development Indicators (WDI) database; and (d) the instruments for 2SLS regression are collected from the World Bank Information and Communication Technologies (ICT) indicators database.

Given the trade-off between bank data availability (e.g., missing data in some banks' samples and the extent to which the bank's specialization covers the traditional banking business) and cross-country sample coverage, I measure technology-based lending volume data over 92 countries from 2013 to 2019 and match the country-year of technology-based lending volume data with that of the bank-level and other country-specific data. My dataset includes 4200 commercial, cooperative, and savings banks (24439 bank-year observations), representing 49%, 34%, and 17% of the full sample I applied in the estimation. Bank-level data is derived from unconsolidated bank reports. Appendix 1.8 contains the specifics of the bank-level data selections.

The Cornelli et al. (2020) database contains data for 99 countries. Given the objectives of this paper, I exclude countries with few yearly technology-based lending observations, a lack of historical data in the World Bank, and significant missing bank data. For example, only one year of technology-based lending data is available for Luxembourg and Iceland, and the WorldBank database does not include macro-related data for Taiwan or bank data for Cameroon, Costa Rica, Myanmar, and El Salvador, so these countries and regions are not included in the paper. All data at the country and bank levels are conveyed in millions of dollars ($).

1.3.2 Measuring bank profitability

To estimate the bank's profitability, I use the return of average asset ratio (ROAA) as the primary dependent variable to represent bank profitability, but I also consider the return of average equity ratio (ROAE) and net interest margin ratio (NIM) as dependent variables to investigate the consistency of the potential relationship between technology-based lending and banks with different profitability indicator considerations. ROAA shows the profit earned per dollar of assets and, most importantly, reflects the management's ability to utilize the bank's financial and real investment resources to generate profits (Hassan and Bashir, 2003). For any bank, ROAA depends on the bank's policy decisions and uncontrollable factors relating to the economy and government regulations. Many regulators believe ROAA is the best measure of bank profitability (Hassan and Bashir, 2003). ROAE, on the other hand, reflects how effectively bank management uses its shareholders' funds. Since returns on assets tend to be lower for financial intermediaries, most banks utilize financial leverage heavily to increase return on equity to a competitive level (Hassan and Bashir, 2003)), presenting a higher ROAE. ROAA and ROAE are internal determinants mainly influenced by a bank's management decisions and policy objectives.

The net interest margin ratio (NIM) compares the net interest income generated by a bank from credit products such as loans and mortgages to the outgoing interest paid to savings accounts and certificates of deposit holders. Changes in NIM are conceptually related to the demand and supply of bank credit services and market interest rates for borrowing and lending. As a result, NIM more accurately reflects bank profitability in the face of the change in external interest rates, macroeconomic policy, and the bank's market competition. In contrast, while bank managers consider the external economic environment and macroeconomic policies when developing the bank's development strategy and investment portfolio, ROAA and ROAE primarily reflect the profitability of the bank's internal management and asset allocation.

1.3.3 Technology-based lending

I use the database constructed by Cornelli et al. (2020) to create three proxy variables for technology-based lending credit development: (1) Technology-based lending (sum of FinTech credit and BigTech credit) to GDP ratio, (2) BigTech lending credit to GDP ratio; (3) FinTech lending credit to GDP ratio; and (4) relative volume change of Technology-based lending to GDP ratio. This paper treats technology-based lending to GDP ratio results as primarily important. Applying the technology-based lending-related credit volume to GDP ratio can eliminate the potential biased from differences in country size (Ahmed and Zlate, 2014), and all technology-based lending related variables are transformed into natural logarithm with (1+X) form to serve as the main explanatory variable in my empirical analysis.

Figure 1.3 presents the growth of technology-based lending from 2013 to 2019 based on the database constructed by Cornelli et al. (2020), the full country sample at the aggregate level. I note a dramatic growth trend in technology-based lending volume, increasing 4223.70% from 2013 to 2019, with FinTech lending playing a significant role in a 2146% increase in total volume. This trend shifted after 2017, when the total volume of FinTech lending fell, partly due to the diminishing of peer-to-peer lending; for example, LendingClub and Prosper have not expanded lending volumes over the 2016-2020 period (Berg et al., 2021). BigTech lending, on the other hand, has been rapidly expanding, with its total volume surpassing FinTech lending in mid-2018, demonstrating its consumer popularity. By 2019, BigTech lending's total volume had surpassed FinTech lending's by 156.25%. Based on Figure 1.3 I conclude that even though technology-based lending has grown rapidly overall, consumers' product preferences for different types of technology-based lending have varied over time. Figure 1.3 also provides a potential that FinTech and BigTech lending impact banks differently.

Figure 1.4 depicts a distribution of FinTech and BigTech lending volume according to its mean value with natural logarithm transformation in different countries, using ISO 3-digit code to represent the country's name. China, the United States, and the United Kingdom have the highest levels of technology-based lending. BigTech lending has a higher volume in Asian countries than in others, such as South Korea, Japan, and China. At the same time, many countries have FinTech lending but not BigTech lending. Since most of these countries are low- and middle-income, where BigTech companies do not exist to run lending services, or the capacity of existing BigTech lending is so tiny that it is difficult to quantify the existence of BigTech lending.

1.3.4 Bank- and country-specific characteristics variables

Consistent with prior literature on bank's lending behavior (Allen et al., 2017; Chen et al., 2016; De Haas and Van Lelyveld, 2014; Meriläinen, 2016), I control for six bank characteristics variables: the natural logarithm transformation of the total assets of given banks (Size), liquid assets to total asset ratio (Liquidity), the loan loss provisions to gross loans ratio (Risk), the capital to total assets ratio (Capitalization), non-interest income to operating revenue (Diversification), and equity to total assets ratio (Financial leverage). I winsorize all bank-specific explanatory variables at a 1% level to reduce the effect of extreme values on regression, and Table A.1 illustrates the correlation matrix for every bank-specific characteristic variable.

This paper considers the disparity in the impact of growth shocks from technology-based lending on banks due to the difference in the scale of their own operations. For the alternative channels that the factors that technology-based lending influence and induce the profitability change, I consider two different types of indicators to represent a bank's lending service. First, I use the nonperforming loan ratio (NPL ratio) to represent the bank's lending efficiency to evaluate the bank's efficiency performance change. NPL is the value of nonperforming loans divided by the total value of the loan portfolio. It measures bank health and efficiency by identifying problems with asset quality in the loan portfolio, and a smaller value represents a higher lending efficiency of banks. When considering the demand and supply of bank lending services, I use the bank's gross loan volume as a measure of demand and the bank's total deposit volume as a measure of demand. Banks determine and provide lending services based on customer demand and specific characteristics. In this case, a higher gross loan volume represents a higher demand for this bank's lending service. Simultaneously, banks use consumer savings as the principal for lending activities. So, the higher the volume of bank deposits, the more lending services it can offer customers. I apply the natural logarithm transformation for future estimation for these two dependent variables.

For country-specific variables, I follows Allen et al. (2017) and employ macroeconomics-related control variables, i.e., GDP Growth rate, GDP per capita growth rate, and inflation rate that show, in the most basic way, the attractiveness of a given country as a place for running a financial business. Table A.5 illustrates the detail information for each country with the ISO 3-digit code, region, income level, and the number of banks in each country.

I introduce two indicators to assess a country's credit accessibility: the Depth of Credit Information Index (DCII) and the Strength of Legal Right Index (SLRI). SLRI assesses the extent to which collateral and bankruptcy laws protect borrowers' and lenders' rights and thus facilitate lending. The index is scaled from 0 to 12, with higher scores indicating that these laws are better designed to increase credit availability. The DCII index ranges from 0 to 8 and measures the coverage, scope, and accessibility of credit information available through credit reporting service providers such as credit bureaus or credit registries. Access to and utilization of banking services lowers financing barriers for individuals and businesses, expanding opportunities for all.

Financial services are affected by availability, cost, and quality. Data on borrowers' credit experiences is essential for credit market growth. Credit access can be improved by simplifying collateral agreements and providing more creditworthiness information. In the traditional loan procedure, a customer's current credit score and the collaterals that he or she can use to apply for a loan determine whether or not the loan can be approved and how much the customer can apply for. Banks lend less in developing countries with more credit registries and collateral laws. DCII is constructed based on such consideration and represents the accessibility of credit information in a specific country. Also, laws, regulations, institutional arrangements, and macroeconomic indicators determine a country's economic health and influence the financial market. Measurements of the legislative environment of the financial market in a country include business regulation, outcomes, legal property protection, employment regulation flexibility, and business taxes. The financial market activity requires efficient, accessible, and easy-to-implement rules and regulations, according to this data. As a result, stricter disclosure requirements are sometimes prioritized over simplified regulations, such as a one-stop shop for business startup formalities in related-party transactions. Entrepreneurs may only know or avoid certain processes. In areas with heavy regulation, informal firms grow more slowly, have less access to credit, and employ fewer non-union workers. In such a context, SLRI can help policymakers understand a country's business environment and suggest reforms when combined with other data sources like the World Bank's Enterprise Surveys.

I use one lag term in the baseline model for all bank-specific controls to reduce inverse causality issues, but regressions can also have potential endogeneity issues. When significant breakthroughs in communication technology and network penetration occur nationally, technology-based lending can make remarkable progress and improve quickly. Furthermore, banks may vigorously develop the online banking business to provide customers with better financial services and thus increase their profitability.

To address potential endogeneity concerns, I re-estimate the impact of technology-based lending development on bank profitability using a two-stage least square (2SLS) regression, with the country's penetration of communication technology and the degree of development of financial services as exogenous shocks. In considering the country's communication technology penetration, consistent with prior research, I select its mobile cellular, fixed broadband, and fixed telephone subscriptions to measure its penetration as instrumental variables (Andrianaivo and Kpodar, 2012; Sheng, 2021; Demirgüç-Kunt et al., 2020). When considering the development and completeness of the country's financial services, I select the country's SLRI and DCII as instrumental variables.

1.3.5 Descriptive statistics and Methodology

The empirical analysis aims to examine the potential impact of technology-based lending growth on a bank's profitability. My analysis follows an Ordinary Least Square (OLS) with a fixed effect strategy for bank-year panel data. I use $Ln(1+X)$ for each natural logarithm transformation to keep the zero value and positive trend for some specific variables. In the database constructed by Cornelli et al. (2020), there are only 31 countries have non-zero volume of BigTech lending as of 2019.

I first start with a general empirical specification to show the impact of technology-based lending growth β_1 on the bank's profitability:

$$y_{i,c,t} = \beta_1 Tech_{c,t} + \beta_2 Bank_{i,c,t-1} + \beta_3 Macro_{c,t} + \alpha_t + \gamma_c + \varepsilon_{i,c,t} \qquad (1.1)$$

where the dependent variable $y_{i,c,t}$ is either ROAA, ROAE or NIM of bank i in country c at year t. $Tech_{c,t}$ is the natural logarithm transformation of one plus technology-based lending's volume to GDP ratio at country c in year t, α_t represents the year fixed effect, controlling inter alia for other macroeconomic and time-varying global business cycle effects. $Bank_{i,c,t-1}$ is a vector of bank-specific characteristic variables for bank i in country c at year $t-1$ including bank's size, liquidity, risk, capitalization, diversification, and financial leverage. I lag bank-level explanatory variables by one period to mitigate the potential problem of reverse causality (Dushnitsky et al., 2016) . $Macro_{c,t}$ is also a vector that present country-specific characteristic control variables for country c in year t, including GDP growth rate, GDP per capita growth rate, and inflation rate. α_t represents the year fixed effect, controlling for factors changing each year for that common to all countries for a given year. γ_c represents the country fixed effect, controlling the baseline differences between countries. I employ country and year fixed effects because the overall empirical analysis focuses on cross-country differences rather than individual banks within a country. In this instance, country fixed effects take precedence over bank fixed effects. In addition, most of the literature about financial institutions prefers to control for the main characteristics of banks rather than adding bank fixed effects, as bank fixed effects may absorb these potential effects, thereby preventing the examination of the impact of bank-related variables on the dependent variable under consideration. I also consider bank clustering standard errors to account for other possible time-varying factors unique to each bank but not observable in the primary characteristics. A significant and negative

β_1 in Equation 1.1 support hypothesis I.

Going further, I apply the interaction term with technology-based lending and investigate the coefficient change with and without intervention. Formally, the estimation procedure can be written as:

$$y_{i,c,t} = \beta_1 Tech_{c,t} + \beta_2 Tech_{c,t} \times Character_b + \beta_3 Bank_{i,c,t-1} + \beta_4 Macro_{c,t} + \alpha_t + \gamma_c + \varepsilon_{i,c,t} \tag{1.2}$$

or

$$y_{i,c,t} = \beta_1 Tech_{c,t} + \beta_2 Tech_{c,t} \times Character_c + \beta_3 Bank_{i,c,t-1} + \beta_4 Macro_{c,t} + \alpha_t + \gamma_c + \varepsilon_{i,c,t} \tag{1.3}$$

where the indicator variable $Character_b$ is considered either bank-specific characteristics, and $Character_b$ considers country-specific characteristics.

When considering bank's own characteristics, the indicator variable is either 1) $Character_b$ equals to 1 when the bank's size is small (total asset is smaller than 1 billion USD), 2) $Character_b$ equals to 1 when bank's number of total branches is small (total branches is smaller than or equal to 10), or 3) $Character_b$ equals to 1 when a small-size bank has few branches.

When considering country's characteristics, I consider the development, competition, and credit accessibility level of of the country that a bank locate in. In this situation, indicator variable $Character_b$ equals to 1 when country c is a developed country[1] when investigating how the development level of a specific country impact the potential relationship. Also, $Character_b$ equal to 1 (or 2) when the bank-

[1]The World Bank database classified the income level of a specific country into four different levels: low-income, lower-middle income, upper-middle income, and high income. In this paper, I identify a country with a high income as a developed country.

ing sector of country c is moderated (highly) concentrated [2]. When considering the credit accessibility of a specific country, $Character_b$ equal to 1 either 1) if the country c is a low credit information country (DCII smaller than 7), 2)equals to 1 when country c is low legal right country (SLRI is smaller than 6), or 3) $Character_b$ equals to 1 when country c is a low credit accessibility country (both DCII and SLRI is small) based on different specifications. Equation 1.2 have same bank-specific, country-specific control variables, and fixed effect setting as Equation 1.1. Equation 1.2 estimate hypothesis II of this paper: If the sign of estimated β_1 changes based on different factor considerations, hypothesis II will be accepted.

To investigate hypothesis III, I separate technology-based lending to BigTech and FinTech lending to reestimate equation 1.1. In this case, the target coefficient β_1 change to a vector, representing the influence of BigTech and FinTech lending on banking system.

To examine the potential channels through which the factors influenced by technology-based lending growth induce a decrease in a bank's profitability, I consider the demand side of the bank's lending service and lending efficiency as the dependent variables. in this case, the estimation model change to:

$$y_{i,c,t} = \beta_1 Tech_{c,t} \times Character + \beta_2 Bank_{i,c,t-1} + \beta_3 Macro_{c,t} + \alpha_t + \gamma_c + \varepsilon_{i,c,t} \quad (1.4)$$

where the dependent variable $y_{i,c,t}$ is either 1) the natural logarithm transformation of gross loan of bank i in country c at year t or 2)total deposit volume of bank i in country c at year t in demand and supply of bank's lending service consider-

[2]The competitiveness level of the banking sector in a specific country is determinded by the Herfindahl-Hirschman Index (HHI). I consider the specific banking sector in which the HHI is between 1500 and 2500 to be moderately concentrated, and consider specific banking sector in which the HHI is in excess of 2500 to be highly concentrated.

ation, or 3) the natural logarithm transformation of non-performing loan (NPL) of bank i in country c at year t or 4) non-performing loan ratio (NPL ratio) of bank i in country c at year t in bank's lending efficiency consideration.

1.4 Main results

1.4.1 Baseline model

In this section, combining bank- and country-level variables, I estimate whether technology-based lending growth enhances or impedes bank-level profitability. The estimated results present the detrimental effect of technology-based lending growth on the bank's profitability, decreasing 5.623% of the bank's ROAA. This result is statistically significant (5% level) and economically meaningful, representing an average $2,135 billion reduction in the bank's net income.

Table 3 displays the baseline model results estimating the impact of technology-based lending growth on bank profitability with various control considerations in Columns (1) to (3) and BigTech and FinTech lending separately in Columns (4) to (6). The results across both considerations are consistent and provide further supporting evidence that the increase in technology-based lending has a detrimental effect on banks' profitability. Such impact will be reduced by further consideration of country macro- and bank-specific control. Estimates of β_1 in Equation 1.1 without the bank-specific and country-specific characteristics considerations, reported in Table 3 Column (1),show a 7.032% raw and significant decrease in ROAA for each percent increase in technology-based lending, without considering bank-specific and country-specific controls. After controlling for all the bank-specific and country-specific characteristics in Column (3), the impact of technology-based lending on the bank's ROAA alls to 5.623%. This influence is statistically significant and economically significant: on average, it represents around $2,135,462 million loss in the bank's net income[3]. Such results identify which bank and country to control the technology-based lending growth shock, from a raw 7.032% to 5.623%. Column (4) divides technology-based lending into two categories: BigTech and FinTech lending, and re-estimates Equations 1.1 to show that BigTech lending dominates the negative impact of technology-based lending growth on bank ROAA, whereas the impact of FinTech lending volume growth eventually diminishes once the bank- and country-specific control variables are added. The coefficient of relative volume change of technology-based lending to GDP ratio in Column (7) is not statistically significant, indicating that changes in the growth rate of Technology-based lending do not significantly impact bank profitability.

I next estimate the negative impact of technology-based lending growth shock on other indicators of bank profitability. These specifications also confirm a significant decrease in bank profitability even accounting for preferences in the bank's internal management and external lending volume or policy influence. As illustrated in Panel A of Table 1, ROAA is the least volatile of the three indicators representing bank profitability, while ROAE is the most volatile, implying that the impact of technology-based lending growth on various bank profitability indicators may also be diverse. Comparing the empirical results of Table 4 and Table 3,I conclude that the negative impact of technology-based lending growth on bank profitability is robust to different profitability indicators. Following the 1% increase in technology-based lending leads to an 86.71% decrease in ROAE (Column (2)) and a 14.42% decrease in NIM (Column (6)) for banks after including bank- and country-specific controls.

When technology-based lending is separated into BigTech and FinTech lending, the results for ROAE and ROAA are similar, with BigTech lending having a

[3]See Table A.7, which runs the net income in Million USD instead of the ratio representing profitability on the same controls.

more significant impact on bank profitability. In contrast, FinTech lending has a more significant impact on NIM. In addition, unlike ROAA and ROAE, NIM is reduced as the growth rate of technology-based lending increases. Based on the overall baseline model estimation, I conclude that external macroeconomic conditions and legislation inclined to protect traditional financial intermediaries from maintaining the stability of the financial market, combined with the overall external conditions, the impact of banks from technology-based lending will decrease. In addition, when banks face the rapid growth of technology-based lending, the internal investment or development decisions made by banks in the context of the external economic environment and other macro factors will further undermine the profitability of banks.

1.4.2 Endogeneity

The specification to address the potential endogeneity issue is a conventional two-stage least squares (2SLS) panel regression. The first stage specification for the instrumented technology-based lending volume satisfies:

$$Tech_{c,t} = \beta_1 Infrastructure_{c,t-1} + \beta_2 Accessibility_{c,t} + \alpha_t + \gamma_c + \varepsilon_{i,c,t} \tag{1.5}$$

where $Tech_{c,t}$ represents the technology-based lending volume at country c in year t in $Ln(1+x)$ logarithm transformation form. $Infrastructure_{c,t-1}$ illustrates the technology-related infrastructure development at country c in year $t-1$, including mobile cellular, fixed broadband, and telephone subscriptions per 100 people with logarithm transformed both in regression. $Accessibility_{c,t}$ is the level of credit accessibility at country, c in year t, including the Depth of Credit Information Index (DCII) and Strength of Legal Right Index (SLRI). Both time fixed effect

α_t, country fixed effect γ_c follow previous setup.

Results of the first stage IV regression are shown in the Table A.8. The selected instruments $Infrastructure_{c,t-1}$ and $Accessibility_{c,t}$ enter significantly with a positive sign at a 1% confidence level. The second stage specification is estimated as follow:

$$y_{i,c,t} = \beta_1 \hat{Tech}_{ct} + \beta_2 Bank_{i,c,t-1} + \beta_3 Macro_{c,t} + \alpha_t + \gamma_c + \varepsilon_{i,c,t} \tag{1.6}$$

where the dependent variable $y_{i,c,t}$ is either 1) return of average asset ratio (ROAA ratio) of bank i in country c at year t, 2) the return of average equity ratio (ROAE ratio) of bank i in country c at year t or 3) the net interest margin ratio (NIM ratio) of bank i in country c at year t.

Table 5 shows the 2SLS regression estimates of the effect of technology-based lending growth on cross-country bank profitability, while Table A.8 shows the first-stage estimation. I present the estimates without and with various instrumental variables taken into account in Table 5. All estimates of the coefficient of technology-based lending credit volume are significantly negative, consistent with the baseline regression, according to all columns. In the first stage, I use information technology penetration variables in Column (1) of Table A.8, and the F-test values are greater than the critical values for the Stock and Yogo (2005) weak instrument test, indicating that the information technology penetration instruments are strong. Therefore I reject the null hypothesis that the instruments are weak. Column (2) of Table A.8 further considers instrumental variables representing the degree of completeness of national financial services, and then the F-test values continue to exhibit strong instrument trend. Columns (3) and (4) investigate whether these instrumental variables have a delayed effect on technology-based lending and confirm that there is a significant relationship between national information technology penetration and the completeness of financial services on the growth of technology-based lending, with both immediacy and latency impact.

1.5 What factor affects banks more?

After demonstrating the negative impact of technology lending on bank profitability, I investigate the potential factors that expose or protect banks from additional negative shocks. This enables us to examine the changing impact of the increasing adoption of new forms of lending on traditional financial intermediaries, as well as gain some insight into how traditional financial intermediaries respond to new forms of competition. Indeed, access to and selection of financial services is relatively limited from the customer's perspective. People can replace some bank-provided lending services with technology-based lending to meet their original expectation, as can the difficulty of accessing traditional financial services in some countries and the lack of a well-developed lending environment to support technology-based lending to meet the demand for financial services. Based on such hypotheses, this section defines the factors that banks are most likely to be affected by shocks to the growth of technology-based lending, taking into account potential influencing factors such as the size of their operations, as well as the level of development and credit availability in the countries where they are located.

1.5.1 Bank's operating scale

The potential variability of the negative impact of increased technology-based lending on large and small banks' profitability remains uncertain. On the one hand, larger banks have a broad range of business coverage, relatively well-developed types of business, and a substantial and well-established customer base. Given such considerations, larger banks are unlikely to be replaced by technology-based lending. Small banks, on the other hand, are more effective than large banks in credit assessment and loan management by using soft information about customers (Bernanke, 2011; Hughes et al., 2022), and they are more likely to reach potential customers covered by technology lending in their daily business, such as those with no or low credit. In this case, some parts of large banks' business lines may be replaced by technology-based lending due to less targeted business for small volume customers, affecting the bank's profitability. Meanwhile, the size and accessibility of the bank's branches are both positively associated with the bank's operating scale. The larger the bank's operating scale and the more extensive its branches, the higher the coverage of financial services offered by the bank and the less likely it is to be replaced by technology-based lending. Based on the empirical results, I conclude from the regression results that banks with few branches are more prone to more severe profitability damage caused by the rise of technology-based lending than banks with small-sized.

Table 6 provides results from specifications that adapt Equation 1.2 to include interaction terms for representing the bank's operating scale and the country's development level that the bank locates in. I observe that there is still a significant and negative effect of technology-based lending growth on a bank's profitability across the regressions with different interaction terms. At the same time, the additional negative shocks to banks in developed countries due to their small operating scale are relatively insignificant compared to those in developing countries. Column (1) of Table 6 shows that small banks in developing countries reap an additional 3.2% decrease in ROAA relative to other banks from technology-based lending growth. At the same time, such a phenomenon is insignificant for small banks in developed countries. Comparing Column (2) with Column (1) of Table 6, the additional 12.76% decrease in ROAA of banks with fewer branches is more significant than that of small banks. According to Column (3) of Table 6, I observe that when a bank is defined as small based on operating scale, it does not suffer any additional negative shocks from the technology-based lending growth. Such results may be because most banks with fewer branches focus more on online banking operations and development than others; some are even fully online banks, such as American Express. In such a context, these banks are in the same market as technology-based lending and are, therefore, more likely to receive additional negative shocks. On the other hand, the small operating scale banks do not receive the additional negative impact of the growth in technology-based lending because their outreach is less extensive and more targeted to the customers and regions they serve than other banks.

I then further explore the heterogeneity of the influence of technology-based lending on the bank's profitability by its operating scale. Figure 1.5 plots the coefficient on the technology-based lending in Equation 1.1 estimated separately for 10 deciles of bank size groups. The estimation results shown in Figure 1.5a illustrate that the influence of technology-based lending on the bank's ROAA stays negative and insignificant as banks' total assets increase. In contrast, the impact of the growth of technology-based lending tends to be zero when the bank's total asset is in the upper 30% quantile.. When considering the heterogeneity in the impact of the growth of technology-based lending on bank profitability due to the number of branches a bank has, Figure 1.5b shows that banks with zero branch receive the most negative shocks.

To further explore the consistency of the effect of bank size on the growth of technology-based lending on bank profitability, I re-estimate Equation 1.2 with the same interaction term, replace the dependent variables with the ROAE and NIM, and report the results in Table A.9. Panel A of Table A.9 shows that banks with fewer branches have an additional 146.2% impact on their ROAE ratio due to the growth of technology-based lending. In comparison, small-sized banks reap a minor negative shock of 70.44%. Compared to the estimation results considering the ROAA ratio, the additional negative shock is significant when banks are located in developed countries with more than 10 branches. Such result represents that the ROAE ratio of larger banks in developed countries is more negatively affected by the growth of technology-based lending than in developing countries. Comparing Panel B of Table A.9 with Table 6, I find that banks' ROAA and NIM ratios receive similar additional adverse shocks caused by the difference in the bank's operating scale from technology-based lending growth.

Taken together, these results paint a clear picture: banks are not necessarily experiencing an additional negative impact from the growth of technology-based lending due to their potential disadvantage in operating scale. At the same time, banks in developed countries do not have a significant and negative impact on their profitability due to the growth of technology-based lending. Adding to the interaction, banks with fewer branches receive more negative shocks than other banks because their business mainly concentrates on online banking, which competes with technology-based lending at the same stage. Meanwhile, mega-banks are least affected and less volatile, while smaller banks face increased volatility and potential negative impacts.

1.5.2 Competition level of banking sectors

Besides the potential for individual banks to receive an additional negative impact on profitability from the growth of technology-based lending because of the difference in operating scale, the competitiveness level of the banking sector in the countries in which the bank operates can also be a potential factor. If a bank is in a country where the banking sector is highly competitive, it may face a different outcome in the presence of the growth of technology-based lending. On the one hand, the bank can maintain the competitiveness of its financial services due to prolonged and active competition, even if new technology-based lending enters the credit market. On the other hand, banks may see their market share squeezed by new competitors and be more easily squeezed out of the market due to their lack of volume.

The estimation results in Table 7 how the growth of technology-based lending in different competitive level markets affects banks' profitability. Column (1) results show that when a bank in a developing country is in a competitive market, the growth of technology-based lending causes a significant negative shock of 4.328% to the bank's ROAA, and such negative impact is even more profound if the bank is in a moderated concentrated market, reaching 130.2%. The growth of technology-based lending will not significantly impact banks' profitability if the banking sector in which they operate is highly concentrated. When considering the bank's ROAE as the dependent variable, Column (2) yields a similar result to Column (1), except a larger coefficient, indicating that the bank's ROAE is affected more than the ROAA ratio, consistent with previous results. Column (3) results indicate that, when considering NIM, only banks in developing nations and competitive markets experience a decrease of 14.77% when technology-based lending increases. In contrast, banks in developed countries do not experience a significant negative impact from the increase in technology-based lending across different market competitiveness levels, except in markets where the ROAE receive a significant negative impact, as illustrated in Column (2).

Table 7 demonstrates that when technology-based lending enters the credit market as a new option for customers, banks in highly concentrated markets dominated by large banks will not be impacted by the growth of technology-based lending due to their significant financial services volume and market share. In contrast, if banks operate in a highly competitive market, the emergence of technology-based lending makes the credit market more competitive. It gives customers more options, reducing banks' profitability significantly. However, such a situation is rare in developed countries, where banks do not suffer from an increase in technology-based lending due to the competitive strength of their national banks and, thus, their profitability.

1.5.3 Credit Accessibility

As mentioned before, existing literature points out that technology-based lending helps under-served markets, whereas the stringency of regulations can hinder their further development (Kowalewski and Pisany, 2022b). Building on the previous analysis, I further examine country characteristics of where the lending service provider operates to understand better what may have contributed to the differences in the impact on banks in developed and developing countries in the face of the growth of technology-based lending. There are several aspects of access to financial services for individuals or businesses: availability, costs, and quality of service. The development and growth of the credit market depend on the availability of timely, reliable, and accurate data about borrowers' credit experiences. Access to credit can be improved by making it easier to create and execute collateral agreements and by increasing information about the creditworthiness of potential borrowers. Lenders look at a borrower's credit history and collateral. Credit registries and effective collateral laws still need to be improved, as in many developing countries; thus, banks in developing countries may offer fewer loans. Indicators covering access to credit include the SLRI and DCII. I again adapt Equation 1.2 including interaction terms with characteristics country's accessibility of credit, to empirically test for the difference in the shock banks received from the technology-based lending growth in different countries. In this paper, I define that the DCII and SLRI are low when the index is smaller than 7 and 6, respectively, and a specific country with low credit accessibility when both DCII and SLRI are low.

Table 8 summarizes the initial results from this analysis. I focus on results relative to the influence of change in technology-based lending growth for brevity. I observe that the proxy for low accessibility of credit in a specific country (both low DCII and SLRI based on identification) has a negative trend on the bank's profitability from a ROAA ratio perspective. The result in Column (3) of Table 8 shows that a 1% increase in technology-based lending will cause a 21.26% decrease in banks' ROAA in developing countries with low accessibility of credit. Such results provide a similar argument to the conclusion of Hodula (2022) that technology-based lending has a higher potential to compete with the traditional banking sector in the immature credit market and that banks are additionally undermined by such competition. Columns (1) and (2) results show that banks in developing countries with low SLRI indexes are more vulnerable to additional adverse shocks to profitability, with a 1% increase in technology-based lending decreasing the bank's ROAA ratio by 4.713%. The results of Table 8 also show that banks in developed countries do not significantly influence the growth of technology-based lending due to their countries' low credit accessibility.

Table A.10 examines whether banks' other profitability indicators, ROAE and NIM ratio, are subject to adverse shocks because of the increased volume of technology-based lending in countries with low credit accessibility. Panel A in Table A.10 shows, consistent with previous results, that the growth of technology-based lending has the most significant impact on banks' ROAE. Column (3) also shows that even if the bank is in a developed country, its ROAE decreases by 332% due to its low credit accessibility. Column (4) results show that banks in developed countries are reduced their NIM by 12.38% due to an increase in technology-based lending because of their low SLRI index. The remaining results are similar to Table 8, with banks in developed countries being less exposed to the impact of technology-based lending on profitability and banks in low SLRI countries being less exposed to the effect when compared to low DCII.

1.5.4 Overall discussion

Based on the previous results, I conclude that banks in developing countries are more likely to decline in profitability significantly due to the growth of technology-based lending than in developed countries. The disparity of the emerging lending models across countries is also a topical discussion in the relevant literature. High penetration of financial services in high-income countries requires technology-based lending to compete directly with the financial services offered by traditional financial intermediaries or to serve a complementary role to the customer base under-served by banks. In comparison, banks and BigTech and FinTech firms are likely to cohabitate in emerging economies. Banks in emerging economies may need to be more technologically advanced, serving customers through more traditional channels, while FinTech and BigTech firms play a role in developing digital, remote customer lending channels. Banks in emerging economies may compete fiercely with technology-based lending in this context. Also, the impact that the growth of technology-based lending will have on a bank depends on the specific characteristics of the country in which it is located. In comparison, banks in countries with higher levels of development, greater credit accessibility, and less competition in the banking sector are less susceptible to the adverse effects on profitability of the rise of technology-based lending.

When considering different indicators of bank profitability, banks' ROAE ratios are most negatively impacted by the growth of technology-based lending. Panel A of Table 1 demonstrates that the banks in the dataset present a higher ROAE than ROAA, indicating that these banks have a high leverage ratio (lower equity). In this case, when banks face the impact of technology-based lending, they will receive more detrimental effects on their ROAE than on their ROAA due to their higher leverage. As mentioned above, NIM reflects the revenue from interest on operating activities and the impact of external interest rates and macroeconomic policies on the bank's profitability, whereas the ROAA ratio reflects the impact of the bank's internal asset management, decision-making, and allocation on its profitability. Based on previous results, the NIM ratio is also more susceptible to shocks from the growth of technology-based lending than the ROAA ratio, indicating that the growth of technology-based lending has a greater impact on the operating income of banks' daily operations.

1.6 Robustness check

1.6.1 With and without top five high volume countries

In Figure 1.4, I can observe that technology-based lending is mainly distributed among some countries, such as China, the United States, the United Kingdom, Japan and Russia. Meanwhile, banking system development in these four countries is relatively mature and well-established compared to the rest of the world, and there is concern that these countries drive the regression results in the baseline model while having no significant effect on other countries. To address this concern, I subgroup the data, re-regress the baseline model, and compare the pattern of β_2 in the full sample and across sub-samples to determine whether the leading countries with top development of technology-based volume drive the initial regression results. I identify the high-volume countries subgroup, including the top five high technology-based lending volume countries, including China, the United States, the United Kingdom, Japan and Russia. The non-high-volume countries would be the subgroup without these five countries.

Results shown in Table A.11 indicate that the negative impact of technology-based lending growth on bank's profitability is robust in high and non-high volume countries. Comparing the regression results, I can see that banks in countries with high and low technology-based lending volumes suffer from the same negative impact trend on profitability due to their growth. In contrast, banks in non-high-volume countries are more negatively affected.

1.6.2 Commercial and non-commercial banks

Table 1 shows that commercial banks constitute a significant proportion of the total dataset. Meanwhile, commercial banks have a pivotal position in the banking sector, providing the majority of financial and credit services. In such a context, the estimation of this paper will raise another concern, namely that commercial banks drive the results of the estimation of the baseline model. Therefore, as in the previous robustness check, I divided the whole dataset into only commercial banks and subsamples of banks other than commercial banks for comparison to see whether the results of the baseline model were robust in the subgroups of different bank classifications.

Table A.12 shows that the negative impact of technology-based loan growth on bank profitability is robust in commercial and non-commercial banks in ROAA and ROAE ratio respective. Column (9) of Table A.12 indicates that the NIM ratio of non-commercial banks is profitable from the growth of technology-based lending, with a 42.45% increase in the NIM ratio for every 1% increase in technology-based lending. Comparing the regression results, I can see that commercial and non-commercial banks have the same negative trend in profitability due to technology-based lending growth, the same trend as the estimation result of the baseline model. In contrast, non-commercial banks are more negatively affected, and the regression results have a higher R Square.

1.7 Implication of technology-based lending growth

1.7.1 Alternative channels

Until now, I have already found the robust impact that technology-based lending development negatively exerts on a bank's profitability. Following that, I investigate the alternative channels that connect technology-based lending to bank profitability, that is, which variables of technology-based lending have been affected, resulting in a decline in bank profitability. The potential factors that can affect a bank's profitability are many and complex. Still, the most fundamental of them is the demand for the bank's own business, that is, the supply and demand for its lending service, and the quality of its existing loans and the risk of default, that is, lending efficiency. This section will discuss these two factors separately and explore the alternative channels behind the main results.

I first estimate the effect of technology-based lending growth on the demand and supply of the bank's lending service. Table 9 reports the estimates of the relationship between technology-based lending growth and the demand for bank lending services, measuring banks' demand for lending services by the logarithm of the bank's gross loan. In the regressions, I also consider the variability of the impact of technology-based lending growth on bank lending service demand under different country-specific and bank-specific interaction term considerations. Column (1) results show a significant detrimental effect of the increase in technology-based loan volume on demand for banks' lending services. Each 1% increase in technology-based lending will result in a 2.886% decrease in total bank loans. The results of Columns (2) to (5) show that the demand for lending services is affected to a greater extent by banks in developed countries, countries with high banking sector concentration, low credit accessibility, or banks with medium to large operating scales.

Table 10 depicts the relationship between technology-based lending and the supply of bank lending services, with the latter represented by the total deposit of natural logarithm transformation. Column (1) shows that for every one percent increase in the volume of technology-based lending, bank total deposits fall by 3.906%. This result demonstrates the negative impact of technology-based lending on bank lending service supply. Similar to the effect on the demand for bank lending services, Columns (2) to (5) show that the detrimental effect is more profound when banks are developed, moderately concentrated, or low credit accessibility countries or when the banks' operating scale is not small. Comparing the results of Table 1 with those of Table 2, I conclude that the negative impact on the supply of bank lending services is more profound than that on demand for bank lending services.

In many cases, there is no single choice of financial services, and customers tend to choose a limited number of lending services provided by different financial intermediaries to meet their needs. When technology-based lending enters the credit market as a new product, some banks may be substituted, and the volume of their lending services may decline significantly. However, it is also possible that customers will have more lending services than before because of the easy-to-access nature of technology-based lending, in which case the bank's lending service will not decline but rather increase its potential default risk. The high default risk also significantly affects the profitability of banks.

Table 11 shows the impact of the growth of technology-based lending on the lending efficiency of banks. I first use the bank's non-performing loan in natural logarithm transformation to represent the bank's lending efficiency, and the higher the non-performing loan, the lower the lending efficiency of the bank. Consistent with previous estimations, I also consider the variability of the impact of technology-based lending growth on bank lending efficiency under different country-specific and bank-specific interaction term considerations. The regression results indicate that the increase in technology-based lending also significantly negatively impacts banks' lending efficiency; for every 1% increase in technology-based lending, the bank's non-performing loan increases by 13.93%. After considering the bank-specific and country-specific interaction terms, I find that the growth of technology-based lending significantly increases the bank's non-performing loans when the bank is in a developing or competitive banking sector or a country with low credit accessibility, or when the bank has a small operating scale.

For consistency, I use the non-performing loan (NPL) ratio as a representation of default risk in existing loans in the following analysis, re-examining the relationship between the growth of technology-based lending and bank lending service efficiency. The Column (1) results of Table 12 show a significant increase in the NPL ratio for banks with every 1% increase in technology-based lending; specifically, the NPL ratio rises by 45.19%. Furthermore, the observations in Columns (2) – (5) highlight some nuanced interactions under different conditions. It has been discovered that banks in developed countries, those operating in a competitive banking sector, those in countries with non-trivial credit accessibility, and larger scale banks all experience a heightened negative impact on their NPL ratio as a result of technology-based lending. This suggests that the rise of technology-based lending may disproportionately increase default risk in these contexts, necessitating a more in-depth examination of these factors' impact on lending service efficiency.

When comparing the results from Tables 9 to 12 with the regression results from Section 1.5, I conclude that the growth of technology-based lending increases banks' current non-performing loans, reducing lending efficiency and significantly impairing profitability. Non-performing loans are those in which the borrower cannot make scheduled payments, resulting in interest or principal repayments that are not made. Thus, an increase in the ratio of NPLs directly impacts the bank's revenue stream and, ultimately, its profitability, as Salas and Saurina (2002) demonstrated in their study of Spanish banks. Also, non-performing loans necessitate provisioning, which is a direct cost to the bank and reduces available capital for other profit-generating activities. Boudriga et al. (2009) also argued that a decrease in the capital may limit the bank's ability to extend more loans, reducing both interest and non-interest income.

A decrease in the demand or supply of lending services, on the other hand, does not always imply a decrease in profitability. Banks can adjust interest rates, fees or engage in other banking activities to compensate for reduced lending. They can diversify their income streams and become less reliant on lending services. Stiroh (2004) supports this viewpoint, finding that diversification into non-traditional banking activities can improve bank performance. Furthermore, banks can manage their lending supply by focusing on high-quality borrowers when low demand allows them to maintain profitability while reducing risk. Berger and DeYoung (1997) demonstrated the link between credit quality and profitability.

Section 1.5 results concluded that, compared to developed countries, banks in developing countries are more vulnerable and receive additional adverse effects due to the growth of technology-based lending. By comparing every column (2) from Tables 9 to 12, I observe that banks in developed countries also see a decrease in demand and supply for their lending services as technology-based lending grows. This reduction reduces the capacity of existing gross loans, raising their NPL ratio. However, the increase in technology-based lending does not result in more non-performing loans among existing loans. On the other hand, banks in developing countries are also experiencing a significant drop in demand and supply for their lending services and a significant increase in non-performing loans, resulting in a decrease in profitability.

Such results reveal an interesting phenomenon: as technology-based lending develops and becomes a new type of financial service to meet customer needs, banks in developing countries tend to compete with technology-based lending and, to reclaim lost market share, are willing to issue higher-risk loans, causing significant harm to their profitability. Even though banks' lending services in developed countries have been significantly impacted, they continue to exercise relatively strict risk control over existing and potential loans, thereby controlling the increase in non-performing loans. As a result, the increase in technology-based lending has little effect on their profitability.

When considering the competitiveness of the banking sector in which a bank operates, I observe that banks in more competitive banking sectors experience a significant increase in NPLs, affecting their profitability. According to the results in Section 1.5.2, when a bank operates in a highly concentrated banking sector, its profitability have no influence due to technology-based lending growth. Based on the empirical results in Column (3) from Tables 9 to 12, demand for lending services falls significantly when a bank is located in a country with a highly concentrated banking sector. However, lending service on the supply side, non-performing loan levels and non-performing loan ratios remain uninfluenced. On the contrary, when banks operate in a competitive market, there is a significant decrease in the supply and demand of lending services and efficiency. Such findings imply that when banks operate in competitive markets, their loan service market share declines significantly due to the rise of technology-based lending. Banks may issue riskier loans in order to maintain their market share, decreasing in lending efficiency and profitability.

According to the empirical results, when a bank's operation scale is small, an increase in technology-based lending reduces demand for its lending services without affecting supply, lending efficiency, or profitability. When considering the sophistication of the credit market in the country where the bank operates, the result shows that when credit accessibility is low in the country where the bank operates, both the demand and supply sides of the bank's loan services suffer significantly. The non-performing loan ratio has also increased significantly in this situation.

1.7.2 FinTech and BigTech lending

Previous regression findings demonstrated the impact of technology-based lending on bank profitability in various ways. While the two major classifications in technology-based lending, which are FinTech and BigTech lending, have each emerged as compelling forces offering innovative financial solutions that challenge the status quo of the traditional banking sector, many scholars have begun to investigate the similarities and differences in the impact of these two types of lending on traditional banks. Understanding the differences of the impact of these two lending modalities on banks can assist regulators in developing effective policies, and help banks strategically address the competitive challenges of FinTech and BigTech lending.

Based on such consideration, I separate technology-based lending into FinTech and BigTech lending and reestimate the baseline model. In addition, I also consider the country-specific and bank-specific interaction terms to investigate whether BigTech and FinTech lending have different effects on bank profitability depending on certain conditions. The empirical results in Column (1) of Table 13 indicate that FinTech lending has no significant impact on bank profitability. On the contrary, a 1% increase in BigTech lending significantly reduces the bank's ROAA ratio by 6.835%. Such results suggest that an increase in BigTech lending has a more substantial and negative impact on technology-based lending on bank profitability compared with FinTech lending. Furthermore, the results in Column (2) show that the increase in BigTech lending significantly and negatively impacts profitability for banks in developing countries. This finding is consistent with previous findings indicating that neither BigTech nor FinTech lending significantly impacts developed-country bank profitability.

Columns (3) and (4) further confirm that BigTech lending has a significant impact on bank profitability, revealing that the growth of BigTech lending has a significant negative impact on banks' ROAA ratios when they operate in a country with a highly competitive or moderately concentrated banking industry, or when the banks themselves operate on a small scale. It is worth noting that Column (5) of Table 13 shows that when banks are located in countries with low credit accessibility, the growth of BigTech lending does not cause a significant additional negative impact on banks. However, each percent increase in FinTech lending will reduce the number of banks with low credit accessibility by 18.5%. As a result of the findings, BigTech lending is less prevalent when the level of country's credit accessibility is low, and the negative impact of FinTech lending on bank profitability is more negatively significant.

Such results are consistent with Stulz (2019) view and provide strong empirical results. Based on Stulz (2019)'s arguments, BigTech firms already have the customer base necessary to operate a platform bank. A platform bank competes with banks in all customer-facing activities, including deposits, payments, and wealth management. Generally, FinTech companies rely on traditional banks in their current operations. They deposit cash into bank accounts, have credit lines, and make payments through banks. With a platform bank, a BigTech company doesn't need to rely on an existing bank. It can have its subsidiary bank, through which it can offer deposit accounts, credit cards, and electronic cash to its customers. Additionally, BigTech firms can provide customers with various financial services from third parties, and have a potential advantage over banks and FinTech that could be enormous. Compared with FinTech firms, BigTech firms possess the technical expertise and cutting-edge systems that FinTech companies covet, have the same customer size as the largest banks, can access data not available to banks and FinTech companies, and do not have any legacy problems and organizational issues that banks initially have.

1.8 Conclusion

This paper pointed out the potential impact of growing technology-based lending on the profitability of traditional banks. I illustrated the potential relationship from a cross-country perspective based on the novel database constructed by Cornelli et al. (2020), focusing on the period from 2013 to 2019. Besides the baseline model, this paper also considers the factors that can strengthen such a relationship and the potential channels that may induce such a relationship.

As predicted by the model, I found that the increase in technology-based lending negatively influences banks' profitability from a cross-country perspective. Bank's ROAA ratio saw a 5.263% decrease after a 1% increase in technology-based lending, indicating an average $2,135 billion loss in the bank's net income. Such a relationship is also robust when considering other profitability indicators, such as ROAE and NIM ratio, and instrumental variables regression to address potential endogeneity. Compared with FinTech lending, BigTech lending dominated the influence of technology-based lending on bank profitability, showing a higher threat to banks. I also confirm that the characteristics of the country that the bank is located in are more accessible to induce banks to receive more negative shock from technology-based lending growth than the bank's operating characteristics. Based on the potential channel analysis, I find that technology-based lending lowers a bank's profitability mainly by increasing the non-performing loan volume.

My results imply that when facing the increase in technology-based lending, banks are more prone to reduce their profitability when they tend to increase loan issuance to maintain their market share of lending services. Meanwhile, the impact of FinTech lending on banks' profitability is insignificant because it is similar to banks' traditional lending business. In contrast, BigTech lending has a more significant impact on banks' profitability because it relies on BigTech firm's customer base and business to produce the service that banks cannot replicate easily.

Figures

Figure 1.1: Interaction between FinTech lending, customers and banks

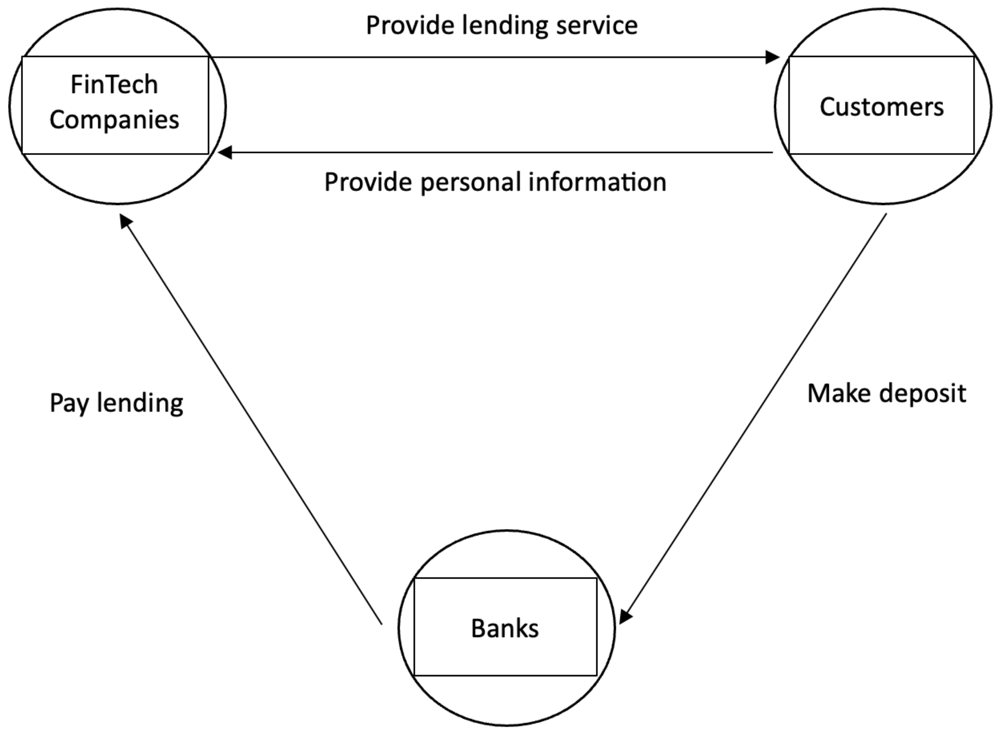


Figure 1.2: Interaction between BigTech lending and customers

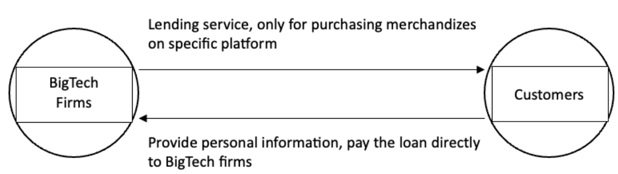
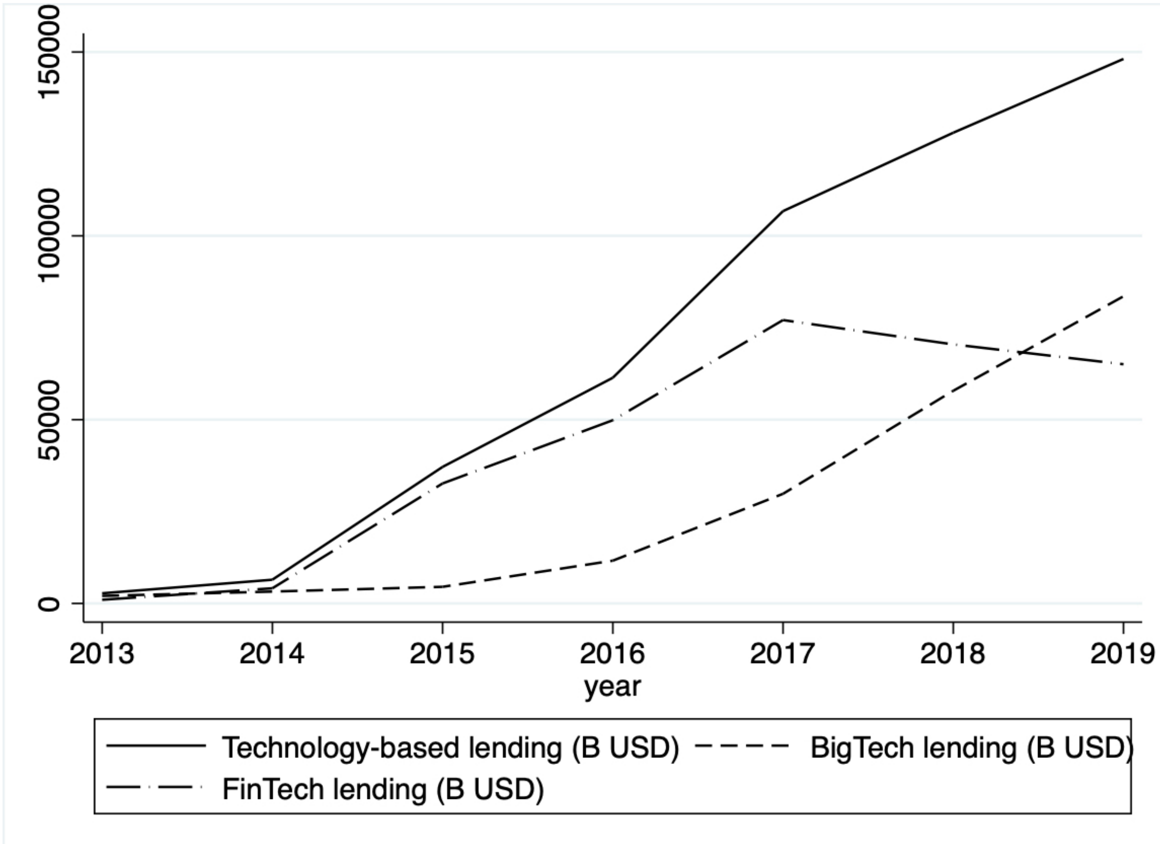
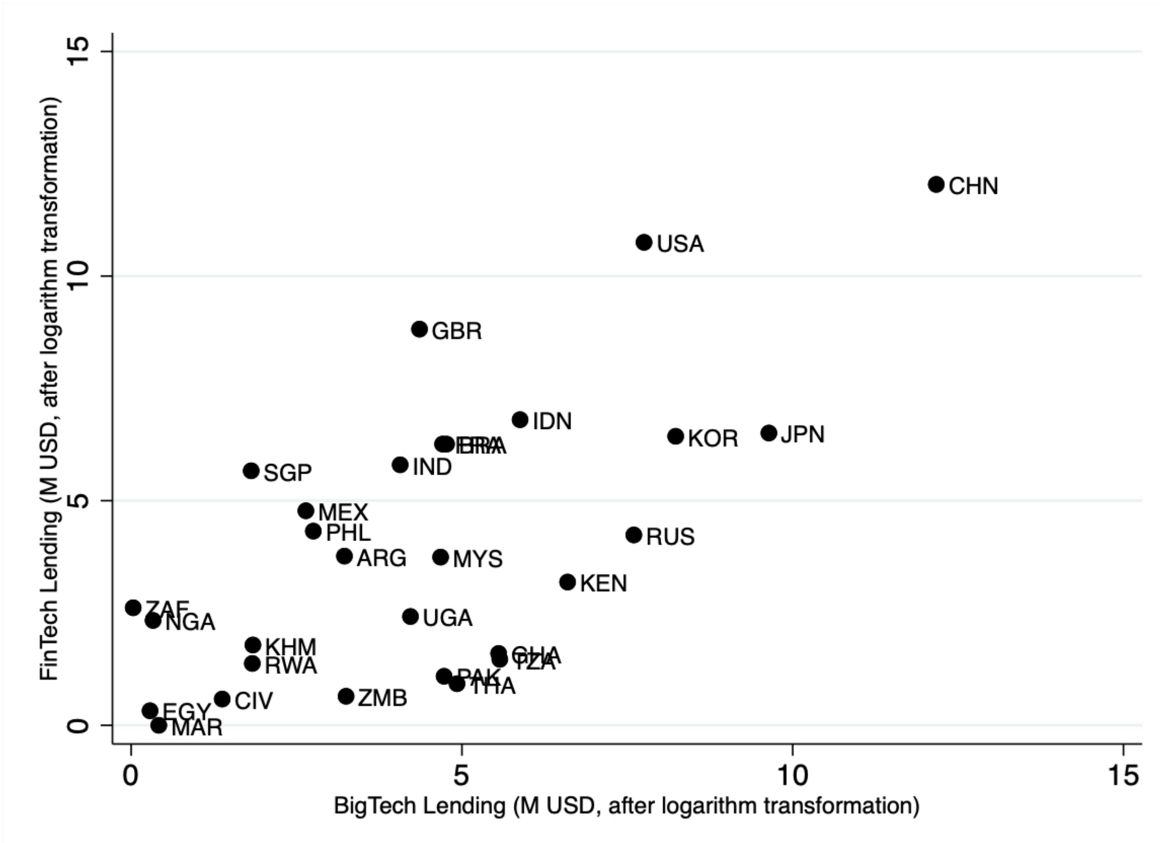


Figure 1.3: The total volume of technology-based lending in aggregate level



Note: This figure shows the growth trend of total volume of BigTech and FinTech credit based on the databased constructed by Cornelli et al. (2020).

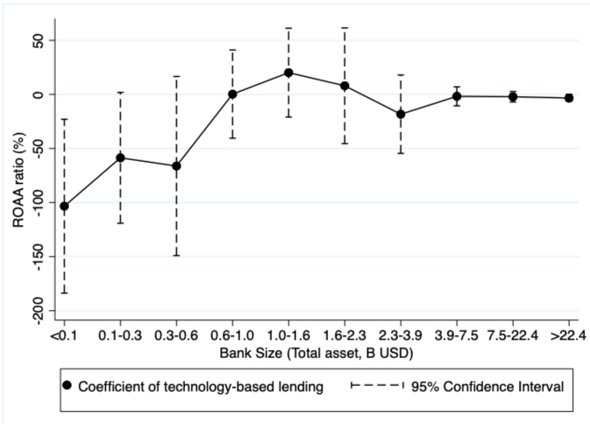
Figure 1.4: FinTech and BigTech lending volume in different countries



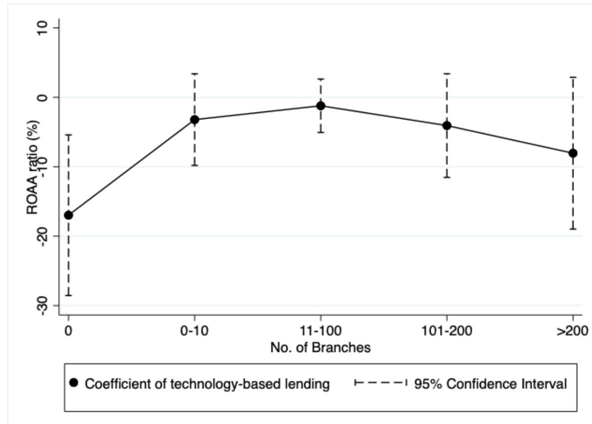
Note: This figure shows the distribution of BigTech and FinTech credit volume after logarithm transformation in a specific country. This figure omits countries without BigTech lending credit during 2013 to 2019. Country names are abbreviated to the ISO 3-digit code.

Figure 1.5: Heterogeneity in the influence of technology-based lending growth in bank's profitability by bank's operating scale

(a) Bank size



(b) Number of branches of banks



Note: These figures show the heterogeneity influence of technology-based lending growth in bank's ROAA ratio by bank's size and number of branches based on Equation 1.1. Figure (1.5a) plots the subsample coefficient β_1 estimation of the technology-based lending in the bank's ROAA ratio perspective, where the regression is run separately by 10 quantile subgroup of bank's total asset based on full sample. Figure (1.5b) plots the subsample coefficient β_1 estimation of the technology-based lending in the bank's ROAA ratio perspective, where the regression is run separately by 5 subgroup of the number of the bank's branches on full sample.

Tables

Table 1: Summary statistics

Variable	# of obs	# of countries	Mean	SD	Min	Max
Panel I: Dependent variables						
ROAA	19695	92	.646	.957	-2.89	4.45
ROAE	19695	92	5.841	7.472	-24.664	30.274
NIM	19695	92	3.084	2.821	.55	17.588
Panel II: Technology-based variables						
Ln(Tech / GDP)	19695	92	.002	.007	0	.043
Ln(BigTech / GDP)	19695	92	.001	.004	0	.035
Ln(FinTech / GDP)	19695	92	.001	.003	0	.029
Ln(Tech change / GDP)	18778	92	.001	.002	-.002	.019
Panel III: Bank-specific characters						
Size	19695	92	21.309	1.924	16.994	26.409
Liquidity	19695	92	23.062	17.06	2.186	77.445
Risk	19695	92	.584	1.426	-2.232	8.872
Capitalization	19695	92	11.105	5.748	3.007	42.88
Diversification	19695	92	30.58	16.244	-3.915	83.058
Financial leverage	19695	92	10.939	5.717	3.54	41.239
# of branch	19695	92	89.29	616.096	0	23682
Ln(Loan)	19695	92	20.745	1.967	16.013	25.749
Ln(Deposit)	19695	92	20.945	1.967	16.213	25.918
Ln(NPL)	19695	92	21.552	2.189	12.101	29.039
NPL ratio	19695	92	5.25	8.474	.02	59.085
Panel IV: Country-specific factors						
DCII	19695	92	7.119	1.337	0	8
SLRI	19695	92	5.828	2.624	0	12
GDP growth	19695	92	2.289	1.976	-27.995	24.37
Per Capita growth	19695	92	1.67	1.84	-29.922	23.201
Inflation rate	19695	92	2.217	3.082	-3.749	53.55
Ln(Broadband)	19553	92	3.196	.858	.001	3.869
Ln(Cellular)	19690	92	4.808	.195	3.477	5.66
Ln(Fixed telephone)	19679	92	16.694	1.834	0	19.403

Note(s): This table shows the total sample summary statistics for the bank-specific variables, technology-based lending related variables, macroeconomic variables and the variables that are used as instruments in the instrumental variable regressions throughout the paper. Detailed definitions and the sources of the variables are provided in Appendix Table A.3. The full sample contains 19,695 bank-year observations. Coverage: 2013-19.

Table 3: Impact of technology-based lending growth on bank's profitability in ROAA ratio perspective

	Dep Var: ROAA ratio						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ln(Tech/GDP)	-7.032***	-6.547***	-5.623***				
	(1.342)	(1.364)	(1.379)				
Ln(BigTech/GDP)				-8.291***	-7.807***	-6.835***	
				(1.585)	(1.598)	(1.602)	
Ln(FinTech/GDP)				-3.810**	-3.331*	-2.529	
				(1.842)	(1.919)	(1.951)	
Ln(Relative Growth/GDP)							-0.423
							(2.606)
Constant	0.661***	-0.962***	-1.182***	0.659***	-0.964***	-1.184***	-1.194***
	(0.010)	(0.179)	(0.181)	(0.010)	(0.179)	(0.181)	(0.181)
Observations	19,695	19,695	19,695	19,695	19,695	19,695	18,778
R-squared	0.303	0.328	0.333	0.303	0.328	0.333	0.329
Cluster variable	Bank	Bank	Bank	Bank	Bank	Bank	Bank
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Countries FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bank Characteristics	No	Yes	Yes	No	Yes	Yes	Yes
Macro Factors	No	No	Yes	No	No	Yes	Yes

Note(s): This table presents coefficient estimates for baseline models estimating the impact of technology-based lending volume growth on bank's ROAA ratio as the profitability measure, progressively adding the controls. Ln(Tech / GDP) denotes a continuous variable that the volume of technology-based lending to given country's GDP ratio in Ln(1+X) logarithm transformation. Columns (1) and (4) control for country and year fixed effects only. Columns (2) and (5) add bank-specific characteristic control variables, including bank's size, risk, liquidity, capitalization, diversification and financial leverage. Columns (3) and (6) add country-specific control variables, including GDP growth rate, GDP per capita growth rate and inflation rate. Detailed information for all control variables are included in Table 1. Columns (4), (5) and (6) separate the technology-based lending to two parts: FinTech and BigTech lending, represented by Ln(FinTech/GDP) and Ln(BigTech/GDP) respectively. Standard errors are clustered at the bank–level and in parentheses *** p<0.01, ** p < 0.05, * p < 0.10.

Table 4: Impact of technology-based lending growth on bank's profitability in ROAE and NIM ratio perspective

	Panel A: ROAE ratio				Panel B: NIM ratio			
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ln(Tech/GDP)	-95.49***	-86.71***			-16.82***	-14.42***		
	(15.06)	(14.20)			(2.791)	(2.586)		
Ln(BigTech/GDP)			-106.6***				-11.95***	
			(16.17)				(2.918)	
Ln(FinTech/GDP)			-36.81**				-19.64***	
			(17.82)				(3.702)	
Ln(Relative Growth/GDP)				21.54				-14.74***
				(23.98)				(4.664)
Constant	6.039***	-5.526***	-5.552***	-5.627***	3.119***	5.154***	5.156***	5.193***
	(0.0843)	(1.526)	(1.526)	(1.511)	(0.0257)	(0.386)	(0.386)	(0.391)
Observations	19,695	19,695	19,695	18,778	19,695	19,695	19,695	18,778
R-squared	0.301	0.334	0.334	0.330	0.610	0.689	0.689	0.688
Cluster variable	Bank	Bank	Bank	Bank	Bank	Bank	Bank	Bank
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Countries FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bank Characteristics	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Macro Factors	No	Yes	Yes	Yes	No	Yes	Yes	Yes

Note(s): This table presents coefficient estimates for baseline models estimating the impact of technology-based lending volume growth on bank's alternative profitability measure, such as ROAE and NIM ratio, with different controls consideration. Ln(Tech / GDP), Ln(BigTech / GDP), Ln(FinTech / GDP) and Ln(Relative Growth/GDP) denotes continuous variables that the volume of technology-based, BigTech, FinTech lending and relative technology-based lending change to given country's GDP ratio in Ln(1+X) logarithm transformation. Columns (1) and (5) control for country and year fixed effects only. Columns (2) and (6) add bank-specific and country-specific characteristic control variables as provided in Table 1. Columns (3) and (7) separate the technology-based lending to BinTech lending and FinTech lending, respectively. Columns (4) and (8) consider the relative technology-based lending change instead of the total credit volume. The dependent variable of Columns (1) to (4) is ROAE ratio, and Columns (5) to (8) is NIM ratio, respectively. Standard errors are clustered at the bank–level and in parentheses *** $p<0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 5: Addressing endogeneity: 2SLS regression approach

	Dep Var: ROAA ratio			
	(1)	(2)	(3)	(4)
Ln(Tech/GDP)	-12.698***	-17.115***	-12.058**	-17.540***
	(4.690)	(4.502)	(4.671)	(4.823)
Observations	19,548	19,548	18,641	18,641
K–P LM Statistics	237.580	292.684	212.830	280.095
Hansen J Statistics	20.127	25.239	24.659	34.685
C-D Wald F Statistics	3333.635	2555.929	2523.116	2026.050
R-squared	0.332	0.333	0.331	0.331
Cluster variable	Bank	Bank	Bank	Bank
Year FEs and countries FE	Yes	Yes	Yes	Yes
Infrastructure Instrument	Yes	Yes	Yes	Yes
Financial Environment	No	Yes	No	Yes
Bank Characteristics	Yes	Yes	Yes	Yes
Macro Factors	Yes	Yes	Yes	Yes

Note: This table illustrates the comparison of the estimation results between the β_1 in Equation 1.1 and 1.6 considering the instrument variables for technology-based lending growth. K-P LM statistics represents the Kleibergen-Paap rk LM statistic, and the null hypothesis is the equation is under-identified. The null hypothesis of Hansen J statistic is the over-identifying restrictions are valid. C-D Wald F statistics represents the Cragg-Donald Wald F statistic, the the null hypothesis is the instrument variables are weak instruments. Standard errors are clustered at the bank-level and in parentheses *** p<0.01, **p<0.05, *p<0.10.

Table 6: Effect of technology-based lending growth on bank's profitability by bank's operating scale

	Dep Var: ROAA ratio		
	(1)	(2)	(3)
Ln(Tech/GDP)	-59.25**	-4.545***	-6.022***
	(26.28)	(1.350)	(1.461)
Developed × Ln(Tech/GDP)	-6.015	-21.88*	-13.37
	(20.49)	(11.63)	(11.52)
Small Bank × Ln(Tech/GDP)	-3.199**		
	(1.368)		
Developed × Small Bank × Ln(Tech/GDP)	-11.98		
	(10.90)		
Few Branches × Ln(Tech/GDP)		-12.76***	
		(3.802)	
Developed × Few Branches × Ln(Tech/GDP)		16.49	
		(14.23)	
Small Scale ×Ln(Tech/GDP)			-2.130
			(3.633)
Developed × Small Scale × Ln(Tech/GDP)			27.30*
			(16.58)
Constant	-1.102***	-1.204***	-1.156***
	(0.186)	(0.184)	(0.181)
Observations	19,695	19,695	19,695
R-squared	0.336	0.334	0.333
Cluster variable	Bank	Bank	Bank
Year FEs	Yes	Yes	Yes
Countries FEs	Yes	Yes	Yes
Bank Characteristics	Yes	Yes	Yes
Macro Factors	Yes	Yes	Yes

Note(s): This table presents coefficient estimates β_1 for models estimating the impact of technology-based lending volume growth on bank's profitability by different financial inclusion identification. Ln(Tech / GDP) denotes a continuous variable that the volume of technology-based lending to given country's GDP ratio in Ln(1+X) logarithm transformation. Small Bank indicator variable equals to 1 when the total asset of bank i in country c at year t is smaller than 1 Billion USD. Few Branch indicator variable equals to 1 when the total branches of bank i in country c at year t is smaller than or equal to 10. Small Scale indicator variable equals to 1 when the total branches of bank i in country c at year t is smaller than or equal to 10 and the total asset is also smaller than 1 Billion USD.The baseline regression observations are banks in developing countries without small banks nor few branches characteristics. Detailed information for all control and indicator variables are included in Table 1. Standard errors are clustered at the bank–level and in parentheses *** p<0.01, ** p < 0.05, * p < 0.10.

Table 7: Effect of technology-based lending growth on bank's profitability by bank's competition level

	Dep var: ROAA ratio	Dep var: ROAE ratio	Dep var: NIM ratio
	(1)	(2)	(3)
Competitive × Ln(Tech/GDP)	-4.328***	-79.78***	-14.77***
	(1.282)	(13.99)	(2.555)
Competitive × Developed × Ln(Tech/GDP)	-14.59	-301.7***	-33.74
	(11.55)	(92.98)	(21.06)
Moderated × Ln(Tech/GDP)	-130.2***	-960.7***	-11.30
	(37.53)	(261.6)	(47.91)
Moderated × Developed × Ln(Tech/GDP)	25.89	90.36	-120.7
	(47.31)	(308.9)	(81.41)
Concentrated × Ln(Tech/GDP)	-67.02	-278.2	-11.85
	(77.69)	(344.4)	(98.40)
Concentrated × Developed × Ln(Tech/GDP)	-16.80	-89.76	22.89
	(15.99)	(138.9)	(34.58)
Constant	-1.171***	-5.349***	5.169***
	(0.182)	(1.531)	(0.388)
Observations	19,695	19,695	19,695
R-squared	0.335	0.336	0.689
Cluster variable	Bank	Bank	Bank
Year FEs	Yes	Yes	Yes
Countries FEs	Yes	Yes	Yes
Bank Characteristics	Yes	Yes	Yes
Macro Factors	Yes	Yes	Yes

Note(s): This table presents coefficient estimates β_1 for models estimating the impact of technology-based lending volume growth on bank's profitability by the competition level of the banking sector in the country that the bank locate in. Ln(Tech / GDP) denotes a continuous variable that the volume of technology-based lending to given country's GDP ratio in Ln(1+X) logarithm transformation. Competitive indicator variable equals to 1 when the Herfindahl–Hirschman index (HHI index) of the country c that bank i locate in is smaller than 1500. Moderated indicator variable equals to 1 when the HHI index of the country c that bank i locate in is larger than 1500 and smaller than 2500. Concentrated indicator variable equals to 1 when the HHI index of the country c that bank i locate in is lager than 2500. Detailed information for all control and indicator variables are included in Table 1. Standard errors are clustered at the bank−level and in parentheses *** p<0.01, ** p < 0.05, * p < 0.10.

Table 8: Effect of technology-based lending growth on bank's ROAA ratio by countries' credit accessibility level

	Dep Var: ROAA ratio		
	(1)	(2)	(3)
Ln(Tech/GDP)	-27.55	-6.329***	-7.227***
	(16.75)	(1.746)	(1.841)
Developed × Ln(Tech/GDP)	-7.749	-6.803	-3.820
	(15.96)	(15.76)	(15.61)
Low SLRI × Ln(Tech/GDP)	-4.713***		
	(1.220)		
Low SLRI × Developed × Ln(Tech/GDP)	-7.642		
	(11.92)		
Low DCII × Ln(Tech/GDP)		-13.39	
		(8.797)	
Low DCII × Developed × Ln(Tech/GDP)		-2.210	
		(9.393)	
Low credit accessibility × Ln(Tech/GDP)			-21.26***
			(7.762)
Low credit accessibility × Developed × Ln(Tech/GDP)			-5.821
			(9.381)
Constant	-1.181***	-1.182***	-1.179***
	(0.182)	(0.182)	(0.182)
Observations	19,695	19,695	19,695
R-squared	0.333	0.333	0.333
Cluster variable	Bank	Bank	Bank
Year FEs	Yes	Yes	Yes
Countries FEs	Yes	Yes	Yes
Bank Characteristics	Yes	Yes	Yes
Macro Factors	Yes	Yes	Yes

Note(s): This table presents coefficient estimates β_1 for models estimating the impact of technology-based lending volume growth on bank's ROAA ratio by different level of accessibility of credit in different countries. Ln(Tech / GDP) denotes a continuous variable that the volume of technology-based lending to given country's GDP ratio in Ln(1+X) logarithm transformation. Developed indicator variable equals to 1 when the country that the bank locate in is a developed country. Low DCII (The Depth of Credit Information Index) indicator variable equals to 1 when the DCII in country c at year t is smaller than or equal to 6. Low SLRI (The Strength of Legal Right Index) indicator variable equals to 1 when the SLRI in country c at year t is smaller than or equal to 5. Low DCII and SLRI indicator variable equals to 1 when theDCII in country c at year t is smaller than or equal to 6 and the SLRI in country c at year t is smaller than or equal to 5. The baseline observation of regression are banks in the developing countries that have high SLRI and DCII. Detailed information for all control and indicator variables are included in Table 1. Standard errors are clustered at the bank−level and in parentheses *** $p<0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 9: Alternative mechanisms estimation: bank's lending service in demand side

	Dep Var: Ln(Loan)				
	(1)	(2)	(3)	(4)	(5)
$Ln(Tech/GDP)$	-2.886***	-2.868***	-2.690***	-3.037***	-3.185***
	(0.556)	(0.557)	(0.558)	(0.547)	(0.656)
DEVELOPED $\times Ln(Tech/GDP)$		-15.70***			
		(4.271)			
MODERATED CONCENTRATED $\times Ln(Tech/GDP)$			-13.85**		
			(5.617)		
HIGHLY CONCENTRATED $\times Ln(Tech/GDP)$			-12.89***		
			(4.785)		
SMALL SCALE $\times Ln(Tech/GDP)$				-1.571	
				(1.555)	
LOW ACCESSIBILITY $\times Ln(Tech/GDP)$					-5.772**
					(2.852)
CONSTANT	-0.057	-0.048	-0.056	-0.055	-0.054
	(0.076)	(0.076)	(0.076)	(0.0755)	(0.076)
OBSERVATIONS	19,695	19,695	19,695	19,695	19,695
R-SQUARED	0.979	0.979	0.979	0.979	0.979
CLUSTER VARIABLE	BANK	BANK	BANK	BANK	BANK
YEAR FES	YES	YES	YES	YES	YES
COUNTRIES FES	YES	YES	YES	YES	YES
BANK CHARACTERISTICS	YES	YES	YES	YES	YES
MACRO FACTORS	YES	YES	YES	YES	YES

Note: This table estimates the alternative mechanism of the impact of technology-based lending growth on bank's lending service. The dependent variables $Ln(Loan)$ is the logarithm of bank's gross loan in Columns (1) and (2), and $Ln(Deposit)$ is the logarithm of bank's total deposit volume in Columns (3) to (4). All regression control for year and countries fixed effect. All bank-specific and country-specific control variables follow the setup of baseline model. Robust standard errors clustered at the bank level are reported in parentheses.Standard errors are clustered at the bank–level and in parentheses *** p<0.01, ** p < 0.05, * p < 0.10.

Table 10: Alternative mechanisms estimation: bank's lending service in supply side

	Dep Var: Ln(Deposit)				
	(1)	(2)	(3)	(4)	(5)
$Ln(Tech/GDP)$	-3.906***	-3.884***	-3.756***	-3.968***	-4.016***
	(0.521)	(0.520)	(0.517)	(0.539)	(0.621)
Developed $\times Ln(Tech/GDP)$		-19.63***			
		(4.953)			
Moderated Concentrated $\times Ln(Tech/GDP)$			-15.96***		
			(5.612)		
Highly Concentrated $\times Ln(Tech/GDP)$			-8.489		
			(7.395)		
Small Scale $\times Ln(Tech/GDP)$				-3.367**	
				(1.676)	
Low accessibility $\times Ln(Tech/GDP)$					-4.968*
					(2.561)
Constant	1.088***	1.098***	1.089***	1.088***	1.089***
	(0.123)	(0.124)	(0.123)	(0.124)	(0.124)
Observations	19,695	19,695	19,695	19,695	19,695
R-squared	0.963	0.963	0.963	0.963	0.963
Cluster variable	Bank	Bank	Bank	Bank	Bank
Year FEs	Yes	Yes	Yes	Yes	Yes
Countries FEs	Yes	Yes	Yes	Yes	Yes
Bank Characteristics	Yes	Yes	Yes	Yes	Yes
Macro Factors	Yes	Yes	Yes	Yes	Yes

Note: This table estimates the alternative mechanism of the impact of technology-based lending growth on bank's lending service. The dependent variables $Ln(Deposit)$ is the logarithm of bank's total deposit volume in Columns (1) to (5). All regression control for year and countries fixed effect. All bank-specific and country-specific control variables follow the setup of baseline model. Robust standard errors clustered at the bank level are reported in parentheses.Standard errors are clustered at the bank–level and in parentheses *** p<0.01, ** p < 0.05, * p < 0.10.

Table 11: Alternative mechanisms estimation: bank's lending efficiency in non-performing loan volume perspectives

	DV = Ln(Non-performing loan)				
	(1)	(2)	(3)	(4)	(5)
$Ln(Tech/GDP)$	13.93***	13.98***	13.99***	15.30***	14.39***
	(1.958)	(1.962)	(1.949)	(2.006)	(2.363)
Developed $\times Ln(Tech/GDP)$		-15.85			
		(16.56)			
Moderated Concentrated $\times Ln(Tech/GDP)$			0.699		
			(27.01)		
Highly Concentrated $\times Ln(Tech/GDP)$			19.60		
			(21.65)		
Small Scale $\times Ln(Tech/GDP)$				2.087	
				(4.396)	
Low accessibility $\times Ln(Tech/GDP)$					18.32*
					(9.973)
Constant	1.230***	1.250***	1.231***	1.218***	1.227***
	(0.279)	(0.280)	(0.279)	(0.279)	(0.280)
Observations	19,695	19,695	19,695	19,695	19,695
R-squared	0.766	0.766	0.766	0.766	0.766
Cluster variable	Bank	Bank	Bank	Bank	Bank
Year FEs	Yes	Yes	Yes	Yes	Yes
Countries FEs	Yes	Yes	Yes	Yes	Yes
Bank Characteristics	Yes	Yes	Yes	Yes	Yes
Macro Factors	Yes	Yes	Yes	Yes	Yes

Note: This table estimates the alternative mechanism of the impact of technology-based lending growth on bank's lending service. The dependent variable Ln(Non-performing loan) is a non-performing loan in natural logarithm transformation, representing bank's lending efficiency. Column (1) illustrates the estimation results of baseline model, changing to the dependent variable indicating bank's profitability to bank's non-performing loan after transformation. The baseline observation of Columns (2) to (5) are banks in developing countries, banks in a country with a competitive banking sector, banks with non small operating scale, and banks in the country with non low credit accessibility, respectively. All bank-specific and country-specific control variables follow the setup of baseline model. Robust standard errors clustered at the bank level are reported in parentheses.Standard errors are clustered at the bank–level and in parentheses *** p<0.01, ** p < 0.05, * p < 0.10.

Table 12: Alternative mechanisms estimation: bank's lending efficiency in NPL ratio perspective

	DV = NPL ratio				
	(1)	(2)	(3)	(4)	(5)
$Ln(Tech/GDP)$	45.19***	44.87***	48.40***	51.51***	44.84***
	(9.539)	(9.448)	(8.573)	(9.963)	(12.63)
Developed $\times Ln(Tech/GDP)$		270.4**			
		(119.1)			
Moderated Concentrated $\times Ln(Tech/GDP)$			-132.1		
			(355.3)		
Highly Concentrated $\times Ln(Tech/GDP)$			-119.3		
			(242.2)		
Small Scale $\times Ln(Tech/GDP)$				-9.686	
				(20.15)	
Low accessibility $\times Ln(Tech/GDP)$					41.78
					(47.18)
Constant	10.81***	10.65***	10.82***	10.75***	10.81***
	(1.624)	(1.622)	(1.627)	(1.623)	(1.626)
Observations	19,695	19,695	19,695	19,695	19,695
R-squared	0.426	0.426	0.426	0.426	0.426
Cluster variable	Bank	Bank	Bank	Bank	Bank
Year FEs	Yes	Yes	Yes	Yes	Yes
Countries FEs	Yes	Yes	Yes	Yes	Yes
Bank Characteristics	Yes	Yes	Yes	Yes	Yes
Macro Factors	Yes	Yes	Yes	Yes	Yes

Note: This table estimates the alternative mechanism of the impact of technology-based lending growth on bank's lending service. The dependent variables NPL ratio is a non-performing loan to gross loan ratio, representing bank's lending efficiency. Column (1) illustrates the estimation results of baseline model, separating the technology-based lending into FinTech and BigTech lending. The baseline observation of Column (2) to (5) are banks in developing countries, banks in a country with a competitive banking sector, banks with non small operating scale, and banks in the country with non low credit accessibility, respectively. All regression control for year and countries fixed effect. All bank-specific and country-specific control variables follow the setup of baseline model. Robust standard errors clustered at the bank level are reported in parentheses.Standard errors are clustered at the bank–level and in parentheses *** p<0.01, ** p < 0.05, * p < 0.10.

Table 13: The impact of FinTech and BigTech lending growth on bank profitability in ROAA ratio perspective

	Dep Var: ROAA ratio				
	(1)	(2)	(3)	(4)	(5)
$Ln(FinTech/GDP)$	-2.529	-2.460	-0.972	-1.741	-4.054*
	(1.951)	(2.040)	(1.503)	(2.173)	(2.338)
$Ln(BigTech/GDP)$	-6.835***	-6.865***	-5.202***	-8.514***	-8.318***
	(1.602)	(1.615)	(1.499)	(1.823)	(1.996)
Developed × $Ln(FinTech/GDP)$		-4.230			
		(17.71)			
Developed × $Ln(BigTech/GDP)$		-5.909			
		(13.13)			
Moderate concentrated × $Ln(FinTech/GDP)$			85.01**		
			(37.04)		
Highly concentrated × $Ln(FinTech/GDP)$			-39.64		
			(34.32)		
Moderate concentrated × $Ln(BigTech/GDP)$			-142.5***		
			(39.83)		
Highly concentrated × $Ln(BigTech/GDP)$			20.20		
			(25.07)		
Small scale × $Ln(FinTech/GDP)$				-6.357	
				(10.30)	
Small scale × $Ln(BigTech/GDP)$				4.715	
				(5.770)	
Low accessibility × $Ln(FinTech/GDP)$					-18.50**
					(8.840)
Low accessibility × $Ln(BigTech/GDP)$					-6.000
					(10.06)
Constant	-1.184***	-1.184***	-1.177***	-1.184***	-1.180***
	(0.181)	(0.182)	(0.181)	(0.181)	(0.182)
Observations	19,695	19,695	19,695	19,695	19,695
R-squared	0.333	0.333	0.335	0.333	0.333
Cluster variable	Bank	Bank	Bank	Bank	Bank
Year FEs	Yes	Yes	Yes	Yes	Yes
Countries FEs	Yes	Yes	Yes	Yes	Yes
Bank Characteristics	Yes	Yes	Yes	Yes	Yes
Macro Factors	Yes	Yes	Yes	Yes	Yes

Note: This table estimates the alternative mechanism of the impact of technology-based lending growth on bank's lending service. The dependent variables ROAA ratio is the main dependent variable that represent the profitability of the bank. The main independent variables are $Ln(FinTech/GDP)$ and $Ln(BigTech/GDP)$, which are the natural logarithm transformation the FinTech lending and BigTech lending volume of a specific country. Column (1) illustrates the estimation results of baseline model, separating the technology-based lending into FinTech and BigTech lending. The baseline observation of Columns (2) to (5) are banks in developing countries, banks in a country with a competitive banking sector, banks with non small operating scale, and banks in the country with non low credit accessibility, respectively. All regression control for year and countries fixed effect. All bank-specific and country-specific control variables follow the setup of baseline model. Robust standard errors clustered at the bank level are reported in parentheses.Standard errors are clustered at the bank−level and in parentheses *** p<0.01, ** p < 0.05, * p < 0.10.

Appendix Tables

Table A.1: Correlation matrix for bank-specific characteristic variables

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
(3) NIM ratio	0.434	0.288	1.000											
(1) ROAA ratio	1.000													
(2) ROAE ratio	0.886	1.000												
(4) Size	0.075	0.225	-0.271	1.000										
(5) Liquidity	0.125	0.099	0.169	-0.042	1.000									
(6) Risk	-0.044	-0.097	0.454	-0.040	0.179	1.000								
(7) Capitalization	0.252	0.008	0.504	-0.409	0.158	0.173	1.000							
(8) Diversification	-0.026	-0.051	-0.163	-0.088	0.114	0.031	0.039	1.000						
(9) Financial Leverage	0.303	0.031	0.529	-0.397	0.198	0.169	0.922	0.037	1.000					
(10) # of Branch	0.012	0.039	-0.010	0.234	-0.008	0.049	-0.052	0.002	-0.055	1.000				
(11) Ln(Loan)	0.067	0.215	-0.272	0.981	-0.158	-0.053	-0.411	-0.118	-0.404	0.228	1.000			
(12) Ln(Deposit)	0.054	0.216	-0.289	0.977	-0.071	-0.070	-0.463	-0.113	-0.448	0.227	0.964	1.000		
(13) Ln(NPL)	0.022	0.101	-0.094	0.773	0.052	0.167	-0.275	-0.041	-0.265	0.214	0.771	0.741	1.000	
(14) NPL ratio	-0.048	-0.157	0.294	-0.238	0.337	0.366	0.276	0.113	0.286	-0.004	-0.272	-0.273	0.207	1.000

Table A.3: Variable Definitions an Source

Variable	Definition	Source
Panel I: Dependent variables		
ROAA	Return of Average Asset (%)	BankFocus
ROAE	Return of Average Equity (%)	BankFocus
NIM	Net Interest Margin (%)	BankFocus
Panel II: Technology-based variables		
Ln(Tech / GDP)	Natural logarithm of total Technology-based lending credit / Country GDP, % (USD)	Cornelli et al. (2020)
Ln(BigTech / GDP)	Natural logarithm of BigTech lending Credit / Country GDP, % (USD)	Cornelli et al. (2020)
Ln(FinTech / GDP)	Natural logarithm of FinTech lending Credit / Country GDP, % (USD)	Cornelli et al. (2020)
Ln(Tech change / GDP)	Natural logarithm of total Technology-based lending credit change / Country GDP, % (USD)	Cornelli et al. (2020)
Panel III: Bank-specific characters		
Size	Natural logarithm of of bank's total asset (USD)	BankFocus
Liquidity	Liquid Asset / Total Asset (%)	BankFocus
Risk	Loan loss provision to gross loan (%)	BankFocus
Capitalization	Total capital to Total Asset (%)	BankFocus
Diversification	Non-Interest Income to Operating Revenue (%)	BankFocus
Financial leverage	Total equity to Total Asset (%)	BankFocus
# of branch	The number of branches that the bank own	BankFocus
Ln(Loan)	Natural logarithm of the volume of bank's gross loan (USD)	BankFocus
Ln(Deposit)	Natural logarithm of the volume of bank's total deposit (USD)	BankFocus
Ln(NPL)	Natural logarithm of the volume of bank's non-performing loan (USD)	BankFocus
NPL ratio	Non-Performing Loan Ratio (%)	BankFocus
Panel IV: Macroeconomic factors		
DCII	Depth of Credit Information Index (0 = low to 8 = High)	WorldBank
SLRI	Strength of Legal Right Index (0 = weak to 10 = strong)	WorldBank
GDP growth	GDP growth rate (%)	WorldBank
Per Capita growth	GDP per capita growth rate (%)	WorldBank
Inflation rate	Inflation rate (%)	WorldBank
Panel V: Instrument variables		
Ln(Broadband)	Natural Logarithm of fixed broadband subscriptions (per 100 people)	WorldBank
Ln(Cellular)	Natural Logarithm of mobile cellular subscriptions (per 100 people)	WorldBank
Ln(Fixed telephone)	Natural logarithm of Fixed telephone subscriptions (per 100 people)	WorldBank

Table A.5: Number of banks and income level of countries

ISO code	Country Name	Region	Income Level	# banks	Commercial bank	Cooperative bank	Savings bank
ARE	United Arab Emirates	Middle East & North Africa	High income	10	10		
ARG	Argentina	Latin America & Caribbean	Upper middle income	23	20	1	1
AUS	Australia	East Asia & Pacific	High income	55	20	34	
AUT	Austria	Europe & Central Asia	High income	11	5	3	2
BDI	Burundi	Sub-Saharan Africa	Low income	1	1		
BEL	Belgium	Europe & Central Asia	High income	12	8	1	3
BFA	Burkina Faso	Sub-Saharan Africa	Low income	3	3		
BGD	Bangladesh	South Asia	Lower middle income	35	35		
BGR	Bulgaria	Europe & Central Asia	Upper middle income	14	13		1
BHR	Bahrain	Middle East & North Africa	High income	4	4		
BOL	Bolivia	Latin America & Caribbean	Lower middle income	13	13		
BRA	Brazil	Latin America & Caribbean	Upper middle income	58	39	18	
CAN	Canada	North America	High income	34	20	12	2
CHE	Switzerland	Europe & Central Asia	High income	4	3		
CHL	Chile	Latin America & Caribbean	High income	13	11	1	
CHN	China	East Asia & Pacific	Upper middle income	144	140	3	1
CMR	Cameroon	Sub-Saharan Africa	Lower middle income	1	1		
COD	Congo	Sub-Saharan Africa	Lower middle income	1	1		
COL	Colombia	Latin America & Caribbean	Upper middle income	15	13	1	
CPV	Cabo Verde	Sub-Saharan Africa	Lower middle income	5	5		
CRI	Costa Rica	Latin America & Caribbean	Upper middle income	8	7		
DEU	Germany	Europe & Central Asia	High income	7	7		

ISO code	Country Name	Region	Income Level	# banks	Commercial bank	Cooperative bank	Savings bank
DNK	Denmark	Europe & Central Asia	High income	5	5		
ECU	Ecuador	Latin America & Caribbean	Upper middle income	29	17	11	1
EGY	Egypt	Middle East & North Africa	Lower middle income	20	20		
ESP	Spain	Europe & Central Asia	High income	35	14	18	3
EST	Estonia	Europe & Central Asia	High income	7	6		
FIN	Finland	Europe & Central Asia	High income	5	2	1	1
FRA	France	Europe & Central Asia	High income	81	18	58	3
GBR	United Kingdom	Europe & Central Asia	High income	31	24		
GEO	Georgia	Europe & Central Asia	Upper middle income	9	9		
GHA	Ghana	Sub-Saharan Africa	Lower middle income	19	17		2
GTM	Guatemala	Latin America & Caribbean	Upper middle income	7	7		
HKG	Hong Kong	East Asia & Pacific	High income	22	19		
IDN	Indonesia	East Asia & Pacific	Lower middle income	77	77		
IND	India	South Asia	Lower middle income	42	38	4	
IRL	Ireland	Europe & Central Asia	High income	4	4		
IRQ	Iraq	Middle East & North Africa	Upper middle income	1	1		
ISR	Israel	Middle East & North Africa	High income	9	9		
ITA	Italy	Europe & Central Asia	High income	309	60	236	9
JOR	Jordan	Middle East & North Africa	Upper middle income	13	13		
JPN	Japan	East Asia & Pacific	High income	363	97	263	
KAZ	Kazakhstan	Europe & Central Asia	Upper middle income	13	13		
KEN	Kenya	Sub-Saharan Africa	Lower middle income	19	18	1	
KHM	Cambodia	East Asia & Pacific	Lower middle income	7	7		
KOR	South Korea	East Asia & Pacific	High income	8	7		
LAO	Lao	East Asia & Pacific	Lower middle income	1	1		

ISO code	Country Name	Region	Income Level	# banks	Commercial bank	Cooperative bank	Savings bank
LBN	Lebanon	Middle East & North Africa	Upper middle income	22	22		
LBR	Liberia	Sub-Saharan Africa	Low income	2	2		
LTU	Lithuania	Europe & Central Asia	High income	2	2		
LVA	Latvia	Europe & Central Asia	High income	7	7		
MAR	Morocco	Middle East & North Africa	Lower middle income	5	5		
MDG	Madagascar	Sub-Saharan Africa	Low income	1	1		
MEX	Mexico	Latin America & Caribbean	Upper middle income	44	33	1	
MLI	Mali	Sub-Saharan Africa	Low income	5	5		
MNG	Mongolia	East Asia & Pacific	Lower middle income	2	2		
MOZ	Mozambique	Sub-Saharan Africa	Low income	8	8		
MWI	Malawi	Sub-Saharan Africa	Low income	4	4		
MYS	Malaysia	East Asia & Pacific	Upper middle income	23	22		
NGA	Nigeria	Sub-Saharan Africa	Lower middle income	13	13		
NLD	Netherlands	Europe & Central Asia	High income	11	9		
NOR	Norway	Europe & Central Asia	High income	81	8		73
NZL	New Zealand	East Asia & Pacific	High income	9	7	2	
PAK	Pakistan	South Asia	Lower middle income	20	20		
PAN	Panama	Latin America & Caribbean	Upper middle income	39	32		
PER	Peru	Latin America & Caribbean	Upper middle income	16	11		5
PHL	Philippines	East Asia & Pacific	Lower middle income	19	16		3
POL	Poland	Europe & Central Asia	High income	15	14	1	
PRT	Portugal	Europe & Central Asia	High income	91	13	5	72
PRY	Paraguay	Latin America & Caribbean	Upper middle income	8	8		
PSE	Palestine	Middle East & North Africa	High income	2	2		
RUS	Russia	Europe & Central Asia	Upper middle income	260	258		1

ISO code	Country Name	Region	Income Level	# banks	Commercial bank	Cooperative bank	Savings bank
RWA	Rwanda	Sub-Saharan Africa	Low income	6	6		
SAU	Saudi Arabia	Middle East & North Africa	High income	6	6		
SEN	Senegal	Sub-Saharan Africa	Lower middle income	3	3		
SGP	Singapore	East Asia & Pacific	High income	4	3		
SLE	Sierra Leone	Sub-Saharan Africa	Low income	1	1		
SLV	El Salvador	Latin America & Caribbean	Lower middle income	1	1		
SVK	Slovakia	Europe & Central Asia	High income	8	8		
SVN	Slovenia	Europe & Central Asia	High income	9	7	1	1
SWE	Sweden	Europe & Central Asia	High income	7	5		1
SWZ	Eswatini	Sub-Saharan Africa	Lower middle income	1	1		
TGO	Togo	Sub-Saharan Africa	Low income	1			
THA	Thailand	East Asia & Pacific	Upper middle income	6	4		
TUR	Turkey	Europe & Central Asia	Upper middle income	25	24	1	
TZA	Tanzania	Sub-Saharan Africa	Lower middle income	20	20		
UGA	Uganda	Sub-Saharan Africa	Low income	14	13		1
USA	United States	North America	High Income	693	400	1	52
VNM	VietNams	East Asia & Pacific	Lower middle income	6	6		
ZAF	South Africa	Sub-Saharan Africa	Upper middle income	12	7		
ZMB	Zambia	Sub-Saharan Africa	Lower middle income	12	12		
ZWE	Zimbabwe	Sub-Saharan Africa	Lower middle income	12	10		1

Table A.7: The Net income (in M USD) as a function of bank-specific and country-specific characteristics

	DV: Net Income				
	(1)	(2)	(3)	(4)	(5)
Ln(Tech/GDP)	-2017676**	-2169504***	-2135462**		
	(834,483)	(828,201)	(829,878)		
Ln(Relative Growth/GDP)				-647852.1	
				(419,317)	
Ln(BigTech/GDP)					-1979805**
					(771,856)
Ln(FinTech/GDP)					-2434181**
					(978,414)
Size		19,144***	19,168***	19,271***	19,169***
		(3,201)	(3,202)	(3,213)	(3,202)
Liquidity		-1,156**	-1,043**	-1,012*	-1,045**
		(521.5)	(516.4)	(533.8)	(516.7)
Risk		75.41	70.00	78.43	69.52
		(167.7)	(167.8)	(169.3)	(167.9)
Diversification		440.3	461.4	446.3	459.9
		(447.4)	(447.5)	(411.4)	(447.3)
Capitalization		225.5**	224.6**	198.1**	225.1**
		(97.38)	(97.32)	(96.17)	(97.22)
Financial leverage		900.6*	883.9*	935.5**	885.0*
		(477.9)	(477.2)	(474.1)	(477.2)
GDP growth rate			2,281***	1,551***	2,281***
			(536.3)	(391.3)	(536.5)
Inflation rate			157.8	-100.5	150.3
			(204.3)	(198.0)	(203.7)
GDP per cap growth rate			-96.50	267.8	-97.39
			(623.4)	(369.8)	(623.8)
Constant	19,790***	-409,540***	-415,498***	-419,326***	-415,350***
	(3,658)	(69,337)	(69,626)	(70,538)	(69,610)
Observations	19,695	19,695	19,695	18,778	19,695
R-squared	0.067	0.108	0.108	0.104	0.108
Cluster variable	Bank	Bank	Bank	Bank	Bank
Year FEs	Yes	Yes	Yes	Yes	Yes
Countries FEs	Yes	Yes	Yes	Yes	Yes
Bank Characteristics	No	Yes	Yes	Yes	Yes
Macro Factors	No	No	Yes	Yes	Yes

Note: This table provides estimates for coefficients in Equation 1.1, where the left hand side variable is net income instead of ratio representing bank's profitability, such as ROAA, ROAE and NIM ratio. Column (1) controls for country and year fixed effect only. Column (2) add bank specific control variables and has the same controls as Column (2). Columns (3) to (5) consider both country and bank specific control variables in Table 3, and has the same controls as Column (3) in Table 3. Standard errors are clustered at the bank−level and in parentheses *** p<0.01, ** p < 0.05, * p < 0.10.

Table A.8: Fist-stage estimates of 2SLS regression

	Dep Var: Ln(Tech/GDP)			
	(1)	(2)	(3)	(4)
Broadband	0.018***	0.018***		
	(0.001)	(0.001)		
Cellular	0.015***	0.014***		
	(0.001)	(0.001)		
Telephone	-0.003***	-0.002***		
	(0.001)	(0.001)		
DCII		0.001***		
		(0.0001)		
SLRI		-0.001***		
		(0.0001)		
L.Broadband			0.019***	0.018***
			(0.001)	(0.001)
L.Cellular			0.009***	0.010***
			(0.0006)	(0.0005)
L.Telephone			-0.003***	-0.002***
			(0.0004)	(0.0004)
L.DCII				0.001***
				(0.0001)
L.SLRI				-0.001***
				(0.0001)
Observations	19548	19548	18641	18641
F test of excluded instruments	104.82	79.85	93.57	94.30

Note: This table presents the first-stage of estimation results of 2SLS regressions on Table 5. I employ the technology information infrastructure penetration, and financial service completeness two categories as instrument variables. Standard errors are clustered at the bank-level and in parentheses *** p<0.01, **p<0.05, *p<0.10.

Table A.9: Effect of technology-based lending growth on bank's profitability by bank's operating scale

	Panel A: Dep Var = ROAE ratio			Panel B: Dep Var = NIM ratio		
	(1)	(2)	(3)	(4)	(5)	(6)
Ln(Tech/GDP)	-445.2**	-77.63***	-87.51***	-40.38*	-13.17***	-14.05***
	(187.6)	(14.46)	(14.69)	(21.74)	(2.771)	(2.780)
Small Bank × Ln(Tech/GDP)	-70.44***			-13.22***		
	(14.54)			(2.530)		
Developed × Ln(Tech/GDP)	-96.99	-308.5***	-224.1**	-50.14	-40.75	-36.06
	(149.3)	(94.70)	(90.03)	(43.89)	(25.21)	(25.11)
Developed × Small Bank × Ln(Tech/GDP)	-213.7**			-13.03		
	(88.92)			(20.92)		
Few Branches × Ln(Tech/GDP)		-146.2***			-22.65***	
		(30.01)			(5.792)	
Developed × Few Branches × Ln(Tech/GDP)		45.45			10.63	
		(111.4)			(32.42)	
Small Scale ×Ln(Tech/GDP)			-78.47**			-17.46***
			(30.68)			(6.174)
Developed × Small Scale × Ln(Tech/GDP)			116.0			44.95
			(129.9)			(39.21)
Constant	-5.068***	-5.679***	-5.265***	5.261***	5.127***	5.206***
	(1.568)	(1.542)	(1.531)	(0.404)	(0.394)	(0.386)
Observations	19,695	19,695	19,695	19,695	19,695	19,695
R-squared	0.337	0.335	0.335	0.689	0.689	0.689
Cluster variable	Bank	Bank	Bank	Bank	Bank	Bank
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Countries FEs	Yes	Yes	Yes	Yes	Yes	Yes
Bank Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Macro Factors	Yes	Yes	Yes	Yes	Yes	Yes

Note(s): This table presents coefficient estimates β_1 for models estimating the impact of technology-based lending volume growth on bank's other profitability indicators, which are ROAE ratio and NIM ratio, by different financial inclusion identification. Ln(Tech / GDP) denotes a continuous variable that the volume of technology-based lending to given country's GDP ratio in Ln(1+X) logarithm transformation. Small Bank indicator variable equals to 1 when the total asset of bank *i* in country *c* at year *t* is smaller than 1 Billion USD. Few Branch indicator variable equals to 1 when the total branches of bank *i* in country *c* at year *t* is smaller than or equal to 10. Small Scale indicator variable equals to 1 when the total branches of bank *i* in country *c* at year *t* is smaller than or equal to 10 and the total asset is also smaller than 1 Billion USD. The base observation of the regression without any interaction is the non-small banks with more than 10 branches in developing countries. Detailed information for all control and indicator variables are included in Table A.6. Standard errors are clustered at the bank–level and in parentheses *** p<0.01, ** p < 0.05, * p < 0.10.

Table A.10: Effect of technology-based lending growth on bank's other profitability indicators by countries' credit accessibility level

	Panel A: Dep Var = ROAE ratio			Panel B: Dep Var = NIM ratio		
	(1)	(2)	(3)	(4)	(5)	(6)
Ln(Tech/GDP)	-197.1*	-101.4***	-109.6***	-63.97***	-20.41***	-18.31***
	(107.4)	(16.44)	(17.37)	(24.28)	(3.593)	(3.155)
Developed × Ln(Tech/GDP)	-2.588	-20.00	1.905	-6.167	-18.22	-24.37
	(131.5)	(122.7)	(120.9)	(34.90)	(33.40)	(33.29)
Low SLRI × Ln(Tech/GDP)	-82.26***			-12.38***		
	(13.97)			(2.477)		
Low SLRI × Developed × Ln(Tech/GDP)	-267.2***			-37.30*		
	(96.14)			(21.90)		
Low DCII × Ln(Tech/GDP)		-239.2***			-79.28***	
		(70.82)			(22.55)	
Low DCII × Developed × Ln(Tech/GDP)		-300.1***			-23.26	
		(84.27)			(17.76)	
Low credit accessibility × Ln(Tech/GDP)			-304.4***			-52.78***
			(69.00)			(13.54)
Low credit accessibility × Developed × Ln(Tech/GDP)			-332.4***			-12.06
			(85.15)			(17.28)
Constant	-5.442***	-5.387***	-5.378***	5.165***	5.175***	5.164***
	(1.533)	(1.534)	(1.534)	(0.388)	(0.388)	(0.388)
Observations	19,695	19,695	19,695	19,695	19,695	19,695
R-squared	0.334	0.334	0.334	0.689	0.689	0.689
Cluster variable	Bank	Bank	Bank	Bank	Bank	Bank
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Countries FEs	Yes	Yes	Yes	Yes	Yes	Yes
Bank Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Macro Factors	Yes	Yes	Yes	Yes	Yes	Yes

Note(s): This table presents coefficient estimates β_1 for models estimating the impact of technology-based lending volume growth on bank's other profitability indicators by different level of accessibility of credit in different countries. Ln(Tech / GDP) denotes a continuous variable that the volume of technology-based lending to given country's GDP ratio in Ln(1+X) logarithm transformation. Developed indicator variable equals to 1 when the country that the bank locate in is a developed country. Low DCII (The Depth of Credit Information Index) indicator variable equals to 1 when the DCII in country c at year t is smaller than or equal to 6. Low SLRI (The Strength of Legal Right Index) indicator variable equals to 1 when the SLRI in country c at year t is smaller than or equal to 5. Low DCII and SLRI indicator variable equals to 1 when theDCII in country c at year t is smaller than or equal to 6 and the SLRI in country c at year t is smaller than or equal to 5.Detailed information for all control and indicator variables are included in Table A.6. Standard errors are clustered at the bank–level and in parentheses *** p<0.01, ** p < 0.05, * p < 0.10.

Table A.11: Effect of technology-based lending on high and non-high volume countries

	DV: ROAA ratio			DV: ROAE ratio			DV: NIM ratio		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	All	High volume	Non High volume	All	High volume	Non High volume	All	High volume	Non High volume
Ln(Tech/GDP)	-5.623***	-4.022***	-39.88***	-86.71***	-62.02***	-279.1***	-14.42***	-14.69***	-32.33
	(1.379)	(1.422)	(14.87)	(14.20)	(15.90)	(98.85)	(2.586)	(3.300)	(20.22)
Constant	-1.182***	-0.996***	-1.304***	-5.526***	-3.500	-6.685***	5.154***	7.864***	4.415***
	(0.181)	(0.322)	(0.220)	(1.526)	(2.469)	(1.896)	(0.386)	(0.899)	(0.361)
Observations	19,695	5,511	14,184	19,695	5,511	14,184	19,695	5,511	14,184
R-squared	0.333	0.244	0.377	0.334	0.296	0.349	0.689	0.668	0.716
Cluster variable	Bank	Bank	Bank	Bank	Bank	Bank	Bank	Bank	Bank
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Countries FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bank Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Macro Factors	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table illustrates the comparison of the estimation results between the $\beta_1$in Equation 1.1 with the full sample, high technology-based lending volume countries subgroup, and non-high technology-based volume countries subgroup. I identify the top-four highest technology-based lending volume countries as "High volume", including China, the United State, the United Kingdom, and Japan. Rest of countries including in full sample are "Non-high volume" countries. Standard errors are clustered at the bank-level and in parentheses *** p<0.01, **p<0.05, *p<0.10.

Table A.12: Effect of technology-based lending on commercial and non-commercial banks

	DV: ROAA ratio			DV: ROAE ratio			DV: NIM ratio		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	All	Commercial	Non-commercial	All	Commercial	Non-commercial	All	Commercial	Non-commercial
Ln(Tech/GDP)	-5.623***	-5.122***	-30.15***	-86.71***	-78.57***	-470.6***	-14.42***	-21.70***	42.45***
	(1.379)	(1.614)	(9.179)	(14.20)	(15.56)	(127.6)	(2.586)	(3.081)	(14.51)
Constant	-1.182***	-2.408***	0.0381	-5.526***	-15.20***	5.380***	5.154***	7.073***	3.993***
	(0.181)	(0.329)	(0.133)	(1.526)	(2.656)	(1.278)	(0.386)	(0.702)	(0.240)
Observations	19,695	9,602	10,093	19,695	9,602	10,093	19,695	9,602	10,093
R-squared	0.333	0.220	0.573	0.334	0.243	0.376	0.689	0.623	0.845
Cluster variable	Bank	Bank	Bank	Bank	Bank	Bank	Bank	Bank	Bank
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Countries FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bank Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Macro Factors	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table illustrates the comparison of the estimation results between the $\beta_1$in Equation 1.1 with the full sample, commercial banks subgroup, and non-commercial banks subgroup. Standard errors are clustered at the bank-level and in parentheses *** p<0.01, **p<0.05, *p<0.10.

Data Build

Bank data

The bank-related variables sample was restricted to the valid banks between January 1st, 2012, and December 31st, 2019, and 3,327,430 entities were selected. In this paper, I select three types of bank specialization: commercial, cooperative, and savings. After selecting the data, it has 13,899 banks in selection.

I have the following restrictions when choosing a sample of banks. First, I only choose banks that are active from 2013 to 2019. After this, BankFocus will have 3,364,296 banks are selected. Also, many banks report both financial statements that include and exclude their own subsidiaries or branches.To prevent duplicate estimates, I select data only considering bank statement that not integrating the statements of the possible controlled subsidiaries or branches of the concerned bank with no consolidated companion, which the consolidation code in BankFocus is U1. Following this, I select four types of bank specialization, namely bank holding companies, commercial banks, cooperative banks and savings banks. After filtering, the dataset is left with 116,189 banks. A large proportion of these selected banks had serious data deficiencies. Therefore, I set three restrictions to further filter the severely missing bank samples. First, I remove observations that have any of the selected variables missing. After that, I further remove bank samples that have less than three observations. Finally, after merging with country-specific variables, I remove samples for which macro data do not exist for the country in which the bank is located. Banks in different countries initially represent their financial data on BankFocus in the fiat currency of their country. For uniformity in subsequent regression analyses, I use BankFocus's conversion method, which is based on the U.S. dollar exchange rate data on the day they report data to BankFocus.

FinTech and BigTech lending classification

The data on FinTech lending credit of Cornelli et al. (2020) comes from the Cambridge Centre for Alternative Finance's (CCAF) Global Alternative Finance Database. FinTech lending credit includes all loan-based business models. In short, the FinTech lending credit is constituted by:

- Peer-to-peer (P2P) or marketplace lending to consumers, businesses, or for the property.
- Balance sheet lending to consumers, businesses, or the property; invoice trading.
- Debt-based securities (debentures and bonds).
- Mini-bonds.

FinTech lending credit does not include equity-based, donation-based, or reward-based crowdfunding. Profit-sharing crowdfunding, community shares, pension-led funding, and real estate crowdfunding are also not included because they fall under the broader alternative finance category.

Data on BigTech lending credits are gathered through contacts at central banks and BigTech firms and through various open sources. The former are collected under agreements that protect the confidentiality of company-specific numbers. Because most BigTech firms do not report loan flows by country or year, and many do not even submit annual reports, Cornelli et al. (2020) rely on incomplete data and assumptions to build BigTech lending volume, which includes:

- When only the year-end stock of outstanding credit is available, we estimate credit flows as the difference between these stocks plus loans assumed to mature during the year based on average loan maturity estimates.

- If a company only reports total loans in a few countries, Cornelli et al. (2020) allocate this loan volume in proportion to its revenues in those countries, or the absence of data, in proportion to those countries' GDP.
- In the absence of loan volume data for 2019, we assume that growth is proportional to the total number of users. If this information is unavailable, Cornelli et al. (2020) project the volume for 2019 using the growth rate from 2017 or 2018.

Airtel, Amazon, Alibaba/Ant Group, Apple, Au Jibun Bank, Baidu/Du Xiaoman, BKash Facebook, Fuse, Go-Jek, Google, and others have credit activity in 31 countries as of 2019. Not all of these businesses engage in lending.

2 Determinant of the herding behavior of cryptocurrency market: What is the bellwether?

Abstract

This paper provides empirical evidence of herding behavior and its sensitivity to external influences and market spillover effects in the cryptocurrency market. Analyzing the cryptocurrency market with the CSAD model, I confirm that internal cryptocurrency market movement and investor sentiment have an insignificant impact on herding behavior in the cryptocurrency market, whereas external sentiment proxies significantly influence herding behavior, especially during a bullish market. Further analysis revealed that market movement in external financial markets does not induce herding or volatility spillover effects on the cryptocurrency market. Overall empirical results demonstrate that investors are more responsive to external market sentiment proxies when cryptocurrencies rise, and traditional financial markets influence cryptocurrency markets primarily through the transmission of fear. This paper enhances comprehension of cryptocurrency market dynamics and highlights the significance of prerequisites when considering herding behavior.

2.1 Introduction

Since Bitcoin was launched in January 2009[4], cryptocurrencies have been widely scrutinized by investors, regulators, and academics. Compared with other broadly recognized fiat currencies and financial instruments, cryptocurrency relies on unique blockchain technology to support transactions without going through financial institutions, promoting the decentralization principle. According to historical data from Coinmarketcap, the total market capitalization of the cryptocurrency market has increased more than 400%, from $7.14 billion at the beginning of 2016 to $2.21 trillion at the beginning of 2022. In such a context, many scholars have started to devote their efforts to investigating the driving forces behind the volatile return of cryptocurrency and if the influence of investor behavior can be the driving force of such high volatility and risk of the market return.

Herding behavior, which is widely discussed in behavioral finance, is one of the most notable considerations when assessing how investor behavior influences cryptocurrency returns and stability (Bouri et al., 2019; Kaiser and Stöckl, 2020; Papadamou et al., 2021). Herding behavior is defined by individuals who "suppress their own beliefs and base their investment decisions solely on the collective actions of the market, even when they disagree with its predictions" (Christie and Huang, 1995). Many papers have confirmed herding behavior's existence and significant influence on equity and bond financial markets (Chang et al., 2000; Chiang and Zheng, 2010; Galariotis et al., 2016). Furthermore, these behaviors have been shown to cause market inefficiencies and bull or bear market tendencies and bubbles (Fama, 1998; Shiller, 2003). Therefore, there are three potential implications of shifting the spotlight to the cryptocurrency market. First, studying the potential herding behavior in cryptocurrency allows for analyzing the interaction between

[4]See white paper by Nakamoto (2009)

investors' behaviors and market fluctuations. In addition, as part of the source of market instability, significant herding behavior can be treated as a market signal that advises investors and policymakers so they can avoid potential risks when constructing investment strategies. Furthermore, comparing the similarity of herding behavior in cryptocurrencies to that of traditional financial markets provides further insight into whether investors' behavior differs across markets.

The current research about the herding behavior in cryptocurrency is concentrated on the existence of the behavior (Bouri et al., 2019; Kyriazis, 2020; Papadamou et al., 2021; Vidal-Tomás et al., 2019), the impact of herding on price dynamics (Gurdgiev and O'Loughlin, 2020; King and Koutmos, 2021), and how the herding behavior is influenced by exogenous factors (Philippas et al., 2020). Nevertheless, there are still some research gaps. First, previous research has focused on whether there is herding behavior in the cryptocurrency market and its impact on the market. Still, few studies have examined cross-herding behavior, such as market changes from external financial markets and sentiment contagion. Second, most data analyzed in previous papers are from before 2019. In contrast, cryptocurrencies have entered a considerable growth period since 2018, with total market capitalization rising from nearly 830 million USD in January 2018 to 2 trillion USD in January 2022[5]. Therefore, the increasing maturity of the cryptocurrency market and the potential change in herding behavior with the influx of more capital is also worth further study. Third, much research has examined how external factors affect herding behavior without further categorizing the external factors to determine which type of information affects investors' market behavior more. Further investigating the sensitivity of herding to different external shocks can explain investors' information preferences and reactions in the face of exter-

[5]Data source: https://coinmarketcap.com/charts/

nal information shocks, which can further explore the potential driving force of cryptocurrency market instability.

This paper concentrates on three aspects of herding behavior in the cryptocurrency by analyzing daily price data for the top 59 cryptocurrencies from November 10th, 2017, to December 31st, 2021.[6] Consistent with Chang et al. (2000), I first examine if herding behavior in the cryptocurrency market exists only in internal market fluctuations, with market movement, extreme market conditions, and time-varying considerations. Philippas et al. (2020) and Ballis and Drakos (2020) argue that there is significant herding behavior in the cryptocurrency market. In contrast, Bouri et al. (2019) found no herding behavior, which contradicts the results in many previous papers, further proving that herding behavior is only present and significant for a short period under time-varying considerations. Such findings imply that investors do not engage in persistent and significant herding behavior due to cryptocurrency market changes and that if they do, it is only in the short term.

Second, I test whether external sentiment proxies induce herding behavior in cryptocurrency. Unlike other studies, this paper categorizes the sentiment proxies as internal and external market information-related to explore which information shock is more likely to induce herding behavior in the cryptocurrency market. I consider four sentiment proxies: the Royalton CRIX Crypto Index (CRIX) and Google Trends Index (GT index), which contain internal information about the cryptocurrency market; and the CBOE Volatility Index (VIX) and Global Economic Policy Uncertainty Index (Global EPU), which have information about external economic and financial market uncertainty. Consistent with earlier studies, the herding behavior estimation with sentiment proxies' intervention considers the conditions in the cryptocurrency market.

[6] The ranking of cryptocurrency is based on the market capitalization as 12/31/2021.

The empirical results show that herding behavior is significantly present in the cryptocurrency market when considering external sentiment proxies, and the herding phenomenon gradually increases as the index rises. Furthermore, the upper quantile market return that the cryptocurrency is in, the more likely the indexes represented by the fear of external financial markets (VIX) and economic policy uncertainty (global EPU) will exacerbate the herding behavior. Such findings imply that investors do not regard the cryptocurrency market as an independent market unaffected by external factors. Concerns about the movement of traditional financial market return and policy uncertainty can induce a significant herding behavior in the cryptocurrency market, indicating that herding behavior in the cryptocurrency market is emotionally contagious from investor sentiment in conventional financial markets. When the indices representing investors' searching attention within the cryptocurrency market are considered, the results show that increasing investor attention in the cryptocurrency market can also induce herding behavior within the cryptocurrency market when the attention is in a lower quantile. Such findings suggest that when investors' attention to cryptocurrencies is low, the market has relatively little information and discussion about cryptocurrencies. People make investment decisions based on limited information in this situation, and herding behavior occurs as attention increases. When the level of information and discussion about cryptocurrencies is already high, investors have enough information to make investment decisions; thus, even if the level of attention rises again, herding behavior will not be encouraged.

The third research question addressed in this paper is whether there are herding effects and volatility spillovers to cryptocurrency markets due to fluctuations in traditional financial markets and the contagion of investment sentiment between markets. In the previous discussion of the cryptocurrencies' characteristics, most of the attention has been focused on their fiat (Czapliński and Nazmutdinova, 2019; Levulytė and Šapkauskienė, 2021) and financial attributes (Ariefianto, 2020). This paper focuses on the US stock and major FOREX markets when considering external financial market interventions. However, unlike previous studies, I consider market conditions in both cryptocurrencies and traditional financial markets when examining herding behavior and volatility spillovers and further explore the extent to which they are more sensitive to movements in either market. The results show that both stock market and FOREX market volatility do not have a significant herding and volatility spillover effect on the cryptocurrency market, even considering different market conditions for both the stock and FOREX markets and the cryptocurrency market. Combined with the previous empirical results, I conclude that volatility in the traditional financial markets does not directly affect investor behavior in the cryptocurrency markets and has a spillover effect. The link between the two markets exists only regarding the mutual transmission of investor sentiment.

This paper contributes three ways to the cryptocurrency market's behavioral finance literature. First, by delineating the herding dynamic of an emerging market with high return and volatility, I provide evidence that internal market movements within the cryptocurrency market do not lead to significant herding behavior by investors. Second, this study offers new insights into the relationship between cryptocurrency and traditional financial markets. Consistent with Asante Gyamerah (2021), I observe that the herding effect is exacerbated by an increase in the external market and policy uncertainty sentiment and reduced by an increased attention level of internal attention. Such results suggest that external market sentiments are an essential factor affecting herding within the cryptocurrency market and are mainly concentrated in up markets. Third, I demonstrate that there is no significant herding and volatility spillover effect of traditional financial markets, such as the FOREX market and the US stock market, on cryptocurrencies, proving that the influence of conventional financial markets on investor behavior in cryptocurrencies exists in the transmission of investor fear.

The remainder of this paper is structured as follows. Section 2.2 provides literature review, and section 2.3 introduce the methodology. Section 2.4 discusses the data. Section 2.5 presents the results and Section 2.6 presents the conclusion.

2.2 Literature Review

2.2.1 Herding behavior in the cryptocurrency market

Since the introduction of Bitcoin (Nakamoto, 2009), the cryptocurrency market has attracted increased attention, particularly its possible applications and influence in the financial sector. Due to cryptocurrency's decentralization and the high transparency of transactions[7], users can effectively avoid the involvement of central financial institutions. Such unique procedures grant both parties a high degree of autonomy and significantly reduce the administrative or intermediary fees usually charged by financial institutions in traditional transactions. In 2015 the Commodity Futures Trading Commission (CFTC) officially defined Bitcoin as a commodity in the United States, covered by the Commodity Exchange Act, alongside gold and oil. As of November 9th, 2021, the total market capitalization of cryptocurrencies exceeded $3 trillion [8]. The considerable capital inflow and the gradually increasing debate by official legislatures are evidence of the popularity of the cryptocurrency market among investors and its potential for exploration.

[7]Because of the cryptocurrency blockchain's decentralized nature, all transactions can be transparently viewed by having a personal node or using blockchain explorers that allow anyone to see transactions taking place in real-time.

[8]Data source: https://coinmarketcap.com/charts/

Compared with other financial markets such as equity, FOREX, and commodity markets, cryptocurrencies have the unique characteristic of a lack of fundamentals. As a result, the intrinsic value of cryptocurrencies is not associated with the value of a tangible object. Therefore, the driving force behind the market's volatility needs to be further explored. Some scholars attribute the instability of cryptocurrencies and market and policy uncertainty originating from other financial markets (Corbet et al., 2020a; Davidovic, 2021; Elsayed et al., 2020, 2022; Hsu et al., 2021; Jiang et al., 2021), while others believe that the irrational investors' behavior is the main cause of market destabilization. Dass et al. (2008) stated that irrational investors concentrate on investing in identical "hot" assets in a speculative period and induce herding behavior which leads not only to creating the market bubble but also to bursting the existing bubble and declining prices, which has high potential to explain the cryptocurrency market bubble of 2018, where prices climbed from November 2017 to a peak in January 2018 and fell by 75 percent market capitalization at the end of January [9], showing strong potential for the existence of herding behavior in the cryptocurrency market.

According to Hwang and Salmon (2004), herding is an irrational behavior that occurs when investors imitate the investment decisions of others without reference to fundamentals. Irrational exuberance, which leads to over-enthusiasm and the formation of asset price bubbles, is intimately linked to the prevalence of herding. Trading in the same direction as others, following the trend of previous trades, and mimicking or correlating one's behavior to others are all herding behaviors. Typically, herding makes inexperienced investors prone to becoming risk-takers without understanding the risks they face. Lack of knowledge about economic situations and extreme market conditions, such as during an unstable market pe-

[9] Data source: https://coinmarketcap.com/charts/

riod, may lead to irresponsible behavior when herding exists. The existence and influence of herding in traditional financial asset markets have been studied by many scholars (See Babalos and Stavroyiannis (2015); Chang et al. (2000); Chiang and Zheng (2010); Galariotis et al. (2016)). Investors may be attracted by the potential for high returns of cryptocurrencies and are gradually incorporating them into their investment alternatives (Borri, 2019; Klein et al., 2018; Leung and Nguyen, 2019; Wang et al., 2020), showing the opportunity for herding behavior within the cryptocurrency market.

Many studies have explored herding behavior in the cryptocurrency market, although the results vary. On the one hand, some research finds the existence of herding behavior in cryptocurrency, and the phenomenon is even more significant in up markets (Ballis and Drakos, 2020; Kallinterakis and Wang, 2019; Kyriazis, 2020; Manahov, 2021), when uncertainty increases (Bouri et al., 2019) and under extreme market return conditions (da Gama Silva et al., 2019; Kumar, 2021). On the other hand, other researchers have found anti-herding evidence (Coskun et al., 2020; Kurt Gümüş et al., 2019). Therefore, there is no unified agreement on herding behavior within the cryptocurrency market, which needs to be analyzed.

The primary goal of this paper is to explore the potential for herding behavior in the cryptocurrency market. Such research target is critical to understand because the cryptocurrency market entered a phase of tremendous growth after the crash of 2018, and different countries have since clarified their regulatory attitudes towards cryptocurrencies, further influencing the development of the market. Most of the data analyzed in the earlier papers was mainly concentrated on the period before 2019; however, the popularity of cryptocurrencies and the trading volumes have increased rapidly since then, so the results may change due to the development of the market. Second, cryptocurrencies have changed dramatically over time. Many "strong" cryptocurrencies with volatile returns and high market capitalization at the initial market stage are lagging in the rankings according to market capitalization; some are inactive already. However, some of the newer cryptocurrencies launched because of their new blockchain protocol or their strong ties to other markets, such as Stablecoin [10], have surged in market capitalization in a short period and have taken a significant position.

Herding behavior is also correlated with the maturity of the financial market. Teng (2018) stated that investors' herding tendency increases as the maturity level of an investment portfolio decreases. Yao et al. (2014) also proved that effective regulation can improve information efficiency and market integration, thus diminishing herding behavior. These papers also show the effect of market development and maturation on herding behavior, which can eliminate previously significant herding behavior. Thus, with the fast-growing and gradually maturing characteristic of the cryptocurrency market, this paper tests the following hypothesis:

Hypothesis I: Herding behavior is insignificant in the cryptocurrency market in both unconditional and internal market condition considerations.

2.2.2 Herding in the cryptocurrency market with sentiment proxies intervention

As many papers have demonstrated, sentiment can trigger herding behavior in the market. The most widely discussed factor is the impact of market sentiment due to uncertainty and extreme market conditions on herding within the whole market, especially during the market crisis period (Avery and Zemsky, 1998; Caparrelli et al., 2004; Cui et al., 2019). When the market stays in a crisis period, unanticipated information and market shocks would similarly affect on the senti-

[10] Stablecoin is a type of cryptocurrency pegged to an external asset, such as gold or fiat currency, leading to its early reputation for reliability.

ment and future market expectations of the market investors, thus increasing the likelihood of herding behavior (Kurz and Kurz-Kim, 2013; Schmitt and Westerhoff, 2017). Economou et al. (2015) found significant herding in the equity market during the European debt crisis in Belgium, the Netherlands, and Portugal. Blasco et al. (2017) found that organizational and environmental issues could affect herding in the financial market.

Researchers analyze the impact of market uncertainty and crisis periods on herding behavior within the market in two directions. The first direction is considering a short-term market expectation index based on underlying market information, such as the widely used VIX index and global Economic Policy Uncertainty (EPU) index (Cui et al., 2019; Humayun Kabir and Shakur, 2018). This type of index uses market transaction data and economic policies and volatility which are strong indicators of investors' expectations of the financial markets. The second direction was based on the internet and media data, which can directly influence or be influenced by investors' market behavior. Kim et al. (2022) presented evidence that customers overreact to negative news, leading to adverse herding. Also, search volume data, such as the GT index, is widely used in herding detection research (Hsieh et al., 2020; Li et al., 2021; Wanidwaranan and Padungsaksawasdi, 2022). The internet and media data directly represent investors' awareness of the market and their attitudes, so it is a more explicit proxy for investors' sentiment about the market status.

Prior research has assessed how sentiment affects herding behavior within the cryptocurrency market. On the one hand, some argue that investor sentiment can predict the return and volatility of cryptocurrencies (Corbet et al., 2020b; Gurdgiev and O'Loughlin, 2020; Pathak and Kakkar, 2020), indicating a direct impact of herding and anchoring biases. On the other hand, some believe that herding behavior originates from a contagion effect that emerges during periods of high instability in financial markets (Corbet et al., 2022; Matkovskyy and Jalan, 2019). All prior research demonstrates the strong connections between sentiment and the cryptocurrency market and that financial markets uncertainty and the potential of market crises influence investor sentiment and introduce the possibility of herding in the cryptocurrency market.

Also, some scholars have categorized sentiment to explore the similarities or differences in the impact of different types of sentiment on herding behavior. Chen et al. (2019) suggested that a cryptocurrency-related lexicon can be used to assess investor sentiment and predict cryptocurrency market returns, and investor sentiment also can positively predict excess Cryptocurrency Index (CRIX) return. Another study finds that investor attention has more impact on the cryptocurrency return in the bearish market in the short term (Li et al., 2021). For market and public sentiment, (Anastasiou et al., 2021) argue that cryptocurrency's price crash risk is positively related to the market sentiment index. Demir et al. (2018) illustrated the predictive power of EPU on Bitcoin returns and revealed that Bitcoin returns are essentially inversely related to the EPU. Furthermore, Kristoufek (2013) indicated that increased interest in Google searches on Bitcoin led to higher prices and higher search volume; however, Urquhart (2018) found that the Google search index is not a suitable volatility predictor.

Consistent with prior research on the impact of sentiment on the potential for herding in the cryptocurrency market, this paper uses the VIX, CRIX, and GT index with different search terms and global EPU as proxies for sentiment. This paper re-estimates indicators to evaluate their impact on herding behavior changes. In addition, I also categorize the proxy's representation of information sources to estimate whether herding behavior in the cryptocurrency market is more sensitive to movements in the external financial market or to information-related movements of the cryptocurrency market itself. Specifically, based on the data source and methodologies of the index, I define VIX and Global EPU as sentiment proxies integrated from the cryptocurrency market's external financial and economic market information, and CRIX and GT index as sentiment proxies obtained from internal information and attention of the cryptocurrency market itself. Since herding has the potential to be different in up and down market conditions, we should also consider the market condition when considering the impact of sentiment proxies.

The period for estimation selected in this paper includes the 2018 crypto crash[11] phase and the COVID-19 pandemic. Contrary to prior studies, I do not consider these two events separately because the primary consideration of this paper is still the impact of the sentiment proxy and not the impact of the two events. Thus, I test the following hypothesis:

Hypothesis II: Sentiment proxies have the potential to influence the herding in the cryptocurrency market, but results might vary according to proxy classification and market conditions.

2.2.3 Herding and volatility spillover from external financial markets

With the continuing expansion and sophistication of financial markets, the relationship between markets is strengthened by the interaction between the return and volatility within different financial markets. One explanation of such phenomena is the spillover effect, which means one economy is influenced by a seemingly unrelated event from another economy[12]. from one market to the other. A signif-

[11] See:Smith, Noah (8 February 2018). "Crypto Cynics Stand to Profit the Most". Bloomberg. Retrieved from: https://www.bloomberg.com/view/articles/2018-02-08/crypto-cynics-stand-to-profit-the-most

[12] Source: Corporate Finance Institute, https://corporatefinanceinstitute.com/resources/knowledge/other/spill effect/

icant spillover can demonstrate the connectedness between the two markets and provide further insight into the origin of changes in market return.

Several studies have shown the spillover effect between financial markets. Hung and Vo (2021) illustrated the significance of information spillover among the crude oil, S&P 500, and gold markets, indicating that return transmissions were more apparent during the COVID-19 crisis. Both papers assess the significance of the spillover effect in the presence of market crises. They find that the driving force of the spillover formation between markets is related to the investors' market strategy in the context of extreme market conditions. When a specific market is uncertain and highly volatile, some investors will make investment choices based on private information or use the external market as a temporary safe haven to avoid risk. At the same time, their pessimistic investment sentiment may affect their judgment of the other market, thus negatively affecting the other market.

In the context of the high volatility in the cryptocurrency market, prior research has been dedicated to exploring whether the driving force behind it arises from the spillover effect in other financial markets. Cao and Xie (2022) found significant evidence that negative spillovers are more influential than positive spillovers between cryptocurrency and the Chinese stock market. Also, Ji et al. (2019) illustrated that the nature of information spillovers changes over time, with cryptocurrencies becoming more connected and more prominent within the financial system. Furthermore, Frankovic et al. (2021) indicated unidirectional return spillover and weak volatility spillover from the cryptocurrency market to crypto-related stocks in Australia. In addition, crypto-related stocks with high involvement in blockchain technology display more vital connectedness to the cryptocurrency market through return spillover.

These papers demonstrate that there are different types of spillover effects between traditional financial markets and the cryptocurrency market. A limited amount of research focuses on the induction and impact of spillover effects between markets from an investor behavior perspective. As the leading players in financial markets, investors' judgments about the market and investment decisions can trigger the spillover effect. In extreme market conditions, investors are more likely to mimic others' choices and ignore their private information to avoid losses, leading to herding behavior. At the same time, a significant herding spillover between different markets has been documented (BenMabrouk and Litimi, 2018; Yasir and Önder, 2022); therefore, based on the significant results obtained from the literature above, testing the spillover effect between the cryptocurrency market and traditional financial market is meaningful.

This paper focuses on the herding and volatility spillover effects of the U.S. stock market and the major FOREX market on cryptocurrencies. The reason for taking volatility spillover into account is that both the U.S. stock and major FOREX market received the most attention in considering the search for spillover effects, and much evidence has been generated to demonstrate the impact of market volatility on herding behavior in both traditional financial and cryptocurrency markets (Asante Gyamerah, 2021; BenMabrouk and Litimi, 2018; Economou et al., 2011; Park, 2011). At the same time, I mainly follow the current discussions of the financial (Ariefianto, 2020) and currency (Levulytė and Šapkauskienė, 2021) nature of cryptocurrencies, selecting the U.S stock market and FOREX market as targets to investigate how the spillover effect comes from financial markets influence the herding behavior in the cryptocurrency market. At the same time, to further explore whether variations in different markets may have different effects on the potential spillover, this paper also incorporates market movements of two traditional financial markets into the estimation. Thus, the third hypothesis tested is:

Hypothesis III: There is a significant herding and volatility spillover effect from the traditional financial market to the cryptocurrency market, but results may vary based on different market conditions.

2.3 Methodology

Prior research uses two methods to estimate investor herding behavior. The first methodology was proposed by Lakonishok et al. (1992), focusing on transaction history of investors in a specific security to document the existence of herding in the stock market. The second direction to estimate the herding behavior is concentrated on the aggregate market perspective, and notable methodologies are the standard cross-sectional dispersion (CSSD) model of Christie and Huang (1995) and the improved version, the cross-sectional absolute deviation (CSAD) model of Chang et al. (2000). This paper follows the CSAD model to measure the herding behavior at the aggregate market level with daily price data. This method has been widely employed in the studies on the herding behavior of the cryptocurrency market (Bouri et al., 2019; Kyriazis, 2020; Papadamou et al., 2021; Yarovaya et al., 2021), and using the identical methodology but different data considerations can further uncover the variation in herding behavior over time periods. A stronger argument for the possibility of herding behavior in cryptocurrencies may be made using CSAD model estimation with more recent data and the inclusion of more cryptocurrencies in market portfolio consideration. It can also show the change and sensitivity of herding behavior in the cryptocurrency. By using this method my results can be compared to prior results.

2.3.1 Herding behavior in the cryptocurrency market

Chang et al. (2000) proposed the following:

$$CSAD_{crypto,t} = \frac{\sum_{i=1}^{n} |r_{i,t} - r_{crypto,t}|}{n-1} \quad (2.1)$$

where the $r_{i,t}$ is the return of the specific cryptocurrency i at time t, $r_{crypto,t}$ is the equal-weighted market return of cryptocurrency at time t, and n is the number of active cryptocurrencies included in the cross-section at time t. Some cryptocurrencies were not launched in an earlier period or not reported in the trading data on all days; such context will be counted as "inactive". All missing weekend, holiday, and inactive-trading data are excluded in market portfolio calculation to be consistent with further external financial market consideration. Daily and market return is calculated as follows:

$$r_{i,t} = 100(ln(P_{i,t}) - ln(P_{i,t-1})) \quad (2.2)$$

$$r_{crypto,t} = \frac{\sum_{i=1}^{n} r_{i,t}}{n} \quad (2.3)$$

where the $P_{i,t}$ and $P_{i,t-1}$ are the daily closing price of the specific cryptocurrency i at time t and $t-1$. I use equally weighted cryptocurrency returns consistent with prior research (Kumar, 2021; Rubbaniy et al., 2021; Vidal-Tomás et al., 2019).

Herding exists when investors decide to follow the market instead of their own beliefs or strategies and this will appear in the data analysis as a nonlinear relationship between CSAD and market returns. This paper follows the CSAD model to estimate the unconditional herding (Chang et al., 2000; Chiang and Zheng, 2010) and adjusts the standard error according to Newey and West (1987):

$$CSAD_{crypto,t} = \beta_0 + \beta_1|R_{crypto,t}| + \beta_2 R^2_{crypto,t} + u_t \quad (2.4)$$

In the absence of unconditional herding, the relationship between the return dispersion (CSAD) and absolute market returns is expected to be significantly positive, which means a positive and significant β_1 coefficient. However, if herding exists, investors may switch from their strategies to following the market consensus, leading to a lower cross-sectional dispersion of returns compared to the case of rational pricing, then pulling individual returns towards market return. This means lowering the cross-sectional dispersion (CSAD) and getting a significantly negative β_2 coefficient (Chang et al., 2000). A significant and negative β_2 partly reject the hypothesis I that herding behavior is insignificant in unconditional market consideration of cryptocurrency. To further consider the time-varying perspective of potential herding behavior, I set a 30-day, 60-day and 90-day rolling-window estimation for CSAD model, perspectively.

This paper uses an indicator variable to determine whether the results vary across different market conditions. This setting is based on the empirical extension of Chang et al. (2000), which is widely considered in prior research (Ballis and Drakos, 2020; King and Koutmos, 2021; Kumar, 2021; Yarovaya et al., 2021). This paper follows the same setting, when the market return $R_{m,t}$ is positive (negative), then the cryptocurrency market at time t is in the upward (downward) market and equal to one (zero). Thus, the equation 2.4 can be extended to:

$$\begin{aligned} CSAD_{crypto,t} = & \beta_0 + D_{up,crypto} * (\beta_1 |R_{crypto,t}| + \beta_2 R^2_{crypto,t}) \\ & + (1 - D_{up,crypto}) * (\beta_3 |R_{crypto,t}| + \beta_4 R^2_{crypto,t}) + u_t \end{aligned} \tag{2.5}$$

which estimates the herding behavior in upward and downward market conditions. A significant negative value of β_2 (β_4) would represent a herding behavior in the cryptocurrency up (down) market, partly reject the hypothesis I that herding behavior is insignificant in conditional market movement consideration of cryptocurrency.

2.3.2 Herding in the cryptocurrency market with external factors

While considering the asymmetric effect of sentiment proxies on potential herding behavior, this paper also considers the impact of the external sentiment index on herding behavior in the cryptocurrency market. This allows for an assessment of the effect of externalities on herding behavior by comparing the results with the previous empirical analysis. To capture asymmetries in herding behavior caused by external sentiment index factors (Chiang and Zheng, 2010), equation 2.4 can be extended as follows:

$$CSAD_{crypto,t} = \beta_0 + \beta_1 |R_{crypto,t}| + \beta_2 R^2_{crypto,t} + \beta_3 Factor_t + u_t \tag{2.6}$$

where the external sentiment index factors, $Factor_t$, in different consideration, is either 1) VIX index, 2) CRIX index, 3) GT indexes with "cryptocurrency" and "bitcoin" search terms, or 4) global EPU index, respectively. Assuming that investors use the CAPM to price assets, under the null hypothesis of no herding, exogenous proxies should not influence CSAD and β_3 should be statistically insignificant. If the null hypothesis is rejected, then conditional herding describes the case where investors respond to market stress and/or increased information flow by uniformly pricing individual assets. An significant and negative $\beta_3$partly support the hypothesis II that sentiment proxies have the potential to influence the herding in the cryptocurrency market. I also expect that the impact of these external factors on herding behavior varies depending on up and down-market conditions. For example, Vidal-Tomás et al. (2019) document herding only during down-market days, whereas Rubbaniy et al. (2021) argue that the presence of herding asymmetry. Thus equation 2.7 is an extension of equation 2.6:

$$
\begin{aligned}
CSAD_{crypto,t} =& \beta_0 + D_{up} * (\beta_1 |R_{crypto,t}| + \beta_2 R^2_{crypto,t} + \beta_3 Factor_t) \\
& + (1 - D_{up}) * (\beta_4 |R_{crypto,t}| + \beta_5 R^2_{crypto,t} + \beta_6 Factor_t) + u_t
\end{aligned} \tag{2.7}
$$

The market condition setting used is consistent with equation 2.5. An significant, negative and different β_3 partly support the hypothesis II that sentiment proxies have the potential to influence the herding in the cryptocurrency market with different market consideration.

In summary, there are three categories of external factors. For the first factor, market sentiment consideration, this paper applies two indices: the VIX index for tracing the volatility from the stock market and the CRIX index for tracing the volatility from the cryptocurrency market. The second factor is social attention, and the GT index will be applied in this consideration. The third factor is the policy and market uncertainty; this paper will apply the Global Economic Policy Uncertainty (EPU) index to trace. At the same time, these four indices are also classified in a second way according to their sources of information. VIX and Global EPU are external market indicators, while GT index and CRIX are internal market indicators because their calculations are based on crypto-related information. Utilizing different classifications can enhance our understanding of which external proxies the herding behavior within the cryptocurrency market is more sensitive to.

2.3.3 Herding and volatility spillover from external financial market

To investigate the connectedness between cryptocurrency and traditional financial markets, this paper also considers the herding and volatility spillover effect of the U.S. stock market and major FOREX market on the herding behavior of the cryptocurrency market. The daily individual return, market return and CSAD are calculated as Equation 2.1, 2.2 and 2.3.

Consistent with Yasir and Önder (2022), I estimate the spillover effect from external financial market as follows:

$$CSAD_{crypto,t} = \beta_0 + \beta_1|R_{crypto,t}| + \beta_2 R^2_{crypto,t} + \beta_3 CSAD_{m,t} + \beta_4 R^2_{m,t} + u_t \quad (2.8)$$

Where $CSAD_{m,t}$ and $R_{m,t}$ represent the cross-sectional dispersion and equally-weight market return of market m either 1) US stock market or 2) major FOREX market at time t. If the coefficient of $CSAD_{m,t}$ is significantly positive, then the result indicates evidence of volatility spillover from market m to the cryptocurrency market. Similarly, a significantly negative coefficient of$R^2_{m,t}$ represent the evidence of herding spillover from market m to the cryptocurrency market.

To consider cryptocurrency and two external financial market conditions when considering the spillover effect on herding behavior. In other words, equation 2.8 can be expanded to:

$$\begin{aligned} CSAD_{crypto,t} =& \beta_0 + D_{up}(\beta_1|R_{crypto,t}| + \beta_2 R^2_{crypto,t} + \beta_3 CSAD_{m,t} + \beta_4 R^2_{m,t}) \\ &+ (1 - D_{up})(\beta_5|R_{crypto,t}| + \beta_6 R^2_{crypto,t} + \beta_7 CSAD_{m,t} + \beta_8 R^2_{m,t}) + u_t \end{aligned} \quad (2.9)$$

WhereD_{up} and D_{down} indicates the cryptocurrency and external financial up and down market condition, respectively.

2.4 Data

Many altcoins with a more influencing position today, such as Ethereum, Binance Coin, have daily transaction data available only after November 9th, 2017. Based on this information, this paper analyzes the 59 top marketcryptocurrencies data from November 9th, 2017, to December 31st, 2021[13]. Economou et al. (2008) stated that herding is more pronounced from a daily perspective. Therefore, this paper utilizes the daily closing price for all estimations. Consistent with prior research, all cryptocurrency-related data are extracted from coinmarketcap.com. According to Vidal-Tomás et al. (2019), even though the smallest cryptocurrencies are herding with the largest ones, the driving factor of cryptocurrency market return change is still the mean return of the largest cryptocurrencies. Stavroyiannis and Babalos (2019) found that the top 8 cryptocurrencies held approximately 95% market capitalization. At the same time, according to coinmarketcap, the top 10 cryptocurrencies consistently constitute 80% of the market capitalization of the entire cryptocurrency market, so selecting the top 59 cryptocurrencies is representative of the entire cryptocurrency market to avoid selection bias[14].

Table A.13 provides information on all selected cryptocurrencies, including the average daily return of each cryptocurrency. The average daily return of most cryptocurrencies is more significant than zero, further illustrating the cryptocurrency market's investment potential. At the same time, some emerging cryptocurrencies with later launch dates have higher average daily returns than average market returns. This table also shows that when studying the cryptocurrency market, we should not only consider the top players with high attention and have already been launched for a long time.

[13]The ranking is derived from the cryptocurrency market capitalization ranking as of December 31, 2021.

[14]Data source from: https://coinmarketcap.com/charts/

External sentiment proxies are retrieved from a variety of sources. The VIX index estimates the expected volatility in the next 30 calendar days in the U.S. market. It applies the S&P 500 options as an underlying benchmark, and the historical data was retrieved from CBOE global market. Wolfgang Karl Härdle developed the CRIX index, which traces the evolution of the diverse and fast-changing crypto market, and the historical data is retrieved from the S&P Dow Jones website. The global EPU index is a GDP-weighted average of national EPU indices from 21 countries[15]. It reflects the relative frequency of own-country newspaper articles that contain a trio of terms about the economy (E), policy (P), and uncertainty (U), and the historical data is retrieved from policyuncertainty.com. The GT index compares the search activity for a particular phrase in a particular time period to the search activity for the same phrase in the specified geography. The GT index will be indexed from 0 to 100, where 100 represents the highest relative attention received by the search phrase for the specific search parameters. The historical data is retrieved from Google Search's official website. GT index may vary depending on the search term and search period in different considerations. To eliminate the potential selection bias, I first adjust the GT index, obtaining the daily data of the GT index in segments and then constructing the weekly data of the whole interval. After adjusting according to the weekly index, I obtained the adjusted daily GT index of the selected period for further estimation. I consider "cryptocurrency" and "bitcoin" as the search term to investigate whether the estimated result varies based on different search terms. Both the CRIX and VIX indexes do not report the weekend and holiday data, and the paper solves the missing value problem for data consistency by skipping the weekend and holiday data and filling in the next

[15]These 21 countries are: Australia, Brazil, Canada, Chile, China, Colombia, France, Germany, Greece, India, Ireland, Italy, Japan, Mexico, the Netherlands, Russia, South Korea, Spain, Sweden, the United Kingdom, and the United States.

missing value with the most recent historical data available for global EPU data.

This paper selects the U.S. stock market and major FOREX market for traditional financial markets' volatility and herding spillover effect. The daily stock price is retrieved from Wharton Research Data Service (WRDS). This paper applies all available major FOREX market daily historical data from tradingview.com for the FOREX market consideration. These financial markets are non-trading at weekends and holidays, so to be consistent with previous data sets, this paper removes all inactive and missing market data Table A.14 shows the summary statistics before empirical analysis for three hypotheses.

2.5 Empirical Results

2.5.1 Internal herding behavior in the cryptocurrency market

I present the regression results of the CSAD model without and with market condition consideration following Equations 2.4 and 2.5, as listed in Table A.15. The results of the baseline CSAD model shown in Column (1) of Table A.15 that the coefficient of $R^2_{m.t}$ is significant in 10% level and positive, indicating that there are no herding behaviors at the aggregate market level. I then further examine whether investors in cryptocurrency market tend to herd in different market conditions of cryptocurrencies. The results of Columns (2) and (3) demonstrate that investors do not show significant herding behavior when the aggregate cryptocurrency market or Bitcoin is in either up or down market. I also consider the particular scenario when Bitcoin has alternative market return compared with aggregate altcoin market. Columns (4) and (5) show the estimation result of this consideration and shows that he coefficient of $R^2_{m.t}$ is insignificant and positive, demonstrating that there is no significant herding behavior. These results indicate the absence of unconditional and conditional herding in the cryptocurrency market, supporting Hypothesis I. These empirical results are consistent with the findings of Stavroyiannis and Babalos (2019); Coskun et al. (2020); Kurt Gümüş et al. (2019).

I further estimate that if investors tend to herd when the cryptocurrency market is in extreme conditions. Hence, I employ the model in Equation 2.5 to detect herding behavior under extreme up and down cryptocurrency markets. Table A.16 reports the estimation results for potential investor herding behavior in different quantiles of overall cryptocurrency market return. Columns (1) to (3) of Table A.16 show that as the overall cryptocurrency market return increases and enters the extreme up market (upper 95% quantile of the market return), the coefficient of $R^2_{m.t}$, which detects herding behavior within the cryptocurrency market, gradually moves to negative, though still not significant. Table A.16's Columns (3) to (6) show that the coefficient of $R^2_{m.t}$ decreases as the overall market return of cryptocurrency decreases, but still remain positive and insignificant. The overall results of Table A.16 suggests that investors do not exhibit significant herding behavior even when the cryptocurrency market is experiencing extreme up or down market conditions. In contrast, the likelihood of herding behavior increases when the cryptocurrency market's return is in the more extreme quantile of market return. Meanwhile, investors are more sensitive to the extreme up market in cryptocurrency than the extreme down market.

Figure 2.1 further illustrates the time-varying potential herding behavior of investors under the rolling-window consideration. By comparing the (a) to (c) in Figure 2.1, I conclude that herding behavior of cryptocurrency investors is only significant in the short term, and herding is not significant most of the time. Meanwhile, when considering smaller time-window, such as 30-day, the coefficient of $R^2_{m.t}$ is more volatile, with a shorter time interval in the negative. Whereas, when considering a 60-day or even 90-day time-window, the time interval when $R^2_{m.t}$ is negative becomes longer, implying a longer duration of herding behavior in the cryptocurrency market.

The overall findings indicate that investors hardly exhibit herding behavior in response to cryptocurrency market movements, and if they do, it only persists in the short term. Kaiser and Stöckl (2020) found that herding behavior in the cryptocurrency market can be explained by the dominating role of irrational individual investors in the most traded cryptocurrency. Such interpretation can be the explanation for previously significant results from other research found by Gurdgiev and O'Loughlin (2020); Bouri et al. (2019). However, with the development of the cryptocurrency market during the selected time period for estimation, investors are more likely to receive market-related information to make their own investment decisions relative to before, so the results of this paper are consistent with the results of Yao et al. (2014) who find that when the market matures with information efficiency and its potential herding behavior gradually disappears.

Also, such insignificant herding behavior results have the potential to be explained by the fact that the potential herding behavior is not sensitive to the swings in the cryptocurrency market itself. The movements of the cryptocurrency market do not affect the herding behavior of investors, which shows that the market conditions do not affect the investment strategies of investors in the market without considering the influence of external factors on cryptocurrencies and that they continue to trade based on their private information. This result may arise because most cryptocurrencies are highly correlated with each other, so the market movement will make the most cryptocurrencies' price move in one direction[16]. Thus, it is difficult for investors to get a better choice than the private information they have, so herding behavior is not significant in this situation.

[16] See Appendix Table A.1.

2.5.2 Conditional herding with external factors intervention

In this section, I estimate whether the external sentiment factors or financial market spillover are related to herding behavior in the cryptocurrency market. Column (1) in Table A.17 represents the estimation result of the baseline CSAD model without external factors for further comparison. The result of Column (2) in Table A.17 exhibits a significantly negative coefficient for the VIX index, indicating the existence of conditional herding in the cryptocurrency market with the intervention of variation of the VIX index. Column (3) of Table A.17 shows similar results, and the global EPU coefficient is also negative and highly significant. Combining the result of Columns (2) and (3) of Table A.17, I conclude that an increase in the sentiment of external financial markets and economic policies will also result in herding within the cryptocurrency market. Such results become insignificant when considering the sentiment index associated with representing the cryptocurrency market. In Columns (4) through (6) of Table A.17, the coefficients of the CRIX index, which represents cryptocurrency market sentiment, and the two GT indexes with different search terms, which represent investor attention, are positive and significant, demonstrating that changes in these indices do not lead to herding behavior among investors in the cryptocurrency market. Such findings support the results of internal herding consideration, which is that investors do not herd due to market movements and sentiment in cryptocurrency market.

Investor sentiment in the market fluctuates as a result of market instability. Columns (7) and (8) in Table A.17 illustrates the estimation results of the potential impact of market fluctuation rather than sentiment in the FOREX and stock markets on the herding behavior of the cryptocurrency market. Column (7) reveals that the coefficient of $CSAD_{forex,t}$ is significant and negative, and the coefficient of $R^2_{forex,t}$ is positive and insignificant, while the all coefficients in Column (8), representing the influence from stock market, are both not significant. Such results confirm that instantaneous market fluctuations in traditional financial markets like FOREX and stocks do not lead investors to herd in the cryptocurrency market. And only then will such sentiment impact the cryptocurrency market, leading to herding behavior within the cryptocurrency market as investors become more anxious and concerned about the direction of financial markets and economic policies in the future.

Table A.18 investigates further the potential impact of changes in the VIX index on investors' herding behavior under different market conditions in cryptocurrency. Column (1) shows that when the market is in a down market, the coefficient of $R^2_{crypto,t}$ is positive and significant, while the coefficient of $Ln(VIX)$ is negative and significant. When the market is upmarket, the coefficient of $R^2_{crypto,t}$ is negative but insignificant. At the same time, the coefficient of $Ln(VIX)$ is also negative and more prominent in the upmarket compared to the down market. The results of Column (1) are consistent with the previous results, showing that the index measuring investment sentiment in the external market significantly influences herding behavior within the cryptocurrency market and that this influence is more substantial when the price of cryptocurrencies is rising.

The results in Columns (2) through (7) in Table A.18 further support these findings. The results of Column (2) demonstrate that the coefficient of $R^2_{crypto,t}$ is significant and negative at the 5% confidence interval and the coefficient of $Ln(VIX)$ is larger than that of the up market alone when the return of the cryptocurrency market is at 75% of the upper quantile. This result suggests that the cryptocurrency market exhibits significant herding behavior and that the external sentiment index has a greater impact on investors' herding behavior. Comparing Columns (3) and (4) with Column (2) of Table A.18, I observe that when the cryptocurrency market is in a more extreme up market, the change in the VIX index has a greater impact on the market's herding behavior and on the investors' herding behavior. When turning to the extreme down market at different levels of the cryptocurrency market, the movement of the VIX index still impacts the herding behavior of the cryptocurrency market, but not to the same extent as the up market. Meanwhile, the coefficient of $R^2_{crypto,t}$ is positive for all three levels of down market and is insignificant only when the cryptocurrency is in the lower quantile down market of 5%. This result proves that the external market sentiment has less influence on the herding behavior within the cryptocurrency market when the cryptocurrency is in the down market than in the up market.

Table A.19 illustrates the estimation results for considering whether varying levels of external sentiment indices would also produce comparable outcomes. Similar to Table A.18, Column (1) of Table A.19 demonstrates that changes in $Ln(VIX)$ significantly impact herding behavior in the cryptocurrency market, with a greater impact when the VIX index is on an upward trend. Comparing Columns (2) through (4) of Table A.18 to Columns (5) through (7), when market panic, as measured by the VIX index, is greater, that is, when the VIX index is greater, the increased VIX index results in greater cryptocurrency market herding. In contrast, when market panic is lower, for example, when the VIX index is smaller, the influence of the VIX index on herding behavior gradually decreases, and even when the VIX index is at the lower quantile of 5%, an increase in the VIX index can reduce herding behavior in the cryptocurrency market. Compared to Table A.18, Table A.19 demonstrates that changes in the VIX index have a smaller impact on herding behavior when only considering the movement of VIX index itself than considering the market movement of cryptocurrency.

Considering the sentiment index on economic policies, as shown in Table A.2, the results demonstrate that the increase of the EPU index in up markets is more susceptible to investor herding behavior than in down markets, and the more extreme the up market, the greater the impact. The more stable the economic policy, that is, the smaller the EPU index, the less impact the increase in EPU has on herding behavior when considering only the change in EPU index. Such result is consistent with Table A.18 and A.19, which shows that investor sentiment indexes on traditional financial markets and economic policies have the same impact on herding behavior in cryptocurrency markets.

When shedding light on the sentiment index of external financial markets and economic policies to an index that reflects the sentiment and concerns within the cryptocurrency market, I find that when the higher CRIX index is associated with less herding activity in the cryptocurrency market. The results of Column (1) of Table A.3 shows that the coefficient of $Ln(CRIX)$ is significant and positive in both up and down markets, which proves that an increase in the CRIX index reduces the herding behavior of investors in the cryptocurrency market. Also, the coefficient of $Ln(CRIX)$ is more extensive in the down market than in the up market, implying that CRIX has a more significant effect on reducing herding behavior in the down market than in the up market. Such findings are further confirmed by considering different quantile market scenarios in cryptocurrency. Columns (2) through (7) of Table A.3 show that the coefficient of $Ln(CRIX)$ becomes progressively smaller and less significant when the cryptocurrency is in the more extreme up market, and progressively more extensive and more significant when it is in the more extreme down market. Columns (8) to (14) of Table A.3 consider the effect of $Ln(CRIX)$ in different quantiles on herding behavior. Similar results are obtained, that is, an increased CRIX index mitigates herding behavior in the cryptocurrency market.

Tables A.4 and A.5 consider the effect of herding behavior when the internet's attention to cryptocurrency is in different quantiles. Similar to the results in Table A.3, when the cryptocurrency market is in different quantile ranges, the increase in GT index leads to a decrease in herding behavior, especially when the cryptocurrency is in the extreme up (95% quantile) and down (5%) market. However, when considering that the GT index is in a different interval, the regression results are different compared to the previous CRIX index consideration. When the GT index is in the lower extreme quantile, the coefficients of the GT index are negative and significant. In contrast, the coefficients are positive and significant when the GT index is in the upper extreme quantile. Such result indicates that when cryptocurrencies are in a stable period or a period of low attention, investors will strengthen their herding behavior due to the increase in internet attention. When the cryptocurrency market receives great attention, i.e., high volatility or very high or very low returns, the further increase in internet attention will bring investors' behavior back to rational and reduce herding behavior.

The VIX index, which reflects the fear in financial markets, and the EPU index, which reflects economic policy uncertainty, tend to rise due to unfavorable market conditions. When the cryptocurrency market is extremely high, investor sentiment is more likely to be infected by the external financial market or economic policy fears and, therefore, more prone to herding behavior. In comparison, the impact of a change in fear from outside the cryptocurrency market on investors' herding behavior is lower. Whereas the CRIX index reflects the general performance of the cryptocurrency market, and the GT index reflects the public's interest in cryptocurrencies, these two indices will only rise when the cryptocurrency market is in an upward or highly volatile phase. When the performance and attention indices within the cryptocurrency market increase, it often means that the cryptocurrency market is in a high return or rising phase, when investors are making different choices due to the high return of cryptocurrencies and therefore herding behavior decreases with the increase of these indices.

I next further consider whether fluctuations in external traditional financial markets would cause different levels of spillover effect on the cryptocurrency market under different market conditions. Column (1) of Table A.20 shows that the coefficients of $R^2_{market,t}$ are positive for both up and down markets, and the coefficients of $CSAD_{market,t}$ are negative. Such result proves that the FOREX market does not have a herding and volatility spillover effect on the cryptocurrency market under different market scenarios. Column (2) considers the market condition of the FOREX market, but neither the up nor the down market has herding and volatility spillover effects on the cryptocurrency market. Columns (3) and (4) of Table A.20 illustrates the empirical results of whether there is a herding and volatility spillover effect in the US stock market under different conditions of cryptocurrency or the stock market. The regression results show that all $CSAD_{market,t}$ and $R^2_{market,t}$ are insignificant, proving that the stock market does not have significant volatility and spillover effects on the cryptocurrency market. Combining the previous results, I conclude that traditional financial markets' influence on investors' behavior in the cryptocurrency market exists only in the transmission of fear.

2.6 Conclusion

This paper provides empirical evidence on cryptocurrency market herding behavior and how the herding behavior is influenced by external factor and spillover effect from other markets. I analyze daily closing prices for the top 59 cryptocurrencies from November 10th, 2017, to December 31st, 2021, investigating how the potential of herding behavior is influenced by internal market, external sentiment proxies and external financial market spillover effects consideration with well-known CSAD model.

In the internal market consideration, this paper shows insignificant herding behavior in the cryptocurrency market in the baseline CSAD model, in both unconditional and market conditional consideration. When considering the time-varying format, the herding behavior only exists in a short-term period. This finding is consistent with previous literature that anti-herding inside the cryptocurrency market (Coskun et al., 2020; Stavroyiannis and Babalos, 2019). In external sentiment proxies' consideration, this paper confirms a significant herding behavior in the cryptocurrency market but shows the different impact according to proxy classification. The VIX and Global EPU indexes, which are constructed with the market and policy-related information external to the cryptocurrency market, have a significant impact on herding behavior in both the up and down markets, with a stronger tendency in the up market, and the rise in both VIX and Global EPU indexes will lead to an increase in herding behavior in the cryptocurrency market. When considering the CRIX and GT Index, the increase in internal market performance and attention of the cryptocurrency market will lead to a decrease in the herding behavior in the cryptocurrency market. Therefore, this paper concludes that herding behavior in cryptocurrency markets is more sensitive to market external information and up market conditions.

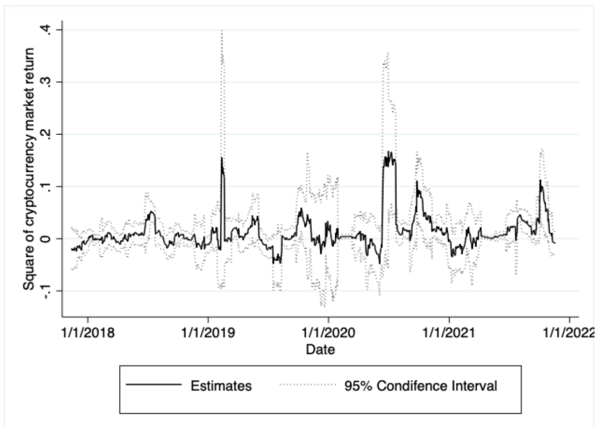
In external financial market spillover effect consideration, this paper selects U.S. stock market and major FOREX market to analyze and investigate how the external financial market's volatility and herding spillover effect influences the herding behavior in cryptocurrency market. Empirical results indicate that both these two financial markets have no significant herding and volatility spillover effect on the cryptocurrency market, even considering both the cryptocurrency and financial market movements. Combining the previous results, I further conclude that traditional financial markets' influence on investors' behavior in cryptocurrency exists only in the transmission of fear.

The results of this study have important implications for policymakers, academics, and investors in the cryptocurrency market. The first implication arises in the form of a better understanding of the current cryptocurrency market, highlighting the importance of prerequisites when considering herding. Second, this research shows the connectedness between cryptocurrency and the traditional financial market, suggesting a further potential investment portfolio construction.

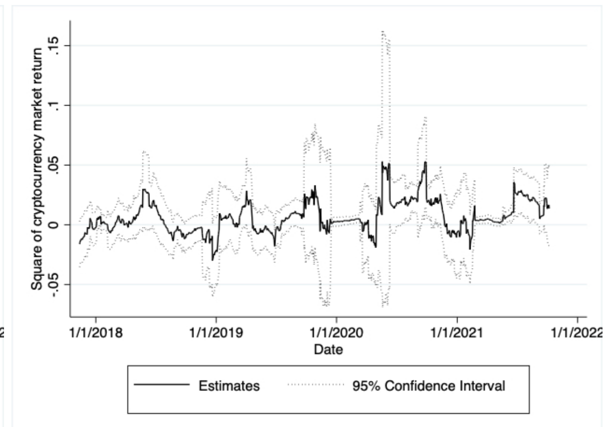
Figures

Figure 2.1: Time-varying rolling window estimation of internal herding

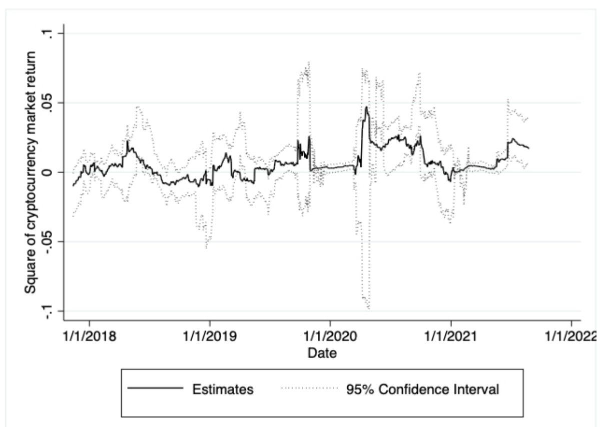
(a) 30-day rolling window



(b) 60-day rolling window



(c) 90-day rolling window



Note(s): This figure present the empirical results of the following equation with time varying rolling-window estimation:
$CSAD_{m,t} = \beta_0 + \beta_1 |R_{m,t}| + \beta_2 R^2_{m,t} + u_t$.
$CSAD_{m,t}$ denotes the Cross-Sectional Absolute Deviation of the cryptocurrency market; $|R_{m,t}|$ is the absolute value of market return on day t; $R^2_{m,t}$ is their equal-weight square average market return on day t. Figure (a), (b) and (c) present the 30-day, 60-day and 90-day rolling-window estimation for CSAD model, respectively.

Table A.13: Selected Cryptocurrencies for Estimations

Name	Time Period	Average Return (%)	Name	Time Period	Average Return (%)
Bitcoin (BTC)	(11/10/17 - 12/31/21)	0.124	Stellar (XLM)	(11/10/17 - 12/31/21)	0.126
Ethereum (ETH)	(11/10/17 - 12/31/21)	0.161	Decentraland (MANA)	(11/10/17 - 12/31/21)	0.355
Tether (USDT)	(11/10/17 - 12/31/21)	0.367	Bitcoin BEP2 (BTCB)	(06/18/19 - 12/31/21)	0.179
Binance Coin (BNB)	(11/10/17 - 12/31/21)	0.000	Hedera (HBAR)	(09/17/19 - 12/31/21)	0.140
USD Coin (USDC)	(10/09/18 - 12/31/21)	-0.001	Ethereum Classic (ETC)	(11/10/17 - 12/31/21)	0.058
XRP (XRP)	(11/10/17 - 12/31/21)	0.089	The Sandbox (SAND)	(08/14/20 - 12/31/21)	0.891
Terra (LUNA1)	(07/27/19 - 12/31/21)	0.468	Monero (XMR)	(11/10/17 - 12/31/21)	0.042
Cardano (ADA)	(11/10/17 - 12/31/21)	0.245	Elrond (EGLD)	(09/04/20 - 12/31/21)	0.539
Solana (SOL)	(04/11/20 - 12/31/21)	0.823	Waves (WAVES)	(11/10/17 - 12/31/21)	0.072
HEX (HEX)	(12/18/19 - 12/31/21)	0.996	Filecoin (FIL)	(12/13/17 - 12/31/21)	0.074
Binance USD (BUSD)	(09/21/19 - 12/31/21)	0.000	VeChain (VET)	(08/03/18 - 12/31/21)	0.135
Avalanche (AVAX)	(07/14/20 - 12/31/21)	0.656	Fantom (FTM)	(10/30/18 - 12/31/21)	0.402
Polkadot (DOT)	(08/21/20 - 12/31/21)	0.446	Axie Infinity (AXS)	(11/04/20 - 12/31/21)	01.531
Dogecoin (DOGE)	(11/10/17 - 12/31/21)	0.317	Klaytn (KLAY)	(03/30/20 - 12/31/21)	0.380
Terra (UST)	(11/26/20 - 12/31/21)	0.000	Frax (FRAX)	(12/27/20 - 12/31/21)	0.004
SHIBA INU (SHIB)	(08/01/20 - 12/31/21)	1.350	THETA (THETA)	(01/17/18 - 12/31/21)	0.225
Wrapped Bitcoin (WBTC)	(01/30/19 - 12/31/21)	0.243	Tezos (XTZ)	(11/10/17 - 12/31/21)	0.062
Polygon (MATIC)	(04/28/19 - 12/31/21)	0.649	THORChain (RUNE)	(07/23/19 - 12/31/21)	0.671
Dai (DAI)	(11/22/19 - 12/31/21)	0.001	Helium (HNT)	(06/08/20 - 12/31/21)	0.861
Crypto.com Coin (CRO)	(12/14/18 - 12/31/21)	0.301	BitTorrent (BTT)	(01/31/19 - 12/31/21)	0.162
Cosmos (ATOM)	(03/14/19 - 12/31/21)	0.157	Zcash (ZEC)	(11/10/17 - 12/31/21)	-0.038
Litecoin (LTC)	(11/10/17 - 12/31/21)	0.054	Wrapped BNB (WBNB)	(09/29/20 - 12/31/21)	0.629
NEAR Protocol (NEAR)	(10/14/20 - 12/31/21)	0.567	Aave (AAVE)	(10/02/20 - 12/31/21)	1.363
TRON (TRX)	(11/10/17 - 12/31/21)	0.229	EOS (EOS)	(11/10/17 - 12/31/21)	0.0062
Chainlink (LINK)	(11/10/17 - 12/31/21)	0.296	Maker (MKR)	(11/20/17 - 12/31/21)	0.143
Uniswap (UNI1)	(09/17/20 - 12/31/21)	0.339	IOTA (MIOTA)	(11/10/17 - 12/31/21)	0.062
UNUS SED LEO (LEO)	(05/21/19 - 12/31/21)	0.135	Flow (FLOW)	(01/27/21 - 12/31/21)	0.073
Bitcoin Cash (BCH)	(11/10/17 - 12/31/21)	-0.028	The Graph (GRT1)	(12/17/20 - 12/31/21)	0.443
Lido stETH (STETH)	(12/23/20 - 12/31/21)	0.490	PancakeSwap (CAKE)	(09/29/20 - 12/31/21)	0.514
Algorand (ALGO)	(06/21/19 - 12/31/21)	-0.072	**Total**	**(11/10/17 – 12/31/21)**	**0.332**

Note(s): This table reports all selected 59 cryptocurrencies in empirical analysis. The name column includes the full name of each cryptocurrency, and the abbreviation is inside the bracket. The time period column illustrates the time period with available data, and the format is (MM/DD/YY – MM/DD/YY). If the start date of specific cryptocurrency is later than 11/10/19, it means the launch date of this cryptocurrency is later than the start of estimated time period, then this paper uses the daily closing data from the launch date to 12/31/21. The average return column reports the daily average return of each selected cryptocurrency with activate trading day. The right side of last column reports the whole time period for empirical results, and the average daily return the cryptocurrency market.

Table A.14: Summary Statistics: equally weighted portfolios return Rm and CSAD in cryptocurrency, FOREX, stock market and stock market sectors

Market Portfolio	Obs	Variables	Mean	Min.	Max.	Std. dev.	Skew.	Kurt.	J-B Stat.
Panel A: Whole Market									
Cryptocurrency	59	Rm	0.0580	-43.0812	19.6450	5.0970	-1.5110	13.0760	4355.5***
		CSAD	3.5374	0.9274	18.5784	1.7740	2.9356	17.1490	9345.5***
FOREX	1847	Rm	0.0018	-0.2645	0.2589	0.0361	-0.0098	10.9780	2209.1***
		CSAD	0.1747	0.0532	0.0532	0.0634	2.8107	17.3246	8218.7***
Stock	4088	Rm	-0.0035	-5.5075	3.4674	0.6393	-1.6684	17.9643	10196***
		CSAD	1.0576	0.6701	3.7926	0.3037	3.5039	22.8460	19214***
Panel B: External Factors									
VIX			19.78	9.15	82.69	8.7867	2.6232	13.5143	8704.5***
CRIX			1616.2	274.9	6434.4	1583.173	1.4110	3.5649	522.14***
GT index (cryptocurrency)			7.261	0.12	100.00	12.706	4.071	20.957	1700***
GT index (bitcoin)			6.580	0.42	93	8.622	4.648	33.615	4400***
Global EPU			242.1	123.8	430.1	65.3681	0.6293	3.02390	99.893***

Note(s): This table represents data sample's properties and summary statistics of key variables. The Obs means the maximum number of components of the market portfolio in the time duration from 11/10/2017 to 12/31/2021. Missing value won't be counted in market portfolio calculation. Key variables include Rm (the equal-weighted average market return in daily closing price consideration) and CSAD (the Cross-Sectional Absolute Deviation of the daily market portfolio return). *** indicates the statistical significance in 0.001 level.

Table A.15: Herding in cryptocurrency market

	Dep Var: $CSAD_{crypto,t}$				
	(1)	(2)	(3)	(4)	(5)
$R^2_{crypto,t}$	0.002*				
	(0.001)				
$Down_{crypto} \times R^2_{crypto,t}$		0.005***			
		(0.001)			
$Up_{crypto} \times R^2_{crypto,t}$		0.001			
		(0.005)			
$Down_{Bitcoin} \times R^2_{crypto,t}$			0.004***		
			(0.001)		
$Up_{Bitcoin} \times R^2_{crypto,t}$			0.006		
			(0.005)		
$Down_{Bitcoin} \times Up_{altcoin} \times R^2_{crypto,t}$				0.041	
				(0.058)	
$(1 - Down_{Bitcoin} \times Up_{altcoin}) \times R^2_{crypto,t}$				0.002**	
				(0.001)	
$Up_{Bitcoin} \times Down_{altcoin} \times R^2_{crypto,t}$					0.098**
					(0.046)
					0.002*
					(0.001)
Constant	2.605***	2.577***	2.628***	2.543***	2.617***
	(0.072)	(0.072)	(0.074)	(0.074)	(0.075)
Observations	1,040	1,040	1,040	1,040	1,040
R-squared	0.333	0.374	0.353	0.374	0.374

Note(s): This table present the empirical results of the following equations:
$CSAD_{m,t} = \beta_0 + \beta_1|R_{m,t}| + \beta_2 R^2_{m,t} + u_t$.(Column (1))
$CSAD_{m,t} = \beta_0 + D_{up} * (\beta_1|R_{m,t}| + \beta_2 R^2_{m,t}) + (1 - D_{up}) * (\beta_3|R_{m,t}| + \beta_4 R^2_{m,t}) + u_t$ (Column (2) to (4))$CSAD_{m,t}$ denotes the Cross-Sectional Absolute Deviation of the cryptocurrency market; $|R_{m,t}|$ is the absolute value of market return on day t; $R^2_{m,t}$ is their equal-weight square average market return on day t. $Up_{crypto,t}$ ($Down_{crypto,t}$) is a dummy variable and equal to 1 if the market return of cryptocurrency market is larger (smaller) than 0. $Up_{Bitcoin,t}$ ($Down_{Bitcoin,t}$) is a dummy variable and equal to 1 if the return of Bitcoin is larger (smaller) than 0. $Down_{Bitcoin,t} \times Up_{other_{crypto},t}$ is a dummy variable and equal to 1 if the market return of Bitcoin is smaller than 0 but other cryptocurrency is larger than 0. I only report the estimated coefficient of $R^2_{m,t}$ for table abbreviation. The sample period of equations estimation is from 11/10/2017 to 12/31/2021, all missing weekend and holiday data are excluded. Parentheses include the corrected standard errors according to Newey and West (1987), and ***, **, *, and . represent 0.1%, 1%, 5% and 10% significance level respectively.

Table A.16: Herding in cryptocurrency market in extreme market condition

	Dep Var: $CSAD_{crypto,t}$					
	(1)	(2)	(3)	(4)	(5)	(6)
$Up_{crypto,75\%} \times R^2_{crypto,t}$	0.002					
	(0.005)					
$(1 - Up_{crypto,75\%}) \times R^2_{crypto,t}$	0.005***					
	(0.001)					
$Up_{crypto,90\%} \times R^2_{crypto,t}$		-1.27e-05				
		(0.006)				
$(1 - Up_{crypto,90\%}) \times R^2_{crypto,t}$		0.004***				
		(0.001)				
$Up_{crypto,95\%} \times R^2_{crypto,t}$			-0.004			
			(0.008)			
$(1 - Up_{crypto,95\%}) \times R^2_{crypto,t}$			0.003***			
			(0.001)			
$Down_{crypto,25\%} \times R^2_{crypto,t}$				0.004***		
				(0.001)		
$(1 - Down_{crypto,25\%}) \times R^2_{crypto,t}$				0.002		
				(0.005)		
$Down_{crypto,10\%} \times R^2_{crypto,t}$					0.005***	
					(0.001)	
$(1 - Down_{crypto,10\%}) \times R^2_{crypto,t}$					0.007	
					(0.006)	
$Down_{crypto,5\%} \times R^2_{crypto,t}$						3.68e-05
						(0.002)
$(1 - Down_{crypto,5\%}) \times R^2_{crypto,t}$						0.007**
						(0.003)
Constant	2.699***	2.725***	2.696***	2.523***	2.538***	2.735***
	(0.066)	(0.067)	(0.067)	(0.079)	(0.089)	(0.080)
Observations	1,040	1,040	1,040	1,040	1,040	1,040
R-squared	0.373	0.357	0.349	0.371	0.360	0.343

Note(s): This table present the empirical results of the following equation:
$CSAD_{m,t} = \beta_0 + D_{up} * (\beta_1|R_{m,t}| + \beta_2 R^2_{m,t}) + (1 - D_{up}) * (\beta_3|R_{m,t}| + \beta_4 R^2_{m,t}) + u_t$ with extreme cryptocurrency market consideration
$CSAD_{m,t}$ denotes the Cross-Sectional Absolute Deviation of the cryptocurrency market; $|R_{m,t}|$ is the absolute value of market return on day t; $R^2_{m,t}$ is their equal-weight square average market return on day t. $Up_{crypto,q\%}$ ($Down_{crypto,q\%}$) is a dummy variable and equal to 1 if the market return of cryptocurrency market is larger (smaller) than 0 and above (below) the q% quantile. I only report the estimated coefficient of $R^2_{m,t}$ for table abbreviation. The sample period of equations estimation is from 11/10/2017 to 12/31/2021, all missing weekend and holiday data are excluded. Parentheses include the corrected standard errors according to Newey and West (1987), and ***, **, *, and . represent 0.1%, 1%, 5% and 10% significance level respectively.

Table A.17: Herding in the cryptocurrency market with external factors and market

	Dep Var: $CSAD_{crypto,t}$							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$R^2_{crypto,t}$	0.002*	0.002**	0.003**	0.002**	0.003*	0.003*	0.002*	0.002*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$Ln(VIX)$		-0.605***						
		(0.159)						
$Ln(EPU_{global})$			-0.913***					
			(0.202)					
$Ln(CRIX)$				0.380***				
				(0.047)				
$Ln(GT_{cryptocurrency})$					0.490***			
					(0.066)			
$Ln(GT_{bitcoin})$						0.645***		
						(0.073)		
$CSAD_{forex,t}$							-0.624**	
							(0.309)	
$R^2_{forex,t}$							2.199	
							(1.756)	
$CSAD_{stock,t}$								0.031
								(0.072)
$R^2_{stock,t}$								0.001
								(0.004)
Observations	1,040	1,040	1,040	1,040	1,040	1,040	1,040	1,040
R-squared	0.333	0.348	0.350	0.365	0.412	0.430	0.335	0.333

Note(s): This table present the empirical results of the following equations:
$CSAD_{m,t} = \beta_0 + \beta_1|R_{m,t}| + \beta_2 R^2_{m,t} + \beta_3 Factor_{i,t} + u_t$.(Column (1) to (6))
$CSAD_t = \beta_0 + \beta_1|R_{m,t}| + \beta_2 R^2_{m,t} + \beta_3 CSAD_{j,t} + \beta_4 R^2_{j,t} + u_t$ (Column (7) and (8)). $CSAD_{m,t}$ denotes the Cross-Sectional Absolute Deviation of the cryptocurrency market; $|R_{m,t}|$ is the absolute value of market return on day t; $R^2_{m,t}$ is their equal-weight square average market return on day t. $Ln(VIX)$, $Ln(CRIX)$, $Ln(GT_{cryptocurrency})$, $Ln(GT_{bitcoin})$ and $Ln(EPU_{global})$ represent the natural logarithm transformation of VIX index, CRIX index, Google Search with "cryptocurrency" search term, Google Search with "bitcoin" search term, and global Economic Policy Uncertainty index, respectively. I only report the estimated coefficient of $R^2_{m,t}$ for table abbreviation. The sample period of equations estimation is from 11/10/2017 to 12/31/2021, all missing weekend and holiday data are excluded. Parentheses include the corrected standard errors according to Newey and West (1987), and ***, **, *, and . represent 0.1%, 1%, 5% and 10% significance level respectively.

Table A.18: Herding in the cryptocurrency market with VIX index in different cryptocurrency market condition

	Dep Var: $CSAD_{crypto,t}$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$Down_{crypto} \times R^2_{crypto,t}$	0.006***						
	(0.001)						
$Up_{crypto} \times R^2_{crypto,t}$	-0.001						
	(0.005)						
$Down_{crypto} \times Ln(VIX)$	-0.571***						
	(0.153)						
$Up_{crypto} \times Ln(VIX)$	-0.646***						
	(0.159)						
$Up_{crypto,75\%} \times R^2_{crypto,t}$		-0.020**					
		(0.008)					
$Up_{crypto,75\%} \times Ln(VIX)$		-0.974***					
		(0.224)					
$Up_{crypto,90\%} \times R^2_{crypto,t}$			-0.043***				
			(0.015)				
$Up_{crypto,90\%} \times Ln(VIX)$			-2.102***				
			(0.533)				
$Up_{crypto,95\%} \times R^2_{crypto,t}$				-0.053***			
				(0.020)			
$Up_{crypto,95\%} \times Ln(VIX)$				-3.259***			
				(0.944)			
$Down_{crypto,25\%} \times R^2_{crypto,t}$					0.006***		
					(0.001)		
$Down_{crypto,25\%} \times Ln(VIX)$					-0.569***		
					(0.157)		
$Down_{crypto,10\%} \times R^2_{crypto,t}$						0.007***	
						(0.001)	
$Down_{crypto,10\%} \times Ln(VIX)$						-0.415**	
						(0.185)	
$Down_{crypto,5\%} \times R^2_{crypto,t}$							0.001
							(0.002)
$Down_{crypto,5\%} \times Ln(VIX)$							-0.443***
							(0.171)
Observations	1,040	1,040	1,040	1,040	1,040	1,040	1,040
R-squared	0.391	0.399	0.394	0.390	0.388	0.377	0.355

Note(s): This table present the empirical results of the following equations:
$CSAD_{m,t} = \beta_0 + D_{up} * (\beta_1 |R_{m,t}| + \beta_2 R^2_{m,t} + \beta_3 factor_{i,t}) + (1 - D_{up}) * (\beta_4 |R_{m,t}| + \beta_5 R^2_{m,t} + \beta_6 factor_{i,t}) + u_t$ with extreme cryptocurrency market consideration.
For abbreviation this table only report the estimated coefficients β_2 and β_3. Parentheses include the corrected standard errors according to Newey and West (1987), and ***, **, *, and . represent 0.1%, 1%, 5% and 10% significance level respectively.

Table A.19: Herding in the cryptocurrency market with VIX index in different VIX index movement condition

	Dep Var: $CSAD_{crypto,t}$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$Down_{VIX} \times R^2_{crypto,t}$	0.002						
	(0.003)						
$Up_{VIX} \times R^2_{crypto,t}$	0.001						
	(0.002)						
$Down_{VIX} \times Ln(VIX)$	-0.555***						
	(0.156)						
$Up_{VIX} \times Ln(VIX)$	-0.635***						
	(0.164)						
$Up_{VIX,75\%} \times R^2_{crypto,t}$		0.004***					
		(0.001)					
$Up_{VIX,75\%} \times Ln(VIX)$		-1.138***					
		(0.258)					
$Up_{VIX,90\%} \times R^2_{crypto,t}$			0.004***				
			(0.001)				
$Up_{VIX,90\%} \times Ln(VIX)$			-0.615***				
			(0.204)				
$Up_{VIX,95\%} \times R^2_{crypto,t}$				0.006***			
				(0.002)			
$Up_{VIX,95\%} \times Ln(VIX)$				-0.522***			
				(0.178)			
$Down_{VIX,25\%} \times R^2_{crypto,t}$					-0.006		
					(0.004)		
$Down_{VIX,25\%} \times Ln(VIX)$					-0.673***		
					(0.187)		
$Down_{VIX,10\%} \times R^2_{crypto,t}$						-0.017**	
						(0.007)	
$Down_{VIX,10\%} \times Ln(VIX)$						-0.072	
						(0.220)	
$Down_{VIX,5\%} \times R^2_{crypto,t}$							-0.024
							(0.015)
$Down_{VIX,5\%} \times Ln(VIX)$							0.680*
							(0.353)
Observations	1,040	1,040	1,040	1,040	1,040	1,040	1,040
R-squared	0.353	0.363	0.350	0.351	0.360	0.414	0.472

Note(s): This table present the empirical results of the following equations:
$CSAD_{m,t} = \beta_0 + D_{up} * (\beta_1|R_{m,t}| + \beta_2 R^2_{m,t} + factor_{i,t}) + (1 - D_{up}) * (\beta_3|R_{m,t}| + \beta_4 R^2_{m,t} + factor_{i,t}) + u_t$ with extreme cryptocurrency market consideration.
For abbreviation this table only report the estimated coefficients β_2 and β_3. Parentheses include the corrected standard errors according to Newey and West (1987), and ***, **, *, and . represent 0.1%, 1%, 5% and 10% significance level respectively.

Table A.20: Herding in the cryptocurrency market with external market spillover

	Dep Var: $CSAD_{crypto,t}$			
	Panel A: FOREX market		Panel B: Stock market	
	(1)	(2)	(3)	(4)
$Down_{crypto} \times R^2_{crypto,t}$	0.005***		0.005***	
	(0.001)		(0.001)	
$Up_{crypto} \times R^2_{crypto,t}$	-0.001		-0.001	
	(0.005)		(0.005)	
$Down_{crypto} \times CSAD_{market,t}$	-0.408		0.037	
	(0.308)		(0.078)	
$Up_{crypto} \times CSAD_{market,t}$	-0.940**		-0.018	
	(0.373)		(0.086)	
$Down_{crypto} \times R^2_{market,t}$	1.964		0.004	
	(1.794)		(0.004)	
$Up_{crypto} \times R^2_{market,t}$	4.634*		-0.004	
	(2.542)		(0.012)	
$Down_{market} \times R^2_{crypto,t}$		0.003**		0.001
		(0.001)		(0.001)
$Up_{market} \times R^2_{crypto,t}$		-0.003		0.0002
		(0.002)		(0.004)
$Down_{market} \times CSAD_{market,t}$		-0.687*		0.034
		(0.401)		(0.083)
$Up_{market} \times CSAD_{market,t}$		-0.750**		0.081
		(0.305)		(0.078)
$Down_{market} \times R^2_{market,t}$		1.583		0.002
		(1.427)		(0.005)
$Up_{market} \times R^2_{market,t}$		5.837***		-0.013
		(2.044)		(0.012)
Observations	1,040	1,040	1,040	1,040
R-squared	0.378	0.339	0.376	0.334

Note(s): This table present the empirical results of the following equations:

$$CSAD_t = \beta_0 + D_{up}(\beta_1|R_{m,t}| + \beta_2 R^2_{m,t} + \beta_3 CSAD_{j,t} + \beta_4 R^2_{j,t}) + (1 - D_{up})(\beta_5|R_{m,t}| + \beta_6 R^2_{m,t} + \beta_7 CSAD_{j,t} + \beta_8 R^2_{j,t}) + u_t$$

$CSAD_{m,t}$ denotes the Cross-Sectional Absolute Deviation of the cryptocurrency market and external financial market at time t; $|R_{m,t}|$ is the absolute value of market return on day t; $R^2_{m,t}$ is their equal-weight square average market return on day t. Up_{crypto} ($Down_{crypto}$) is a dummy variable and equal to 1 if the market return of cryptocurrency market is larger (smaller) than 0. Up_{market} ($Down_{market}$) is a dummy variable and equal to 1 if the market return of financial market identified as panel name is larger (smaller) than 0 . I only report the estimated coefficient of $R^2_{m,t}$ and external financial markets' $CSAD_{m,t}$ and $R^2_{m,t}$ for table abbreviation. The sample period of equations estimation is from 11/10/2017 to 12/31/2021, all missing weekend and holiday data are excluded. Parentheses include the corrected standard errors according to Newey and West (1987), and ***, **, *, and . represent 0.1%, 1%, 5% and 10% significance level respectively.

Appendix Tables

Table A.1: Corrlation matrix of top 10 cryptocurrencies based on market capitalization

Cryptocurrency	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) BTC	1.000									
(2) ETH	0.920	1.000								
(3) BNB	0.916	0.958	1.000							
(4) USDC	-0.371	-0.317	-0.292	1.000						
(5) USDT	-0.130	-0.125	-0.123	0.244	1.000					
(6) XRP	0.557	0.662	0.597	-0.213	-0.053	1.000				
(7) LUNA1	0.717	0.851	0.812	-0.152	-0.153	0.678	1.000			
(8) ADA	0.886	0.943	0.912	-0.295	-0.114	0.692	0.726	1.000		
(9) SOL	0.664	0.868	0.785	-0.141	-0.151	0.660	0.913	0.742	1.000	
(10) HEX	0.591	0.766	0.675	-0.133	-0.110	0.637	0.777	0.802	0.839	1.000

Table A.2: Herding in the cryptocurrency market with EPU index in different market and movement condition

	Dep Var: $CSAD_{crypto,t}$													
	Panel A: $Up_{m,t} = Up_{crypto,t}$							*Panel B:* $Up_{m,t} = Up_{EPU,t}$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
$Down_{m,t} \times R^2_{crypto,t}$	0.006***							0.003**						
	(0.001)							(0.001)						
$Up_{m,t} \times R^2_{crypto,t}$	0.0002							0.066**						
	(0.005)							(0.027)						
$Down_{m,t} \times Ln(EPU)$	-0.875***							-0.896***						
	(0.196)							(0.202)						
$Up_{m,t} \times Ln(EPU)$	-0.897***							-0.985***						
	(0.199)							(0.211)						
$Up_{m,t,75\%} \times R^2_{crypto,t}$		-0.013							0.004***					
		(0.008)							(0.001)					
$Up_{m,t,75\%} \times Ln(EPU)$		-0.992***							-1.577***					
		(0.218)							(0.325)					
$Up_{m,t,90\%} \times R^2_{crypto,t}$			-0.030							0.005***				
			(0.019)							(0.001)				
$Up_{m,t,90\%} \times Ln(EPU)$			-1.410***							-1.011***				
			(0.415)							(0.251)				
$Up_{m,t,95\%} \times R^2_{crypto,t}$				-0.046							0.014			
				(0.029)							(0.020)			
$Up_{m,t,95\%} \times Ln(EPU)$				-2.032**							-0.922***			
				(0.809)							(0.227)			
$Down_{m,t,25\%} \times R^2_{crypto,t}$					0.005***							0.001		
					(0.001)							(0.002)		
$Down_{m,t,25\%} \times Ln(EPU)$					-0.882***							-0.913***		
					(0.199)							(0.221)		
$Down_{m,t,10\%} \times R^2_{crypto,t}$						0.006***							-0.005	
						(0.002)							(0.007)	
$Down_{m,t,10\%} \times Ln(EPU)$						-0.817***							0.211	
						(0.210)							(0.187)	
$Down_{m,t,5\%} \times R^2_{crypto,t}$							0.002							-0.006
							(0.002)							(0.009)
$Down_{m,t,5\%} \times Ln(EPU)$							-0.696***							0.565**
							(0.181)							(0.231)
Observations	1,040	1,040	1,040	1,040	1,040	1,040	1,040	1,040	1,040	1,040	1,040	1,040	1,040	1,040
R-squared	0.391	0.393	0.378	0.370	0.388	0.377	0.362	0.354	0.363	0.355	0.351	0.353	0.389	0.432

Table A.3: Herding in the cryptocurrency market with CRIX index in different market and diffferent movement condition

	Dep Var: $CSAD_{crypto,t}$													
	Panel A: $Up_{m,t}= Up_{crypto,t}$							Panel B: $Up_{m,t}= Up_{CRIX,t}$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
$Down_{\mathrm{m,t}} \times R^2_{crypto,t}$	0.005***							0.004***						
	(0.001)							(0.001)						
$Up_{m,t} \times R^2_{crypto,t}$	0.002							0.005						
	(0.006)							(0.006)						
$Down_{m,t} \times Ln(CRIX)$	0.367***							0.382***						
	(0.049)							(0.050)						
$Up_{m,t} \times Ln(CRIX)$	0.351***							0.355***						
	(0.048)							(0.048)						
$Up_{m,t,75\%} \times R^2_{crypto,t}$		-0.012*							0.001					
		(0.007)							(0.001)					
$Up_{m,t,75\%} \times Ln(CRIX)$		0.263***							0.740***					
		(0.060)							(0.134)					
$Up_{m,t,90\%} \times R^2_{crypto,t}$			-0.011							0.016				
			(0.009)							(0.011)				
$Up_{m,t,90\%} \times Ln(CRIX)$			0.223*							0.489***				
			(0.135)							(0.058)				
$Up_{m,t,95\%} \times R^2_{crypto,t}$				-0.007							0.023**			
				(0.011)							(0.010)			
$Up_{m,t,95\%} \times Ln(CRIX)$				0.327							0.393***			
				(0.266)							(0.056)			
$Down_{m,t,25\%} \times R^2_{crypto,t}$					0.005***							0.004***		
					(0.001)							(0.001)		
$Down_{m,t,25\%} \times Ln(CRIX)$					0.361***							0.262***		
					(0.051)							(0.065)		
$Down_{m,t,10\%} \times R^2_{crypto,t}$						0.005***							0.004***	
						(0.002)							(0.001)	
$Down_{m,t,10\%} \times Ln(CRIX)$						0.377***							0.424***	
						(0.065)							(0.065)	
$Down_{m,t,5\%} \times R^2_{crypto,t}$							0.005**							-0.022
							(0.002)							(0.019)
$Down_{m,t,5\%} \times Ln(CRIX)$							0.693***							0.326***
							(0.086)							(0.069)
Observations	1,040	1,040	1,040	1,040	1,040	1,040	1,040	1,040	1,040	1,040	1,040	1,040	1,040	1,040
R-squared	0.402	0.405	0.391	0.381	0.399	0.390	0.410	0.382	0.378	0.375	0.369	0.370	0.367	0.366

Table A.4: Herding in the cryptocurrency market with GT index with "cryptocurrency" search term in different movement condition

	Dep Var: $CSAD_{crypto,t}$													
	Panel A: $Up_{m,t} = Up_{crypto,t}$							*Panel B: $Up_{m,t} = Up_{GT,cryptocurrency,t}$*						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
$Down_{m,t} \times R^2_{crypto,t}$	0.004***							-0.001						
	(0.001)							(0.005)						
$Up_{m,t} \times R^2_{crypto,t}$	0.003							0.003**						
	(0.005)							(0.001)						
$Down_{m,t} \times Ln(GT_{cryptocurrency})$	0.336***							0.428***						
	(0.087)							(0.085)						
$Up_{m,t} \times Ln(GT_{cryptocurrency})$	0.557***							0.580***						
	(0.077)							(0.081)						
$Up_{m,t,75\%} \times R^2_{crypto,t}$		0.007							0.002					
		(0.005)							(0.002)					
$Up_{m,t,75\%} \times Ln(GT_{cryptocurrency})$		0.724***							0.665***					
		(0.129)							(0.087)					
$Up_{m,t,90\%} \times R^2_{crypto,t}$			0.005							0.0002				
			(0.005)							(0.003)				
$Up_{m,t,90\%} \times Ln(GT_{cryptocurrency})$			0.980***							0.802***				
			(0.250)							(0.110)				
$Up_{m,t,95\%} \times R^2_{crypto,t}$				0.006							0.002			
				(0.006)							(0.004)			
$Up_{m,t,95\%} \times Ln(GT_{cryptocurrency})$				1.550***							0.939***			
				(0.418)							(0.148)			
$Down_{m,t,25\%} \times R^2_{crypto,t}$					0.005***							-0.014		
					(0.001)							(0.011)		
$Down_{m,t,25\%} \times Ln(GT_{cryptocurrency})$					0.278***							-0.186**		
					(0.103)							(0.085)		
$Down_{m,t,10\%} \times R^2_{crypto,t}$						0.004***							-0.005	
						(0.001)							(0.010)	
$Down_{m,t,10\%} \times Ln(GT_{cryptocurrency})$						0.262							-0.487***	
						(0.189)							(0.118)	
$Down_{m,t,5\%} \times R^2_{crypto,t}$							0.005**							0.013
							(0.002)							(0.008)
$Down_{m,t,5\%} \times Ln(GT_{cryptocurrency})$							1.092***							-0.626***
							(0.152)							(0.182)
Observations	1,040	1,040	1,040	1,040	1,040	1,040	1,040	1,040	1,040	1,040	1,040	1,040	1,040	1,040
R-squared	0.450	0.451	0.440	0.442	0.447	0.438	0.469	0.415	0.457	0.485	0.469	0.461	0.464	0.458

Table A.5: Herding in the cryptocurrency market with GT index with "bitcoin" search term in different market and movement condition

	Dep Var: $CSAD_{crypto,t}$													
	Panel A: $Up_{m,t} = Up_{crypto,t}$							Panel B: $Up_{m,t} = Up_{GT,bitcoin,t}$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
$Down_{\mathrm{m,t}} \times R^2_{crypto,t}$	0.004***							0.003						
	(0.001)							(0.007)						
$Up_{m,t} \times R^2_{crypto,t}$	0.003							0.003**						
	(0.005)							(0.001)						
$Down_{m,t} \times Ln(GT_{bitcoin})$	0.516***							0.594***						
	(0.094)							(0.082)						
$Up_{m,t} \times Ln(GT_{bitcoin})$	0.666***							0.699***						
	(0.076)							(0.096)						
$Up_{m,t,75\%} \times R^2_{crypto,t}$		0.007							0.004**					
		(0.006)							(0.002)					
$Up_{m,t,75\%} \times Ln(GT_{bitcoin})$		0.832***							0.776***					
		(0.135)							(0.089)					
$Up_{m,t,90\%} \times R^2_{crypto,t}$			0.007							-0.002				
			(0.005)							(0.003)				
$Up_{m,t,90\%} \times Ln(GT_{bitcoin})$			1.299***							0.799***				
			(0.286)							(0.120)				
$Up_{m,t,95\%} \times R^2_{crypto,t}$				0.006							0.0001			
				(0.006)							(0.004)			
$Up_{m,t,95\%} \times Ln(GT_{bitcoin})$				2.065***							1.027***			
				(0.620)							(0.171)			
$Down_{m,t,25\%} \times R^2_{crypto,t}$					0.003***							-0.026		
					(0.001)							(0.016)		
$Down_{m,t,25\%} \times Ln(GT_{bitcoin})$					0.423***							-0.134		
					(0.112)							(0.127)		
$Down_{m,t,10\%} \times R^2_{crypto,t}$						0.004***							-0.012	
						(0.001)							(0.014)	
$Down_{m,t,10\%} \times Ln(GT_{bitcoin})$						0.358*							-0.955***	
						(0.183)							(0.196)	
$Down_{m,t,5\%} \times R^2_{crypto,t}$							0.005***							-0.008
							(0.002)							(0.011)
$Down_{m,t,5\%} \times Ln(GT_{bitcoin})$							1.394***							-0.902**
							(0.173)							(0.420)
Observations	1,040	1,040	1,040	1,040	1,040	1,040	1,040	1,040	1,040	1,040	1,040	1,040	1,040	1,040
R-squared	0.460	0.461	0.459	0.461	0.458	0.451	0.489	0.433	0.464	0.477	0.471	0.457	0.469	0.473

3 Do Americans Feel Inflation?

Abstract

Do Americans feel inflation? In this paper, we use state-level inflation data to estimate the effect of inflation on household's attention on inflation, which is measured by Google Trends data. We find that a 1% increase in inflation raises the number of Google search in that state by 2%. Results from threshold regression suggest that household's attention become more responsive to inflation rates from tradable goods than those of non-tradable goods. The thresholds for inflation rates from tradables and non-tradables are 0.47% and 3.2%, respectively. Furthermore, the effects of inflation rates from tradables and non-tradables when inflation exceeds the threshold are 0.21% and 0.16%, respectively. This paper suggests that Google trends can be used as a real-time indicator of state-level inflation rates across the United States, state with higher inflation rate in average will have a higher inflation attention threshold.

Keywords: inflation, google trends

JEL classification numbers: E30

3.1 Introduction

Understanding how households and businesses feel the inflation and its expectation is crucial. On the one hand, survey-based inflation expectations of professionals and households are useful in forecasting future inflation (Ang et al., 2007). On the other hand, businesses' and householders' perceptions of inflation and expectations will further influence businesses' decisions on price and wage levels and householders' own consumption-savings decisions. This excessive focus on future inflation and changes in decision-making can also change people's behavior, resulting in the internalization of inflation, which further leads to the acceleration of the inflationary regime and promotes hyperinflation (Evans et al., 1992).

In this paper, we investigate the impact of changes in inflation on people's attention to inflation and further explore if people's attention to inflation changes below and above the threshold. We use the Google Search Volume Index (GSVI) of the "inflation" search term as a proxy to represent the households' and businesses' attention to inflation at the national and state-level in the United States. Using national-level data, we first estimate whether different inflation rates have similar effects on people's attention to inflation and which inflation rates people pay more attention to and are more sensitive to changes. Then, we use threshold regressions to identify potential thresholds for changes in the relationship between attention and actual inflation. When inflation is low, people do not pay excessive attention to it. When it is high and exceeds people's inflation expectations, for example, their inflation threshold, they tend to pay more attention to understanding current or future inflation trends and adjust their consumption or business behavior accordingly.

Going further, we estimate the potential heterogeneity of attention threshold between different states. Weber et al. (2022) found heterogeneity in realized inflation, with low-income, low-education, and black households experiencing greater increases in realized inflation than other households. Hayo and Neumeier (2022) also confirmed that people who do not earn their primary income in the labor market are less interested in forming inflation expectations or less responsive to expected inflation rates, and that those who live in towns (rather than cities or villages) are more likely to have a recollection of higher inflation rates and a desire to learn about inflation. Such results imply that there is potential heterogeneity in people's attention to inflation across regions. In this paper, we re-estimate the threshold regression model using state-level data to investigate if there is significant heterogeneity between states regarding inflation rate attention, and the relationship between inflation rate attention and the actual inflation rate for each state below and above the estimated threshold. We investigate whether there is a potential relationship between the state's inflation rate attention threshold and three important macroeconomic indicators: state average GDP, inflation rate, and unemployment rate.

According to the regression results, we find that at the national level, there is evidence of a significant positive relationship between inflation and people's attention to inflation, indicating that people can be aware of the changes in inflation. Meanwhile, there is a threshold where people's attention on inflation increases significantly when inflation exceeds the threshold of 3.418% for headline inflation and 2.498% for core inflation. The difference in the threshold between headline and core inflation illustrated that people's attention to goods and services with low inflation, such as headline inflation, has a lower threshold. In comparison, the attention to goods and services with relatively volatile inflation, such as core inflation, has a higher threshold. Still, the attention to inflation increases faster after exceeding the threshold.

Similar results appear when we consider state-level inflation, and people pay a different level of attention to inflation growth below and above the estimated threshold. People are more sensitive to changes in inflation for tradable goods than for non-tradable goods, where the threshold for non-tradable goods is 0.477%, and the threshold for tradable goods is 3.205%. However, not every state experiences a significant increase in attention when inflation exceeds the attention threshold. When considering macroeconomic indicators, we find that the average state inflation rate is positively correlated with its inflation attention threshold, indicating that states with higher inflation levels are less sensitive to changes in inflation and have a higher attention threshold. Regarding overall inflation, states with a high GDP, inflation rate, and unemployment rate also have higher attention thresholds. Particularly, the attention threshold of inflation for non-tradable goods is negatively correlated with the state's average GDP, and the attention threshold of tradable goods' inflation is also negatively correlated with the state's unemployment rate.

This paper has three contributions to the current literature. First, we confirm that there is a significant threshold for awareness of the change in the inflation rate in the United States and that people's concern about the inflation rate significantly increases when it exceeds the threshold. Second, we find the threshold heterogeneity, which reveals that individuals have different inflation attention thresholds for various categories of goods. Third, the major macroeconomic indicators all potentially impact the inflation attention threshold. Our results can provide direction for macroeconomists to consider and contribute to effectively controlling inflation in the ideal range.

The rest of the paper is organized as follows. Section 3.2 discusses the methodology and data applied in this paper. Section 3.3 provides the estimation results in both national-level and state-level consideration, and Section 3.4 provides an overall conclusion.

3.2 Methodology and data

3.2.1 Methodology

The empirical goal of this paper is to determine whether there is a significant change in people's attention to inflation when it is above or below their attention threshold. In line with such objectives, we first empirically assess the overall influence of the change of inflation on household and business attention in both national and state level consideration and then further explore the existence of different attention regimes for different types of inflation. We consider the baseline linear regression model:

$$GSVI_t = \beta_0 + \beta_1 Inflation_t + \gamma_t + \varepsilon \tag{3.1}$$

where the $GSVI_t$ is the Google Search Volume Index after seasonality adjustments at month t. When considering the national level data, the $Inflation_t$ is either 1) the headline inflation calculated by headline CPI index after seasonality adjustments, or 2) the core inflation calculated by core CPI index after seasonality adjustments. When investigating state-level data, the $Inflation_t$ is either 1) the tradable inflation, or 2) the non-tradable inflation. γ_t is the time fixed effect vector including year fixed effect. ε represents the error term of regression. Going further, we estimate the threshold model to investigate the potential change of inflation attention:

$$GSVI_t = \begin{cases} \beta_0 + \beta_1 Inflation_t + \varepsilon & Inflation_t < \alpha \\ \beta_0 + \beta_2 Inflation_t + \varepsilon & otherwise \end{cases} \tag{3.2}$$

where the threshold variable is also the different type of inflation rate at month t like we introduced above. α is the threshold parameter that separate household and business attention into two regimes: the first regime occurs when if $Inflation_t < \alpha$ and has a coefficient of β_1; the second regime occurs when $Inflation_t \geq \alpha$ and has a coefficient of β_2.

Combined the two different regimes in a same equation, the threshold regression model can be specified as:

$$GSVI_{i,t} = \beta_0 + \beta_1 Inflation_{i,t} I(q_{i,t} \leq \gamma) + \beta_2 Inflation_{i,t} I(q_{i,t} > \gamma) + \varepsilon_{i,t} \tag{3.3}$$

where $I(.)$ is the indicator function, q_{it} is the threshold variable, and γ is the threshold value that separates the model into two regimes with coefficient β_1 and β_2. In this paper, we set the inflation rate with different considerations as the threshold value, and add fixed effects to absorb individual effects. The compact format of Equation 3.3 is:

$$GSVI_{i,t} = \beta_0 + \beta Inflation_{i,t}(\gamma) + \varepsilon_{i,t} \tag{3.4}$$

and conditional least squares estimator for β for any given γ is:

$$\hat{\beta}(\gamma) = (Inflation^*(\gamma)' Inflation^*(\gamma))^{-1} Inflation^*(\gamma)' GSVI^* \tag{3.5}$$

with the vector of residuals $\hat{\varepsilon}^*(\gamma) = GSVI^* - Inflation^*(\gamma)\hat{\beta}(\gamma)$ and the sum of squared errors

$$S(\gamma) = \hat{\varepsilon}(\gamma)'\hat{\varepsilon}(\gamma) \tag{3.6}$$

The optimal threshold value γ is calculated by conditional least square based on Chan (1993) and Hansen (1999) by minimizing the concentrated sum of squared errors of the threshold regression model estimation, and the least-squares estimators of γ is

$$\hat{\gamma} = argmin_{\gamma} S(\gamma)$$

Based on the methodology constructed by Hansen (1999), we proceed the grid (1%, 1.25%, 1.50%, ..., 99.0%) which contains 393 quantiles to search the threshold value. To test whether the threshold exist in the relationship between inflation attention and actual inflation rate, the null and alternative hypothesis are:

$$H_0 : \beta_1 = \beta_2$$

$$H_1 : \beta_1 \neq \beta_2$$

and the likelihood ratio test of H_0 is testing the existence of the threshold effect:

$$F = \frac{S_0 - S(\hat{\gamma})}{\frac{S(\hat{\gamma})}{n(T-1)}}$$

where n is the observation of the state, and n=1 when considering national level data. T is the monthly data for national level, and quarterly data for state level. If we accepts the null hypothesis, the threshold does not exist and the relationship between inflation attention and actual inflation rate is linear. Based on the argument of Hansen (1999) that the asymptotic distribution of F is non-standard and standard inferences are invalid, we rely on the bootstrapping methods that Hansen (1999) suggests to simulate the asymptotic distribution of the likelihood ratio test and estimate the p-value of during the bootstrap.

3.2.2 Data

The target dependent variable in our analysis is the data related to Google searches. Google search volume index (GSVI) is derived from the Google Trends tool[17]. Compared to other search engines and social media inside-searches, Google search is the world's leading search site, accounting for approximately 90% of global internet queries, and has previously been used as a predictor of economic indicators (Gomes and Taamouti, 2016). GSVI calculates the volume of searches for a keyword or combination of keywords entered by users at a specific time divided by the highest search volume of the keyword during the selected time interval. Data from Google Trends dates back to 2004, but the full service only began in 2006. The frequency of the data varies from monthly to daily, depending on the length of the selected time interval. It is worth noting that GSVI does not provide the total search volume for keywords but rather an index from 0 to 100, and the index obtained will change depending on the selected interval.

The GSVI measures the relative search volume for a keyword in a given region at a given time and ranges from 0 to 100, with higher values representing higher search volume at that time. Since GSVI represents the relative search volume and its value changes depending on the time period chosen, just like Figure 3.1 shown, and shows weekly data if only select the whole-year period. Our first goal in adjusting GSVI is to convert the relative search volume into a change in total search volume using overlap term calculation.

[17] Check website: https://www.google.com/trends

Google Trend implemented an update to the geographical assignment on January 1, 2011, therefore we use 2011 as the base year to adjust for the relative volume of searches in other years compared to 2011. Based on the identification from Google Trend website, we can get the GSVI is calculated by:

$$GSVI_{t,T} = \frac{SV_t}{MSV_T} \tag{3.7}$$

where the $GSVI_t$ indicates the GSVI of a relative search volume of a specific search term at time t within the selected time period T, with a range from 0 to 100. $SV_t$is the actual search volume of a specific search term at time t, and $MSV_T$is the maximum search volume during the time period T that be selected initially. To compare the difference between the maximum search volume of two selected search period, we can calculate as below:

$$Change_{t,T,T+1} = \frac{GSVI_{t,T+1}}{GSVI_{t,T}} \tag{3.8}$$

where $Change_{t,T,T+1}$ indicates the maximize search volume change from time period T to $T+1$ of a specific search term. In this paper we consider two-overlap term, then we calculate the average change to adjust GSVI. For the year before the base year, we select the the time period T from the first day of year y to the date of second Monday of year $y+1$, and the time period $T+1$ from the first day of year $y+1$ to the date of second Monday of year $y+2$. For the year after the base year, we select the the time period T from the last second Monday of year y to the last day of year $y+1$, and the time period $T+1$ from the last second Monday of year $y+1$ to the last day of year $y+2$. In this case, the actual GSVI can be calculated as below:

$$GSVI_{t+1} = Change_{t,T,T+1} \times GSVI_{t,T+1} \tag{3.9}$$

After applying the overlap-term adjustment, we also apply X-13ARIMA-SEATS seasonal adjustment[18] after overlap-term adjustment to eliminate the potential bias from seasonal trend. It is worth mentioning that the X-13ARIMA-SEATS method only adjusts monthly and quarterly data, while the GSVI adjusted by the overlap-term method is weekly data. Therefore, we use monthly and quarterly averages as their monthly and quarterly data after adjusting GSVI for further seasonal adjustment.

The national-level analysis is based on monthly Consumer Price Index (CPI) calculated by Bureau of Labor Statistics (BLS). We consider headline and core CPI, calculated with and without food and energy, respectively, from 2004 to 2022, with 228 observations comprising the full data sample. The headline CPI index considers all items and represents the purchasing habits of all urban consumers. This index accounts for approximately 88% of the total population, including wage earners, clerical workers, technical workers, self-employed, short-term workers, unemployed, retirees, and those not in the labor force. The BLS also adjusts the index seasonally to eliminate the effects of seasonal changes such as weather, school year, production cycles, and holidays. Significant increases in the CPI over a short period may indicate inflation, while significant decreases over a short period may indicate deflation. The CPI, however, may not be a reliable measure of inflationary and deflationary periods because it includes volatile food and oil prices. We also consider the seasonally adjusted core CPI, which excludes food and energy prices when calculating the consumer index, for more accurate detection. In the

[18] X-13ARIMA-SEATS is seasonal adjustment software produced, distributed, and maintained by the Census Bureau, check the website for more information: https://www.census.gov/data/software/x13as.html

empirical analysis, I also calculate food and energy inflation based on their CPI to investigate the estimation disparity. We calculated the inflation rate based on the formula:

$$Inflation_t = (\frac{CPI_t}{CPI_{t-12}} - 1) \times 100$$

where the CPI_t and CPI_{t-12} are either 1) seasonally adjusted headline CPI index, or 2) seasonally adjusted core CPI index at time t and $t-12$.

We obtain quarterly state-level inflation from database constructed by Hazell et al. (2022) to investigate how household and business attention varies in response to different types of inflation. Our sample includes 1488 quarterly state-level data over 2006-2017 period. According to Jenkins et al. (2011), non-tradable goods are not traded internationally. They include services where the demander and producer must be in the same location and commodities with a low value relative to their weight or volume. Non-tradable goods typically include electricity, water supply, all public services, hotel accommodations, real estate, construction, and local transportation; commodities with extremely high transportation costs, such as gravel; and commodities produced to meet the country's specific customs or conditions. Based on such identification, Hazell et al. (2022) classified tradable and non-tradable [19] goods more precisely, and all Elementary-Level Items (ELIs) that are not included in non-tradable are tradable goods. Johnson (2017) indicates that tradable goods have lower inflation than non-tradable goods, and such a phenomenon also implies differences in the sensitivity of households and businesses to changes in the inflation rate of tradable and non-tradable goods.

The state-level inflation data constructed by Hazell et al. (2022) have only 34

[19] Hazell et al. (2022)identified education services, telephone services, food away from home, other personal services, home services, medical services, recreational services, and transportation services as non-tradable goods. Check the detail in the online appendix: https://eml.berkeley.edu/~enakamura/papers/StateLevelCPIs_Appendix.pdf

states[20] from 1989 to 2017. Meanwhile, in the early state-level GSVI, because of the relatively low internet penetration and imperfect statistics, the GSVI for "inflation" in many states fluctuated between 0 and 100 directly in 2004 and 2005, which is inconsistent with the national-level GSVI. Also, the GSVI for inflation has been relatively low and less volatile in some states, such as Alaska, Hawaii, and Mississippi. Based on these constraints, we exclude Alaska, Hawaii, and Mississippi from Hazell et al. (2022)'s database and apply quarterly state-level inflation and GSVI data for the remaining 31 states from 2006 to 2017. Consistent with the previous national-level GSVI, the state-level GSVI is also processed by the overlap term method and seasonal adjustment to convert the relative search volume into its actual variation and eliminate the data's seasonality to reduce its potential bias.

Table A.6 provides a statistical overview of the data presented in the paper. The data in Table A.6 show that headline inflation is more volatile than core inflation, and the inflation of non-tradable goods is more volatile than that of tradable goods. Such results demonstrate that there is a variation in the underlying inflation rate of different goods, and therefore there may be potential differences in people's attention to the inflation rate of different goods.

3.3 Results

3.3.1 National-Level Estimates

We first examine how changes in inflation under different categories affect household and business attention to inflation and the potential differences in their effects using Equation 3.1. It is worth mentioning that the BLS releases CPI data with a

[20]These states are: Alabama, Alaska, Arkansas, California, Colorado, Connecticut, District of Columbia, Florida, Georgia, Hawaii, Illinois, Indiana, Kansas, Louisiana, Maryland, Massachusetts, Michigan, Minnesota, Mississippi, Missouri, New Jersey, New York, North Carolina, Ohio, Oklahoma, Oregon, Pennsylvania, South Carolina, Tennessee, Texas, Utah, Virginia, Washington and Wisconsin.

delay; for example, the CPI index for April is released in May. Therefore, in the regression, we also consider the one-month-lag term of inflation to compare the impact of the change in inflation on the attention of households and businesses in the time interval between the actual change in inflation and the release of inflation data. To avoid the potential multi-collinearity issue, we regress the current and 1-month-lag term of inflation separately instead of regressing them together.

The empirical results shown in Table A.7 illustrate that an increase in inflation significantly increases the attention of households and businesses to inflation, and there is a significant awareness of real-time changes in inflation. Comparing Columns (1) and (2) in Table A.7, we can conclude that the change in core inflation has a more significant impact on the attention of households and businesses about inflation, with a 0.087% increase in GSVI for every 1% increase in core inflation, whereas 1% increase in GSVI for headline inflation rises by 0.047% of the GSVI only. Columns (3) and (4) in Table A.7 show that the two commodities with high price volatility, which are food and energy, have a weaker relationship with inflation attention, with a 1% increase in food or energy inflation leading to a 0.04 % and 0.003 % increase in GSVI, respectively. Combining these regression results, we can summarize that people can feel the inflation changes but are less sensitive to the inflation of originally more volatile commodities. Even though food and energy are essential to daily life, their real-time price changes and inflation have received little additional attention due to their high price volatility. Threshold regressions obtain similar results, with core inflation, which has a more stable price, possessing a lower threshold of only 2.498%. Headline inflation, which includes commodities with high price volatility, has a higher threshold of 3.418%, 36% higher than the threshold for core inflation. When considering food and energy inflation individually, they have even higher thresholds of 4.589% and 22.352%, respectively. After comparing people's attention to the inflation above and below the threshold, Columns (5) to (8) in Table A.7 show that the lower the threshold of people's attention to specific inflation, the higher the change in their attention above and below the threshold. Figure 3.3 compares Columns (5) to (8) of Table A.7 graphically.

We then investigate the attention thresholds of households and businesses on different types of inflation with 1-month-lag consideration an re-regress Equation 3.1 and 3.2.. Comparing estimation results of Table A.7 with Table A.8, we can observe that businesses and households have similar attention level to changes in inflation during the time interval of actual changes in inflation and the time interval of macro data release. Comparing the results in Columns (1) through (4) of Tables A.7 and A.8, we can see that although the coefficients obtained from the 1-lag-term regression results are slightly larger than those from the current term regression, they are still at a similar level. In viewing the results of the threshold regression, comparing Columns (5) through (8) of Tables A.7 and A.8, we find that the thresholds obtained from the 1-month-lag regression are almost identical to those of the current term.

These preliminary findings can be interpreted as follows: People's attention threshold for inflation is equivalent to their expectation of current inflation. When core inflation is below 2.13% or headline inflation is below 3.41%, people do not or pay too much attention to inflation. Therefore, there is little correlation between these two types of inflation and the frequency of searches for inflation. When the inflation threshold is exceeded, people start paying attention to inflation, reflecting that the number of searchers increases as inflation rises; therefore, there is a strong positive correlation between inflation and search behavior. Because food and energy are more volatile than other goods and services, the threshold of headline inflation is higher for households and businesses. However, once the threshold is exceeded, the degree of inflation concern changes significantly. In contrast, core inflation is more stable than headline inflation, so households and businesses have a lower threshold of concern for core inflation. However, the effect of exceeding the core inflation threshold on inflation search volume is comparably smaller than that of headline inflation. Based on the headline and core CPI, the estimated threshold is relatively higher than the Federal Reserve's target inflation rate of 2%.

3.3.2 State-Level Estimates

After applying fixed effect and threshold regressions on national-level data help obtain general trends in household and business attention to inflation across countries, we further investigate the potential heterogeneity of thresholds at the state level. Due to the differences in urbanization, price, income, and taxation levels of households and businesses in each state, there may be disparities in the attention paid to inflation. Columns (1) and (2) of Table A.9 show the panel regression results to determine whether there is a significant general trend between inflation growth and people's attention on inflation. Column (1) shows that the coefficient of inflation is not significant, whereas Column (2) illustrates that every 1% increase in inflation of tradable goods reduces people's attention level by 0.003% in 10% of the significance level. This finding suggests that the overall trend of rising inflation has little impact on people's attention to inflation and that people are more sensitive to changes in the inflation of tradable goods than non-tradable goods.

The results in Table A.9 are consistent with the previous results in that households and businesses are more sensitive to price changes in goods and services with lower inflation rates than others, and a threshold exists where households and businesses are significantly more attentive to inflation when the inflation rate exceeds this threshold. By comparing Columns (3) through (5) in Table A.9, we observe that the threshold value of the inflation rate of non-tradable goods, which is 3.205%, is more in line with the threshold value of the overall inflation rate, which is 2.738%. This result illustrates that the inflation of non-tradable goods dominates the result of overall inflation and that households and businesses pay more attention to the inflation of non-tradable goods. The threshold of Column (4) is considerably smaller compared with that of Columns (3) and (5), which is only 0.477%. Yet, there is a significant effect of tradable goods inflation on GSVI both below and above the threshold. When the inflation of tradable goods is below the threshold value, the GSVI of inflation decreases significantly by 0.022% for every 1% increase. When the inflation of tradable goods above the threshold, people's attention to inflation changes sharply, and a 1% increase in tradable goods is associated with a significant increase in GSVI of 0.021%, such a result proves that people are more sensitive to price changes of tradable goods, which are goods with lower inflation.

Table A.10 shows the regression results to consider whether there is a latency in people's perception of inflation changes. We use the one-lag term as the independent variable to re-estimate. Column (1) of Table A.10 shows that every 1% increase in the inflation rate will significantly increase people's attention on inflation in the next quarter by almost 0.01%. Comparing Column (1) of Table A.9, we conclude that there is a latency in people's attention to inflation under the state-level panel regression. The regression results in Columns (2) to (5) of Table A.10 show that there is also a significant threshold for people's postponed perception of inflation. When inflation exceeds the threshold, people's attention to inflation increases significantly. similar to the results in Table A.9, people are more sensitive to changes in the inflation of tradable goods when considering the postponed perception of inflation; however, people's postponed attention to inflation is generally more significant and intense than their immediate attention.

After discussing the baseline panel regressions, we further explore the potential heterogeneity in people's attention to inflation and consider the difference in the attention thresholds on inflation for each state. To explore this, we re-evaluate each state's equations 3.1 and 3.2. Figure 3.4 shows the results of the baseline linear regression considering different states, that is, the increase in the inflation rate changes the inflation attention of the residents in different states. Figure 3.4a shows that the inflation coefficient is positive with inflation attention, i.e., GSVI, in 15 states and negative in 16 states, and no significant coefficient exists. In contrast, when considering the inflation of tradable goods, Figure 3.4b confirms that only the growth of tradable goods' inflation in 6 states is positively correlated with GSVI, and only two states have significant coefficient[21]. Figure 3.4c shows that non-tradable goods' inflation is positively correlated with GSVI in 17 states, and two states have significant coefficient [22]. Figure 3.4 shows no significant and stable relationship between people's attention to inflation and changes in state inflation in most states. After comparing the estimated coefficients, we confirm that there is heterogeneity in inflation attention across states and that the growth of inflation of non-tradable goods is negatively related to inflation attention in most states.

Figure 3.5 shows the estimated results of the threshold regression for each state and the effect of the inflation change below and above the threshold on the inflation attention. In Figure 3.5, the x-axis represents the estimated coefficient of the impact of inflation on attention when inflation is below the threshold, and the y-axis represents the effect of inflation on attention when inflation is above the threshold. Each scatter in the figure represents the individual state's regression results. For a clear view we also add blue dashed lines at y=x, y=-x, x =0, and

[21] these two state are: FL (Florida) and IN (Indiana)

[22] These two states are: OH (Ohio) and SC (South Carolina).

y=o to each subfigure. When the scatter point at y=x is below the dashed line, the state's attention to inflation decreases if inflation exceeds the threshold. When the scatter is above the dotted line at y=x, the state is more interested in inflation when it exceeds the threshold.

Comparing the three plots in Figure 3.5, we further obtain the heterogeneity of the threshold value and inflation attention change in individual states. When we consider the attention threshold for each state for changes in overall inflation, Figure 3.5a shows that when inflation exceeds the threshold, 11 states increase their attention to inflation [23]. Figure 3.5a shows the heterogeneity of changes in individual states' attention to inflation before and after the threshold; not all states pay more attention to inflation after inflation exceeds the threshold. Figure 3.5b illustrates that only 9 states have reduced attention to inflation when it is above the threshold [24]. At the same time, no states have a positive coefficient of inflation when it is below the threshold and change to negative when it is above the threshold, indicating that these states do not experience a significant decrease in inflation attention even after the threshold is reached. The results in Figure 3.5c are similar to those considering overall inflation, showing that 14 states increase for inflation attention when inflation is above the threshold [25].

After we confirmed the heterogeneity of the inflation attention threshold across states, we further explored the impact of three economic indicators, which are gross domestic product (GDP) growth rate, unemployment rate, and inflation rate, on the attention threshold of each state's inflation.

[23] These states are: AL (Alabama), CT (Connecticut), IL (Illinois), MD (Maryland), MN (Minnesota), MO (Missouri), NJ (New Jersey), NY (New York), OH (Ohio), PA (Pennsylvania), and TX (Texas).

[24] These 9 states are: CA (California), IL (Illinois), MD (Maryland), MO (Missouri), NY (New York), OH (Ohio), OR (Oregon), TX (Texas) and WI (Wisconsin).

[25] These 14 states are: CO (Colorado), CT (Connecticut), DC (District of Columbia), FL (Florida), GA (Georgia), IL (Illinois), IN (Indiana), MD (Maryland), MA (Massachusetts), MO (Missouri), NC (North Carolina), OH (Ohio), OK (Oklahoma), and TN (Tennessee).

We start by considering the impact of GDP growth on the inflation attention threshold. When a state's GDP growth rate is high and shows that the economy is growing fast, people will tend to be more confident and optimistic about the economy's growth and stability, thus lowering the attention on the inflation rate. In Figure 3.6, we investigate this aspect by examining the relationship between the estimated threshold and each state's average GDP growth rate. Figure 3.6a reveals that, relative to the overall inflation rate, an increase in the GDP growth rate of a given state results in a slight increase in the state's inflation threshold. Examining the overall inflation rate separately for tradable and non-tradable goods, Figure 3.6b and 3.6c present that the increase in GDP growth rate also slightly raises the inflation attention threshold for tradable goods, whereas lowering the inflation threshold for non-tradable goods in the state.

Initially, the demand for non-tradable goods tends to increase in states experiencing rapid economic growth. These goods, typically from local producers and are difficult to replace through trade, include housing, healthcare, and services. Demand for non-tradable goods exceeds supply as the economy expands, resulting in price pressures and inflationary tendencies. As the attention threshold for non-tradable goods increases, states with a high GDP growth become more sensitive to inflation in these goods. On the other hand, the supply dynamics of tradable versus non-tradable goods play a role. Typically, the supply of tradable goods is more elastic than non-tradable goods because they can be produced in multiple locations and transported with relative ease. Non-tradable goods, on the other hand, are frequently subject to local supply constraints, such as limited land availability for housing or a lack of skilled labor in specific service sectors. These supply constraints can increase the inflation sensitivity of non-tradable goods in states with rapid economic expansion.

After analyzing the impact of a state's GDP growth rate on the inflation attention threshold, we next investigate the impact of the state's inflation rate level on the attention threshold. Intuitively, a state with relatively high inflation levels would have a higher inflation threshold than a state with lower inflation levels. This is due to the concept of anchoring or the persistence of inflation. When a state experiences a prolonged period of high inflation, public expectations, and the economic decision-making process become entrenched. In contrast, states with lower inflation rates tend to develop lower inflation concern thresholds, reflecting expectations of a relatively stable price level. Figure 3.7 illustrate the relationship between the estimated threshold and each state's average inflation rate. All three graphs in Figure 3.7 show a significant positive relationship between the inflation attention threshold and the average inflation rate for a individual state. When considering inflation separately for tradable and non tradable goods, the relationship between the inflation attention threshold of these two goods and the average inflation value of the state is still similar.

The classic Phillip Curve model demonstrates a negative relationship between the inflation rate and the unemployment rate, and it is possible that such a relationship also exists between the inflation rate attention threshold and the unemployment rate in a cross-state perspective. When the state's unemployment rate is high, unemployment or employment difficulties may lead to lower income, while inflation may lead to higher cost of living, and people's attention threshold for inflation will be lower. However, the results in Figure 3.8 shows that only the inflation attention threshold for tradable goods is negatively related to the average unemployment rate from a cross-state perspective. Combined with the previous regression results, we can conclude that people are more attentive to the inflation rate of tradable goods, which is reflected in a lower attention threshold; therefore, when the average unemployment rate in a state is higher, people's income and cost of living are affected, and the inflation rate attention threshold of tradable goods decreases.

3.4 Conclusion

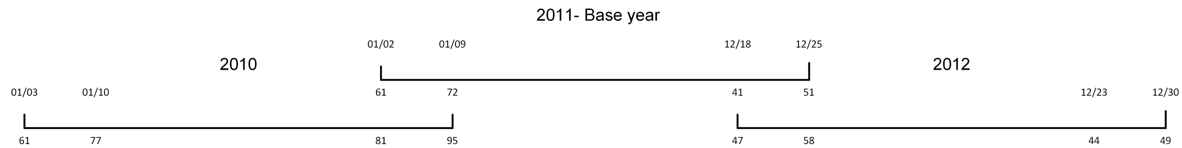
This study investigates the effect of the inflation rate change on the public's attention to inflation, establishing the existence of an attention threshold and analyzing attention behavior before and after the potential threshold. Using the Google Search Volume Index (GSVI) of "inflation" as a proxy, we investigate both the national and state-level interest in inflation in the United States.

Based on our empirical results, we find a substantial relationship between inflation and public attention, demonstrating that households and businesses are susceptible to changes in inflation, with heightened awareness when inflation exceeds a certain threshold. The threshold disparity between headline and core inflation indicates a lower attention threshold for goods and services with higher inflation (headline inflation) than those with more stable inflation rates (core inflation). At the state level, there is a discernible attention threshold, with individuals exhibiting greater sensitivity to changes in inflation for tradable goods than non-tradable goods. The correlation between macroeconomic indicators and the attention threshold for inflation demonstrates that states with higher inflation rates, GDP, and unemployment rates also have higher attention thresholds.

This research offers three notable insights: it affirms a attention threshold for perceiving inflation rate changes in the U.S., identifies threshold heterogeneity across various goods categories, and underscores macroeconomic indicators' influence on the inflation attention threshold. These findings provide macroeconomists valuable guidelines for effectively controlling inflation within the desired range.

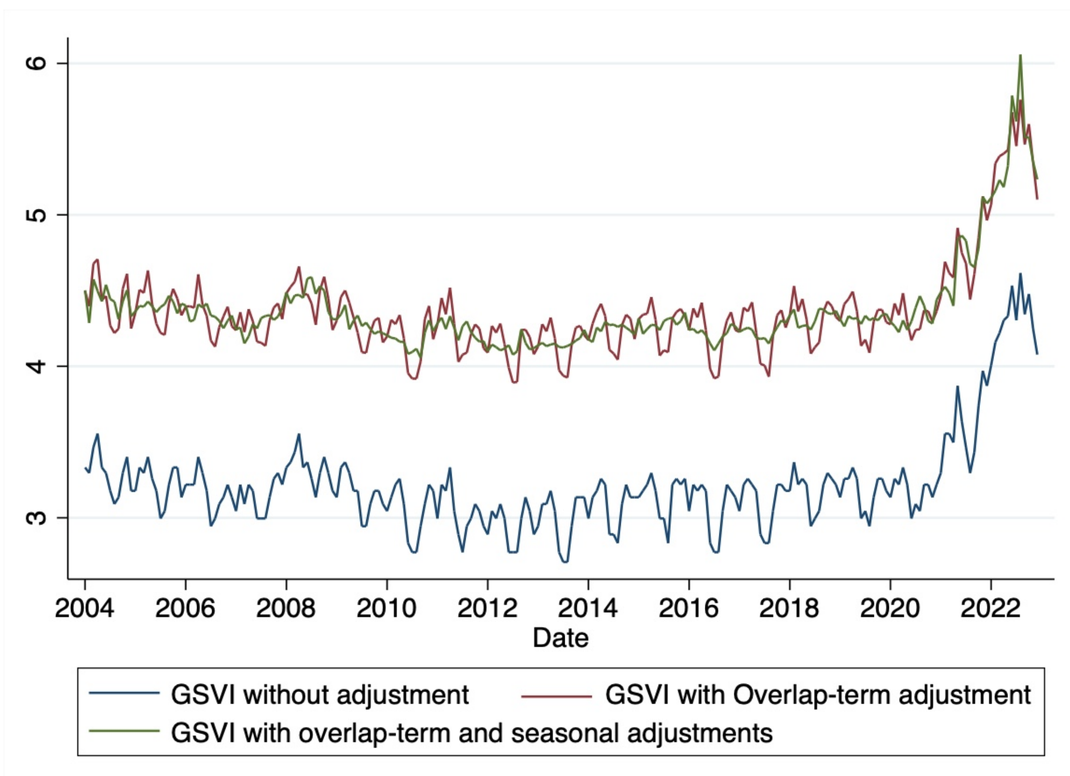
Figures

Figure 3.1: overlap term GSVI



Note(s): This figure illustrates the difference in GSVI based on the selection of different time periods.

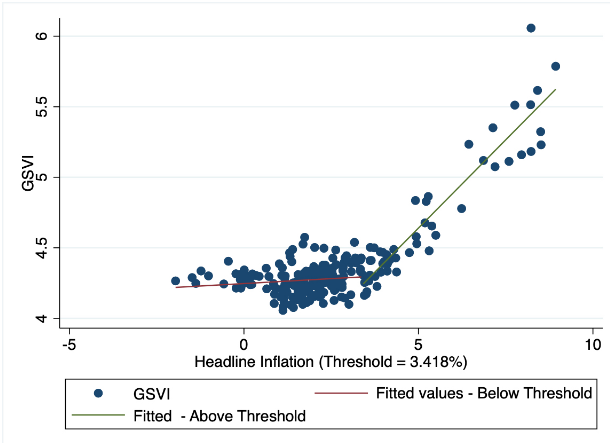
Figure 3.2: GSVI with "inflation" search term with and without adjustment



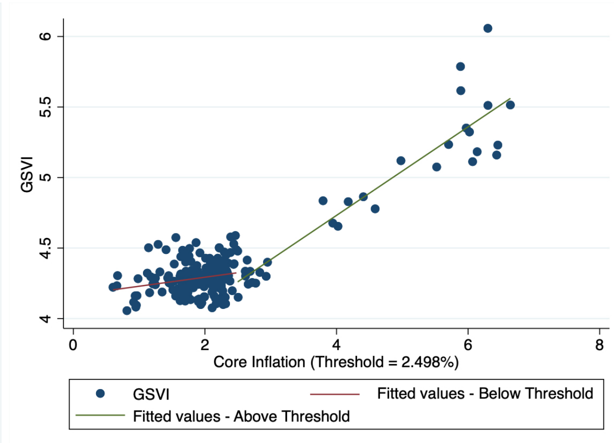
Note(s): All data are shown after logarithm transformation.

Figure 3.3: Threshold model fit for national data

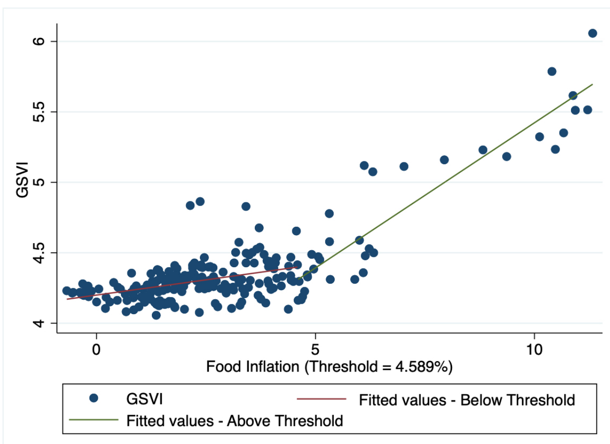
(a) Model fit for headline inflation



(b) Model Fit for core inflation



(c) Model fit for food inflation



(d) Model fit for energy inflation

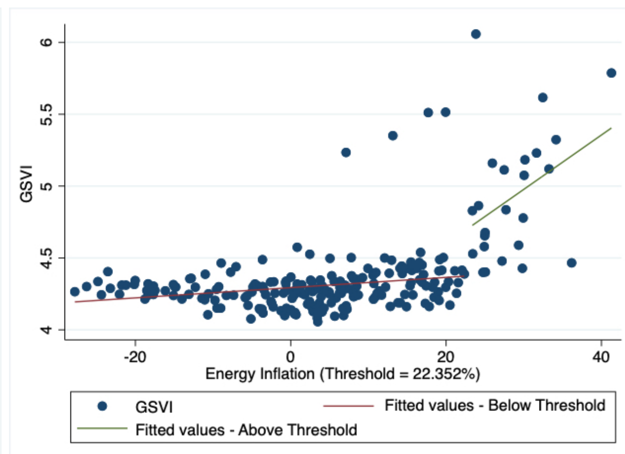
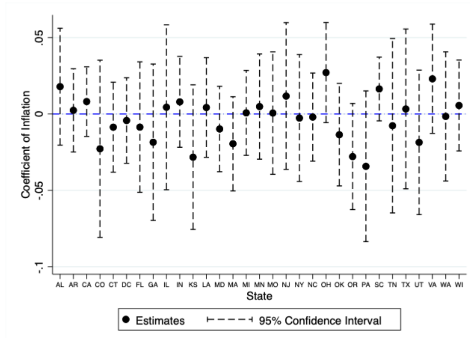
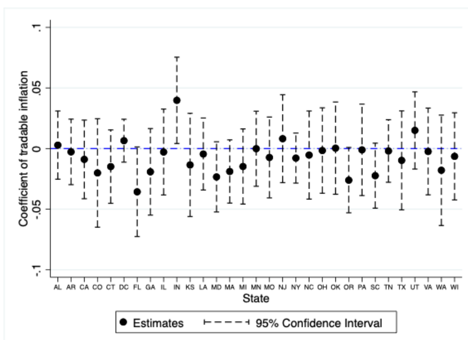


Figure 3.4: Linear regression estimation results by states

(a) Overall inflation



(b) Inflation of tradable goods



(c) Inflation of non-tradable goods

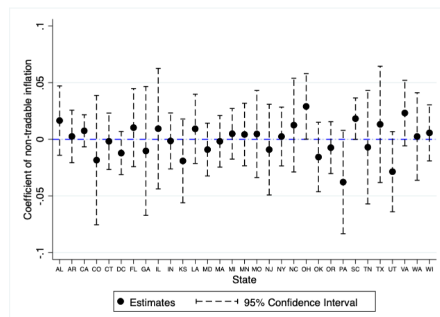
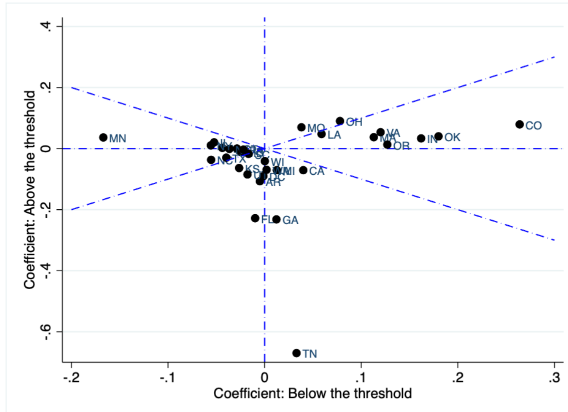
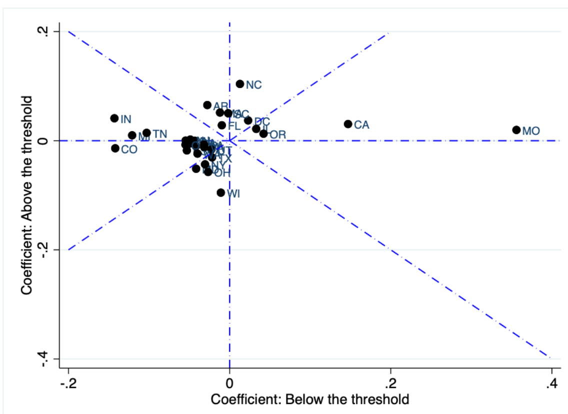


Figure 3.5: Estimated coefficients of threshold regression by state

(a) Overall inflation



(b) Inflation of tradable goods



(c) Inflation of non-tradable goods

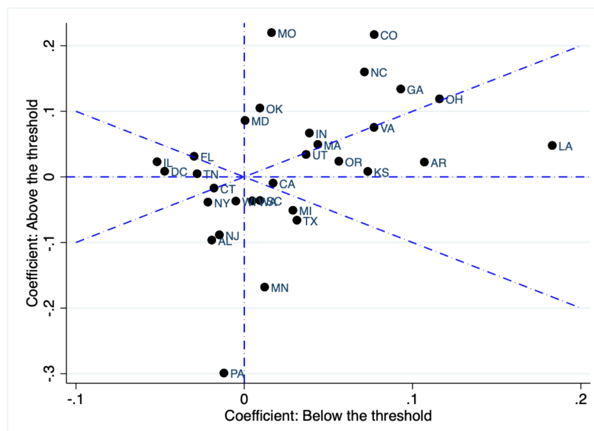
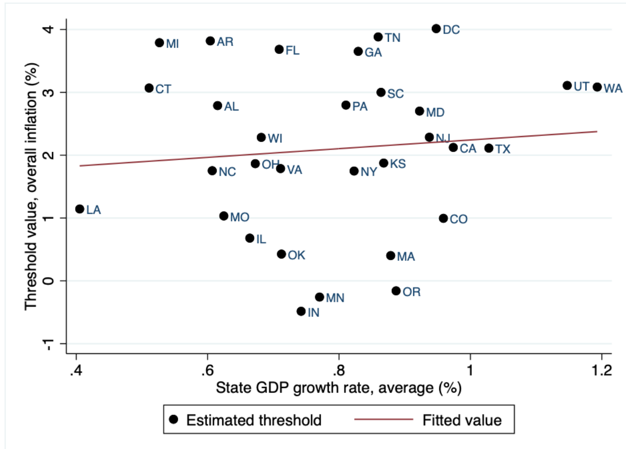
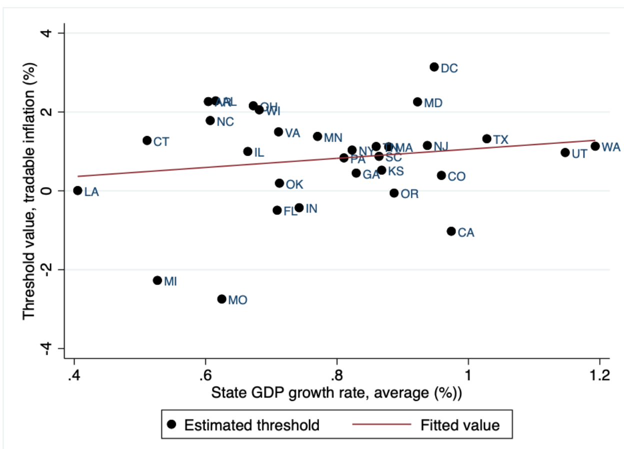


Figure 3.6: Relationship between the state's inflation attention threshold and GDP

(a) Overall inflation



(b) Inflation of tradable goods



(c) Inflation of non-tradable goods

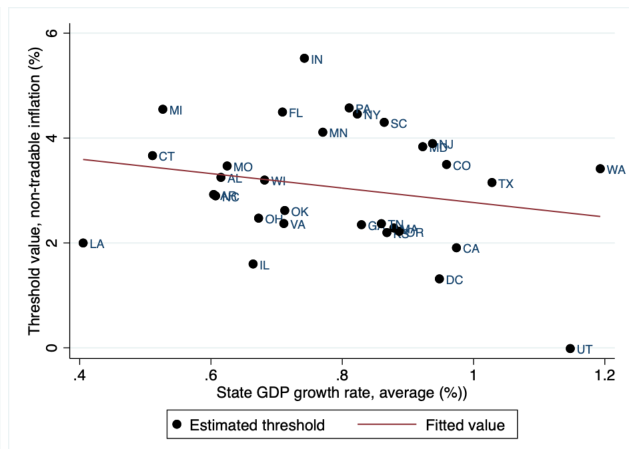
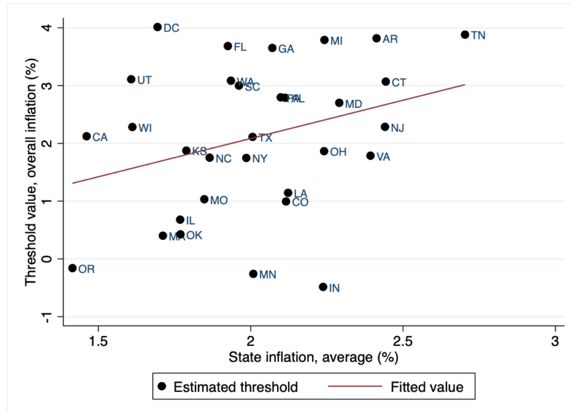
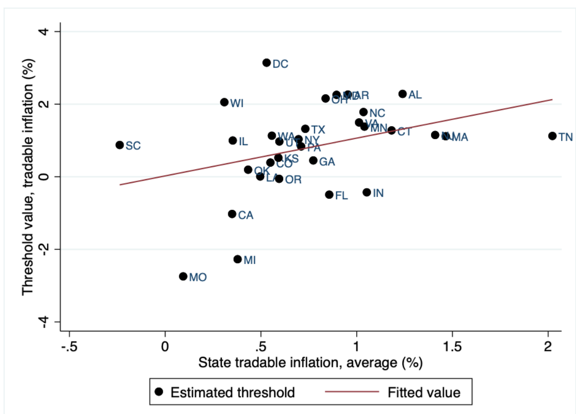


Figure 3.7: Relationship between the state's inflation attention threshold and average inflation level

(a) Overall inflation



(b) Inflation of tradable goods



(c) Inflation of non-tradable goods

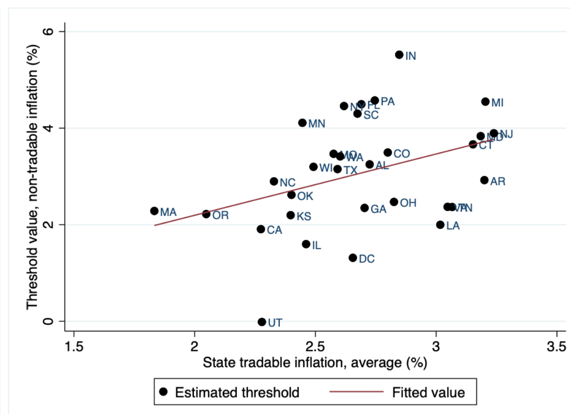
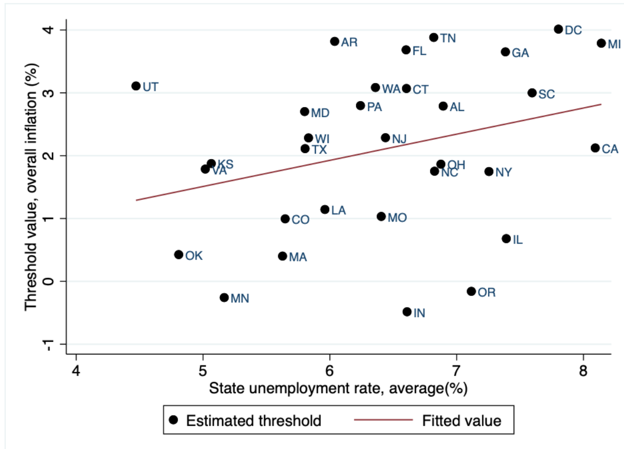
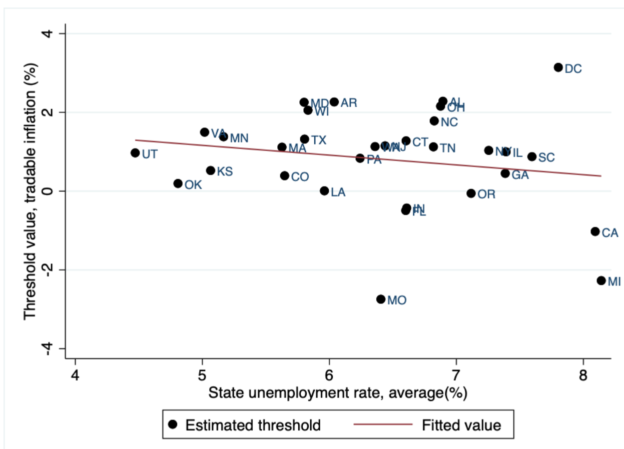


Figure 3.8: Relationship between the state's inflation attention threshold and unemployment rate

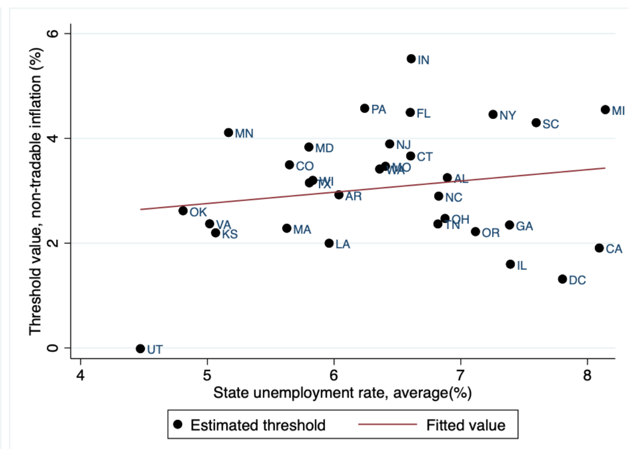
(a) Overall inflation



(b) Inflation of tradable goods threshold



(c) Inflation of non-tradable goods



Tables

Table A.6: Summary Statistics

	Observation	Mean	SD	Min	Max
Panel A: National level					
Headline inflation, %	228	2.489	1.934	-1.959	8.933
Core inflation, %	228	2.244	1.114	0.603	6.643
Food inflation, %	228	2.845	2.272	–0.6780	11.330
Energy inflation, %	228	5.115	13.880	-27.797	41.289
GSVI	228	4.367	0.298	4.056	6.058
Panel B: State Level					
Overall Inflation, %	1488	2.009	1.636	-4.034	7.922
Tradable goods inflation, %	1488	0.758	2.115	-5.997	8.042
Non-tradable goods inflation, %	1488	2.681	1.827	-5.090	13.140
GSVI	1488	3.770	0.147	3.477	4.216
GDP, B USD	1488	466.229	439.962	86.979	2814.209
Unemployment rate, %	1488	6.410	2.141	2.4	13.833

Table A.7: National level of inflation attention estimation

	Dep Var: GSVI							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Headline Inflation	0.047***							
	(0.008)							
Core Inflation		0.087***						
		(0.016)						
Food Inflation			0.040***					
			(0.007)					
Energy Inflation				0.003***				
				(0.001)				
Headline Inflation:Region 1					0.014*			
					(0.008)			
Headline Inflation: Region 2					0.251***			
					(0.009)			
Core Inflation: Region 1						0.067***		
						(0.023)		
Core Inflation: Region 2						0.320***		
						(0.015)		
Food Inflation:Region 1							0.043***	
							(0.008)	
Food Inflation: Region 2							0.206***	
							(0.010)	
Energy Inflation:Region 1								0.004***
								(0.001)
Energy Inflation: Region 2								0.038***
								(0.010)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	228	228	227	227	228	228	227	227
R-squared	0.912	0.908	0.909	0.904				
Threshold Variable					Headline Inflation	Core Inflation	Food Inflation	Energy Inflation
Threshold Value					3.418	2.498	4.589	22.352
SSR					2.970	3.433	4.061	10.851
AIC					-981.737	-948.688	-910.355	-686.273

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.8: National level of inflation attention estimation

	Dep Var: GSVI							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
L.Headline Inflation	0.050***							
	(0.008)							
L.Core Inflation		0.074***						
		(0.017)						
L.Food Inflation			0.030***					
			(0.007)					
L.Energy Inflation				0.004***				
				(0.001)				
L.Headline Inflation:Region 1					0.014			
					(0.009)			
L.Headline Inflation: Region 2					0.255***			
					(0.009)			
L.Core Inflation: Region 1						0.051		
						(0.033)		
L.Core Inflation: Region 2						0.289***		
						(0.010)		
L.Food Inflation:Region 1							0.045***	
							(0.009)	
L.Food Inflation: Region 2							0.206***	
							(0.011)	
L.Energy Inflation:Region 1								0.003***
								(0.001)
L.Energy Inflation: Region 2								0.046***
								(0.009)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	227	227	227	227	227	227	227	227
R-squared	0.911	0.905	0.905	0.906				
Threshold Variable					L.Headline Inflation	L.Core Inflation	L.Food Inflation	L.Energy Inflation
Threshold Value					3.418	2.133	4.589	22.352
SSR					2.982	3.707	4.891	9.541
AIC					-975.432	-926.017	-863.132	-711.452

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.9: State-level fixed effect panel threshold regression estimates

	Dep Var: GSVI				
	(1)	(2)	(3)	(4)	(5)
inflation	0.0005				
	(0.002)				
tradable		-0.003*			
		(0.002)			
Non-tradable		0.001			
		(0.001)			
Inflation: Region 1			-0.005		
			(0.004)		
Inflation: Region 2			0.020***		
			(0.002)		
tradable: Region 1				-0.022***	
				(0.004)	
tradable: Region 2				0.020***	
				(0.003)	
Non-tradable: Region 1					0.002
					(0.004)
Non-tradable: Region 2					0.016***
					(0.002)
Constant	3.769***	3.769***	3.749***	3.734***	3.742***
	(0.004)	(0.004)	(0.006)	(0.006)	(0.007)
Observations	1,488	1,488	1,488	1,488	1,488
R-squared	0.726	0.727	0.070	0.042	0.053
Year FEs	Yes	Yes			
State FEs	Yes	Yes			
Threshold Variable			Inflation	Tradable	Non-tradable
Threshold Value			2.738	0.477	3.205
Threshold Interval			(2.700, 2.743)	(-0.281, 0.491)	(3.167, 3.216)
# of State			31	31	31

* p<0.10, ** p<0.05, *** p<0.01.

Table A.10: State-level fixed effect panel threshold regression estimates

	Dep Var: GSVI				
	(1)	(2)	(3)	(4)	(5)
inflation	0.009***				
	(0.002)				
tradable		0.008***			
		(0.002)			
Non-tradable		0.004***			
		(0.001)			
Inflation: Region 1			0.001		
			(0.004)		
Inflation: Region 2			0.023***		
			(0.002)		
tradable: Region 1				-0.014***	
				(0.004)	
tradable: Region 2				0.019***	
				(0.003)	
Non-tradable: Region 1					0.004
					(0.004)
Non-tradable: Region 2					0.020***
					(0.002)
Constant	3.748***	3.749***	3.738***	3.737***	3.732***
	(0.004)	(0.004)	(0.006)	(0.006)	(0.008)
Observations	1,457	1,457	1,457	1,457	1,457
R-squared	0.757	0.759	0.079	0.032	0.076
Year FEs	Yes	Yes			
State FEs	Yes	Yes			
Threshold Variable			Inflation	Tradable	Non-tradable
Threshold Value			2.743	0.552	3.251
Threshold Interval			(2.674, 2.749)	(-0.469, 0.577)	(3.207, 3.259)
# of State			31	31	31

* p<0.10, ** p<0.05, *** p<0.01.

Appendix Tables

Table A.1: State-level inflation statistics and rank information

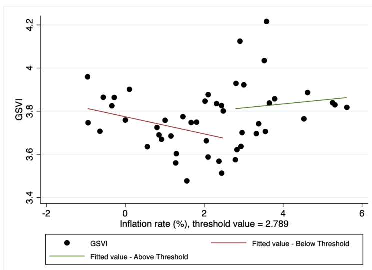
State	Overall Inflation (%)	Tradable Inflation (%)	Non-tradable Inflation (%)	Rank (GDP)	Rank (Inflation)
Alabama	2.113	1.240	2.725	24	12
Arkansas	2.413	0.954	3.200	31	4
California	1.461	0.350	2.274	1	30
Colorado	2.116	0.549	2.800	19	11
Connecticut	2.443	1.1824	3.152	22	2
District of Columbia	1.694	0.530	2.655	30	27
Florida	1.925	0.857	2.690	4	20
Georgia	2.071	0.773	2.703	9	14
Illinois	1.768	0.353	2.461	5	25
Indiana	2.237	1.0523	2.847	16	9
Kansas	1.789	0.591	2.397	28	23
Louisiana	2.123	0.496	3.017	23	10
Maryland	1.712	1.465	1.833	15	26
Massachusetts	2.291	0.895	3.184	12	6
Michigan	2.242	0.378	3.204	13	7
Minnesota	2.009	1.041	2.446	17	15
Missouri	1.848	0.094	2.575	21	22
New Jersey	1.865	1.036	2.328	8	21
New York	2.441	1.410	3.239	3	3
North Carolina	1.985	0.696	2.618	10	17
Ohio	2.241	0.838	2.825	7	8
Oklahoma	1.769	0.433	2.401	27	24
Oregon	1.415	0.595	2.048	26	31
Pennsylvania	2.098	0.710	2.746	6	13
South Carolina	1.961	-0.238	2.674	25	18
Tennessee	2.703	2.022	3.066	18	1
Texas	2.006	0.732	2.592	2	16
Utah	1.608	0.596	2.278	29	29
Virginia	2.393	1.013	3.047	11	5
Washington	1.935	0.557	2.602	14	19
Wisconsin	1.612	0.309	2.492	20	28

Appendix Figures

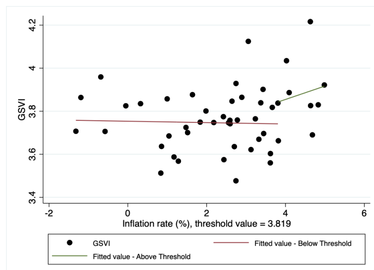
Attention increases when the inflation is above the threshold

Figure B.1: Threshold model results for individual state: Attention increases when the inflation is above the threshold ($\beta_2 > \beta_1$)

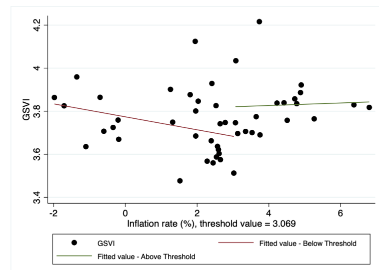
(a) Alabama



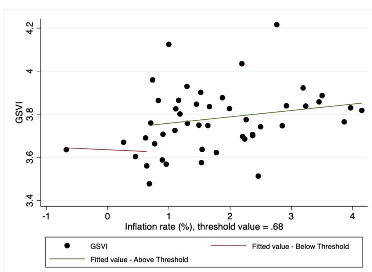
(b) Arkansas



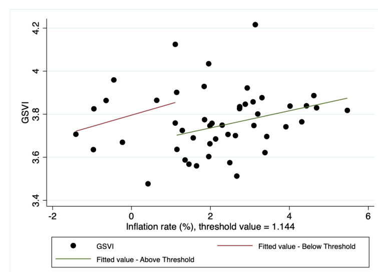
(c) Connecticut



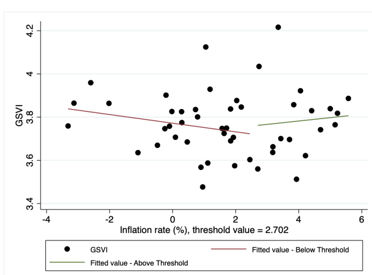
(d) Illinois



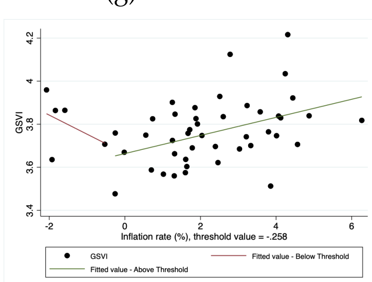
(e) Louisiana



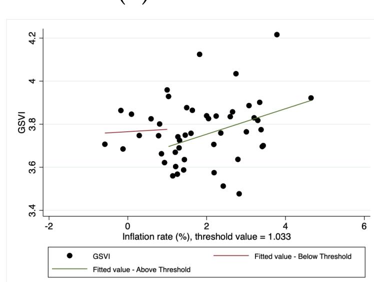
(f) Maryland



(g) Minnesota



(h) Missouri



(i) New Jersey

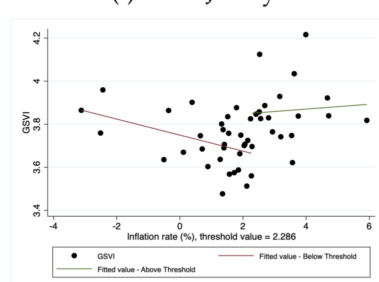
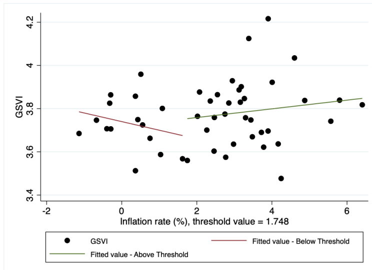
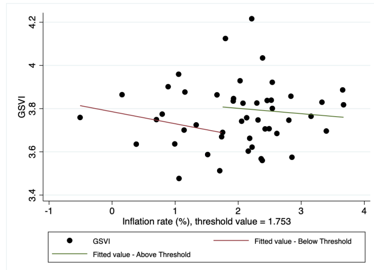


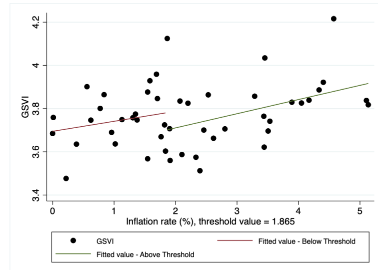
Figure B.1: Threshold model results for individual state: Attention increases when the inflation is above the threshold ($\beta_2 > \beta_1$)



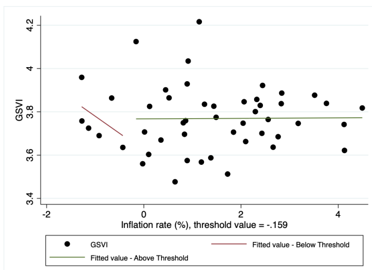
(j) New York



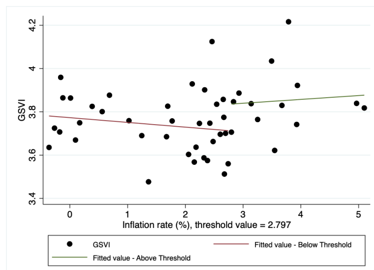
(k) North Carolina



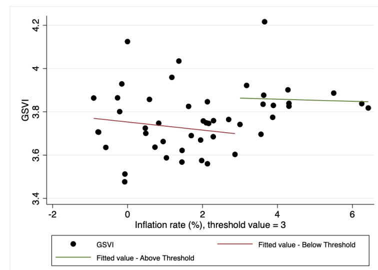
(l) Ohio



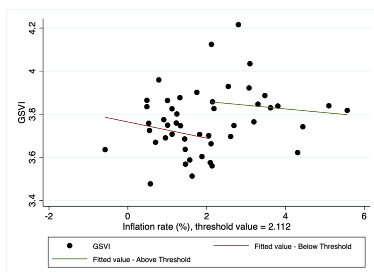
(m) Oregon



(n) Pennsylvania



(o) South Carolina

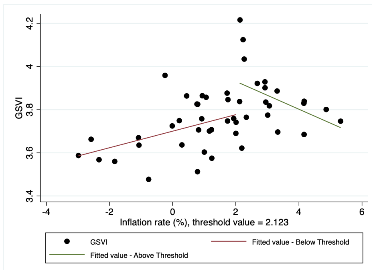


(p) Texas

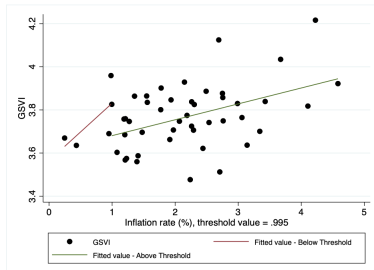
3.4.1 Attention decreases when the inflation is above the threshold

Figure B.2: Threshold model results for individual state: Attention decreases when the inflation is above the threshold ($\beta_2 \leq \beta_1$)

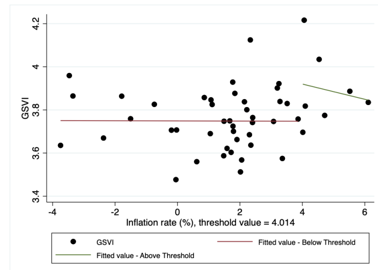
(a) California



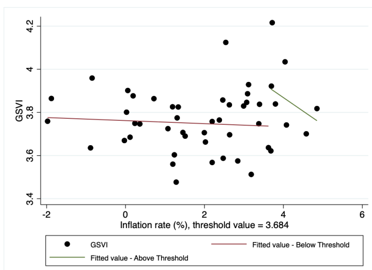
(b) Colorado



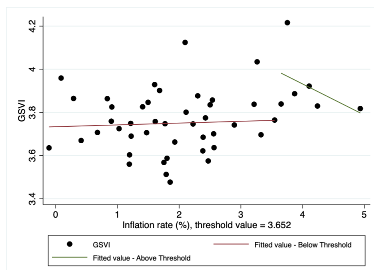
(c) District of Columbia



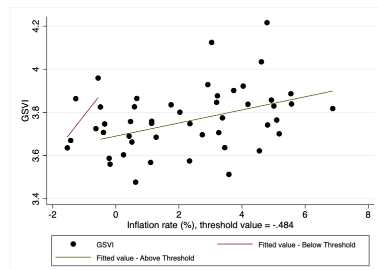
(d) Florida



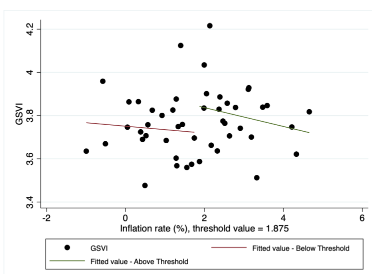
(e) Georgia



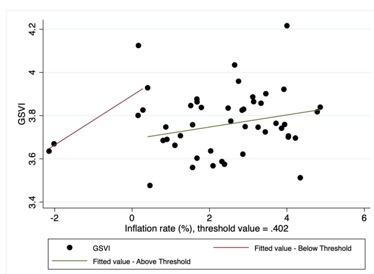
(f) Indiana



(g) Kansas



(h) Massachusetts



(i) Michigan

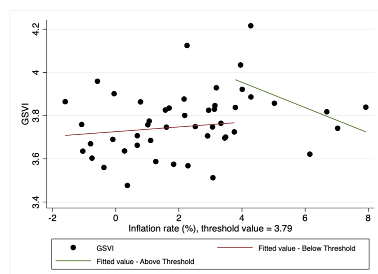
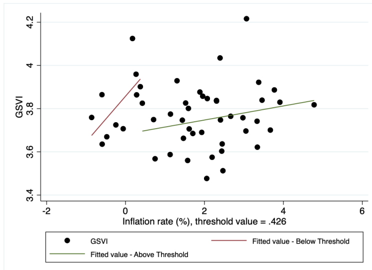
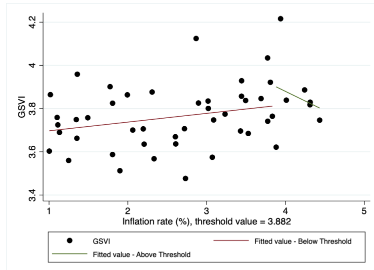


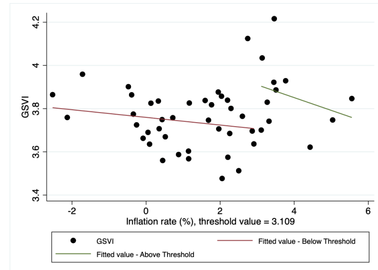
Figure B.2: Threshold model results for individual state: Attention decreases when the inflation is above the threshold ($\beta_2 \leq \beta_1$)



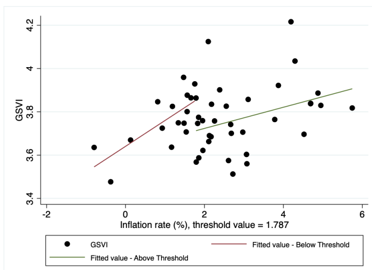
(j) Oklahoma



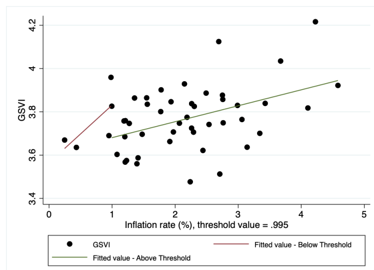
(k) Tennessee



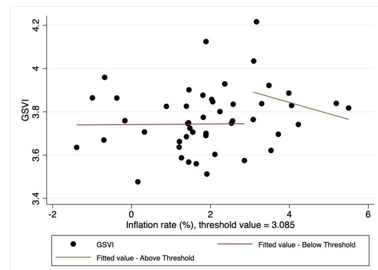
(l) Utah



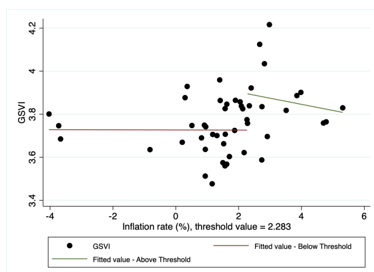
(m) Virginia



(n) Colorado



(o) Washington



(p) Wisconsin

References

Ahmed, Shaghil and Andrei Zlate (2014) "Capital flows to emerging market economies: A brave new world?" *Journal of International Money and Finance*, 48 (PB), 221–248, https://EconPapers.repec.org/RePEc:eee:jimfin:v:48:y:2014:i:pb:p:221-248.

Allen, Franklin, Krzysztof Jackowicz, Oskar Kowalewski, and Łukasz Kozłowski (2017) "Bank lending, crises, and changing ownership structure in Central and Eastern European countries," *Journal of Corporate Finance*, 42, 494–515, https://doi.org/10.1016/j.jcorpfin.2015.05.001.

Anastasiou, Dimitrios, Antonis Ballis, and Konstantinos Drakos (2021) "Cryptocurrencies' Price Crash Risk and Crisis Sentiment," *Finance Research Letters*, 42, 101928, https://doi.org/10.1016/j.frl.2021.101928.

Andrianaivo, M. and K. Kpodar (2012) "Mobile phones, financial inclusion, and growth," *Review of Economics and Institutions*, 3 (2), 1 – 30, https://www.scopus.com/inward/record.uri?eid=2-s2.0-84878914897&partnerID=40&md5=f4385d84e272e467025c888e647c714a, Cited by: 75.

Ang, Andrew, Geert Bekaert, and Min Wei (2007) "Do macro variables, asset markets, or surveys forecast inflation better?" *Journal of Monetary Economics*, 54 (4), 1163–1212, https://doi.org/10.1016/j.jmoneco.2006.04.006.

Ariefianto, Moch Doddy (2020) "Assessing Qualification of Crypto Currency as A Financial Assets: A Case Study on Bitcoin,", 934–939: IEEE, 10.1109/ICIMTech50083.2020.9211133.

Asante Gyamerah, Samuel (2021) "COVID-19 PANDEMIC AND HERDING BE-

HAVIOUR IN CRYPTOCURRENCY MARKET," *Applied Finance Letters*, 10, 58 – 66, 10.24135/afl.v10i.443.

Avery, Christopher and Peter Zemsky (1998) "Multidimensional Uncertainty and Herd Behavior in Financial Markets," *The American Economic Review*, 88 (4), 724–748, http://www.jstor.org/stable/117003.

Babalos, V. and S. Stavroyiannis (2015) "Herding, anti-herding behaviour in metal commodities futures: a novel portfolio-based approach," *Applied Economics*, 47 (46), 4952–4966, 10.1080/00036846.2015.1039702, Export Date: 11 March 2022.

Ballis, Antonis and Konstantinos Drakos (2020) "Testing for herding in the cryptocurrency market," *Finance Research Letters*, 33, 101210, https://doi.org/10.1016/j.frl.2019.06.008.

Bejar, Pablo, Kotaro Ishi, Takuji Komatsuzaki, Ippei Shibata, Jasmin Sin, and Suchanan Tambunlertchai (2022) "Can Fintech Foster Competition in the Banking System in Latin America and the Caribbean?" *Latin American Journal of Central Banking*, 3 (2), 100061, https://doi.org/10.1016/j.latcb.2022.100061.

BenMabrouk, Houda and Houda Litimi (2018) "Cross herding between American industries and the oil market," *The North American Journal of Economics and Finance*, 45, 196–205, https://doi.org/10.1016/j.najef.2018.02.009.

Berg, Tobias, Andreas Fuster, and Manju Puri (2021) "FinTech Lending," *National Bureau of Economic Research Working Paper Series*, No. 29421, 10.3386/w29421.

Berger, Allen N. and Robert DeYoung (1997) "Problem loans and cost efficiency in commercial banks," *Journal of Banking & Finance*, 21 (6), 849–870, https://doi.org/10.1016/S0378-4266(97)00003-4.

Berger, Allen N., Nathan H. Miller, Mitchell A. Petersen, Raghuram G. Rajan, and Jeremy C. Stein (2005) "Does function follow organizational form? Evidence from the lending practices of large and small banks," *Journal of Financial Economics*, 76 (2), 237–269, https://doi.org/10.1016/j.jfineco.2004.06.003.

Bernanke, Ben S. (2011) "Community Banking in a Period of Recovery and Change : a speech at the Independent Community Bankers of America National Convention, San Diego, California, March 23, 2011," Speech 658, Board of Governors of the Federal Reserve System (U.S.).

Blasco, Natividad, Pilar Corredor, and Sandra Ferreruela (2017) "Can agents sensitive to cultural, organizational and environmental issues avoid herding?" *Finance Research Letters*, 22, 114–121, https://doi.org/10.1016/j.frl.2017.01.006.

Boot, Arnoud W. A. and Anjan V. Thakor (2000) "Can Relationship Banking Survive Competition?" *The Journal of Finance*, 55 (2), 679–713, https://doi.org/10.1111/0022-1082.00223.

Borri, Nicola (2019) "Conditional tail-risk in cryptocurrency markets," *Journal of Empirical Finance*, 50, 1–19, https://doi.org/10.1016/j.jempfin.2018.11.002.

Boudriga, Abdelkader, Neila Boulila, and Sana Jellouli (2009) "Does bank supervision impact nonperforming loans : cross-country determinants using agregate data ?" *MPRA Paper*.

Bouri, Elie, Rangan Gupta, and David Roubaud (2019) "Herding behaviour in cryptocurrencies," *Finance Research Letters*, 29, 216–221, https://doi.org/10.1016/j.frl.2018.07.008.

Buchak, Greg, Gregor Matvos, Tomasz Piskorski, and Amit Seru (2018) "Fintech, regulatory arbitrage, and the rise of shadow banks," *Journal of Financial Economics*, 130 (3), 453–483, https://doi.org/10.1016/j.jfineco.2018.03.011.

Campanella, Francesco, Maria Rosaria Della Peruta, and Manlio Del Giudice (2017) "The Effects of Technological Innovation on the Banking Sector," *Journal of the Knowledge Economy*, 8 (1), 356–368, 10.1007/s13132-015-0326-8.

Cao, Guangxi and Wenhao Xie (2022) "Asymmetric dynamic spillover effect between cryptocurrency and China's financial market: Evidence from TVP-VAR based connectedness approach," *Finance Research Letters*, 49, 103070, https://doi.org/10.1016/j.frl.2022.103070.

Caparrelli, Franco, Anna Maria D'Arcangelis, and Alexander Cassuto (2004) "Herding in the Italian Stock Market: A Case of Behavioral Finance," *Journal of Behavioral Finance*, 5 (4), 222–230, 10.1207/s15427579jpfm0504_5, doi: 10.1207/s15427579jpfm0504$_5$.

Chan, K. S. (1993) "Consistency and Limiting Distribution of the Least Squares Estimator of a Threshold Autoregressive Model," *The Annals of Statistics*, 21 (1), 520 – 533, 10.1214/aos/1176349040.

Chang, Eric C., Joseph W. Cheng, and Ajay Khorana (2000) "An examination of herd behavior in equity markets: An international perspective," *Journal of Banking & Finance*, 24 (10), 1651–1679, https://doi.org/10.1016/S0378-4266(99)00096-5.

Chen, Cathy YiâHsuan, Romeo Despres, Li Guo, and Thomas Renault (2019) "What Makes Cryptocurrencies Special? Investor Sentiment and Return Predictability During the Bubble," *SSRN Electronic Journal*, 10.2139/ssrn.3398423.

Chen, Yan-Shing, Yehning Chen, Chih-Yung Lin, and Zenu Sharma (2016) "Is there a bright side to government banks? Evidence from the global financial crisis," *Journal of Financial Stability*, 26, 128–143, https://doi.org/10.1016/j.jfs.2016.08.006.

Chiang, Thomas C. and Dazhi Zheng (2010) "An empirical analysis of herd behavior in global stock markets," *Journal of Banking & Finance*, 34 (8), 1911–1921, https://doi.org/10.1016/j.jbankfin.2009.12.014.

Christie, William G. and Roger D. Huang (1995) "Following the Pied Piper: Do Individual Returns Herd around the Market?" *Financial Analysts Journal*, 51 (4), 31–37, http://www.jstor.org/stable/4479855.

Coase, R. H. (1995) "The Nature of the Firm," in Estrin, Saul and Alan Marin eds. *Essential Readings in Economics*, 37–54, London: Macmillan Education UK, 10.1007/978-1-349-24002-9_3.

Corbet, Shaen, Yang Hou, Yang Hu, and Les Oxley (2022) "Financial contagion among COVID-19 concept-related stocks in China," *Applied Economics*, 54 (21), 2439–2452, 10.1080/00036846.2021.1990844, (Greg) doi: 10.1080/00036846.2021.1990844.

Corbet, Shaen, Charles Larkin, Brian M. Lucey, Andrew Meegan, and Larisa Yarovaya (2020a) "The impact of macroeconomic news on Bitcoin returns," *The European Journal of Finance*, 26 (14), 1396–1416, 10.1080/1351847X.2020.1737168.

Corbet, Shaen, Charles Larkin, Brian Lucey, Andrew Meegan, and Larisa Yarovaya (2020b) "Cryptocurrency reaction to FOMC Announcements: Evidence of heterogeneity based on blockchain stack position," *Journal of Financial Stability*, 46, 100706, https://doi.org/10.1016/j.jfs.2019.100706.

Cornelli, Giulio, Jon Frost, Leonardo Gambacorta, P Raghavendra Rau, Robert Wardrop, and Tania Ziegler (2020) "Fintech and big tech credit: a new database."

Coskun, Esra Alp, Chi Keung Marco Lau, and Hakan Kahyaoglu (2020) "Uncertainty and herding behavior: evidence from cryptocurrencies," *Research in International Business and Finance*, 54, 101284, https://doi.org/10.1016/j.ribaf.2020.101284.

Cui, Yueting, Bartosz Gebka, and Vasileios Kallinterakis (2019) "Do closed-end fund investors herd?" *Journal of Banking & Finance*, 105, 194–206, https://doi.org/10.1016/j.jbankfin.2019.05.015.

Czapliński, Tomasz and Elena Nazmutdinova (2019) "Using FIAT currencies to arbitrage on cryptocurrency exchanges," *Journal of International Studies*, 12 (1), 184–192, 10.14254/2071-8330.2019/12-1/12.

Dass, Nishant, Massimo Massa, and Rajdeep Patgiri (2008) "Mutual Funds and Bubbles: The Surprising Role of Contractual Incentives," *The Review of Financial Studies*, 21 (1), 51–99, 10.1093/rfs/hhm033.

Davidovic, Milivoje (2021) "From pandemic to financial contagion: High-frequency risk metrics and Bayesian volatility analysis," *Finance Research Letters*, 42, 101913, https://doi.org/10.1016/j.frl.2020.101913.

De Haas, Ralph and Iman Van Lelyveld (2014) "Multinational Banks and the Global Financial Crisis: Weathering the Perfect Storm?" *Journal of Money, Credit and Banking*, 46 (s1), 333–364, https://doi.org/10.1111/jmcb.12094.

Demir, Ender, Giray Gozgor, Chi Keung Marco Lau, and Samuel A. Vigne (2018) "Does economic policy uncertainty predict the Bitcoin returns? An empirical investigation," *Finance Research Letters*, 26, 145–149, https://doi.org/10.1016/j.frl.2018.01.005.

Demirgüç-Kunt, Asli, Leora Klapper, Dorothe Singer, Saniya Ansar, and Jake Hess (2020) "The Global Findex Database 2017: Measuring Financial Inclusion and Opportunities to Expand Access to and Use of Financial Services," *World Bank Economic Review*, 34 (Supplement_1), S2–S8, 10.1093/wber/lhz013.

Diamond, Douglas W. (1984) "Financial Intermediation and Delegated Monitoring," *The Review of Economic Studies*, 51 (3), 393–414, 10.2307/2297430.

Dushnitsky, Gary, Massimiliano Guerini, Evila Piva, and Cristina Rossi-Lamastra (2016) "Crowdfunding in Europe: Determinants of Platform Creation across Countries," *California Management Review*, 58 (2), 44–71, 10.1525/cmr.2016.58.2.44.

Economou, Fotini, Konstantinos Gavriilidis, Abhinav Goyal, and Vasileios Kallinterakis (2015) "Herding dynamics in exchange groups: Evidence from Euronext," *Journal of International Financial Markets, Institutions and Money*, 34, 228–244, https://doi.org/10.1016/j.intfin.2014.11.013.

Economou, Fotini, Nikolaos Philippas, and Guglielmo Maria Caporale (2008) "Herding behaviour in extreme market conditions: the case of the Athens Stock Exchange," *Economics Bulletin*, 7, 1–13.

Economou, Fotini, Nikolaos Philippas, and Alexandros Kostakis (2011) "Cross-country effects in herding behaviour: Evidence from four south European markets," *Journal of International Financial Markets, Institutions and Money*, 21, 443–460, 10.1016/j.intfin.2011.01.005.

Elsayed, Ahmed H., Giray Gozgor, and Chi Keung Marco Lau (2020) "Causality and dynamic spillovers among cryptocurrencies and currency markets," *International Journal of Finance & Economics*, n/a (n/a), https://doi.org/10.1002/ijfe.2257.

Elsayed, Ahmed H., Giray Gozgor, and Larisa Yarovaya (2022) "Volatility and return connectedness of cryptocurrency, gold, and uncertainty: Evidence from the cryptocurrency uncertainty indices," *Finance Research Letters*, 102732, https://doi.org/10.1016/j.frl.2022.102732.

Evans, G.W., G. Ramey, and London School of Economics (1992) *Expectation Calculation, Hyperinflation and Currency Collapse*: London School of Economics, https://books.google.com/books?id=NK9QzwEACAAJ.

Fama, E.F. (1980) "Agency problems and the theory of the firm," *Journal of Political Economy*, 88 (2), 288 – 307, https://www.scopus.com/inward/record.uri?eid=2-s2.0-84992955136&partnerID=40&md5=2900b88b4f09dce4fa7ee9b7674086f3, Cited by: 42.

Fama, Eugene F. (1998) "Market efficiency, long-term returns, and behavioral finance1The comments of Brad Barber, David Hirshleifer, S.P. Kothari, Owen Lamont, Mark Mitchell, Hersh Shefrin, Robert Shiller, Rex Sinquefield, Richard Thaler, Theo Vermaelen, Robert Vishny, Ivo Welch, and a referee have been helpful. Kenneth French and Jay Ritter get special thanks.1," *Journal of Financial Economics*, 49 (3), 283–306, https://doi.org/10.1016/S0304-405X(98)00026-9.

Frankovic, Jozo, Bin Liu, and Sandy Suardi (2021) "On spillover effects between cryptocurrency-linked stocks and the cryptocurrency market: Evidence from Australia," *Global Finance Journal*, 100642, https://doi.org/10.1016/j.gfj.2021.100642.

Frost, Jon, Leonardo Gambacorta, Yi Huang, Hyun Song Shin, and Pablo Zbinden (2019) "BigTech and the changing structure of financial intermediation," *Economic Policy*, 34 (100), 761–799, 10.1093/epolic/eiaa003.

Fuster, Andreas, Matthew Plosser, Philipp Schnabl, and James Vickery (2019) "The role of technology in mortgage lending," *The Review of Financial Studies*, 32 (5), 1854–1899.

Galariotis, Emilios C., Styliani-Iris Krokida, and Spyros I. Spyrou (2016) "Bond market investor herding: Evidence from the European financial crisis," *International Review of Financial Analysis*, 48, 367–375, https://doi.org/10.1016/j.irfa.2015.01.001.

da Gama Silva, Paulo Vitor Jordão, Marcelo Cabus Klotzle, Antonio Carlos Figueiredo Pinto, and Leonardo Lima Gomes (2019) "Herding behavior and contagion in the cryptocurrency market," *Journal of Behavioral and Experimental Finance*, 22, 41–50, https://doi.org/10.1016/j.jbef.2019.01.006.

Gomes, Pedro and Abderrahim Taamouti (2016) "In search of the determinants of European asset market comovements," *International Review of Economics & Finance*, 44, 103–117, 10.1016/j.iref.2016.03.005.

Greenbaum, Stuart I, Anjan V Thakor, and Arnoud Boot (2019) *Contemporary financial intermediation*: Academic Press.

Grossman, Sanford J. and Oliver D. Hart (1986) "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy*, 94 (4), 691–719, 10.1086/261404.

Gurdgiev, Constantin and Daniel O'Loughlin (2020) "Herding and anchoring in cryptocurrency markets: Investor reaction to fear and uncertainty," *Journal of Behavioral and Experimental Finance*, 25, 100271, https://doi.org/10.1016/j.jbef.2020.100271.

Hansen, Bruce E. (1999) "Threshold effects in non-dynamic panels: Estimation, testing, and inference," *Journal of Econometrics*, 93 (2), 345–368, https://doi.org/10.1016/S0304-4076(99)00025-1.

Hart, O. and J. Moore (1990) "Property rights and the nature of the firm," *Journal of Political Economy*, 98 (6), 1119 – 1158, https://www.scopus.

com/inward/record.uri?eid=2-s2.0-19044375813&partnerID=40&md5=e844c3895c6c169da79d9eb84b691084, Cited by: 56.

Hassan, M Kabir and Abdel-Hameed M Bashir (2003) "Determinants of Islamic banking profitability," in *10th ERF annual conference, Morocco*, 7, 2–31.

Hayo, Bernd and Florian Neumeier (2022) "Households' inflation perceptions and expectations: survey evidence from New Zealand," *International Economics and Economic Policy*, 19 (1), 185–217, 10.1007/s10368-021-00524-3.

Hazell, Jonathon, Juan Herreño, Emi Nakamura, and Jón Steinsson (2022) "The Slope of the Phillips Curve: Evidence from U.S. States*," *The Quarterly Journal of Economics*, 137 (3), 1299–1344, 10.1093/qje/qjac010.

Hodula, Martin (2022) "Does Fintech credit substitute for traditional credit? Evidence from 78 countries," *Finance Research Letters*, 46, 102469, https://doi.org/10.1016/j.frl.2021.102469.

Hsieh, Shu-Fan, Chia-Ying Chan, and Ming-Chun Wang (2020) "Retail investor attention and herding behavior," *Journal of Empirical Finance*, 59, 109–132, https://doi.org/10.1016/j.jempfin.2020.09.005.

Hsu, Shu-Han, Chwen Sheu, and Jiho Yoon (2021) "Risk spillovers between cryptocurrencies and traditional currencies and gold under different global economic conditions," *The North American Journal of Economics and Finance*, 57, 101443, https://doi.org/10.1016/j.najef.2021.101443.

Hughes, Joseph P., Julapa Jagtiani, and Choon-Geol Moon (2022) "Consumer lending efficiency: commercial banks versus a fintech lender," *Financial Innovation*, 8 (1), 38, 10.1186/s40854-021-00326-1.

Humayun Kabir, M. and Shamim Shakur (2018) "Regime-dependent herding behavior in Asian and Latin American stock markets," *Pacific-Basin Finance Journal*, 47, 60–78, https://doi.org/10.1016/j.pacfin.2017.12.002.

Hung, Ngo Thai and Xuan Vinh Vo (2021) "Directional spillover effects and time-frequency nexus between oil, gold and stock markets: Evidence from pre and during COVID-19 outbreak," *International Review of Financial Analysis*, 76, 101730, https://doi.org/10.1016/j.irfa.2021.101730.

Hwang, Soosung and Mark Salmon (2004) "Market stress and herding," *Journal of Empirical Finance*, 11 (4), 585–616, https://EconPapers.repec.org/RePEc:eee:empfin:v:11:y:2004:i:4:p:585-616.

Iyer, Rajkamal, Asim Ijaz Khwaja, Erzo F. P. Luttmer, and Kelly Shue (2016) "Screening Peers Softly: Inferring the Quality of Small Borrowers," *Management Science*, 62 (6), 1554–1577, 10.1287/mnsc.2015.2181.

Jagtiani, Julapa, Lauren Lambie-Hanson, and Timothy Lambie-Hanson (2021) "Fintech lending and mortgage credit access," *The Journal of FinTech*, 1 (01), 2050004.

Jagtiani, Julapa and Catharine Lemieux (2018) "Do fintech lenders penetrate areas that are underserved by traditional banks?" *Journal of Economics and Business*, 100, 43–54, https://doi.org/10.1016/j.jeconbus.2018.03.001.

Jenkins, Glenn, Chun-Yan Kuo, and Arnold C. Harberger (2011) "Cost-Benefit Analysis for Investment Decisions: Chapter 11 (Economic Prices for Non-Tradable Goods and Services)," report.

Jensen, Michael C. and William H. Meckling (1976) "Theory of the firm: Managerial behavior, agency costs and ownership structure," *Journal of Financial Economics*, 3 (4), 305–360, https://doi.org/10.1016/0304-405X(76)90026-X.

Ji, Qiang, Elie Bouri, David Roubaud, and Ladislav Kristoufek (2019) "Information interdependence among energy, cryptocurrency and major commodity markets," *Energy Economics*, 81, 1042–1055, https://doi.org/10.1016/j.eneco.2019.06.005.

Jiang, Wen, Qiuhua Xu, and Ruige Zhang (2021) "Tail-event driven network of cryptocurrencies and conventional assets," *Finance Research Letters*, 102424, https://doi.org/10.1016/j.frl.2021.102424.

Johnson, Noah (2017) "Tradable and nontradable inflation indexes: replicating New Zealand's tradable indexes with BLS CPI data," *Monthly Labor Review*, 10.21916/mlr.2017.14.

Kaiser, Lars and Sebastian Stöckl (2020) "Cryptocurrencies: Herding and the transfer currency," *Finance Research Letters*, 33, 101214, https://doi.org/10.1016/j.frl.2019.06.012.

Kallinterakis, Vasileios and Ying Wang (2019) "Do investors herd in cryptocurrencies – and why?" *Research in International Business and Finance*, 50, 240–245, https://doi.org/10.1016/j.ribaf.2019.05.005.

Kim, Jikyung, Hang Dong, Jeonghye Choi, and Sue Ryung Chang (2022) "Sentiment change and negative herding: Evidence from microblogging and news," *Journal of Business Research*, 142, 364–376, https://doi.org/10.1016/j.jbusres.2021.12.055, (Jeanne).

King, Timothy and Dimitrios Koutmos (2021) "Herding and feedback trading in cryptocurrency markets," *Annals of Operations Research*, 300 (1), 79–96, 10.1007/s10479-020-03874-4, King, Timothy 1; Koutmos, Dimitrios 2; Email Address: dimitrios.koutmos@tamucc.edu; Affiliations: 1: University of Kent, CT2 7NZ, Canterbury, Kent, UK; 2: Texas AM University - Corpus Christi, 78412, Corpus Christi, TX, USA; Issue Info: May2021, Vol. 300 Issue 1, p79; Thesaurus Term: Cryptocurrencies; Thesaurus Term: Bitcoin; Thesaurus Term: Capitalists financiers; Subject Term: Herding; Subject Term: Animal herds; Author-Supplied Keyword: Feedback trading; Author-Supplied Keyword: Herding behavior; Author-Supplied Keyword: Risk-return tradeoff; Number of Pages: 18p; Document Type: Article.

Klein, Tony, Hien Pham Thu, and Thomas Walther (2018) "Bitcoin is not the New Gold – A comparison of volatility, correlation, and portfolio performance," *International Review of Financial Analysis*, 59, 105–116, https://doi.org/10.1016/j.irfa.2018.07.010.

Kowalewski, Oskar and Paweł Pisany (2022a) "Banks' consumer lending reaction to fintech and bigtech credit emergence in the context of soft versus hard credit information processing," *International Review of Financial Analysis*, 81, 102116, https://doi.org/10.1016/j.irfa.2022.102116.

——— (2022b) "The rise of fintech: A cross-country perspective," *Technovation*, 102642, https://doi.org/10.1016/j.technovation.2022.102642.

Kristoufek, Ladislav (2013) "BitCoin meets Google Trends and Wikipedia: Quantifying the relationship between phenomena of the Internet era," *Scientific Reports*, 3 (1), 3415, 10.1038/srep03415.

Kumar, Ashish (2021) "Empirical investigation of herding in cryptocurrency market under different market regimes," *Review of Behavioral Finance*, 13 (3), 297–308, 10.1108/RBF-01-2020-0014.

Kurt Gümüş, Gülüzar, Yusuf Gümüş, and Ayşegül Çimen (2019) *Herding Behaviour in Cryptocurrency Market: CSSD and CSAD Analysis*, 103–114, Cham: Springer International Publishing, 10.1007/978-3-030-25275-5_6.

Kurz, Claudia and Jeong-Ryeol Kurz-Kim (2013) "What determines the dynamics of absolute excess returns on stock markets?" *Economics Letters*, 118 (2), 342–346, https://doi.org/10.1016/j.econlet.2012.11.029.

Kyriazis, Nikolaos A. (2020) "Herding behaviour in digital currency markets: An integrated survey and empirical estimation," *Heliyon*, 6 (8), e04752, https://doi.org/10.1016/j.heliyon.2020.e04752.

La Porta, Rafael, Florencio Lopez-de Silanes, Andrei Shleifer, and Robert W. Vishny (1997) "Legal Determinants of External Finance," *The Journal of Finance*, 52 (3), 1131–1150, https://doi.org/10.1111/j.1540-6261.1997.tb02727.x.

Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny (1992) "The impact of institutional trading on stock prices," *Journal of Financial Economics*, 32 (1), 23–43, https://doi.org/10.1016/0304-405X(92)90023-Q.

Lee, Chi-Chuan, Xinrui Li, Chin-Hsien Yu, and Jinsong Zhao (2021) "Does fintech innovation improve bank efficiency? Evidence from China's banking industry," *International Review of Economics and Finance*, 74, 468–483, https://doi.org/10.1016/j.iref.2021.03.009.

Lee, Chien-Chiang, Wenjie Ni, and Xiaoming Zhang (2023) "FinTech development and commercial bank efficiency in China," *Global Finance Journal*, 57, 100850, 10.1016/j.gfj.2023.100850.

Leung, Tim and Hung Nguyen (2019) "Constructing cointegrated cryptocurrency portfolios for statistical arbitrage," *Studies in Economics and Finance*, 36 (3), 581–599, 10.1108/SEF-08-2018-0264.

Levulytė, Laura and Alfreda Šapkauskienė (2021) "Cryptocurrency in context of fiat money functions," *The Quarterly Review of Economics and Finance*, 82, 44–54, https://doi.org/10.1016/j.qref.2021.07.003.

Li, Rong, Sufang Li, Di Yuan, and Huiming Zhu (2021) "Investor attention and cryptocurrency: Evidence from wavelet-based quantile Granger causality analysis," *Research in International Business and Finance*, 56, 101389, https://doi.org/10.1016/j.ribaf.2021.101389.

Li, Yue, John W. Goodell, and Dehua Shen (2021) "Comparing search-engine and social-media attentions in finance research: Evidence from cryptocurrencies," *International Review of Economics & Finance*, 75, 723–746, https://doi.org/10.1016/j.iref.2021.05.003.

Manahov, Viktor (2021) "Cryptocurrency liquidity during extreme price movements: is there a problem with virtual money?" *Quantitative Finance*, 21 (2), 341–360, 10.1080/14697688.2020.1788718, doi: 10.1080/14697688.2020.1788718.

Matkovskyy, Roman and Akanksha Jalan (2019) "From financial markets to Bitcoin markets: A fresh look at the contagion effect," *Finance Research Letters*, 31, 93–97, https://doi.org/10.1016/j.frl.2019.04.007.

Meriläinen, Jari-Mikko (2016) "Lending growth during the financial crisis and the sovereign debt crisis: The role of bank ownership type," *Journal of International Financial Markets, Institutions and Money*, 41, 168–182, https://doi.org/10.1016/j.intfin.2015.12.011.

Murinde, Victor, Efthymios Rizopoulos, and Markos Zachariadis (2022) "The impact of the FinTech revolution on the future of banking: Opportunities and risks," *International Review of Financial Analysis*, 81, 102103, https://doi.org/10.1016/j.irfa.2022.102103.

Nakamoto, Satoshi (2009) "Bitcoin: A Peer-to-Peer Electronic Cash System," *Available at https://Bitcoin.org/Bitcoin.pdf*.

Navaretti, Giorgio Barba mname, Giacomo mname Calzolari, Joss Manuel mname Mansilla-Fernandez, and Alberto F. mname Pozzolo (2018) "Fintech and Banking. Friends or Foes?" *SSRN Electronic Journal*, 10.2139/ssrn.3099337.

Newey, Whitney K. and Kenneth D. West (1987) "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55 (3), 703–708, 10.2307/1913610.

Papadamou, Stephanos, Nikolaos A. Kyriazis, Panayiotis Tzeremes, and Shaen Corbet (2021) "Herding behaviour and price convergence clubs in cryptocurrencies during bull and bear markets," *Journal of Behavioral and Experimental Finance*, 30, 100469, https://doi.org/10.1016/j.jbef.2021.100469.

Park, Beum-Jo (2011) "Asymmetric herding as a source of asymmetric return volatility," *Journal of Banking & Finance*, 35 (10), 2657–2665, https://doi.org/10.1016/j.jbankfin.2011.02.025.

Pathak, Soumyajit and Alpana Kakkar (2020) *Cryptocurrency Price Prediction Based on Historical Data and Social Media Sentiment Analysis*, 47–55, Singapore: Springer Singapore, 10.1007/978-981-15-2043-3_7.

Philippas, Dionisis, Nikolaos Philippas, Panagiotis Tziogkidis, and Hatem Rjiba (2020) "Signal-herding in cryptocurrencies," *Journal of International Financial Markets, Institutions and Money*, 65, 101191, https://doi.org/10.1016/j.intfin.2020.101191.

Philippon, Thomas (2015) "Has the US Finance Industry Become Less Efficient? On the Theory and Measurement of Financial Intermediation," *American Economic Review*, 105 (4), 1408–38, 10.1257/aer.20120578.

Rubbaniy, Ghulame, Kienpin Tee, Perihan Iren, and Sonia Abdennadher (2021) "Investors' mood and herd investing: A quantile-on-quantile regression explanation from crypto market," *Finance Research Letters*, 102585, https://doi.org/10.1016/j.frl.2021.102585.

Salas, Vicente and Jesús Saurina (2002) "Credit Risk in Two Institutional Regimes: Spanish Commercial and Savings Banks," *Journal of Financial Services Research*, 22 (3), 203–224, 10.1023/A:1019781109676.

Schmitt, Noemi and Frank Westerhoff (2017) "Herding behaviour and volatility clustering in financial markets," *Quantitative Finance*, 17 (8), 1187–1203, https://EconPapers.repec.org/RePEc:taf:quantf:v:17:y:2017:i:8:p:1187-1203.

Sheng, Tianxiang (2021) "The effect of fintech on banks' credit provision to SMEs: Evidence from China," *Finance Research Letters*, 39, 101558, https://doi.org/10.1016/j.frl.2020.101558.

Shiller, Robert J. (2003) "From Efficient Markets Theory to Behavioral Finance," *Journal of Economic Perspectives*, 17 (1), 83–104, 10.1257/089533003321164967.

Stavroyiannis, Stavros and Vassilios Babalos (2019) "Herding behavior in cryptocurrencies revisited: Novel evidence from a TVP model," *Journal of Behavioral and Experimental Finance*, 22, 57–63, https://doi.org/10.1016/j.jbef.2019.02.007.

Stiglitz, Joseph E. (1975) "Incentives, Risk, and Information: Notes Towards a Theory of Hierarchy," *The Bell Journal of Economics*, 6 (2), 552–579, http://www.jstor.org/stable/3003243.

Stiglitz, Joseph E. and Andrew Weiss (1981) "Credit Rationing in Markets with Imperfect Information," *The American Economic Review*, 71 (3), 393–410, http://www.jstor.org/stable/1802787.

Stiroh, Kevin J. (2004) "Diversification in Banking: Is Noninterest Income the Answer?" *Journal of Money, Credit and Banking*, 36 (5), 853–882, http://www.jstor.org/stable/3839138.

Stock, James and Motohiro Yogo (2005) *Testing for Weak Instruments in Linear IV Regression*, 80–108, New York: Cambridge University Press, http://www.economics.harvard.edu/faculty/stock/files/TestingWeakInstr_Stock%2BYogo.pdf.

Stulz, René M. (2019) "FinTech, BigTech, and the Future of Banks," *Journal of Applied Corporate Finance*, 31 (4), 86–97, https://doi.org/10.1111/jacf.12378, https://doi.org/10.1111/jacf.12378.

Teng, Yin-Pei (2018) "An examination of stock market herding behavior with different market maturity," *Journal of Statistics and Management Systems*, 21 (6), 1121–1130, 10.1080/09720510.2018.1492891, doi: 10.1080/09720510.2018.1492891.

Thakor, Anjan (2020) "Fintech and banking: What do we know?" *Journal of Financial Intermediation*, 41 (C), S104295731930049X, https://EconPapers.repec.org/RePEc:eee:jfinin:v:41:y:2020:i:c:s104295731930049x.

Vidal-Tomás, David, Ana M. Ibáñez, and José E. Farinós (2019) "Herding in the cryptocurrency market: CSSD and CSAD approaches," *Finance Research Letters*, 30, 181–186, https://doi.org/10.1016/j.frl.2018.09.008.

Wang, Gang-Jin, Xin-yu Ma, and Hao-yu Wu (2020) "Are stablecoins truly diversifiers, hedges, or safe havens against traditional cryptocurrencies as their name suggests?" *Research in International Business and Finance*, 54, 101225, https://doi.org/10.1016/j.ribaf.2020.101225.

Wang, Xiaoting, Siyuan Hou, Khine Kyaw, Xupeng Xue, and Xueqin Liu (2023) "Exploring the determinants of Fintech Credit: A comprehensive analysis," *Economic Modelling*, 126, 106422, 10.1016/j.econmod.2023.106422.

Wanidwaranan, Phasin and Chaiyuth Padungsaksawasdi (2022) "Unintentional herd behavior via the Google search volume index in international equity markets," *Journal of International Financial Markets, Institutions and Money*, 77, 101503, https://doi.org/10.1016/j.intfin.2021.101503.

Weber, Michael, Yuriy Gorodnichenko, and Olivier Coibion (2022) "The Expected, Perceived, and Realized Inflation of U.S. Households before and during the COVID19 Pandemic," *National Bureau of Economic Research Working Paper Series*, No. 29640, 10.3386/w29640.

Xie, Ping, Chuanwei Zou, and Haier Liu (2016) "The fundamentals of internet finance and its policy implications in China," *China Economic Journal*, 9 (3), 240–252, 10.1080/17538963.2016.1210366.

Yao, Juan, Chuanchan Ma, and William Peng He (2014) "Investor herding behaviour of Chinese stock market," *International Review of Economics & Finance*, 29, 12–29, https://doi.org/10.1016/j.iref.2013.03.002.

Yarovaya, Larisa, Roman Matkovskyy, and Akanksha Jalan (2021) "The effects of a "black swan" event (COVID-19) on herding behavior in cryptocurrency markets," *Journal of International Financial Markets, Institutions and Money*, 75, 101321, https://doi.org/10.1016/j.intfin.2021.101321.

Yasir, Muhammad and A. Özlem Önder (2022) "Time-varying herding spillover for emerging countries: evidence from BRIC countries and Turkey," *Review of Behavioral Finance*, ahead-of-print (ahead-of-print), 10.1108/RBF-10-2021-0218.

Zhao, Jinsong, Xinghao Li, Chin-Hsien Yu, Shi Chen, and Chi-Chuan Lee (2022) "Riding the FinTech innovation wave: FinTech, patents and bank performance," *Journal of International Money and Finance*, 122, 102552, https://doi.org/10.1016/j.jimonfin.2021.102552.