Essays in Macroeconomics and Finance:

Author: Yushan Hu

Boston College Electronic Thesis or Dissertation, 2020

ESSAYS IN MACROECONOMICS AND FINANCE

YUSHAN HU

A thesis
submitted to the Faculty of
the department of economics
in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

Boston College
Morrissey College of Arts and Sciences
Graduate School

April 2020

Essays in Macroeconomics and Finance

Yushan Hu

Advisor: Fabio Schiantarelli, Ph.D.
Zhijie Xiao, Ph.D.
Peter Ireland, Ph.D.
Philip E. Strahan, Ph.D.

ABSTRACT

This dissertation consists of three essays in macroeconomics and finance. The first and second chapters analyze the impact of the financial shocks and anti-corruption campaign on Chinese firms through the bank lending channel. The third chapter provides a new method to predict the cash flow from operations (CFO) via semi-parametric estimation and machine learning.

The first chapter explores the impact of the financial crisis and sovereign debt crisis on Chinese firms through the bank lending channel and firm borrowing channel. Using new data linking Chinese firms to their bank(s) and four different measurements of exposure to the international markets (international borrowing, importance of lending to foreign listed companies, share of trade settlement, and exchange/income), I find that banks with higher exposure to the international markets cut lending more during the recent financial crisis. In addition, state-owned bank loans are more pro-cyclical compared with private bank loans. Moreover, banks with higher exposure to the international markets cut lending more when there is a negative shock in OECD GDP growth. With regard to firm borrowing channel, I find that firms with higher weighted aggregate exposure to the international markets through banks have lower net debt, cash, employment, and capital investment during the financial crisis. Firms with higher weighted aggregate exposure to

the global markets have higher net debt and lower cash, employment, and capital investment when there is a negative shock in OECD GDP growth. This paper also provides a theoretical model to explain the mechanism in a partially opened economy like China.

The second chapter discusses the impact of the anti-corruption campaign on Chinese firms through the bank lending channel. Using confidential data linking Chinese firms to their bank(s) and prefecture-level corruption index, I find that banks located in more corrupted prefectures offer significantly less credits before the anti-corruption investigation, and this effect changes the direction after the investigation. Moreover, banks located in more corrupted prefectures tend to use higher interest rates, longer maturity, and more collateral before the campaign, all of these effects change the direction after the campaign. This paper suggests that the banks located in more corrupted prefectures have stronger monopoly power (or higher markup, and lower efficiency). This monopoly effect could be proved by that the bank concentration ratio is higher, and the bad loans of the banks are higher in the more corrupted areas, and all of these effects disappear after the campaign.

The third chapter considers the methods of prediction of Cash flow from operations (CFO). Forecasting CFO is an essential topic in financial econometrics and empirical accounting. It impacts a variety of economic decisions, including valuation methodologies employing discounted cash flows, distress prediction, risk assessment, the accuracy of credit-rating predictions, and the provision of value-relevant information to security markets. Existing literature on statistically-based cash-flow prediction has pursued cross-sectional versus time-series estimation procedures in a mutually exclusive fashion. Cumulated empirical evidence indicates that the beta value varies across firms of different sizes, and the cross-sectional regression can not capture an idiosyncratic beta. However, although a time series based predictive model has the advantage of allowing for firm-specific variability in beta, it requires a long enough time series data. In this paper, we extend the literature on statistically-based, cash-flow prediction models by introducing an estimation procedure that, in essence, combine the favorable attributes of both cross-sectional estimation via the use of "local" cross-sectional data for firms of similar size and time-series estimation via the capturing of firm-specific variability in the beta parameters for the independent variables. The local learning approach assumes no a priori knowledge on the constancy of the beta coefficient. It allows the information about coefficients to be represented by only a subset of observations. This feature is particularly relevant in the CFO model, where the beta values are only related to cross-sectional data information that is "local" to its size. We provide empirical evidence that the prediction of cash flows from operations is enhanced by jointly adopting features specific to both cross-sectional and time-series modeling simultaneously.

DEDICATION

To Penglong Zhang and Oliver Zhang

ACKNOWLEDGMENTS

I thank my advisors Fabio Schiantarelli and Zhijie Xiao, for their guidance and persistent help. I thank my committee members Peter Ireland and Philip E. Strahan, for their support and valuable advice. I thank Susanto Basu, Christopher Baum, Ryan Chahrour, Mehmet Ekmekci, Can Erbil, Pablo Guerron, Iftekhar Hasan, Kose John, Tzuo-Hann Law, Arthur Lewbel, Ben G. Li, Jaromir Nosal, Dongho Song, Zheng Sun, Robert Ulbricht, Penglong Zhang, Song Zhang, Tony Weiting Zhang and participants at the 2019 annual meeting of the Financial Management Association, BU/BC Green Line Macro Meeting (2019), 2018 annual meeting of the Southern Economic Association, 2018 annual meeting of the Eastern Economic Association, dissertation workshop, and Macro Lunch for their helpful comments.

Table of Contents

Page

List of Tables

List of Figures

Chapter One

International Exposure and the Transmission of Financial Shocks: Evidence from China

1.1 Introduction

What is the impact of the two crises on the credit supply of Chinese banks? This paper analyzes the effects of the financial crisis in the US and the sovereign debt crisis in European countries on the credit supply of Chinese banks.

Why do we focus on Chinese bank lending channel? Table 1.1 compares the financial development of China and the United States in 2016, column 1 shows that the absolute value of the bank credit in China is even more significant than the bank credit in the United States (15.45 vs. 12.44). When we consider the amount of bank credit as a fraction of the GDP, the difference between China and the United States is significantly amplified (137.95% vs. 67.00%). Moreover, the amount of bank credit is the largest among all financial instruments (stock, fixed income, insurance, and investment funds). Therefore I focus on Chinese bank credit in this paper.

Why do we choose the financial crisis in 2008 as the supply side shock? Figure 1.1

shows the bank assets and ROA of Chinese banks, we could find an obvious turning point in the year 2008, from 2000 to 2008, the ROA of Chinese banks increased dramatically, however, after 2008 the ROA became flat and even decreased later. Thus, something happened in Chinese bank system in the year 2008.

Why do we use the financial crisis as a natural experiment? The recent economic recession happened outside China, but it should significantly affect the Chinese banks that have a high level of openness. In other words, the liquidity problem was experienced by Chinese banks with a higher degree of openness, so that the liquidity and funding shock varied substantially across banks.

Do the financial crisis in the US and sovereign debt crisis in European countries have any impact on Chinese banks? Some people suggest no effects. There are two points to support this hypothesis. First, both the financial and sovereign debt crisis occurred outside China. Secondly, China is almost the most regulated financial market in the main economy. However, this paper suggests financial crisis have an impact on Chinese bank lending channel. We could find three supply-side bank lending channels. First, we have a passive and direct bond market channel. When the international bond markets contracts decrease following the financial crisis, the bank credits of high exposure banks decreases. Secondly, we have an active and indirect stock market channel. The stock price of foreign listed firms declined significantly in the stock market after the recent financial crisis, the bank credits of high exposure banks declined accordingly. Finally, we have an active and indirect goods market channel. When firm exports decrease in the goods market after the financial crisis, the asset side of bank balance sheets worsens in quantity and quality.

Chodorow-Reich (2013) shows the impact of the financial crisis on the bank lending channel of the United States. Similarly, as the financial crisis originated outside China, we could use the dispersion in the exposure to the recent financial crisis as a source of exogenous variation in the availability of credit to borrowers in Chinese financial market. This paper also uses the Khwaja and Mian (2008) technique to examine the bank lending channel and firm borrowing channel simultaneously, in particular, I use the firm-year fixed effects to control the impact endogenous coming from the demand side. Moreover, I also use bank-level control variables to address the issue of bank heterogeneity.

This paper constructs a novel data set which combines information on 281 major Chinese banks and 3302 listed Chinese firms from 2001 to 2016. We also use some variables to measure the degree of the openness of the banks: International borrowing, which assess a bank's openness by dividing the bank's international commercial borrowing and bond by its total assets; Foreign Listed, which is the loans given to B-share, H-share and overseas-listed firms as a fraction of the total loan; Trade Settlement, which is the bank-level trade settlements over the total loans; as well as Exchange/Income, which is the bank's exchange gains as a fraction of its total income.

Acharya et al. (2018) discusses three potential transmission channels of the sovereign debt crisis in Europe. Although I do not focus on the sovereign debt crisis, I will focus on the channels identified in the paper of Acharya et al. (2018). Compared to a typical banking crisis in which the lending supply shock is solely caused by banks' impaired financial health, the impact of the financial crisis and sovereign debt crises on bank lending in China is much more complicated. In particular, there are three channels through which the financial crisis potentially affected banks' lending decisions: two active channels, which worked through a reduction in the bank credit given to foreign listed and international trade firms because of banks' risk-limiting behavior, and one passive channel, which acted via the substantial decline in the bank's existing international borrowings. Moreover, we could use international borrowing to represent the bonds market, foreign listed is a good proxy for the stocks market, and trade settlement is a suitable measurement for goods market. I also compared these three potential channels in three different markets, and find that bonds and stocks markets are more responsive when facing with the recent financial crisis.

In summary, my paper provides several novel findings. I find that banks with higher exposure to the international markets cut lending more following the recent financial crisis. Moreover, banks with higher exposure to the global markets cut credit more when there is a negative shock in the OECD GDP growth rate.

Moreover, this paper considers both the market effect and the government regulation effect. China was experiencing a domestic slow down when it got hammered by the global financial crisis in the fourth quarter of 2008. The Chinese policy response since that time has been extraordinarily vigorous. At the beginning of 2009, China Banking Regulatory Commission (CBRC) and the People's Bank of China (PBOC) put strict financial regulations on the banks with high exposure to the international financial market to avoid the liquidity risk. However, these kinds of banks are highly encouraged before the financial crisis in 2008. This paper groups all the banks into state-owned banks and private banks; thus, state-owned banks were affected more because of the regulation effect. To Sum up, the data suggest that regulatory pressure was substantial. This finding could be seen by observing that the reduction in loans is more significant for state-owned banks on which the Chinese government has more control.

In addition, I find that the financial shock from the international markets also impacts Chinese firm borrowing channel. In particular, firms with higher weighted aggregate exposure to the global markets through the banks have lower net debt, cash, employment and capital investment during the financial crisis, and firms with higher weighted aggregate exposure to the international markets have higher net debt and lower cash, employment and capital investment when there is a negative shock in OECD GDP growth rate.

This paper also develops a model bank in the partially opened economy where banks differ in their exposure to the international market. This general equilibrium model sheds light on the mechanism behind the transmission channels that impact the bank lending channel. The proposed model predicts that when world interest rate increases, banks with higher exposure to the international market would reduce more bank loans through the balance sheet hit the channel. Besides, when firm return decreases, banks with higher international exposure will reduce more bank credits through the risk-limiting behavior channel. The predictions of the model are consistent with crucial empirical results in this paper.

The structure of the paper is as follows. Section 1.2 further discusses our contribution to the existing literature. Section 1.3 presents the background of the Chinese bank system. Section 1.4 illustrates the three potential channels to impact the Chinese bank lending channel. Section 1.5 presents the bank loan effects of the financial crisis. Section 1.6 shows the firm-level financial and real effects of the financial shocks. Section 1.7 studies an incentive conceptual framework of financial inter-mediation to illustrate the lending patterns observed during the recent financial crises. Section 1.8 presents the robustness checks. Section 1.9 concludes the paper.

1.2 Literature Review

Chodorow-Reich (2013) discuss the link between credit supply shocks and employment during the recent financial crisis. He constructs a new data set which combines information on banking relationships and employment at 2000 non-financial firms during the 2008-2009 crisis. Chodorow-Reich (2013) first verifies empirically the importance of banking relationships, which imply a cost to borrowers who switch lenders. To address potential endogeneity issues, he uses the dispersion in lender health following the Lehman crisis as a source of exogenous variation in the availability of credit to borrowers. Similarly, I use the distribution in lender exposure to the international markets following the financial crisis as the source of exogenous variation to exploit the fact that the financial crisis originated outside the Chinese financial market. Finally, he shows that firms that had pre-crisis relationships with less healthy lenders had a lower likelihood of obtaining a loan following the Lehman bankruptcy, paid a higher interest rate if they did borrow, and reduced employment by more compared to pre-crisis clients of healthier lenders.

Since Chodorow-Reich (2013) only consider the cross-sectional regression, he only paid attention to a point instead of a period, including the periods before and after the financial crisis. In this case, he only considers the firm fixed effects. In my paper, I use panel data regression, which means I could estimate the time trend of the impact, and also I could use year fixed effects, or more general, year-firm fixed effects.

Khwaja and Mian (2008) analyze how supply-side bank liquidity shocks get transmitted to the rest of the economy. They examine the impact of liquidity shocks by exploiting cross-bank liquidity variation induced by unanticipated nuclear tests in Pakistan. When Pakistan tested nuclear devices in 1998, the IMF suspended the exchange rate liquidity support. Banks experienced the deposit run with larger dollar deposit accounts. The liquidity shock varied substantially across banks.Khwaja and Mian (2008) estimate the bank lending channel and firm borrowing channel simultaneously, and show that for the same firm borrowing from two different banks, its loan from the bank experiencing a 1 percent larger decline in liquidity drops by an additional 0.6 percent. This paper also uses the fixed effect technique in Khwaja and Mian (2008) to estimate the bank lending channel and firm borrowing channel simultaneously.

Whereas several contributions have analyzed the effects of the financial crisis on bank lending, there is more limited evidence about the resulting impact on sovereign debt crises. Balduzzi et al. (2017) summarize the literature regarding micro evidence on the effects of financial shocks related to the recent financial and sovereign debt crisis. They use Italian firm-bank data (more representative than the data from the Italian Credit Registry) and find that financial market valuations of banks affect firm's investment and employment decisions through their impact on the level and volatility of banks' cost of funding. The identification strategy in their paper relies on the heterogeneous time variation in banks' cost of funding generated by the 2007-2009 financial crisis and by the 2010-2012 sovereign debt crisis, both generated outside the Italian non-financial corporate sector.

Moreover, Acharya et al. (2018) explore the impact of the credit crunch that followed the European debt crisis on the corporate policies of European firms. They show that banks' exposures to impaired sovereign debt and the risk-limiting behavior of under-capitalized banks contributed significantly to the severity of the crisis.

Overall, the specific business cycle model is a hybrid of Gertler and Karadi (2011) that allows for financial inter-mediation and Kiyotaki and Moore (2012) that allows for liquidity risk. Gertler and Kiyotaki (2010) develop a canonical framework to think about credit market frictions and aggregate economic activity in the context of the financial crisis. They use the framework to address two issues in particular: first, how disruptions in financial intermediation can induce a crisis that affects real activity; and second, how various credit market interventions by the central bank might work to mitigate the crisis. In this paper, we allow for the interaction of central bank, private bank as well as the state-owned bank. Song et al. (2011) construct a model consistent with China's economic transition: entrepreneurial firms use more productive technologies, but due to financial imperfections, they must finance investments through internal savings. State-owned firms have low productivity but survive because of better access to credit markets. In this paper, we use homogeneous firms and heterogeneous banks. We assume that state-owned banks focus on risk, thus they have low return and low risk, while private banks focus on profit, therefore they have a high return and high risk. Based on Holmstrom and Tirole (1997), this paper tries to develop an incentive model of financial intermediation to illustrate the lending patterns observed during the recent financial crisis.

This paper also contributes to the academic literature on the Chinese banking system. There is a growing literature that analysis the Chinese financial market. The Chinese banking system has played an important role in the growth of China's economy. (Allen et al., 2005) In the past decade, several papers have been published, analyzing different aspects of the Chinese banking system. García-Herrero et al. (2006), Fu and Heffernan (2009), Lin and Zhang (2009), and Dong et al. (2016) focus on the reform and/or performance of the Chinese banking system. In the 1990s, the banking system in China was dominated by four large state-owned banks. However, these four state-owned big banks faced serious problems, such as high non-performing loans and inefficient operation and management. Thus, the Chinese authorities initiated a series of reforms on the banking system in 2003. The four state-owned banks became joint-stock commercial banks, and they have been listed successively on the Shanghai Stock Exchange since 2006. Reforms were also implemented in other small and medium-sized commercial banks and rural credit cooperatives since 2003. After the reform, the Chinese banking system became more and more comprehensive and diversified, playing a dominant role in the Chinese financial system.

Berger et al. (2009), Ariff and Luc (2008), and Asmild and Matthews (2012) investigate the efficiency of Chinese banks. Bailey et al. (2011) and Fenech et al. (2014) investigate the quality of bank loans and some other characteristic of the Chinese banking system. Fenech et al. (2014) find that the Chinese banking system and its loan quality is directly linked to real estate values and government-supported infrastructure projects. Clearly, in China, the two-way causality of both bank growth and its soundness partly depends on these projects' propensity in generating sufficient cash flows to repay the bank loans. Chen et al. (2014), Wang et al. (2015) and Huang et al. (2015) investigate systemic risk in the Chinese banking system.

There is a growing literature that analyzes the Chinese bank lending channel. Qian et al. (2015) use the event that many Chinese banks implemented reforms that delegated authority to individual loan officers in 2002 and 2003 as a plausibly exogenous shock, and find that the bank's internal risk rating becomes a stronger predictor of loan interest rates and ex-post outcomes after the reform. Gao et al. (2019) analyze the borrowing and defaults of local governments in China. They find that policy bank loans to local governments have significantly lower default rates than commercial bank loans with similar characteristics. The exogenous shock they use is the announcement of the 4-trillion stimulus package in 2009. Li et al. (2018) provide a novel empirical finding that the recent anti-corruption investigations in China are associated with bank loan reallocation from less productive state-owned enterprises (SOEs) to more productive non-SOEs, indicating that the competition effect dominates the contagion effect for non-SOEs. The exogenous shock they use is that the government required the immediate information disclosures of the corruption-related officials to the public, intending to improve the transparency of governance since late 2012.

1.3 Background

According to the Chinese Financial Stability Report (2009-2014), the banking system accounted for more than 90% of the total asset of financial institutions since 2008. According to the China Banking Regulatory Commission (CBRC) and the People's Bank of China (PBOC), the composition of Chinese banks are 3 development banks, 5 large-scale commercial banks, 12 joint-stock commercial banks, 145 city commercial banks, 468 rural commercial banks, 122 rural cooperative banks, 1803 agricultural credit cooperatives, 1134 new rural financial institution, 1 postal savings bank, 92 foreign banks' branches or non-bank financial institutions.

China has achieved remarkable progress in reforming its banking system. There are 117 Chinese banks in the 2015 Top 1000 of banks; three of them (the Bank of China, the Industrial and Commercial Bank of China, and the Agricultural Bank of China) are rated as global systemically important banks. Chinese banks made $292 billion in aggregate pretax profit in 2013, or 32% of total earnings of the world's top 1,000 banks, according to The Banker magazine.

However, there are also lots of severe problems in the Chinese banking system. First off, although the Chinese banking system has become more diversified, it is still dominated by a few big banks. For example, five large-scale commercial banks accounted for

43% of total assets of the Chinese banking system at the end of 2013 and 12 joint-stock commercial banks for 18%.

Moreover, China's banking sector, together with other sectors of strategic importance, has been under intensive monitoring by the government, mainly through its central bank (People's Bank of China, PBOC) and the China Banking Regulatory Commission (CBRC). Qian et al. (2015) states that PBOC limits the movements of interest rates on both deposits and loans by setting base rates along with upper and lower bounds. These rates and bounds vary over business cycles and with loan maturities.

1.4 Transmission Channels

Compared to the financial crisis in which the lending supply shock is solely caused by banks' impaired financial health in the United States, the impact of the financial crisis on bank lending is much more complex in China. In particular, there are three channels through which the financial crisis potentially affected banks' lending decisions: one passive channel, which worked through the substantial decrease on the international borrowing, and two active channels, which acted via the reduction of loans to foreign listed and trade firms because of bank's risk-limiting behavior.

To evaluate the passive channel, I need to determine how strongly a bank was affected by the credit crunch in the international markets. As in Acharya et al. (2018), I construct a variable to measure the exposure to global markets of bank i in year t as follows:

$$International\ Borrowing_{i,t} = \frac{International\ bond_{i,t} + International\ commercial\ borrowing_{i,t}}{Asset_{i,t}} \tag{1.1}$$

In particular, Chinese banks borrow money from three different kinds of institutions; they are foreign commercial banks, other foreign financial institutions, and the World Bank. A primary concern about this measurement is that if most of the international borrowing came from the World Bank, it wouldn't be affected significantly during the financial crisis period. However, there are only two Chinese banks could get money from the World Bank (China Investment Bank and Agricultural Bank of China), and the borrowing amount is minimal. Moreover, the bank-level average international borrowing declined significantly at the beginning of the year 2009, which hits the bank's balance sheet directly.

The risk-limiting motive arises since, as the default risk of both the foreign listed firms and the trade firms increased, the banks had an incentive to decrease the credit given to these firms. There are four compositions of stocks in China: A-share stock, which faces to only Chinese citizens; B-share stock, which faces to foreigners (non-Chinese citizens); H-share, which are listed in Hong Kong, as well as overseas-listed stocks, which are listed in U.S., Japan, Singapore... Since the financial policy and government regulation is utterly different in Hong Kong compared with the mainland, this paper regards the firm listed in Hong Kong as foreign listed firms. Therefore, I could construct another variable to measure the exposure to international markets of bank i in year t as follows:

$$Foreign\ Listed_{i,t} = \frac{\sum_j Loan\ to\ B\ ,\ H\text{-}share\ \&\ overseas\ listed\ Chinese\ firms_{ijt}}{Total\ loan_{i,t}} \tag{1.2}$$

From my data, the impacts of the financial crisis on foreign listed firms are significant. First of all, the sales of foreign listed firms decline significantly relative to A-share listed firms. Secondly, the capital expenditure (investment) of foreign listed firms drop dramatically relative to A-share listed firms. Thirdly, the unemployment rate of foreign listed firms increases significantly relative to A-share listed firms.

Figure 1.2 shows the change in China's GDP, international trade, and foreign exchange reserve from 1976 to 2016. International trade here includes both export and import. We could find a significant drop in trades in the year 2008. Moreover, we could also regard the year 2008 as a turning point. China joined the WTO at the year 2001, and from 2001 to 2008, the international trade of China increased dramatically, however, after the year

2008, the trades trend became flat, and it even decreases later in the year 2013. Therefore, I construct the third international exposure measurement of bank i in year t as:

$$Trade\ Settlement_{i,t} = \frac{Trade\ settlement_{i,t}}{Total\ loan_{i,t}} \tag{1.3}$$

Note that we use trade settlement divided by total loans to control for the size of the bank. In this paper, we focus on the bank loan, so we use total loans instead of assets to define the different sizes of banks. A bank may have lots of assets but only a small amount of loans, it is determined as a small bank in the research background of this paper.

1.5 Bank Loan Effects of the Financial Crisis

Before assessing the importance of the three channels (international borrowing, foreign listed, trade settlement) separately, we first analyze whether, in general, the financial crisis in global markets affected the financial market in China through a change in the banks' lending behavior. All three channels could potentially lead to a reduction in the bank loan supply after the financial crisis, either by reducing a bank's debt capacity (international borrowing) or by risk-limiting behavior (foreign listed and trade settlement). Hence, we expect that banks that are more dependent on the global markets significantly affected by the financial crisis. Moreover, state-owned banks affected by government regulation acted differently in terms of financial decisions compared to private firms after the financial crisis. In Section 1.7, we then analyze which of the three channels were of first-order importance for the adverse financial effects incurred by Chinese banks.

1.5.1 Data

This paper uses a novel data set that contains bank-firm relationships in China, along with detailed firm and bank-specific information. Our sample period is from 2001 to 2016, such that we have an asymmetric time window surrounding the beginning of the financial crisis in the United States. Chinese data in this paper comes three primary data sets: Wind Datafeed Service (referred to as WDS), GTA The China Stock Market and Accounting Research (referred to as GTA CSMAR) Database as well as Almanac of China's Finance and Banking (2001-2016).

Information about bank-firm relationships is from the bank loan data in GTA CSMAR Database. GTA CSMAR access data on the China stock markets and the financial statements of China's listed companies. GTA CSMAR is a unique, comprehensive database of China stock returns, covering all companies listed on the Shanghai Stock Exchange and the Shenzhen Stock Exchange. I collect information on bank loans to all of the listed firms from China.

I augment the data on bank-firm relationships with bank-level and firm-level data taken from WDS. WDS provides historical reference data, real-time market data, and historical intraday market data, covering stocks, bonds, futures, foreign exchanges, funds, indices, warrants, and macro market data as well as descriptions, real-time market data, financial data, dividend data, corporate actions, and historical intraday data.

In addition, I combine the data set with bank-level information (trade settlement) from Almanac of China's Finance and Banking (2001-2016). Almanac of China's Finance and Banking is a high informative yearbook that is supervised by the People's Bank of China, sponsored by China society for finance and banking. Established in 1986, the yearbook has been consecutively published for 29 volumes at the publication frequency of one volume a year.

To get data about the OECD GDP growth rate, I obtain data about the GDP of 35 OECD member countries from information disclosed on the OECD data websites. The definitions of all variables are summarized in Table 1.2.

1.5.2 Summary Statistics

In the following, I present summary statistics and explore whether our identifying assumptions are plausible. Table 1.3 presents summary statistics for the loan level variables in our primary data set. Since our data covers the universe of all business loans of listed firms, there is considerable variation in loan sizes. For example, the average loan size is about 358.828 million yuan, and the standard deviation is 2418.99 million yuan. Given the considerable size variation, I choose to use the log loan volume instead of the loan volume. Table 1.3 also shows that the average of the log loan volume of state-owned banks is similar to that of private banks (18.842 vs. 18.548), and the average of the change of the log loan volume of state-owned banks are very similar to that of private banks (11.4% vs. 12.3%).

Table 1.4 presents summary statistics for the bank-level variables in our data set, the bad loan is a variable used to measure the health of a specific bank i in year t, and it is constructed as:

$$Bad\ Loan_{i,t-1} = \frac{Subprime\ loan_{i,t-1} + Doubt\ loan_{i,t-1} + Loss\ loan_{i,t-1}}{Asset_{i,t-1}} \tag{1.4}$$

The average bad loan rate is higher for the state-owned banks compared with private banks. The reason behind this is that the Chinese government has more control power on state-owned banks. To save some non-profitability or insolvent state-owned firms, the Chinese government would require the state-owned banks to give credit to this kind of firms. The loan managers in state-owned banks don't have any chose but lend money to these firms even though they know that these firms have extremely high default risk. Moreover, I also find that state-owned banks could borrow more money from the central bank, it is obvious that state-owned banks may have more connection or relationships with the Chinese government and the central bank (PBOC).

1.5.3 Empirical Methodology

Following the Chow test for structural breaks, I employ the three-period model, which includes pre-crisis, crisis, and postcrisis periods. I define the indicator variable $Crisis_t$ equals to one only in the period of the financial crisis (2008-2009) and sovereign debt crisis (2010-2012), and $Post_t$ equals to one for fiscal years after 2012. The year 2001 is a particular year for China; this is the year China joined WTO. Moreover, we could regard this year as the first year China opens its door after closing it for thousands of years.

For the level effect of the bank-firm matched loans, I employ the following specification for bank i and firm j in year t:

$$
\begin{aligned}
L_{ijt} \quad = \quad & \beta_1 Exposure_{i,t-1} + \beta_2 Exposure_{i,t-1} \times State\text{-}owned_i + \beta_3 Exposure_{i,t-1} \times Crisis_t \\
& + \beta_4 Exposure_{i,t-1} \times Crisis_t \times State\text{-}owned_i + \beta_5 Exposure_{i,t-1} \times Post_t \\
& + \beta_6 Exposure_{i,t-1} \times Post_t \times State\text{-}owned_i + \gamma X_{i,t-1} + \lambda_{jt} + \epsilon_{ijt}
\end{aligned} \tag{1.5}
$$

$$
L_{ijt} \quad = \quad \beta_1 Exposure_{i,t-1} + \beta_2 Exposure_{i,t-1} \times gr_{Yt}^{OECD} + \gamma X_{i,t-1} + \lambda_{jt} + \epsilon_{ijt} \tag{1.6}
$$

where i represents bank, j represents firm, t represents year, L_{ijt} is the log loan size, $Crisis_t$ is an indicator variable equals to one for fiscal years 2008-2012, $Post_t$ is an indicator variable equals to one for fiscal years after 2012, $State\text{-}owned_i$ is also an indicator variable equals to one if the bank is a state-owned bank, and it equals to zero if the bank is a privately-owned bank, $X_{i,t-1}$ is the bank-level control variables, λ_{jt} is the year-firm fixed effects. I use three foreign exposure measurements: international borrowing, foreign listed and trade settlement. International Borrowing$_{i,t-1}$ could be represented as (International bond$_{i,t-1}$+International commercial borrowing$_{i,t-1}$)/Asset$_{i,t-1}$. Foreign Listed$_{i,t-1}$ is (Loan gives to B and H-share and overseas listed Chinese firms$_{i,t-1}$/Total loan$_{i,t-1}$), Trade Settlement$_{i,t-1}$ is Trade settlements$_{i,t-1}$/ Total loan$_{i,t-1}$.

The term (Exposure$_{i,t-1}\times$ State-owned$_i$) is used to control for the different effects of state-owned and private banks on the dimension of the international exposure measurement. If banks with higher exposure to the international markets cut lending more following the recent financial crisis, higher exposure banks would reduce their loan amount after the financial crisis, that is, I expect β_3 in Eq. (1.5) to be negative. Moreover, if the Chinese government tried to close the door to avoid the substantial loss from the outside world following the financial crisis, the state-owned banks should be affected more compared with the privately-owned banks. Therefore, I expect β_4 in Eq. (1.5) to be negative.

$Crisis_t$ and $Post_t$ are only two dummies. If we want to know something more about the specific changes of the bank credit, we need a variable to measure the international market shocks. It is just like if we want to pull the water into the swimming pool, we need to consider both the velocity of the water and the intersecting surface of the water pipe. If we consider the different measurements of the exposure to international markets as the intersecting surface of the water pipe, the international market shocks should be a good indicator for the velocity of the water. In this paper, I use the OECD GDP growth rate to measure the international market shock for China, China is not one of the 35 OECD member countries, for sure I have to assume that OECD GDP growth rate could represent the financial shocks in the international world from 2001 to 2016.

For the growth rate effect of the bank-firm matched loans, I employ the following specification for bank i and firm j in year t:

$$\begin{aligned}\Delta L_{ijt} \quad = \quad & \beta_1 Exposure_{i,t-1} + \beta_2 Exposure_{i,t-1} \times State\text{-}owned_i + \beta_3 Exposure_{i,t-1} \times Crisis_t \\ & + \beta_4 Exposure_{i,t-1} \times Crisis_t \times State\text{-}owned_i + \beta_5 Exposure_{i,t-1} \times Post_t \qquad (1.7) \\ & + \beta_6 Exposure_{i,t-1} \times Post_t \times State\text{-}owned_i + \gamma X_{i,t-1} + \lambda_{jt} + \epsilon_{ijt}\end{aligned}$$

$$\Delta L_{ijt} \quad = \quad \beta_1 Exposure_{i,t-1} + \beta_2 Exposure_{i,t-1} \times gr^{OECD}_{Yt} + \gamma X_{i,t-1} + \lambda_{jt} + \epsilon_{ijt} \qquad (1.8)$$

where i represents bank, j represents firm, t represents year, ΔL_{ijt} is the growth rate of loan size, gr_{Yt}^{OECD} is the OECD GDP growth rate, $X_{i,t-1}$ is the bank-level control variables, λ_{jt} is the year-firm fixed effects. I also use the three foreign exposure measurements here: international borrowing, foreign listed and trade settlement. International Borrowing$_{i,t-1}$ could be represented as (International bond$_{i,t-1}$+International commercial borrowing$_{i,t-1}$)/Asset$_{i,t-1}$. Foreign Listed$_{i,t-1}$ is (Loan gives to B and H-share and overseas listed Chinese firms$_{i,t-1}$)/Total loan$_{i,t-1}$), Trade Settlement$_{i,t-1}$ is Trade settlements$_{i,t-1}$/ Total loan$_{i,t-1}$.

If banks with higher exposure to the international markets cut lending more when there is a negative shock in OECD GDP growth rate, which means the higher the OECD GDP growth rate, the higher the impact of the exposure to the international markets, that is, I expect β_2 in Eq. (1.8) to be positive.

1.5.4 Identification Strategy

The identification strategy of this paper needs to address four problems: omitted variable, reverse causality, measurement error, and sample selection.

I use the Khwaja and Mian (2008) technique to simultaneously estimate the bank lending and firm borrowing channels stems from identification concerns. Identification concerns arise because events that trigger changes in liquidity supply, such as monetary policy innovations or financial shocks are often accompanied by changes in investment returns and consequently, credit demand. Changes in firm borrowing, therefore, reflect both changes in credit supply as well as credit demand. This paper uses year-firm fixed effects to control credit shocks from the demand side.

One concern is the endogeneity problem of the exposure measurement, if pre-crisis banks that have higher exposure to international markets could costlessly switch to lower exposure banks, with no reason to expect differential outcomes at the pre-crisis level of exposure for different banks. Similar to Chodorow-Reich (2013), I conduct a preliminary test, using demeaned exposure measurement value as dependent variable and $After_t$ (a dummy variable equal to one after 2008) as independent variable, and find that the coefficient before $After_t$ is insignificant, which suggests that there is no systematically changes relative to liquidity supply shock in 2008.

Another concern about the omitted variable problem is the heterogeneity in bank response to financial shocks. Could the lending channel coefficient be driven by inherent differences in how banks respond to the shocks induced by the credit crunch in the international markets? It is possible if there is such response heterogeneity, and it is systematically correlated with a bank's liquidity shock. For example, perhaps the lending channel estimate is picking up differences in how state-owned and private banks react to financial crisis since we know that the Chinese government has more control power on state-owned banks.

At the beginning of 2009, China Banking Regulatory Commission (CBRC) and the People's Bank of China (PBOC) put strict financial regulations on the banks with high exposure to the international financial market to avoid the liquidity risk. However, these kinds of banks are highly encouraged before the financial crisis in 2008. Since state-owned banks should be affected more by this government regulation change, this paper uses an interaction term with state-owned to address the differences.

We also test for such concerns by including various bank characteristics that proxy for such differential lending sensitivity as controls, such as the bank size, the bank's return on assets, the bad loan rate, the amount lend from central bank, tangibility, cash flow, and dummies state-owned, policy, rural and listed banks. These bank-level controls are likely to capture a banks' sensitivity to financial shocks. In particular, I use lagged value to avoid concern about the endogeneity problem. The results below indicate that the lending channel coefficient remains robust to all these bank-level controls.

Although firm fixed effects address the main identification concerns expressed in the literature, there may remain some additional questions. While the fixed effects strategy does not require or make any assumptions about the correlation between liquidity supply and demand shocks, the concern about the reverse causality problem is that if the liquidity supply shocks are anticipated, banks may adjust their lending or firms adjust their borrowing prior to the shock. This would lead to either an under or overestimate of the bank lending channel depending on the direction of the pre-shock loan adjustments. However, in this paper, the natural experiment financial crisis is unanticipated; the financial crisis is happened out of China in the international markets. Therefore, it is difficult for Chinese banks and firms to anticipate this kind of liquidity supply shock. In particular, the identifying assumption for all the level effect regressions in this paper is that year-before financial positions are not positively correlated with unobserved within-bank changes in loans lending following the onset of the crisis.

To reduce the measurement Error, I winsorize all variables at the 1% and 99% level to lessen the influence of outliers. For the sample selection problem, my data set provides more comprehensive coverage of small, micro, and rural banks than other data sets. I include all of the banks, both listed and non-listed banks. A primary concern about the sample selection problem in my paper is that my data set only provides the information of the listed firms, there are also many small and micro firms (non-listed) in China. Gertler and Gilchrist (1994) suggest that since size may proxy for financial constraints, a higher sensitivity of small firms would provide evidence in favor of the "financial accelerator", the view that financial frictions can amplify downturns. Mehrotra et al. (2017) use new and confidential data on income statements and balance sheets of US manufacturing firms to bear on this idea. Thus, my analysis of the impact of the financial crisis on Chinese bank lending channel could be regarded as the "lower bound" impact, since we only consider the listed firms, if listed firms are affected by the financial crisis, small and micro firms should be affected more.

1.5.5 Detection for Structural Breaks

The goal of this analysis is to detect the structural breakpoints for the Chinese bank lending channel.

I conduct the structural Wald test for each period, respectively. Figure 1.3 plots the Wald test statistics of the change-point diagnostics. First, I use international borrowing to measure international exposure, the Wald test statistic is above the 95% critical value for the year 2009 and 2013-2016, implying two structural breakpoints, the year of 2009 and 2013.[1] Secondly, I choose foreign listed to measure external exposure, the Wald test statistic is above the 95% critical value for the year 2008-2009 and 2013-2016, implying two structural breakpoints, the year of 2008 and 2013. Thirdly, I change the exposure measurement to trade settlement, the Wald test statistic is above the 95% critical value for the year 2008 and 2013-2016, implying two structural breakpoints, the year of 2008 and 2013.

1.5.6 Baseline Results

This section is about the empirical results of the analysis. I first provide results for the level effect and then turn to the growth rate effect.

Level Effect

Table 1.8 presents estimates from panel regressions explaining bank-firm-level yearly bank loans from 2001 to 2016. The conditional information set includes Size, State-owned, Policy, Rural, Tangibility, Roa, List, Cash Flow, Lend Central Bank, Bad Loan, Bad Loan $\times$ Crisis, and Bad Loan $\times$ Post. Explicitly speaking, Size is ln(assets), State-owned is an indicator variable equals to one if the bank is a state-owned bank. Policy is an indicator

[1] In the years of 2014-2016, hypothesis H_0 is also rejected. But I drop them because consecutive break points imply the same structural break.

variable that equals one if the bank is a policy bank. Policy bank is something very special in China. The difference between the policy bank and the commercial banks is that the goal of the policy banks is not profit maximization. Their goal is to try to put the government policy into the financial market. In addition, there are lots of differences between the central bank and the policy bank. However, the most crucial difference in this paper is the central bank could not lend money directly to firms, it could only lend money to policy or commercial banks, but policy banks could give loans to firms directly. Rural is also an indicator variable equals one if the bank is located in a rural area. Tangibility is represented by (fixed assets/assets). Roa is the return on assets. List is a dummy variable equals to one if the bank is a listed bank. Cash Flow is the operating income before depreciation over the assets. Lend Central Bank is the credit lent from the central bank over the assets. As noted before, Bad loan is a measurement of the bank's health, I use Bad Loan $\times$ Crisis, and Bad Loan $\times$ Post to capture the different impacts of Bad Loan pre- and post-crisis.

Table 1.8 present the fixed effect estimation in equation (1.5) that provides an unbiased estimate of the bank lending channel coefficient. All regressions include fixed effects. Robust standard errors given below coefficient estimates are clustered at the bank-level. The results indicate a large bank lending channel: column 1 of table 1.8 shows that one percentage point increase of international commercial lending and bond as a fraction of assets could lead to a decrease of private bank loans by 4.725 percentage points more during financial crisis than before, and 10.385 percentage points more decrease in state-owned loans during financial crisis than before. Moreover, one percentage point increase of international borrowing would lead to the rise of private bank loans by 0.893 percentage points more after financial crisis than before, and 7.071 percentage points more increase in state-owned loans after financial crisis than before.

Column 3 of table 1.8 shows that one percentage point increase of loans lend to B-share, H-share and overseas-listed Chinese firms as a fraction of total loans could lead to a decrease of private bank loans by 5.149 percentage points more during financial crisis than before, and 9.831 percentage points more decrease in state-owned loans during financial crisis than before. Column 3 also shows that one percentage point increase of international borrowing would lead to the rise of private bank loans by 0.724 percentage points more after financial crisis than before, and 5.505 percentage points more increase in state-owned loans after financial crisis than before.

Column 5 of table 1.8 shows that one percentage point increase of the bank's trade settlements as a fraction of its total loans could lead to a decrease of private bank loans by 1.773 percentage points more during financial crisis than before, and 2.054 percentage points more decrease in state-owned loans during financial crisis than before. Furthermore, one percentage point increase of international borrowing would lead to the rise of private bank loans by 0.824 percentage points more after financial crisis than before, and 0.988 percentage points more increase in state-owned loans after financial crisis than before. All in all, Banks with higher international exposure tend to cut more lending during the financial crisis, state-owned bank loans are more pro-cyclical.

To show the robustness of our results, we rerun the above regressions using less restrictive specifications. Column 2, 4, 6 presents results for specifications, including the only firm fixed effects and year fixed effects. Importantly, results remain very similar both in terms of economic and statistical significance using either specification. Hence, also the economic magnitude of the impact described is very stable across different specifications.

Since Crisis and Post are two indicator variables, they only address the difference before, during, and after the financial crisis, and there is a very long-time period in my sample (16 years). Thus we need to plot the year-specific effects of these three potential channels.

I employ the following specification for bank i and firm j in year t for the year-specific effects for both state-owned and private banks.

$$L_{ijt} = \beta_1 Exposure_{i,t-1} \times Year\ Dummy_t + \gamma X_{i,t-1} + \lambda_{jt} + \epsilon_{ijt} \quad (1.9)$$

where i represents bank, j represents firm, t represents year, L_{ijt} is the log loan size, Year Dummy equals to one for each specific year, otherwise it equals to zero, $X_{i,t-1}$ is the bank-level control variables, λ_{jt} is the year-firm fixed effects.

Figure 1.5 plot the year-specific effects of the three exposure measurements state above on the level of the log loan volume during 2001-2016. The red line represents the year-specific effect of the state-owned banks, and the blue line represents the year-specific effect of the private banks, robust standard errors are clustered at the bank-level. All of the three figures show that the parallel trends of the year-specific log loan volume are divergent following the onset of the financial crisis in 2008, and they finally converged together when the year moves on.

The baseline result in this paper is consistent with the effect in Russia, Fungáčová et al. (2013) also find that bank ownership affected credit supply during the financial crisis and that the crisis led to an overall decrease in the credit supply, however, the direction is opposite. Fungáčová et al. (2013) suggest relative to Russian private banks, Russian state-controlled banks reduced their credit less, but in my paper compared with Chinese private banks, Chinese state-owned banks reduced their credit more.

Growth Rate Effect

Table 1.12 and 1.13 present the growth rate effect estimates from panel regressions explaining bank-firm-level yearly bank loans from 2001 to 2016. OECD growth rate could be represented as $\sum_{i=1}^{n} dlnGDP_i \times$ GDP share$_i$, i represents for the 35 OECD member countries. The conditional information set includes Cash Flow, Lend Central Bank, Tangibility, Roa, List, Bad Loan, and Bad Loan $\times$ OECD growth rate. Specifically speaking, Cash Flow is cash/assets, Lend Central Bank is log(the money borrowed from the central bank), Tangibility is tangible assets/assets, Roa is the return on assets. List is an indicator variable equal to one if the bank is a listed bank. As noted before, Bad loan is a measurement of the bank's balance sheet health, and I use Bad Loan $\times$ OECD growth rate to capture the different impacts of Bad Loan given different levels of the OECD growth rate.

Table 1.12 and 1.13 present the fixed effect estimation in equation (1.8) that provides an unbiased estimate of the bank lending channel coefficient. All regressions include fixed effects. Robust standard errors given below coefficient estimates are clustered at the bank-level. The results indicate a large bank lending channel: column 1 of table 1.12 shows that one percentage point increase of international commercial lending and bond as a fraction of assets could lead to a decline of the growth rate of bank loans by 0.379 percentage points when the OECD GDP growth rate is 0, and this decline is substantially offset and even reversed when OECD GDP growth rate increases.

Column 3 of table 1.12 shows that one percentage point increase of loans lend to B-share, H-share and overseas-listed Chinese firms as a fraction of total loans could lead to a decline of the growth rate of bank loans by 0.455 percentage points when the OECD GDP growth rate is 0, and this decline is substantially offset and even reversed when OECD GDP growth rate increases.

Column 5 of table 1.12 shows that one percentage point increase of loans lend to trade settlements as a fraction of total loans could lead to a decline of the growth rate of bank loans by 0.609 percentage points when the OECD GDP growth rate is 0, and this decline is substantially offset and even reversed when OECD GDP growth rate increases.

Figure 1.6 presents the average marginal bank credit effects of three different international exposure measurements. Panel A of figure 1.6 shows that the average marginal effects of international borrowing is positive when OECD GDP growth rate is higher than 1.192%, the average marginal effects is negative when the growth rate is less than 1.192%, the average marginal effects equals to zero when growth rate equals to 1.192%. The 95% confidence interval when the average marginal effects of international borrowing equals to 0 is [-0.2%,3.8%].

Panel B of figure 1.6 shows that the average marginal effects of foreign listed is positive when OECD GDP growth rate is higher than 1.750%, the average marginal effects is negative when the growth rate is less than 1.750%, the average marginal effects equals to zero when growth rate equals to 1.750%. The 95% confidence interval when the average marginal effects of foreign listed equals to 0 is [1.5%,2.3%].

Panel C of figure 1.6 shows that the average marginal effects of trade settlement is positive when OECD GDP growth rate is higher than 2.086%, the average marginal effects is negative when the growth rate is less than 2.086%, the average marginal effects equals to zero when growth rate equals to 2.086%. The 95% confidence interval when the average marginal effects of foreign listed equals to 0 is [1.2%,5.8%].

1.5.7 Using All Measures of Exposure Together

Since we have three different international exposure measurements: international borrowing, foreign listed, and trade settlement, while international borrowing could be regarded as a proxy in the bond market, foreign listed is a proxy in the stock market, and trade settlement is also a good measure in the goods market. Therefore, an interesting question could be asked, which effect dominates? The impact in the bond market, stock market, or goods market?

To answer this question, I employ the following specification for bank i and firm j in year t:

$$\begin{aligned} L_{ijt} \quad = \quad & \beta_1 International\ Borrowing_{i,t-1} + \beta_2 International\ Borrowing_{i,t-1} \times Crisis_t \\ & + \beta_3 International\ Borrowing_{i,t-1} \times Post_t + \beta_4 Foreign\ Listed_{i,t-1} \\ & + \beta_5 Foreign\ Listed_{i,t-1} \times Crisis_t + \beta_6 Foreign\ Listed_{i,t-1} \times Post_t \\ & + \beta_7 Trade\ Settlement_{i,t-1} + \beta_8 Trade\ Settlement_{i,t-1} \times Crisis_t \\ & + \beta_9 Trade\ Settlement_{i,t-1} \times Post_t + \gamma X_{i,t-1} + \lambda_{jt} + \epsilon_{ijt} \end{aligned} \tag{1.10}$$

where i represents bank, j represents firm, t represents year, L_{ijt} is the log loan size, $Crisis_t$ is an indicator variable equals to one for fiscal years 2008-2012, $Post_t$ is an indicator variable equals to one for fiscal years after 2012, $X_{i,t-1}$ is the bank-level control variables, λ_{jt} is the year-firm fixed effects. I use three foreign exposure measurements: international borrowing, foreign listed and trade settlement. International Borrowing$_{i,t-1}$ could be represented as (International bond$_{i,t-1}$+International commercial borrowing$_{i,t-1}$)/Asset$_{i,t-1}$. Foreign Listed$_{i,t-1}$ is (Loan gives to B and H-share and overseas listed Chinese firms$_{i,t-1}$/Total loan$_{i,t-1}$), Trade Settlement$_{i,t-1}$ is Trade settlements$_{i,t-1}$/ Total loan$_{i,t-1}$.

Table 1.10 presents the estimates from panel regressions in equation (1.10), we could find that the international borrowing in the bond market and the foreign listed in the stock market are more severely affected by the financial crisis compared with the trade settlement in the goods market.

For the growth rate effect of the bank-firm matched loans, I employ the following specification for bank i and firm j in year t:

$$\begin{aligned}\Delta L_{ijt} \quad = \quad & \beta_1 International\ Borrowing_{i,t-1} + \beta_2 International\ Borrowing_{i,t-1} \times gr^{OECD}_{Yt} \\ & + \beta_3 Foreign\ Listed_{i,t-1} + \beta_4 Foreign\ Listed_{i,t-1} \times gr^{OECD}_{Yt} \\ & + \beta_5 Trade\ Settlement_{i,t-1} + \beta_6 Trade\ Settlement_{i,t-1} \times gr^{OECD}_{Yt} \\ & + \gamma X_{i,t-1} + \lambda_{jt} + \epsilon_{ijt}\end{aligned} \tag{1.11}$$

where i represents bank, j represents firm, t represents year, ΔL_{ijt} is the growth rate of loan size, gr^{OECD}_{Yt} is the OECD GDP growth rate, $X_{i,t-1}$ is the bank-level control variables, λ_{jt} is the year-firm fixed effects. I also use the three foreign exposure measurements here: international borrowing, foreign listed and trade settlement. International Borrowing$_{i,t-1}$ could be represented as (International bond$_{i,t-1}$+International commercial borrowing$_{i,t-1}$)/Asset$_{i,t-1}$. Foreign Listed$_{i,t-1}$ is (Loan gives to B and H-share and overseas listed Chinese firms$_{i,t-1}$)/Total loan$_{i,t-1}$), Trade Settlement$_{i,t-1}$ is Trade settlements$_{i,t-1}$/

Total loan$_{i,t-1}$, $After_t$ is an indicator variable equal to one for fiscal years after 2008.

Table 1.14 shows the estimates from panel regressions in equation (1.11), we find the results in growth rate effect are quite similar to the results in level effect, the international borrowing in bond market and the foreign listed in stock market are more severely affected by the financial crisis compared with the trade settlement in goods market.

1.6 Firm-Level Effects of the Financial Crisis: Loans and Real Outcomes

We have seen that negative shocks to a bank's liquidity supply translate into a drop in its client firms' loans for both state-owned and private firms. However, such bank lending channels may not have any aggregate effect if firms can compensate for bank-specific loan losses by borrowing more from banks with greater liquidity. In this section, we discuss the firm-level effects of the financial crisis as well as the sovereign debt crisis.

1.6.1 Summary Statistics

Table 1.6 presents summary statistics for the firm-level variables in our data set. The average net debt over total assets is 3.8%; the average cash over total assets is 16.6%. There are around 4074 staffs in all the Chinese listed firms on average. The mean value for sales over assets, capital investments over assets, and operating income before depreciation over assets is 60.4%, 4%, and 3.2% respectively.

1.6.2 Empirical Methodology

I also utilize both the fixed effect and the GMM estimates of the firm borrowing channel to argue that I can provide conservative estimates of the impact of the liquidity and funding shock on firm-level outcomes such as a firm's net debt, cash (financial effects) as well as sales, capital investment, and employment (real effects).

Let Y_{jt} be a firm-level attribute of interest in period t (such as a firm's net debt, log cash, log sales, log capital investment, and log employment). The reduced form firm borrowing channel can be determined by estimating the following equation:

$$
\begin{aligned}
Y_{jt} \quad = \quad & \beta_1^F \overline{Exposure}_{j,t-1} + \beta_2^F \overline{Exposure}_{j,t-1} \times Crisis_t + \beta_3^F \overline{Exposure}_{j,t-1} \times Post_t \\
& + \beta_4^F \frac{CF_{j,t-1}}{Assets_{j,t-2}} + \beta_5^F \frac{Sales_{j,t-1}}{Assets_{j,t-2}} + \gamma X_{j,t-1} + \lambda_j + \mu_t + \eta_{jt}
\end{aligned}
\tag{1.12}
$$

$$
\begin{aligned}
Y_{jt} \quad = \quad & \beta_1^F \overline{Exposure}_{j,t-1} + \beta_2^F \overline{Exposure}_{j,t-1} \times gr_{Yt}^{OECD} + \beta_3^F \frac{CF_{j,t-1}}{Assets_{j,t-2}} \\
& + \beta_4^F \frac{Sales_{j,t-1}}{Assets_{j,t-2}} + \gamma X_{j,t-1} + \lambda_j + \mu_t + \eta_{jt}
\end{aligned}
\tag{1.13}
$$

where $\overline{Exposure}_{j,t-1}$ is the weighted aggregate exposure to international markets faced by firm j's banks in period $t-1$, which is measured by international borrowing, foreign listed and trade settlement. $Crisis_t$ equals to one only in the period of financial crisis (2008-2009) and sovereign debt crisis (2010-2012), and $Post_t$ equals to one for fiscal years after 2012, gr_{Yt}^{OECD} is the OECD GDP growth rate, $Cash\ Flow_{j,t-1}$ is calculated by (Operating income before depreciation$_{j,t-1}$/Assets$_{j,t-1}$). $X_{j,t-1}$ is the firm-level control variables. Firm-level controls includes Size, State-owned, Tangibility, Roa. Specifically speaking, Size is ln(assets), State-owned is an indicator variable equals to one if the firm is a state-owned firm. Tangibility is represented by (fixed assets/assets). Roa is the return on assets.

λ_j is the firm fixed effects, μ_t is the year fixed effects. If the firm borrowing channel completely insulates a firm from the bank lending channels, then the liquidity shocks should have no net impact on the firm's aggregate outcomes, i.e., β_2^F and β_3^F in (1.12) as well as β_2^F in (1.13) should be zero.

1.6.3 Baseline Results

Table 1.15, 1.16 and 1.17 present the fixed effect estimation in equation (1.12) that provides an unbiased estimate of the firm borrowing channel coefficient. All regressions include fixed effects. Robust standard errors given below coefficient estimates are clustered at the firm-level. The results indicate a large firm borrowing channel: column 1 of table 1.15 shows that one percentage point increase of bank's international borrowing aggregated at the firm's level could lead to a decrease of the firm's net debt by 1.172 percentage points more during the financial crisis than before, Moreover, one percentage point increase of the weighted aggregate international borrowing would lead to a decline of the firm's net debt by 1.333 percentage points more after the financial crisis than before. Column 2, 3, 4 of table 1.15 shows that one percentage point increase of bank's international borrowing aggregated at firm's level could lead to a decrease of firm's cash, employment and capital investment by 2.738, 0.0772 and 3.717 percentage points respectively more during financial crisis than before, Moreover, one percentage point increase of the weighted aggregate international borrowing would lead to an increase of firm's cash, employment and capital investment by 9.046, 9.911 and 19.78 percentage points more after financial crisis than before.

Column 1 of table 1.16 shows that one percentage point increase of the bank foreign listed aggregated at the firm's level could lead to a decrease of the firm's net debt by 0.372 percentage points more during the financial crisis than before. Moreover, one percentage point increase of the weighted aggregate foreign listed would lead to a decrease of the firm's net debt by 1.169 percentage points more after the financial crisis than before. Column 2, 3, 4 of table 1.16 shows that one percentage point increase of the bank foreign listed aggregated at firm's level could lead to a decrease of firm's cash, employment and capital investment by 0.398, 0.153 and 1.144 percentage points respectively more during financial crisis than before, Moreover, one percentage point increase of the weighted aggregate foreign listed would lead to a rise of the firm's cash, employment and capital investment by 7.318, 8.472 and 12.06 percentage points more after financial crisis than before.

Column 1 of table 1.17 shows that one percentage point increase of bank's trade settlement aggregated at the firm's level could lead to a decrease of the firm's net debt by 0.139 percentage points more during the financial crisis than before, Moreover, one percentage point increase of the weighted aggregate trade settlement would lead to a decrease in the firm's net debt by 0.171 percentage points more after the financial crisis than before. Column 2, 3, 4 of table 1.17 shows that one percentage point increase of bank's trade settlement aggregated at firm's level could lead to a decrease of firm's cash, employment and capital investment by 0.366, 0.486 and 1.558 percentage points respectively more during financial crisis than before, Moreover, one percentage point increase of the weighted aggregate trade settlement would lead to a rise of firm's cash, employment and capital investment by 0.877, 0.850 and 2.286 percentage points more after financial crisis than before.

Table 1.18, 1.19 and 1.20 present the fixed effect estimation in equation (1.13) that provides an unbiased estimate of the firm borrowing channel coefficient. All regressions include fixed effects. Robust standard errors given below coefficient estimates are clustered at the firm-level. The results indicate a large firm borrowing channel: column 1 of table 1.18 shows that one percentage point increase of bank's international borrowing aggregated at firm's level could lead to a decline of the firm's net debt by 5.362 percentage points when the OECD GDP growth rate is 0, and this decline is accelerated when OECD GDP growth rate increases. Column 2, 3 and 4 of table 1.18 shows that firms with higher weighted aggregate international borrowing will have lower cash stock, employment, and capital investment when there is a negative shock in OECD GDP growth rate.

Column 1 of table 1.19 shows that firms with higher bank foreign listed aggregated at the firm's level will have higher net debt when there is a negative shock in OECD GDP growth rate. Column 2, 3 and 4 of table 1.19 shows that firms with higher weighted aggregate foreign listed will have lower cash stock, employment, and capital investment when there is a negative shock in OECD GDP growth rate.

Column 1 of table 1.20 shows that firms with higher bank's trade settlement aggregated at the firm's level will have higher net debt when there is a negative shock in OECD GDP growth rate. Column 2, 3 and 4 of table 1.20 shows that firms with higher weighted aggregate foreign listed will have lower cash stock, employment, and capital investment when there is a negative shock in OECD GDP growth rate.

1.7 Conceptual Framework

This section presents a general equilibrium model of bank lending in the partially open financial market that sheds light on the mechanism behind the transmission channels that impact the bank lending channel. The proposed model predicts that when world interest rate increases, banks with higher exposure to the international market will reduce more bank credits through the balance sheet hit the channel. In addition, when firm return decreases, banks with higher international exposure will reduce more bank loans through the risk-limiting behavior channel.

This model is based on the framework of Holmstrom and Tirole (1997). The model has three types of agents: firms, banks, and depositors. All parties are risk-neutral and protected by limited liability so that no one could end up with a negative cash position.

Firms are run by entrepreneurs, who in the absence of proper incentives or outside monitoring may deliberately reduce the probability of success to enjoy a private benefit. This model formalizes this moral hazard problem by assuming that the entrepreneur can privately choose between two versions of the project as described in Figure 1.7. There are two periods, in the first period, financial contracts are signed, and investments are made. In the second period, investment returns are realized, and claims are settled. In period 2, the investment generates a verifiable, financial return equaling either 0 (failure) or R (success).

The function of banks is to monitor firms and thereby alleviate the moral hazard problem. In the case of bank lending, covenants are particularly typical and extensive. Covenants intend to reduce the firm's opportunity cost of being diligent. With that in mind, we assume that a bank could monitor a firm to prevent them from the bad project by cost γ. Thus, bank incentive compatibility condition could be written as

$$\lambda_H R_m - \gamma \geq \lambda_L R_m \tag{1.14}$$

1.7.1 Setup

There is a continuum of capitalists of measure K, who can become bankers or depositors, each capitalist is endowed with one unit of capital, and there is a continuum of potential entrepreneurs who can run firms. Moreover, in developing countries like China, being a banker needs tremendous verification procedures and license given by the central government. Thus we assume that the number of bankers is proportional to total capitalists and fixed at δ.

Bankers have two tasks in the economy: first, they channel capital from depositors to firms. Second, they monitor the firms they lend to at a cost to increase the probability that production is successful. As the suppliers of capital, bankers collect the gross return to capital R in the second period. There are two types of depositors: domestic depositors with the endogenous domestic deposit rate r — foreign depositors with the exogenous foreign deposit rate r^w. Depositors invest their endowments in banks and obtain the endogenously determined return r. Before each banker makes his investment decision, he learns about his efficiency as a banker. Each banker draws a foreign exposure s from a continuous distribution g(s) with support $s \in [0,1]$. Therefore, the financing cost could be written as

$$c(s) = sr^w + (1-s)r = r + s(r^w - r) \quad with \quad r^w < r. \tag{1.15}$$

The lower the financing cost draw the higher the capitalist's efficiency as a banker. The timeline is presented in figure 1.8.

Incentive compatibility requires that the banker's expected return under monitoring is higher than the return without monitoring, which results in the following condition:

$$\lambda_H Rz - \lambda_H c(s)(z - v(s)) - \gamma z \geq \lambda_L Rz - \lambda_L c(s)(z - v(s)) \tag{1.16}$$

where capital input per firm z is fixed. Each capitalist (endowed with one unit of capital) decides whether to become a banker or a depositor. γz is the total monitoring cost, and v is the banker's capital invested in the firm. Thus, in equilibrium (minimizing v, equation (1.16)) holds with equality, banker's own capital in each bank loan is

$$v(s) = (1 - \frac{R}{c(s)} + \frac{\gamma}{\Delta\lambda c(s)})z \tag{1.17}$$

The number of firms that one banker endowed with one unit of capital can monitor is

$$n(s) = 1/v(s) \tag{1.18}$$

Normalize z to 1, a banker with exposure s operates under the leverage:

$$\frac{\text{debt}}{\text{equity}} = \frac{\text{depositor capital}}{\text{bank capital}} = \frac{n(s)(1 - v(s))}{1} = \frac{1}{v(s)} - 1 \tag{1.19}$$

Banker's expected return per firm is

$$\lambda_L R - \lambda_L c(s)(1 - v(s)) = \frac{\gamma \lambda_L}{\Delta\lambda} \tag{1.20}$$

Total bank loans of the bank with exposure s could be represented as

$$I(s) = n(s) = \frac{c(s)}{c(s) - \beta} \tag{1.21}$$

where we assume $c(s) > \beta = R - \gamma/\Delta\lambda > 0$ for all s.

1.7.2 Equilibrium

The equilibrium condition in this model is the market for financial inter-mediation clears.

All active bankers together must intermediate the existing capital in the economy. The economy is endowed with domestic capital of measure K and foreign capital of measure K^*. A banker of type s can supply a measure of $n(s)$ firms with capital. Thus the market clears (financial intermediation) condition is

$$K + K^* = N \int_0^1 I(s) g(s) ds = N \int_s^1 \frac{c(s)}{c(s) - \beta} g(s) ds \tag{1.22}$$

where N is the number of banks. In the regulated financial markets, the total amount of the capital inflow is strictly controlled by the central government. Specifically speaking, in China, it is limited as a proportional to domestic capital; in this model, I assume $K^* = \rho K$. Furthermore, this model assumes that the number of bankers is proportional to total capitalists and fixed at δ, thus

$$(1+\rho)K = \delta K \int_0^1 \frac{c(s)}{c(s) - \beta} g(s) ds \tag{1.23}$$

Which could be simplified as

$$\frac{1+\rho-\delta}{\delta} = \int_0^1 \frac{\beta}{c(s) - \beta} g(s) ds \tag{1.24}$$

For simplicity, my model assumes $g(s)$ follows uniform distribution $s \sim U[0,1]$, thus $g(s) = 1$, equation (1.24) could be rewritten as

$$\frac{1+\rho-\delta}{\delta} = \beta \frac{\ln(r^w - \beta) - \ln(r - \beta)}{r^w - r} \tag{1.25}$$

Take derivative respect to r^w, we have

$$\frac{dr}{dr^w} = \frac{\frac{1+\rho-\delta}{\delta} - \frac{\beta}{r^w - \beta}}{\frac{1+\rho-\delta}{\delta} - \frac{\beta}{r-\beta}} \tag{1.26}$$

Proposition 1 *Assume r^w is large enough i.e.* $r^w > \frac{\beta(1+\rho)}{1+\rho-\delta}$, *then* $0 < \frac{dr}{dr^w} < 1$.

Proof. See appendix A.1. ■

Take derivative respect to R, we have

$$\frac{dr}{dR} = \frac{\frac{1+\rho-\delta}{\beta\delta}(r - r^w) + (\frac{\beta}{r^w-\beta} - \frac{\beta}{r-\beta})}{\frac{1+\rho-\delta}{\delta} - \frac{\beta}{r-\beta}} \tag{1.27}$$

Proposition 2 *Assume r^w is large enough i.e.* $\quad r^w > \frac{\beta(1+\rho)}{1+\rho-\delta}$ *and close to r, then* $0 < \frac{dr}{dR} < 1$.

Proof. See appendix A.2. ■

Proposition 3 *Channel I: Balance Sheet Hit*

first-order effect

$$\frac{\partial I}{\partial r^w} = \underbrace{-\frac{\beta s}{(c-\beta)^2}}_{Direct\ effect} \underbrace{-\frac{\beta(1-s)}{(c-\beta)^2}\frac{dr}{dr^w}}_{General\ equilibrium\ effect} \tag{1.28}$$

$\frac{\partial I}{\partial r^w} < 0$ *under the sufficient condition* $0 < \frac{dr}{dr^w} < 1$.

Asymmetric effect across banks

$$\frac{\partial I^2}{\partial r^w \partial s} = -\beta\frac{(1 - dr/dr^w)(c-\beta) + 2(r - r^w)(s + (1-s)dr/dr^w)}{(c-\beta)^3} \tag{1.29}$$

$\frac{\partial I^2}{\partial r^w \partial s} < 0$ *under the sufficient condition* $0 < \frac{dr}{dr^w} < 1$.

Proof. See appendix A.3. ■

Proposition 4 *Channel II: Risk-limiting Behavior*

first-order effect

$$\frac{\partial I}{\partial R} = \underbrace{\frac{c}{(c-\beta)^2}}_{Direct\ effect} \underbrace{-\frac{\beta(1-s)}{(c-\beta)^2}\frac{dr}{dR}}_{General\ equilibrium\ effect} \tag{1.30}$$

$\frac{\partial I}{\partial R} > 0$ *under the sufficient condition* $0 < \frac{dr}{dR} < 1$.

Asymmetric effect across banks

$$\frac{\partial I^2}{\partial R \partial s} = \frac{\beta dr/dR(c-\beta) + (r - r^w)(c + \beta - 2\beta(1-s)dr/dR)}{(c-\beta)^3} \tag{1.31}$$

$\frac{\partial I^2}{\partial R \partial s} > 0$ *under the sufficient condition* $0 < \frac{dr}{dR} < 1$.

Proof. See appendix A.4. ■

To conclude, this proposed model predicts that there are two channels through which the financial shocks potentially affected banks' lending decisions. Through the balance sheet hit the channel, the higher world interest rate is associated with lower bank loans, banks with higher exposure to the international market would cut lending more when the world interest rate increases. Furthermore, through the risk-limiting behavior channel, the lower firm return is associated with lower bank credits, banks with higher international exposure would cut lending more when the firm return decreases.

1.8 Robustness Analysis

This paper implements several checks to verify the robustness of my results.

1.8.1 Changing from Previous Year to Initial Year

In this part, I use the initial year instead of the previous year to measure the value of exposure to avoid the endogeneity problem of international exposure. Table 1.25 shows that the results are robust.

1.8.2 The Health of the Bank

Chodorow-Reich (2013) focus on the impact of the financial crisis associated with the bank balance sheet's health. In this subsection, I replicate the specification in Chodorow-Reich (2013) using Chinese data. For the level effect of the bank-firm matched loans, I employ the following specification for bank i and firm j in year t:

$$\begin{aligned} L_{ijt} = {} & \beta_1 Bad\ Loan_{i,t-1} + \beta_2 Bad\ Loan_{i,t-1} \times State\text{-}owned_i + \beta_3 Bad\ Loan_{i,t-1} \times Crisis_t \\ & + \beta_4 Bad\ Loan_{i,t-1} \times Crisis_t \times State\text{-}owned_i + \beta_5 Bad\ Loan_{i,t-1} \times Post_t \\ & + \beta_6 Bad\ Loan_{i,t-1} \times Post_t \times State\text{-}owned_i + \gamma X_{i,t-1} + \lambda_{jt} + \epsilon_{ijt} \end{aligned} \tag{1.32}$$

where i represents bank, j represents firm, and t represents year, L_{ijt} is the log loan size, Bad Loan$_{i,t-1}$ is (Subprime loan$_{i,t-1}$+Doubt loan$_{i,t-1}$+Loss loan$_{i,t-1}$)/Asset$_{i,t-1}$, $After_t$ is an indicator variable equal to one for fiscal years after 2008, State-owned is an indicator variable equal to one if the bank is a state-owned bank. $X_{i,t-1}$ represents the bank-level control variables, λ_{jt} is the year-firm fixed effects.

Table 1.27 presents the fixed effect estimation in equation (1.32) that provides an unbiased estimate of the bank lending channel coefficient. The results indicate a large bank lending channel: column 1 shows that one percentage point increase of bad loans as a fraction of assets could lead to a decrease of private bank loans by 14.85 percentage points more during the financial crisis than before, and 22.327 percentage points more decrease in state-owned loans during financial crisis than before. Moreover, one percentage point increase of international borrowing would lead to an increase of private bank loans by 10.31 percentage points more after financial crisis than before, and 30.43 percentage points more increase in state-owned loans after financial crisis than before.

Moreover, for the growth rate effect of the bank-firm matched loans, I employ the following specification for bank i and firm j in year t:

$$\begin{aligned}\Delta L_{ijt} &= \beta_1 Bad\ Loan_{i,t-1} + \beta_2 Bad\ Loan_{i,t-1} \times gr_{Yt}^{OECD} \\ &\quad + \gamma X_{i,t-1} + \lambda_{jt} + \epsilon_{ijt}\end{aligned} \tag{1.33}$$

where i represents bank, j represents firm, and t represents year, ΔL_{ijt} is the growth rate of loan size, Bad Loan$_{i,t-1}$ is (Subprime loan$_{i,t-1}$+Doubt loan$_{i,t-1}$+Loss loan$_{i,t-1}$)/Asset$_{i,t-1}$, gr_{Yt}^{OECD} is the OECD GDP growth rate, $X_{i,t-1}$ represents the bank-level control variables, λ_{jt} is the year-firm fixed effects.

Table 1.29 presents the fixed effect estimation in equation (1.33) that provides an unbiased estimate of the bank lending channel coefficient. The results indicate a large bank lending channel: column 1 of table 1.29 shows that one percentage point increase of bad loan as a fraction of assets could lead to a decline of the growth rate of bank loans by 5.785 percentage points when the OECD GDP growth rate is 0, and this decline is substantially offset and even reversed when OECD GDP growth rate increases.

Figure 1.6 shows that the average marginal effects of bad loan is positive when OECD GDP growth rate is higher than 2.940%, the average marginal effects is negative when the growth rate is less than 2.940%, the average marginal effects equals to zero when growth rate equals to 2.940%. The 95% confidence interval when the average marginal effects of the bad loan equals to 0 is [1.1%,6.0%].

1.8.3 Alternative Identification Strategy for Exposure Measurement

As a further robustness check, I provide one alternative definition for our critical explanatory variable, which is the exchange gains as a fraction of total income, and this alternative measurement could be constructed by

$$(Exchange/Income)_{i,t} = \frac{Exchange\ gains_{i,t}}{Total\ income_{i,t}} \tag{1.34}$$

China maintains a closed capital account, meaning companies, banks, and individuals could not move money in or out of the country except under strict rules. The People's Bank of China (PBOC) and State Administration of Foreign Exchange (SAFE) regulate the flow of foreign exchange in and out of the country and set exchange rates through a managed float system. The scheme was aimed at stopping the flight of foreign currency overseas and also preventing the inflow of foreign capital from affecting the Chinese economy.

Then, for the level effect of the bank-firm matched loans, I employ the following specification for bank i and firm j in year t:

$$
\begin{aligned}
L_{ijt} \quad = \quad & \beta_1(Exchange/Income)_{i,t-1} + \beta_2(Exchange/Income)_{i,t-1} \times State\text{-}owned_i \\
& + \beta_3(Exchange/Income)_{i,t-1} \times Crisis_t \\
& + \beta_4(Exchange/Income)_{i,t-1} \times Crisis_t \times State\text{-}owned_i \\
& + \beta_5(Exchange/Income)_{i,t-1} \times Post_t \\
& + \beta_6(Exchange/Income)_{i,t-1} \times Post_t \times State\text{-}owned_i + \gamma X_{i,t-1} + \lambda_{jt} + \epsilon_{ijt}
\end{aligned} \tag{1.35}
$$

where i represents bank, j represents firm, and t represents year, L_{ijt} is the log loan size, (Exchange/Income)$_{i,t-1}$ is Exchange gains$_{i,t-1}$/Total income$_{i,t-1}$, $After_t$ is an indicator variable equal to one for fiscal years after 2008, State-owned is an indicator variable equal to one if the bank is a state-owned bank. $X_{i,t-1}$ represents the bank-level control variables, λ_{jt} is the year-firm fixed effects.

Table 1.26 presents the results of the bank lending channel regression in equation (1.35). Column 1 shows that one percentage point increase of exchange income as a fraction of total income could lead to a decrease of private bank loans by 2.009 percentage points more during the financial crisis than before, and 8.297 percentage points more decrease in state-owned loans during the financial crisis than before. Column 1 also shows that one percentage point increase of international borrowing would lead to an increase of private bank loans by 1.641 percentage points more after financial crisis than before, and 7.250 percentage points more increase in state-owned loans after financial crisis than before.

Moreover, for the growth rate effect of the bank-firm matched loans, I employ the following specification for bank i and firm j in year t:

$$
\begin{aligned}
\Delta L_{ijt} \quad = \quad & \beta_1(Exchange/Income)_{i,t-1} + \beta_2(Exchange/Income)_{i,t-1} \times gr_{Yt}^{OECD} \\
& + \gamma X_{i,t-1} + \lambda_{jt} + \epsilon_{ijt}
\end{aligned} \tag{1.36}
$$

where i represents bank, j represents firm, and t represents year. ΔL_{ijt} is the growth rate of loan size, $(\text{Exchange/Income})_{i,t-1}$ is $\text{Exchange gains}_{i,t-1}/\text{Total income}_{i,t-1}$, gr^{OECD}_{Yt} is the OECD GDP growth rate, $X_{i,t-1}$ represents the bank-level control variables, λ_{jt} is the year-firm fixed effects.

Table 1.28 presents the results of the bank lending channel regression in equation (1.36). Column 1 of table 1.28 shows that one percentage point increase of exchange gains as a fraction of total income could lead to a decline of the growth rate of bank loans by 1.110 percentage points when the OECD GDP growth rate is 0, and this decline is substantially offset and even reversed when OECD GDP growth rate increases.

Figure 1.6 shows that the average marginal effects of exchange/income is positive when OECD GDP growth rate is greater than 1.961%, the average marginal effects is negative when the growth rate is less than 1.961%, the average marginal effects equals to zero when growth rate equals to 1.961%. The 95% confidence interval when the average marginal effects of exchange/income equals to 0 is [0.5%,5.5%].

1.8.4 Alternative Identification Strategy for Two Periods Model

Following the Duchin et al. (2010), I define the indicator variable $After_t$ equals to one for fiscal years after 2008 to equally divide the main sample period into the pre-crisis period (2001-2008) and the post-crisis period (2009-2016). The year 2001 is a particular year for China; this is the year China joined WTO. Moreover, we could regard this year as the first year China opens its door after closing it for thousands of years.

For the level effect of the bank-firm matched loans, I employ the following specification for bank i and firm j in year t:

$$\begin{aligned} L_{ijt} = \; & \beta_1 Exposure_{i,t-1} + \beta_2 Exposure_{i,t-1} \times State\text{-}owned_i + \beta_3 Exposure_{i,t-1} \times After_t \\ & + \beta_4 Exposure_{i,t-1} \times After_t \times State\text{-}owned_i + \gamma X_{i,t-1} + \lambda_{jt} + \epsilon_{ijt} \end{aligned} \tag{1.37}$$

where i represents bank, j represents firm, t represents year, L_{ijt} is the log loan size, $After_t$ is an indicator variable equals to one for fiscal years after 2008, $State\_owned_i$ is also an indicator variable equals to one if the bank is a state-owned bank, and it equals to zero if the bank is a privately-owned bank, $X_{i,t-1}$ is the bank-level control variables, λ_{jt} is the year-firm fixed effects. I use three foreign exposure measurements: international borrowing, foreign listed and trade settlement. International Borrowing$_{i,t-1}$ could be represented as (International bond$_{i,t-1}$+International commercial borrowing$_{i,t-1}$)/Asset$_{i,t-1}$. Foreign Listed$_{i,t-1}$ is (Loan gives to B and H-share and overseas listed Chinese firms$_{i,t-1}$/Total loan$_{i,t-1}$), Trade Settlement$_{i,t-1}$ is Trade settlements$_{i,t-1}$/ Total loan$_{i,t-1}$.

The term (Exposure$_{i,t-1}\times$ State-owned$_i$) is used to control for the different effects of state-owned and private banks on the dimension of the international exposure measurement. If banks with higher exposure to the international markets cut lending more following the recent financial crisis, higher exposure banks would reduce their loan amount after the financial crisis, that is, I expect β_3 in Eq. (1.37) to be negative. Moreover, if the Chinese government tried to close the door to avoid the substantial loss from the outside world following the financial crisis, the state-owned banks should be affected more compared with the privately-owned banks. Therefore, I expect β_4 in Eq. (1.37) to be negative.

Table 1.30 presents estimates from panel regressions explaining bank-firm-level yearly bank loans from 2001 to 2016. The conditional information set includes Size, State-owned, Policy, Rural, List, Tangibility, Roa, Cash Flow, Lend Central Bank, Bad Loan, Bad Loan $\times$ Crisis, and Bad Loan $\times$ Post.

Table 1.30 presents the fixed effect estimation in equation (1.37) that provides an unbiased estimate of the bank lending channel coefficient. All regressions include fixed effects. Robust standard errors given below coefficient estimates are clustered at the bank-level. The results indicate a large bank lending channel: column 1 of table 1.30 shows that one percentage point increase of international commercial lending and bond as a fraction of assets could lead to a decrease of private bank loans by 0.839 percentage points more after the onset of the financial crisis, and 3.111 percentage points more decrease in state-owned loans after the onset of the financial crisis.

Column 3 of table 1.30 shows that one percentage point increase of loans lend to B-share, H-share and overseas-listed Chinese firms as a fraction of total loans could lead to a decrease of private bank loans by 4.012 percentage points more after the onset of the financial crisis, and 4.519 percentage points more decrease in state-owned loans after the beginning of the financial crisis.

Column 5 of table 1.30 shows that one percentage point increase of the bank's trade settlements as a fraction of its total loans could lead to a decrease of private bank loans by 0.121 percentage points more after the onset of the financial crisis, and 1.723 percentage points more decrease in state-owned loans after the beginning of the financial crisis.

To show the robustness of our results, we rerun the above regressions using less restrictive specifications. Column 2, 4, and 6 present results for specifications, including the only firm fixed effects and year fixed effects. Importantly, results remain very similar both in terms of economic and statistical significance using either specification. Hence, also the economic magnitude of the impact described is very stable across different specifications.

1.9 Conclusion

In this paper, I showed that the credit crunch following the financial crisis in the international markets did impact Chinese bank lending channel. In particular, I found that banks with higher exposure to the global markets cut lending more following the recent financial crisis, and state-owned bank loans are more pro-cyclical compared with private bank loans. I also found that banks with higher exposure to the international markets cut lending more when there was a negative shock in the OECD GDP growth rate.

Moreover, I compared the effect from the international borrowing in bonds market, the foreign listed in the stocks market, as well as the trade settlement in the goods market, and found that the effects of the financial crisis on bonds and stocks market were more significant than the effect on goods market.

Furthermore, I found that the financial shock from the international markets also impacts Chinese firm borrowing channel. Explicitly speaking, firms with higher weighted aggregate exposure to the global markets through the banks had the lower net debt, cash, employment and capital investment during the financial crisis, and firms with higher weighted aggregate exposure to the international markets had higher net debt and lower cash, employment and capital investment when there was a negative shock in OECD GDP growth rate.

In my analysis, I took advantage of a novel data set, covering a large number of small, privately owned and rural banks, with information on firm-bank relationships. This paper is the first paper to analyze the impact of the financial crisis and the sovereign debt crisis on both the Chinese bank lending channel and the firm borrowing channel. To further disclose the bank loan effects of the financial shocks, one may redo my analysis by industries to see which industry is affected most by the exogenous financial shocks.

My findings foster the understanding of the unfolding of the impact of the financial crisis on the Chinese bank lending channel and firm borrowing channel. My results also indicate that both markets oriented and government-oriented effects are critical in the financial market in China.

1.10 Tables

Table 1.1 Financial Development: China and US (USD Trillion) (2016)

Sector	Bank Credit	Stock	Fixed Income	Insurance	Investment Funds
Size (China)	15.45	7.32	10.43	2.19	1.32
Size (US)	12.44	27.35	39.36	8.46	19.20
% GDP (China)	137.95%	64.30%	93.12%	19.55%	11.76%
% GDP (US)	67.00%	147.40%	211.95%	45.56%	103.39%

Table 1.2 Variable Definitions

Variable	Definition
Bank-level	
Dependent Variables (winsorized at the 1% level)	
L	Natural logarithm of bank loans
ΔL	ln(bank loan$_{t+1}$−bank loan$_t$)
Key Explanatory Variables (winsorized at the 1% level)	
$After_t$	Dummy equal to one if the year is after 2008
$Crisis_t$	Dummy equal to one if the year is 2008-2012
gr^{OECD}_{Yt}	OECD GDP growth rate
$Bad\ Loan_{i,t}$	(Subprime loan$_{i,t}$+Doubt loan$_{i,t}$+Loss loan$_{i,t}$)/Asset$_{i,t}$
$International\ Borrowing_{i,t}$	(International bond$_{i,t}$+International commercial borrowing$_{i,t}$)/Asset$_{i,t}$
$Foreign\ Listed_{i,t}$	(Loan to B-, H-share & overseas listed Chinese firms$_{i,t}$)/Total loan$_{i,t}$
$Trade\ Settlement_{i,t}$	Trade settlements$_{i,t}$/ Total loan$_{i,t}$
$(Exchange/Income)_{i,t}$	Exchange gains$_{i,t}$/Total income$_{i,t}$
Control Variables (winsorized at the 1% level)	
$State\text{-}owned_i$	Dummy equals to one if the bank is a state-owned bank
$Policy_i$	Dummy equals to one if the bank is a policy bank
$Rural_i$	Dummy equals to one if the bank is a rural bank
$List_{i,t}$	Dummy equals to one if the bank is a listed bank
$Size_{i,t}$	ln(Assets$_{i,t}$)
$Roa_{i,t}$	Net income$_{i,t}$/((Assets$_{i,t-1}$+Assets$_{i,t}$)/2)
$Tangibility_{i,t}$	Fixed assets$_{i,t}$/Assets$_{i,t}$
$Cash\ Flow_{i,t}$	Operating Income Before Depreciation$_{i,t}$/((Assets$_{i,t-1}$+Assets$_{i,t}$)/2)
$Lend\ Central\ Bank_{i,t}$	Debt lent from central bank$_{i,t}$/Assets$_{i,t}$
Firm-level	
Dependent Variables (winsorized at the 1% level)	
$Net\ Debt$	(Current+Non-Current Liabilities−Cash)/(Total Assets)
$\Delta Cash$	(Cash$_{t+1}$−Cash$_t$)/(Total Asset$_t$)
$Employment\ Growth$	ln(Employment$_t$)−ln(Employment$_{t-1}$)
$CAPX$	(Fixed Assets$_{t+1}$−Fixed Assets$_t$+Depreciation$_t$)/Assets$_t$, set to 0 if negative
Key Explanatory Variables (winsorized at the 1% level)	
$\overline{TL}_{j,t-1}$	$\frac{1}{I}\sum_{i=1}^{I}$Total Loan$_{ijt}$
Control Variables (winsorized at the 1% level)	
$Sales\ Growth$	ln(Sales$_t$)−ln(Sales$_{t-1}$)
$Cash\ Flow_{i,t}$	Operating Income Before Depreciation$_{i,t}$/Assets$_{i,t-1}$

Table 1.3 Summary Statistics (Firm-bank Pairwise)

	State-owned Banks			Private Banks			All Banks		
	mean	sd	count	mean	sd	count	mean	sd	count
loan	502.619	3,717.35	9,840	264.614	780.71	15,018	358.828	2,418.99	24,858
lnloan	18.842	1.38	9,840	18.548	1.22	15,018	18.664	1.29	24,858
dlnloan	0.114	0.91	4,822	0.123	0.77	6,655	0.119	0.83	11,477
Observations	9840			15018			24858		

Notes. This table presents descriptive statistics of firm-bank pairwise dependent variables split into state-owned and private banks. State-owned is an indicator variable equal to one if the bank is a state-owned bank. The sample consists of all firms that are listed in the A-share, B-share, H-share, and oversea stocks market.

Table 1.4 Summary Statistics (Bank-level I)

	State-owned Banks			Private Banks			All Banks		
	mean	sd	count	mean	sd	count	mean	sd	count
Bad Loan	0.029	0.05	101	0.007	0.01	731	0.010	0.02	832
Exchange/Income	0.028	0.08	122	0.008	0.03	1,010	0.010	0.04	1,132
International Borrowing	0.009	0.03	99	0.002	0.01	749	0.002	0.01	848
Trade Settlement	0.086	0.34	100	0.000	0.01	820	0.010	0.12	920
Foreign Listed	0.060	0.15	110	0.052	0.21	617	0.053	0.20	727
Size	28.993	1.24	107	25.728	1.34	772	26.126	1.70	879
Profit	23.676	2.20	107	20.925	1.46	772	21.260	1.81	879
Cash Flow	0.095	0.06	107	0.150	0.04	772	0.143	0.05	879
Roa	0.012	0.01	93	0.011	0.00	568	0.011	0.01	661
List	0.402	0.49	122	0.126	0.33	994	0.156	0.36	1,116
Deposit	28.367	1.83	103	25.333	1.32	817	25.673	1.69	920
EBITDA	25.124	1.71	107	22.118	1.34	773	22.483	1.70	880
Tangibility	0.009	0.01	107	0.007	0.01	772	0.007	0.01	879
Lend Central Bank	0.022	0.06	122	0.002	0.01	1,010	0.004	0.02	1,132
Policy	0.344	0.48	122	0.000	0.00	1,010	0.037	0.19	1,132
Rural	0.000	0.00	122	0.139	0.35	1,010	0.124	0.33	1,132
Observations	122			1010			1132		

Notes. This table presents descriptive statistics of bank-level explanatory variables split into state-owned and private banks. The sample consists of all banks that are located in China.

Table 1.5 Summary Statistics (Bank-level II)

	2001-2008			2009-2012			2013-2016		
	mean	sd	count	mean	sd	count	mean	sd	count
Bad Loan	0.037	0.04	113	0.010	0.02	832	0.005	0.00	466
Exchange/Income	0.004	0.02	227	0.003	0.04	1,132	0.004	0.04	552
International Borrowing	0.010	0.03	111	0.002	0.01	848	0.001	0.00	489
Foreign Listed	0.050	0.16	105	0.053	0.20	727	0.056	0.23	431
Trade Settlement	0.043	0.30	134	0.010	0.12	920	0.003	0.02	512
Size	26.914	1.54	120	26.126	1.70	879	26.018	1.64	495
Profit	21.398	1.85	119	21.260	1.81	879	21.288	1.76	496
Cash	24.514	1.59	120	24.087	1.61	879	24.026	1.54	495
Roa	0.008	0.00	67	0.011	0.01	661	0.011	0.01	390
List	0.232	0.42	211	0.156	0.36	1,116	0.116	0.32	552
Deposit	26.547	1.74	134	25.673	1.69	920	25.530	1.57	517
Cash Flow	0.023	0.01	120	0.027	0.01	879	0.029	0.01	495
Tangibility	0.010	0.01	120	0.007	0.01	879	0.007	0.01	495
Lend Central Bank	0.006	0.04	227	0.004	0.02	1,132	0.005	0.01	552
State-owned	0.256	0.44	227	0.108	0.31	1,132	0.058	0.23	552
Policy	0.079	0.27	227	0.037	0.19	1,132	0.022	0.15	552
Rural	0.026	0.16	227	0.124	0.33	1,132	0.168	0.37	552
Observations	227			1132			552		

Notes. This table presents descriptive statistics of bank-level explanatory variables split into different time periods: 2001-2008, 2009-2012, and 2013-2016. The sample consists of all banks that are located in China.

Table 1.6 Summary Statistics (Firm-level)

Variable	Mean	Std. Dev.	N
Net Debt	0.038	0.206	7213
Cash	0.166	0.114	6245
Employment	4073.728	9104.401	7288
Sales	0.604	0.521	7295
Capital Investment	0.04	0.361	7285
Cash Flow	0.032	0.054	7221

Notes. This table presents descriptive statistics of firm-level explanatory variables. The sample consists of all listed Chinese firms.

Table 1.7 Preliminary Test

	(1) Demean (International Borrowing)	(2) Demean (Foreign Listed)	(3) Demean (Trade Settlement)
Crisis	0.000993	-0.000576	0.00485
	(0.0174)	(0.0220)	(0.0139)
Post	-0.000519	-0.00376	0.00375
	(0.0150)	(0.0166)	(0.0150)
Observations	18396	23202	22150
R^2	0.101	0.071	0.164
Firm Fixed Effect	YES	YES	YES
Clusters at Bank Level	132	177	170

Notes. This table presents the results of a preliminary test regression. The unit of observation is a bank-year. The dependent variable is the exposure measurement. All variables are defined in Table 1.2. All regressions include fixed effects. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Robust standard errors given below coefficient estimates are clustered at the bank-level.

Table 1.8 Bank Lending Channel (Three Periods Level Effect)

Log(Loan)	(1) Exposure (International Borrowing)	(2) Exposure (International Borrowing)	(3) Exposure (Foreign Listed)	(4) Exposure (Foreign Listed)	(5) Exposure (Trade Settlement)	(6) Exposure (Trade Settlement)
Bad Loan	-1.289***	-0.903**	-1.532	-1.720*	-1.504***	-1.291**
	(0.471)	(0.377)	(0.957)	(0.910)	(0.506)	(0.518)
Bad Loan x Post	-2.051	3.273	-3.644	2.008	12.31**	11.37**
	(12.20)	(11.18)	(6.567)	(6.183)	(5.934)	(5.524)
Bad Loan x Crisis	3.514	-0.0869	-1.148	-4.178	-16.92*	-20.64**
	(7.711)	(6.406)	(5.275)	(5.253)	(8.916)	(8.652)
Exposure	-1.827*	-3.619***	0.501	-0.639	0.745	1.361**
	(1.021)	(1.185)	(1.714)	(1.635)	(0.752)	(0.648)
Expo. x State-owned	2.563***	4.400***	5.828***	7.793***	1.513***	1.752***
	(0.954)	(1.160)	(1.885)	(1.802)	(0.344)	(0.325)
Exposure x Crisis	-4.725***	-1.942	-5.149***	-6.046***	-1.773**	-1.820***
	(1.541)	(1.379)	(1.221)	(1.168)	(0.857)	(0.612)
Expo. x Crisis x State-owned	-5.660*	-8.028***	-4.682**	-6.159***	-0.281	-0.381
	(3.347)	(2.978)	(1.999)	(1.908)	(0.802)	(0.489)
Expo. x Post	0.893	1.695	0.724	0.355	0.824	0.0179
	(1.104)	(1.300)	(1.729)	(1.647)	(1.101)	(0.864)
Expo. x Post x State-owned	6.178***	2.909**	4.781**	6.816***	0.164***	0.164***
	(1.266)	(1.458)	(1.888)	(1.805)	(0.0587)	(0.0537)
Observations	13934	13934	18895	18895	17222	17222
R^2	0.686	0.593	0.669	0.575	0.670	0.575
Bank Controls	YES	YES	YES	YES	YES	YES
Firm Fixed Effect	NO	YES	NO	YES	NO	YES
Year Fixed Effect	NO	YES	NO	YES	NO	YES
Firm-Year Fixed Effect	YES	NO	YES	NO	YES	NO
Clusters at Bank Level	132	132	177	177	170	170

Notes. This table presents the results of a Chinese bank lending channel regression. The unit of observation is a bank-firm-year. The dependent variable is the level of log loan volume. Bank Controls includes Size, Policy, Rural, Tangibility, Roa, Cash Flow, Lend Central Bank. All variables are defined in Table 1.2. All regressions include fixed effects. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Robust standard errors given below coefficient estimates are clustered at the bank-level.

Table 1.9 Bank Lending Channel (Three Periods Growth Rate Effect)

dLog(Loan)	(1) Exposure (International Borrowing)	(2) Exposure (International Borrowing)	(3) Exposure (Foreign Listed)	(4) Exposure (Foreign Listed)	(5) Exposure (Trade Settlement)	(6) Exposure (Trade Settlement)
Bad Loan	1.095	0.900	0.136	0.869	0.589	1.098
	(1.420)	(1.477)	(1.811)	(1.083)	(1.379)	(1.321)
Exposure	-0.584	-0.748	-1.141	-0.853	-3.422	-3.764**
	(0.616)	(0.752)	(3.145)	(1.573)	(2.205)	(1.777)
Exposure x Crisis	-0.970	-0.645	-1.192	-0.839	-0.421	-0.0993
	(0.692)	(0.778)	(3.201)	(1.540)	(0.458)	(0.430)
Expo. x Crisis x State-owned	-4.379*	-3.024	-1.017	-0.664	-4.802**	-4.720**
	(2.565)	(1.977)	(0.773)	(1.054)	(2.407)	(1.956)
Expo. x Post	0.581	1.468	1.150	0.872	0.0506	0.0265
	(0.972)	(1.092)	(3.149)	(1.582)	(0.0433)	(0.0294)
Expo. x Post x State-owned	3.030***	2.592**	0.221*	0.314**	3.996*	4.439**
	(1.077)	(1.023)	(0.112)	(0.132)	(2.210)	(1.753)
Observations	6903	6903	9212	9212	5818	5818
R^2	0.426	0.198	0.390	0.184	0.417	0.180
Bank Controls	YES	YES	YES	YES	YES	YES
Firm Fixed Effect	NO	YES	NO	YES	NO	YES
Year Fixed Effect	NO	YES	NO	YES	NO	YES
Firm-Year Fixed Effect	YES	NO	YES	NO	YES	NO
Clusters at Bank Level	109	109	135	135	103	103

Notes. This table presents the results of a Chinese bank lending channel regression. The unit of observation is a bank-firm-year. The dependent variable is the growth rate of the loan volume. Bank Controls includes Size, State-owned, Policy, Rural, Tangibility, Roa, Cash Flow, Lend Central Bank. All variables are defined in Table 1.2. All regressions include fixed effects. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Robust standard errors given below coefficient estimates are clustered at the bank-level.

Table 1.10 Bank Lending Channel (Using All Measures of Exposure Together) (Three Periods Level Effect)

	(1)	(2)	(3)	(4)
	lnloan	lnloan	lnloan	lnloan
Bad Loan	-1.419**	-1.827**	-1.131*	-1.708**
	(0.639)	(0.759)	(0.677)	(0.774)
International Borrowing	1.245***	-0.474	0.966	-0.661
	(0.467)	(0.723)	(0.700)	(0.832)
Foreign Listed	3.026***	0.113**	3.449***	0.129**
	(1.067)	(0.0471)	(1.226)	(0.0515)
Trade Settlement	-2.065	0.545	-1.390	0.486
	(1.485)	(0.745)	(1.676)	(0.653)
International Borrowing x Crisis	-6.266***		-4.906***	
	(1.142)		(1.286)	
Foreign Listed x Crisis	-3.515***		-3.931***	
	(1.112)		(1.324)	
Trade Settlement x Crisis	-0.694		-0.291	
	(1.590)		(1.674)	
International Borrowing x Post	1.629*		2.307***	
	(0.884)		(0.811)	
Foreign Listed x Post	2.882***		3.285***	
	(1.064)		(1.228)	
Trade Settlement x Post	3.380**		2.418	
	(1.402)		(1.561)	
Observations	11987	11987	11987	11987
R^2	0.686	0.684	0.590	0.588
Bank Controls	YES	YES	YES	YES
Firm Fixed Effect	NO	NO	YES	YES
Year Fixed Effect	NO	NO	YES	YES
Firm-Year Fixed Effect	YES	YES	NO	NO
Clusters at Bank Level	128	128	128	128

Notes. This table presents the results of a Chinese bank lending channel regression. The unit of observation is a bank-firm-year. The dependent variable is the level of log loan volume. Bank Controls includes Size, State-owned, Policy, Rural, Tangibility, Roa, Cash Flow, Lend Central Bank. All variables are defined in Table 1.2. All regressions include fixed effects. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Robust standard errors given below coefficient estimates are clustered at the bank-level.

Table 1.11 Bank Lending Channel (OECD GDP Growth Rate)

Log(Loan)	(1) Exposure (International Borrowing)	(2) Exposure (International Borrowing)	(3) Exposure (Foreign Listed)	(4) Exposure (Foreign Listed)	(5) Exposure (Trade Settlement)	(6) Exposure (Trade Settlement)
Bad Loan	-1.352	-1.410	-2.049**	-1.151	-2.095*	-2.335
	(1.351)	(1.854)	(0.960)	(1.222)	(1.080)	(1.861)
Bad Loan x OECD	0.0523	0.392	0.129	0.344	0.120	0.432
	(0.376)	(0.605)	(0.267)	(0.581)	(0.343)	(0.571)
Exposure	-2.194***	-2.515***	-0.252	-0.0503	-0.877	-1.243**
	(0.509)	(0.514)	(0.200)	(0.172)	(0.650)	(0.622)
Exposure x OECD growth rate	0.0139	0.0115	0.0775	0.0376	1.204***	1.333***
	(0.163)	(0.175)	(0.0900)	(0.0771)	(0.324)	(0.316)
Exposure x OECD x State-owned	1.659***	1.883***	0.472***	0.529***	1.073***	1.106***
	(0.360)	(0.359)	(0.171)	(0.201)	(0.190)	(0.188)
Observations	15503	15503	20528	20528	19144	19144
R^2	0.623	0.593	0.606	0.574	0.607	0.573
Bank Controls	YES	YES	YES	YES	YES	YES
Firm Fixed Effect	NO	YES	NO	YES	NO	YES
Year Fixed Effect	NO	YES	NO	YES	NO	YES
Firm-Year Fixed Effect	YES	NO	YES	NO	YES	NO
Clusters at Bank Level	132	132	177	177	170	170

Notes. This table presents the results of a Chinese bank lending channel regression. The unit of observation is a bank-firm-year. The dependent variable is the level of log loan volume. Bank Controls includes Size, Policy, Rural, Tangibility, Roa, Cash Flow, Lend Central Bank. All variables are defined in Table 1.2. All regressions include fixed effects. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Robust standard errors given below coefficient estimates are clustered at the bank-level.

Table 1.12 Bank Lending Channel (Growth Rate Effect)

dLog(Loan)	(1) Exposure (International Borrowing)	(2) Exposure (International Borrowing)	(3) Exposure (Foreign Listed)	(4) Exposure (Foreign Listed)	(5) Exposure (Trade Settlement)	(6) Exposure (Trade Settlement)
Bad Loan	-4.877**	-4.720	-5.479**	-3.509	-6.962**	-3.553
	(2.155)	(3.317)	(2.372)	(3.213)	(2.866)	(3.521)
Bad Loan x OECD growth rate	1.851**	2.615**	1.413**	1.766	2.544**	1.874
	(0.772)	(1.070)	(0.687)	(1.122)	(1.033)	(1.239)
Exposure	-0.379	-0.714*	-0.424***	-0.317**	-0.609	-0.535
	(0.290)	(0.426)	(0.138)	(0.154)	(0.375)	(0.454)
Exposure x OECD growth rate	0.318***	0.505***	0.234***	0.207**	0.292**	0.377**
	(0.0854)	(0.137)	(0.0812)	(0.0925)	(0.133)	(0.167)
Observations	7627	7627	9968	9968	8909	8909
R^2	0.195	0.180	0.182	0.167	0.169	0.151
Bank Controls	YES	YES	YES	YES	YES	YES
Firm Fixed Effect	NO	YES	NO	YES	NO	YES
Year Fixed Effect	NO	YES	NO	YES	NO	YES
Firm-Year Fixed Effect	YES	NO	YES	NO	YES	NO
Clusters at Bank Level	118	118	141	141	133	133

Notes. This table presents the results of a Chinese bank lending channel regression. The unit of observation is a bank-firm-year. The dependent variable is the change in log loan volume. Bank Controls includes Size, Policy, Rural, Tangibility, Roa, Cash Flow, Lend Central Bank. All variables are defined in Table 1.2. All regressions include fixed effects. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Robust standard errors given below coefficient estimates are clustered at the bank-level.

Table 1.13 Bank Lending Channel (Growth Rate Effect) (Continue...)

dLog(Loan)	(1) Exposure (International Borrowing)	(2) Exposure (Foreign Listed)	(3) Exposure (Trade Settlement)
OECD GDP growth rate	0.0271**	0.0242*	0.0494***
	(0.0119)	(0.0145)	(0.0169)
Bad Loan	-1.147	-1.348	-0.335
	(1.052)	(2.591)	(1.642)
Bad Loan x OECD growth rate	0.503	0.403	0.560
	(1.018)	(0.872)	(1.041)
Exposure	-0.255	-0.357**	-0.840**
	(0.254)	(0.153)	(0.403)
Exposure x OECD growth rate	0.187*	0.181**	0.474**
	(0.106)	(0.0920)	(0.194)
Observations	7627	9968	8909
R^2	0.171	0.156	0.141
Bank Controls	YES	YES	YES
Firm Fixed Effect	YES	YES	YES
Year Fixed Effect	NO	NO	NO
Firm-Year Fixed Effect	NO	NO	NO
Clusters at Bank Level	118	141	133

Notes. This table presents the results of a Chinese bank lending channel regression. The unit of observation is a bank-firm-year. The dependent variable is the change in log loan volume. Bank Controls includes Size, Policy, Rural, Tangibility, Roa, Cash Flow, Lend Central Bank. All variables are defined in Table 1.2. All regressions include fixed effects. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Robust standard errors given below coefficient estimates are clustered at the bank-level.

Table 1.14 Bank Lending Channel (Using All Measures of Exposure Together) (Growth Rate Effect)

	(1) dlnloan	(2) dlnloan	(3) dlnloan	(4) dlnloan
Bad Loan	-1.054	-1.039	-3.749	-2.059
	(2.192)	(2.351)	(3.099)	(3.408)
Bad Loan x OECD growth rate	1.416*	1.352	2.071**	1.328
	(0.795)	(0.865)	(0.854)	(0.993)
International Borrowing	-2.858***	-0.753	-3.361***	-1.276
	(1.042)	(0.640)	(1.232)	(0.909)
Foreign Listed	-0.810**	-0.0798	-0.778***	-0.0690
	(0.310)	(0.0711)	(0.281)	(0.0700)
Trade Settlement	-2.788***	-2.705***	-3.263***	-3.062***
	(0.886)	(0.927)	(0.587)	(0.669)
International Borrowing x OECD growth rate	1.572***		1.535***	
	(0.408)		(0.380)	
Foreign Listed x OECD growth rate	0.473**		0.456***	
	(0.203)		(0.173)	
Trade Settlement x OECD growth rate	0.115		0.239	
	(0.180)		(0.231)	
Observations	6543	6543	6543	6543
R^2	0.178	0.177	0.161	0.160
Bank Controls	YES	YES	YES	YES
Firm Fixed Effect	NO	NO	YES	YES
Year Fixed Effect	NO	NO	YES	YES
Firm-Year Fixed Effect	YES	YES	NO	NO
Clusters at Bank Level	113	113	113	113

Notes. This table presents the results of a Chinese bank lending channel regression. The unit of observation is a bank-firm-year. The dependent variable is the change in log loan volume. Bank Controls includes Size, Policy, Rural, Tangibility, Roa, Cash Flow, Lend Central Bank. All variables are defined in Table 1.2. All regressions include fixed effects. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Robust standard errors given below coefficient estimates are clustered at the bank-level.

Table 1.15 Firm Borrowing Channel (Three Periods Effect)

	(1)	(2)	(3)	(4)
	Net Debt	Cash	Employment	Capital Investment
$\overline{\textit{International Borrowing}}$	0.535	-0.201	-2.875	-13.51*
	(0.325)	(2.609)	(1.517)	(5.665)
$\overline{\textit{International Borrowing}}$ x Crisis	-1.172*	-2.738	-0.0772	-3.717
	(0.493)	(3.813)	(2.283)	(8.466)
$\overline{\textit{International Borrowing}}$ x Post	-1.333***	9.046***	9.911***	19.78***
	(0.397)	(3.261)	(1.845)	(7.005)
Sales	0.00739*	0.174***	0.0949***	0.0421
	(0.00375)	(0.0282)	(0.0175)	(0.0772)
Cash Flow	-0.670***	1.851***	0.914***	0.958
	(0.0440)	(0.317)	(0.203)	(0.913)
Observations	6841	7060	7053	4197
R^2	0.751	0.766	0.860	0.640
Firm Controls	YES	YES	YES	YES
Firm Fixed Effect	YES	YES	YES	YES
Year Fixed Effect	YES	YES	YES	YES
Clusters at Firm Level	1967	1992	1991	1657

Notes. This table presents the results of a Chinese firm borrowing channel regression. The unit of observation is a firm-year. The dependent variable is the net debt, cash/asset, log employment, and capital investment/asset. Firm Controls includes Size, State-owned, Tangibility, Roa. All variables are defined in Table 1.2. All regressions include fixed effects. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Robust standard errors given below coefficient estimates are clustered at the firm-level.

Table 1.16 Firm Borrowing Channel (Three Periods Effect)

	(1) Net Debt	(2) Cash	(3) Employment	(4) Capital Investment
$\overline{\textit{Foreign Listed}}$	0.198***	-2.054***	-1.661***	-1.992
	(0.0730)	(0.441)	(0.335)	(1.290)
$\overline{\textit{Foreign Listed}}$ x Crisis	-0.372*	-0.398	-0.153	-1.144
	(0.158)	(1.144)	(0.718)	(2.750)
$\overline{\textit{Foreign Listed}}$ x Post	-1.169***	7.318***	8.472***	12.06***
	(0.221)	(1.378)	(1.005)	(3.932)
Sales	0.0102***	0.149***	0.0676***	0.0153
	(0.00378)	(0.0251)	(0.0173)	(0.0794)
Cash Flow	-0.709***	2.422***	0.948***	0.918
	(0.0456)	(0.290)	(0.207)	(0.974)
Observations	5842	6249	6310	3772
R^2	0.765	0.837	0.875	0.650
Firm Controls	YES	YES	YES	YES
Firm Fixed Effect	YES	YES	YES	YES
Year Fixed Effect	YES	YES	YES	YES
Clusters at Firm Level	1321	1902	1911	1573

Notes. This table presents the results of a Chinese firm borrowing channel regression. The unit of observation is a firm-year. The dependent variable is the net debt, cash/asset, log employment, and capital investment/asset. Firm Controls includes Size, State-owned, Tangibility, Roa. All variables are defined in Table 1.2. All regressions include fixed effects. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Robust standard errors given below coefficient estimates are clustered at the firm-level.

Table 1.17 Firm Borrowing Channel (Three Periods Effect)

	(1) Net Debt	(2) Cash	(3) Employment	(4) Capital Investment
$\overline{\textit{Trade Settlement}}$	0.129***	-0.244	-0.494***	-0.462
	(0.0299)	(0.194)	(0.138)	(0.546)
$\overline{\textit{Trade Settlement}}$ x Crisis	-0.139*	-0.366	-0.486*	-1.558
	(0.0551)	(0.371)	(0.247)	(0.957)
$\overline{\textit{Trade Settlement}}$ x Post	-0.171***	0.877***	0.850***	2.286***
	(0.0354)	(0.244)	(0.172)	(0.648)
Sales	0.00849*	0.175***	0.0957***	0.00991
	(0.00372)	(0.0248)	(0.0159)	(0.0771)
Cash Flow	-0.714***	2.588***	0.543***	0.437
	(0.0444)	(0.284)	(0.189)	(0.942)
Observations	6214	6660	6720	3993
R^2	0.762	0.832	0.888	0.653
Firm Controls	YES	YES	YES	YES
Firm Fixed Effect	YES	YES	YES	YES
Year Fixed Effect	YES	YES	YES	YES
Clusters at Firm Level	1349	1943	1952	1611

Notes. This table presents the results of a Chinese firm borrowing channel regression. The unit of observation is a firm-year. The dependent variable is the net debt, cash/asset, log employment, and capital investment/asset. Firm Controls includes Size, State-owned, Tangibility, Roa. All variables are defined in Table 1.2. All regressions include fixed effects. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Robust standard errors given below coefficient estimates are clustered at the firm-level.

Table 1.18 Firm Borrowing Channel (OECD GDP Growth Rate Effect)

	(1)	(2)	(3)	(4)
	Net Debt	Cash	Employment	Capital Investment
$\overline{International\ Borrowing}$	-5.362**	0.708	0.115	-8.104
	(2.400)	(1.465)	(0.965)	(4.944)
$\overline{International\ Borrowing}$ x OECD GDP Growth Rate	-0.171	0.320**	0.215**	2.083***
	(0.243)	(0.138)	(0.0953)	(0.526)
Sales	0.0577	0.166***	0.0892***	0.0101
	(0.0379)	(0.0246)	(0.0168)	(0.0799)
Cash Flow	-6.625***	2.419***	0.721***	0.0355
	(0.448)	(0.277)	(0.196)	(0.994)
Observations	6841	7060	7053	4197
R^2	0.756	0.823	0.871	0.652
Firm Controls	YES	YES	YES	YES
Firm Fixed Effect	YES	YES	YES	YES
Year Fixed Effect	YES	YES	YES	YES
Clusters at Firm Level	1967	1992	1991	1657

Notes. This table presents the results of a Chinese firm borrowing channel regression. The unit of observation is a firm-year. The dependent variable is the net debt, cash/asset, log employment, and capital investment/asset. Firm Controls includes Size, State-owned, Tangibility, Roa. All variables are defined in Table 1.2. All regressions include fixed effects. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Robust standard errors given below coefficient estimates are clustered at the firm-level.

Table 1.19 Firm Borrowing Channel (OECD GDP Growth Rate Effect)

	(1)	(2)	(3)	(4)
	Net Debt	Cash	Employment	Capital Investment
$\overline{Foreign\ Listed}$	0.0822	-1.748***	-1.014***	-1.605
	(0.0691)	(0.421)	(0.300)	(1.358)
$\overline{Foreign\ Listed}$ x OECD GDP Growth Rate	-0.0578***	0.288**	0.220***	1.484***
	(0.0202)	(0.118)	(0.0801)	(0.448)
Sales	0.00692*	0.146***	0.0653***	0.0322
	(0.00390)	(0.0252)	(0.0171)	(0.0853)
Cash Flow	-0.688***	2.429***	0.665***	0.0135
	(0.0477)	(0.291)	(0.206)	(1.087)
Observations	5842	6249	6310	3772
R^2	0.777	0.840	0.885	0.666
Firm Controls	YES	YES	YES	YES
Firm Fixed Effect	YES	YES	YES	YES
Year Fixed Effect	YES	YES	YES	YES
Clusters at Firm Level	1321	1902	1911	1573

Notes. This table presents the results of a Chinese firm borrowing channel regression. The unit of observation is a firm-year. The dependent variable is the net debt, cash/asset, log employment, and capital investment/asset. Firm Controls includes Size, State-owned, Tangibility, Roa. All variables are defined in Table 1.2. All regressions include fixed effects. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Robust standard errors given below coefficient estimates are clustered at the firm-level.

Table 1.20 Firm Borrowing Channel (OECD GDP Growth Rate Effect)

	(1)	(2)	(3)	(4)
	Net Debt	Cash	Employment	Capital Investment
$\overline{Trade\ Settlement}$	0.0283	0.166	-0.0516	0.285
	(0.0188)	(0.117)	(0.0816)	(0.353)
$\overline{Trade\ Settlement}$ x OECD GDP Growth Rate	-0.0544***	0.320***	0.229***	1.393***
	(0.0191)	(0.114)	(0.0745)	(0.422)
Sales	0.00456	0.177***	0.103***	0.0423
	(0.00383)	(0.0249)	(0.0165)	(0.0835)
Cash Flow	-0.696***	2.560***	0.632***	0.273
	(0.0462)	(0.285)	(0.195)	(1.060)
Observations	6214	6660	6720	3993
R^2	0.772	0.834	0.886	0.663
Firm Controls	YES	YES	YES	YES
Firm Fixed Effect	YES	YES	YES	YES
Year Fixed Effect	YES	YES	YES	YES
Clusters at Firm Level	1349	1943	1952	1611

Notes. This table presents the results of a Chinese firm borrowing channel regression. The unit of observation is a firm-year. The dependent variable is the net debt, cash/asset, log employment, and capital investment/asset. Firm Controls includes Size, State-owned, Tangibility, Roa. All variables are defined in Table 1.2. All regressions include fixed effects. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Robust standard errors given below coefficient estimates are clustered at the firm-level.

Table 1.21 Firm Borrowing Channel (Generalized Method of Moments)

	(1)	(2)	(3)	(4)
	Net Debt	Cash	Employment	Capital Investment
$\overline{\textit{International Borrowing}}$	4.351***	-0.536***	-2.488***	-18.02
	(1.317)	(0.147)	(0.815)	(13.79)
$\overline{\textit{International Borrowing}}$ x OECD GDP Growth Rate	-0.781***	0.0744***	0.400***	4.234*
	(0.247)	(0.0244)	(0.145)	(2.561)
Sales	0.0839***	0.0172***	0.0647***	0.537*
	(0.0231)	(0.00205)	(0.0112)	(0.298)
Cash Flow	-1.324***	0.468***	1.349***	2.332
	(0.153)	(0.0425)	(0.273)	(1.864)
Constant	0.364***	0.119***	7.236***	17.17***
	(0.0577)	(0.0122)	(0.107)	(0.556)
Observations	6841	7060	7053	4197
Firms	1967	1992	1991	1657
Firm Controls	YES	YES	YES	YES
Firm Fixed Effect	YES	YES	YES	YES
Year Fixed Effect	YES	YES	YES	YES
Sargan test (p-value)	0.0369	0.0820	0.117	0.296
Serial correlation test (p-value)	0.624	0.785	0.120	0.390

Notes. This table presents the results of a Chinese firm borrowing channel regression. I conduct the two-step GMM system estimation using lagged 3-4 values as instruments. The unit of observation is a firm-year. The dependent variable is the net debt, cash/asset, log employment, and capital investment/asset. Firm Controls includes Size, State-owned, Tangibility, Roa. All variables are defined in Table 1.2. All regressions include fixed effects. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Windmeijer corrected robust standard errors are given below.

Table 1.22 Firm Borrowing Channel (Generalized Method of Moments)

	(1)	(2)	(3)	(4)
	Net Debt	Cash	Employment	Capital Investment
$\overline{\textit{Foreign Listed}}$	-0.282***	0.513**	2.684***	-0.599
	(0.0542)	(0.250)	(0.416)	(2.129)
$\overline{\textit{Foreign Listed}}$ x OECD GDP Growth Rate	-0.0531***	0.149***	0.309***	1.073***
	(0.0199)	(0.0528)	(0.108)	(0.407)
Sales	0.0208***	0.0141*	0.0167	0.914***
	(0.00779)	(0.00786)	(0.0287)	(0.332)
Cash Flow	-0.469***	2.873***	3.895***	0.401
	(0.0477)	(0.127)	(0.469)	(2.312)
Constant	0.180***	1.462***	7.669***	18.06***
	(0.00727)	(0.0299)	(0.0494)	(0.160)
Observations	5842	6249	6310	3772
Firms	1321	1902	1911	1573
Firm Controls	YES	YES	YES	YES
Firm Fixed Effect	YES	YES	YES	YES
Year Fixed Effect	YES	YES	YES	YES
Sargan test (p-value)	0.753	0.357	0.543	0.362
Serial correlation test (p-value)	0.0303	0.808	0.0983	0.613

Notes. This table presents the results of a Chinese firm borrowing channel regression. I conduct the two-step GMM system estimation using lagged 3-4 values as instruments. The unit of observation is a firm-year. The dependent variable is the net debt, cash/asset, log employment, and capital investment/asset. Firm Controls includes Size, State-owned, Tangibility, Roa. All variables are defined in Table 1.2. All regressions include fixed effects. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Windmeijer corrected robust standard errors are given below.

Table 1.23 Firm Borrowing Channel (Generalized Method of Moments)

	(1)	(2)	(3)	(4)
	Net Debt	Cash	Employment	Capital Investment
$\overline{\textit{Trade Settlement}}$	-0.0696***	0.150***	1.215***	-0.435
	(0.0193)	(0.0465)	(0.157)	(0.285)
$\overline{\textit{Trade Settlement}}$ x OECD GDP Growth Rate	-0.0519***	0.234***	0.477***	0.733**
	(0.0162)	(0.0563)	(0.108)	(0.334)
Sales	0.0161***	0.00927	0.190***	0.356
	(0.00530)	(0.0103)	(0.0217)	(0.223)
Cash Flow	-0.574***	2.823***	6.348***	2.308*
	(0.0570)	(0.277)	(0.482)	(1.399)
Constant	0.230***	1.451***	6.918***	18.74***
	(0.00859)	(0.0482)	(0.0723)	(0.276)
Observations	6214	6660	6720	3993
Firms	1349	1943	1952	1611
Firm Controls	YES	YES	YES	YES
Firm Fixed Effect	YES	YES	YES	YES
Year Fixed Effect	YES	YES	YES	YES
Sargan test (p-value)	0.901	0.0104	0.855	0.796
Serial correlation test (p-value)	0.476	0.808	0.423	0.604

Notes. This table presents the results of a Chinese firm borrowing channel regression. I conduct the two-step GMM system estimation using lagged 3-4 values as instruments. The unit of observation is a firm-year. The dependent variable is the net debt, cash/asset, log employment, and capital investment/asset. Firm Controls includes Size, State-owned, Tangibility, Roa. All variables are defined in Table 1.2. All regressions include fixed effects. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Windmeijer corrected robust standard errors are given below.

Table 1.24 Bank Lending Channel (Using All Measures of Exposure Together) (Growth Rate Effect)

	(1)	(2)	(3)	(4)
	Net Debt	Cash	Employment	Capital Investment
$\overline{\textit{International Borrowing}}$ x OECD GDP Growth Rate	-0.183	0.216***	0.0819	5.141**
	(0.136)	(0.0656)	(0.169)	(2.178)
$\overline{\textit{Foreign Listed}}$ x OECD GDP Growth Rate	-0.310*	0.151***	0.596***	1.556
	(0.175)	(0.0284)	(0.221)	(2.612)
$\overline{\textit{Trade Settlement}}$ x OECD GDP Growth Rate	-0.0383	0.00879	0.0220	3.915***
	(0.0814)	(0.0129)	(0.0455)	(0.874)
Sales	0.0460***	0.0193***	0.0635***	0.315
	(0.0176)	(0.000936)	(0.00417)	(0.380)
Cash Flow	-0.916***	0.420***	0.754***	1.167
	(0.119)	(0.0420)	(0.212)	(1.992)
Constant	0.140***	0.158***	7.015***	17.64***
	(0.0151)	(0.00730)	(0.0455)	(0.198)
Observations	5842	6145	6140	3665
Firms	1321	1890	1889	1546
Firm Controls	YES	YES	YES	YES
Firm Fixed Effect	YES	YES	YES	YES
Year Fixed Effect	YES	YES	YES	YES
Sargan test (p-value)	0.219	0.538	0.772	0.364
Serial correlation test (p-value)	0.449	0.826	0.125	0.859

Notes. This table presents the results of a Chinese bank lending channel regression. The unit of observation is a bank-firm-year. The dependent variable is the change in log loan volume. Bank Controls includes Size, Policy, Rural, Tangibility, Roa, Cash Flow, Lend Central Bank. All variables are defined in Table 1.2. All regressions include fixed effects. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Robust standard errors given below coefficient estimates are clustered at the bank-level.

Table 1.25 Bank Lending Channel (Growth Rate Effect) (Initial Value)

dLog(Loan)	(1) Exposure (International Borrowing)	(2) Exposure (International Borrowing)	(3) Exposure (Foreign Listed)	(4) Exposure (Foreign Listed)	(5) Exposure (Trade Settlement)	(6) Exposure (Trade Settlement)
Bad Loan	-4.704**	-3.651	-6.804***	-2.322	-5.541**	-1.264
	(2.217)	(2.674)	(2.345)	(3.166)	(2.154)	(3.239)
Bad Loan x OECD	1.905**	2.310**	1.907***	1.303	1.388*	0.789
	(0.749)	(0.964)	(0.668)	(1.099)	(0.780)	(1.154)
Exposure	-0.836	-0.0827	-0.420***	-0.371**	-0.113	-0.259***
	(0.587)	(0.471)	(0.154)	(0.153)	(0.0998)	(0.0789)
Exposure x OECD growth rate	0.267*	0.408**	0.0890**	0.0872**	0.0830***	0.0753***
	(0.146)	(0.196)	(0.0384)	(0.0395)	(0.00943)	(0.0280)
Observations	7627	7627	9968	9968	9981	9981
R^2	0.195	0.180	0.182	0.167	0.182	0.168
Bank Controls	YES	YES	YES	YES	YES	YES
Firm Fixed Effect	NO	YES	NO	YES	NO	YES
Year Fixed Effect	NO	YES	NO	YES	NO	YES
Firm-Year Fixed Effect	YES	NO	YES	NO	YES	NO
Clusters at Bank Level	118	118	141	141	133	133

Notes. This table presents the results of a Chinese bank lending channel regression. The unit of observation is a bank-firm-year. The dependent variable is the change in log loan volume. Bank Controls includes Size, Policy, Rural, Tangibility, Roa, Cash Flow, Lend Central Bank. All variables are defined in Table 1.2. All regressions include fixed effects. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Robust standard errors given below coefficient estimates are clustered at the bank-level.

Table 1.26 Bank Lending Channel (Exchange/Income) (Three Periods Level Effect)

	(1) lnloan	(2) lnloan	(3) lnloan	(4) lnloan	(5) lnloan	(6) lnloan
Bad Loan	-2.276**	-2.046**	-2.106**	-2.633***	-2.282***	-2.316***
	(0.912)	(0.907)	(0.906)	(0.864)	(0.860)	(0.859)
Bad Loan x Crisis	-2.817	-1.534	-2.399	2.636	5.541	4.390
	(6.584)	(6.098)	(6.061)	(6.198)	(5.718)	(5.679)
Bad Loan x Post	8.934*	7.893	6.130	4.924	3.018	1.683
	(4.858)	(4.840)	(4.781)	(4.835)	(4.817)	(4.764)
Exchange/Income	-0.0724	0.0207	-0.239	-0.268	-0.129	-0.448
	(0.439)	(0.434)	(0.395)	(0.432)	(0.426)	(0.388)
Exchange/Income x State-owned	-4.006*	0.871	-0.294	-7.551***	0.243	-0.402
	(2.213)	(0.783)	(0.338)	(2.071)	(0.777)	(0.330)
Exchange/Income x Crisis	-2.009***	-1.442**		-2.022***	-0.967	
	(0.752)	(0.712)		(0.756)	(0.711)	
Exchange/Income x Crisis x State-owned	-6.288*	-0.167		-11.88***	-0.603	
	(3.520)	(1.157)		(3.353)	(1.122)	
Exchange/Income x Post	1.641			3.580		
	(2.575)			(2.476)		
Exchange/Income x Post x State-owned	5.609**			9.128***		
	(2.381)			(2.247)		
Observations	19212	19212	19212	19212	19212	19212
R^2	0.668	0.668	0.668	0.575	0.574	0.574
Bank Controls	YES	YES	YES	YES	YES	YES
Firm Fixed Effect	NO	NO	NO	YES	YES	YES
Year Fixed Effect	NO	NO	NO	YES	YES	YES
Firm-Year Fixed Effect	YES	YES	YES	NO	NO	NO
Clusters at Bank Level	177	177	177	177	177	177

Notes. This table presents the results of a Chinese bank lending channel regression. The unit of observation is a bank-firm-year. The dependent variable is the level of log loan volume. Bank Controls includes Size, Policy, Rural, Tangibility, Roa, Cash Flow, Lend Central Bank. All variables are defined in Table 1.2. All regressions include fixed effects. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Robust standard errors given below coefficient estimates are clustered at the bank-level.

Table 1.27 Bank Lending Channel (Bad Loan) (Three Periods Level Effect)

	(1)	(2)	(3)	(4)	(5)	(6)
	lnloan	lnloan	lnloan	lnloan	lnloan	lnloan
Bad Loan	-9.737*	-13.26***	-14.72***	-16.65***	-16.59***	-18.93***
	(5.832)	(4.821)	(3.963)	(5.402)	(4.586)	(3.825)
Bad Loan x State-owned	7.576	10.96**	12.18***	14.15***	14.07***	16.27***
	(5.700)	(4.764)	(3.956)	(5.264)	(4.521)	(3.803)
Bad Loan x Crisis	-14.85*	-4.646		-11.65	-6.254	
	(7.960)	(6.405)		(7.666)	(6.296)	
Bad Loan x Crisis x State-owned	-7.477	-10.54**		-7.210	-3.593	
	(8.750)	(4.348)		(8.097)	(4.015)	
Bad Loan x Post	10.31			17.03*		
	(10.74)			(10.12)		
Bad Loan x Post x State-owned	20.12***			13.16*		
	(7.401)			(7.023)		
Observations	19212	19212	19212	19212	19212	19212
R^2	0.669	0.669	0.669	0.576	0.575	0.575
Bank Controls	YES	YES	YES	YES	YES	YES
Firm Fixed Effect	NO	NO	NO	YES	YES	YES
Year Fixed Effect	NO	NO	NO	YES	YES	YES
Firm-Year Fixed Effect	YES	YES	YES	NO	NO	NO
Clusters at Bank Level	177	177	177	177	177	177

Notes. This table presents the results of a Chinese bank lending channel regression. The unit of observation is a bank-firm-year. The dependent variable is the level of log loan volume. Bank Controls includes Size, State-owned, Policy, Rural, Tangibility, Roa, Cash Flow, Lend Central Bank. All variables are defined in Table 1.2. All regressions include fixed effects. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Robust standard errors given below coefficient estimates are clustered at the bank-level.

Table 1.28 Bank Lending Channel (Exchange/Income) (Growth Rate Effect)

	(1)	(2)	(3)	(4)	(5)
	dlnloan	dlnloan	dlnloan	dlnloan	dlnloan
Bad Loan	-6.647***	-5.862***	-2.982	-2.099	-1.315
	(2.408)	(2.112)	(2.690)	(2.415)	(1.086)
Bad Loan x OECD growth rate	2.339***	2.085***	1.612*	1.268	0.514
	(0.612)	(0.658)	(0.950)	(0.963)	(0.981)
Exchange/Income	-1.110*	-0.707**	-0.352	-0.0851	-0.853*
	(0.566)	(0.356)	(0.466)	(0.295)	(0.493)
Exchange/Income x OECD growth rate	0.566**		0.424**		0.401**
	(0.254)		(0.206)		(0.188)
OECD GDP growth rate					0.0362**
					(0.0147)
Observations	9986	9986	9986	9986	9986
R^2	0.185	0.185	0.168	0.168	0.159
Bank Controls	YES	YES	YES	YES	YES
Firm Fixed Effect	NO	NO	YES	YES	YES
Year Fixed Effect	NO	NO	YES	YES	NO
Firm-Year Fixed Effect	YES	YES	NO	NO	NO
Clusters at Bank Level	141	141	141	141	141

Notes. This table presents the results of a Chinese bank lending channel regression. The unit of observation is a bank-firm-year. The dependent variable is the change in log loan volume. Bank Controls includes Size, State-owned, Policy, Rural, Tangibility, Roa, List, Cash Flow, Lend Central Bank. All variables are defined in Table 1.2. All regressions include fixed effects. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Robust standard errors given below coefficient estimates are clustered at the bank-level.

Table 1.29 Bank Lending Channel (Bad Loan) (Growth Rate Effect)

	(1)	(2)	(3)	(4)	(5)
	dlnloan	dlnloan	dlnloan	dlnloan	dlnloan
Bad Loan	-5.785**	-2.429	-2.059	0.797	-0.566
	(2.317)	(2.024)	(3.120)	(1.868)	(2.547)
Bad Loan x OECD growth rate	1.968***		1.250		0.223
	(0.662)		(1.094)		(0.853)
OECD GDP growth rate					0.0365***
					(0.0123)
Observations	9986	9986	9986	9986	9986
R^2	0.185	0.184	0.168	0.167	0.159
Bank Controls	YES	YES	YES	YES	YES
Firm Fixed Effect	NO	NO	YES	YES	YES
Year Fixed Effect	NO	NO	YES	YES	NO
Firm-Year Fixed Effect	YES	YES	NO	NO	NO
Clusters at Bank Level	141	141	141	141	141

Notes. This table presents the results of a Chinese bank lending channel regression. The unit of observation is a bank-firm-year. The dependent variable is the change in log loan volume. Bank Controls includes Size, State-owned, Policy, Rural, Tangibility, Roa, List, Cash Flow, Lend Central Bank. All variables are defined in Table 1.2. All regressions include fixed effects. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Robust standard errors given below coefficient estimates are clustered at the bank-level.

Table 1.30 Bank Lending Channel (Two Periods Level Effect)

Log(Loan)	(1) Exposure (International Borrowing)	(2) Exposure (International Borrowing)	(3) Exposure (Foreign Listed)	(4) Exposure (Foreign Listed)	(5) Exposure (Trade Settlement)	(6) Exposure (Trade Settlement)
Bad Loan	-1.792***	-1.922***	-1.253**	-1.352	-1.515	-1.398
	(0.645)	(0.577)	(0.537)	(0.909)	(0.929)	(0.883)
Bad Loan x After	6.427	7.155	-8.034	-6.506	-6.399	-7.903*
	(6.286)	(6.408)	(6.155)	(4.257)	(4.419)	(4.324)
Exposure	0.602**	0.576**	4.671**	4.903***	0.280*	0.241*
	(0.303)	(0.264)	(2.011)	(1.028)	(0.167)	(0.141)
Exposure x State-owned	0.165**	0.182***	0.133**	0.121***	0.146***	0.135***
	(0.0647)	(0.0578)	(0.0524)	(0.0259)	(0.0419)	(0.0414)
Exposure x After	-0.839	-0.991	-4.012**	-4.276***	-0.121	-0.180
	(1.592)	(1.375)	(1.978)	(1.047)	(0.576)	(0.558)
Exposure x After x State-owned	-2.272*	-1.540	-0.507***	-0.487**	-1.602**	-1.520**
	(1.313)	(1.159)	(0.171)	(0.201)	(0.753)	(0.739)
Observations	13934	13934	18895	18895	17222	17222
R^2	0.685	0.593	0.669	0.575	0.670	0.575
Bank Controls	YES	YES	YES	YES	YES	YES
Firm Fixed Effect	NO	YES	NO	YES	NO	YES
Year Fixed Effect	NO	YES	NO	YES	NO	YES
Firm-Year Fixed Effect	YES	NO	YES	NO	YES	NO
Clusters at Bank Level	132	132	177	177	170	170

Notes. This table presents the results of a Chinese bank lending channel regression. The unit of observation is a bank-firm-year. The dependent variable is the level of log loan volume. Bank Controls includes Size, Policy, Rural, Tangibility, Roa, Cash Flow, Lend Central Bank. All variables are defined in Table 1.2. All regressions include fixed effects. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Robust standard errors given below coefficient estimates are clustered at the bank-level.

Table 1.31 Bank Lending Channel (Exchange/Income) (Two Periods Level Effect)

	(1)	(2)	(3)	(4)	(5)	(6)
	lnloan	lnloan	lnloan	lnloan	lnloan	lnloan
Bad Loan	-2.099**	-2.094**	-2.107***	-2.297***	-2.287***	-2.314***
	(0.906)	(0.906)	(0.535)	(0.859)	(0.859)	(0.594)
Bad Loan x After	3.022	3.535	-2.191	2.851	3.391	-1.490
	(3.886)	(3.894)	(5.596)	(3.804)	(3.810)	(5.645)
Exchange/Income	-0.260	-0.238	-0.474	-0.446	-0.407	-0.595
	(0.394)	(0.394)	(0.612)	(0.386)	(0.387)	(0.536)
Exchange/Income x After	-0.472	-0.177	-0.571***	-0.553*	-0.206	-0.582***
	(0.306)	(0.341)	(0.167)	(0.300)	(0.334)	(0.137)
Exchange/Income x After x State-owned		-1.549**	-0.330		-1.883**	-0.787
		(0.789)	(1.438)		(0.788)	(1.376)
Exchange/Income x State-owned			0.125***			0.121***
			(0.0321)			(0.0412)
Observations	19212	19212	18895	19212	19212	18895
R^2	0.668	0.668	0.668	0.574	0.574	0.575
Bank Controls	YES	YES	YES	YES	YES	YES
Firm Fixed Effect	NO	NO	NO	YES	YES	YES
Year Fixed Effect	NO	NO	NO	YES	YES	YES
Firm-Year Fixed Effect	YES	YES	YES	NO	NO	NO
Clusters at Bank Level	177	177	177	177	177	177

Notes. This table presents the results of a Chinese bank lending channel regression. The unit of observation is a bank-firm-year. The dependent variable is the level of log loan volume. Bank Controls includes Size, Policy, Rural, Tangibility, Roa, Cash Flow, Lend Central Bank. All variables are defined in Table 1.2. All regressions include fixed effects. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Robust standard errors given below coefficient estimates are clustered at the bank-level.

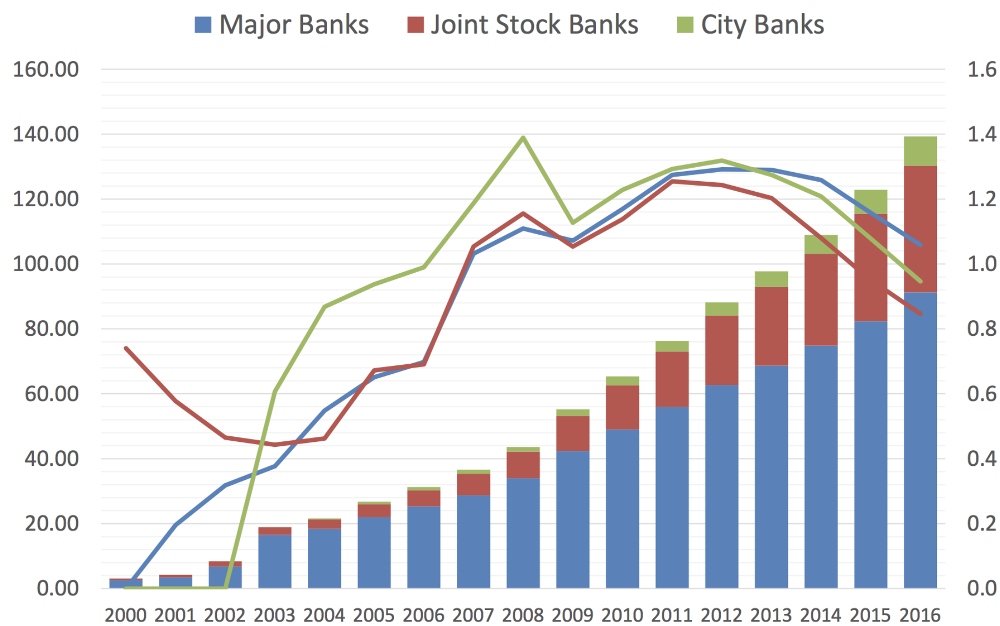
Table 1.32 Bank Lending Channel (Bad Loan) (Two Periods Level Effect)

	(1)	(2)	(3)	(4)	(5)	(6)
	lnloan	lnloan	lnloan	lnloan	lnloan	lnloan
Bad Loan	-2.450***	-2.193**	-7.706	-2.659***	-2.339***	-16.07***
	(0.905)	(0.910)	(5.768)	(0.859)	(0.865)	(5.334)
Bad Loan x After	-9.102**	-1.382	-0.126	-8.305**	-0.399	3.062
	(4.190)	(5.106)	(5.269)	(4.090)	(4.871)	(5.048)
Bad Loan x After x State-owned		-15.38***	-11.67*		-16.70***	-7.995
		(5.816)	(6.964)		(5.590)	(6.508)
Bad Loan x State-owned			5.448			13.54***
			(5.628)			(5.188)
Observations	19212	19212	19212	19212	19212	19212
R^2	0.669	0.669	0.669	0.575	0.575	0.575
Bank Controls	YES	YES	YES	YES	YES	YES
Firm Fixed Effect	NO	NO	NO	YES	YES	YES
Year Fixed Effect	NO	NO	NO	YES	YES	YES
Firm-Year Fixed Effect	YES	YES	YES	NO	NO	NO
Clusters at Bank Level	177	177	177	177	177	177

Notes. This table presents the results of a Chinese bank lending channel regression. The unit of observation is a bank-firm-year. The dependent variable is the level of log loan volume. Bank Controls includes Size, State-owned, Policy, Rural, Tangibility, Roa, Cash Flow, Lend Central Bank. All variables are defined in Table 1.2. All regressions include fixed effects. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Robust standard errors given below coefficient estimates are clustered at the bank-level.

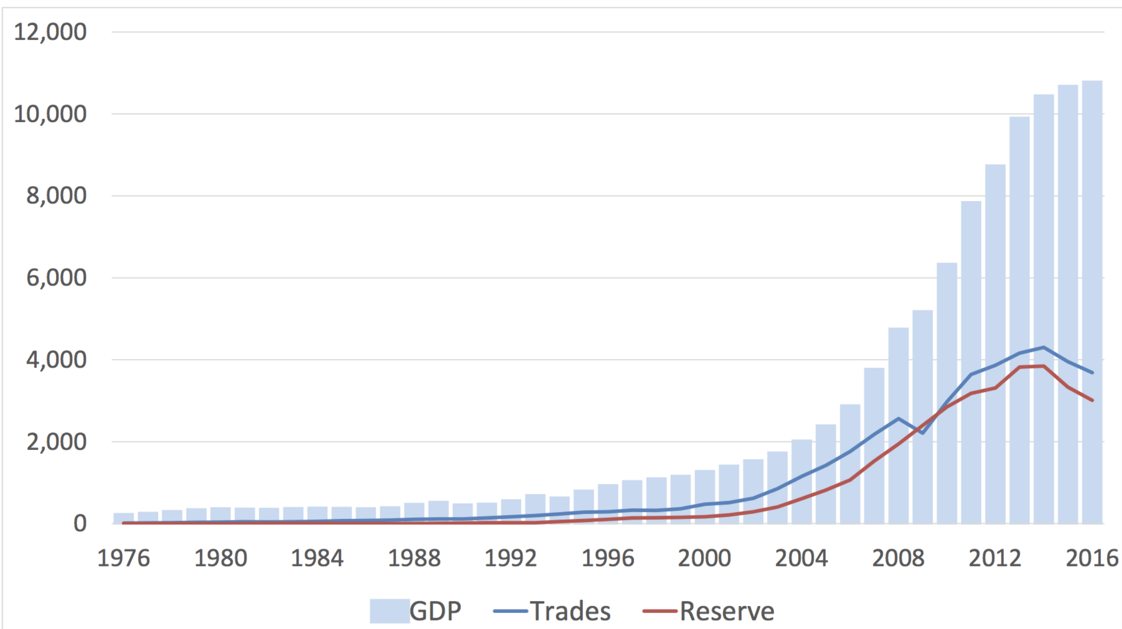
1.11 Figures

Figure 1.1 Bank Assets (RMB Trillion) and ROA (%)



Source: Jiang Wang (MIT), "China's Financial System: Developments and Challenges", MIT Golub Center for Finance and Policy 4th Annual Conference.

Figure 1.2 China's GDP, International Trade and FX Reserve (USD Billion)



Source: Jiang Wang (MIT), "China's Financial System: Developments and Challenges", MIT Golub Center for Finance and Policy 4th Annual Conference.

Figure 1.3 Structural-break Wald Statistic Plot

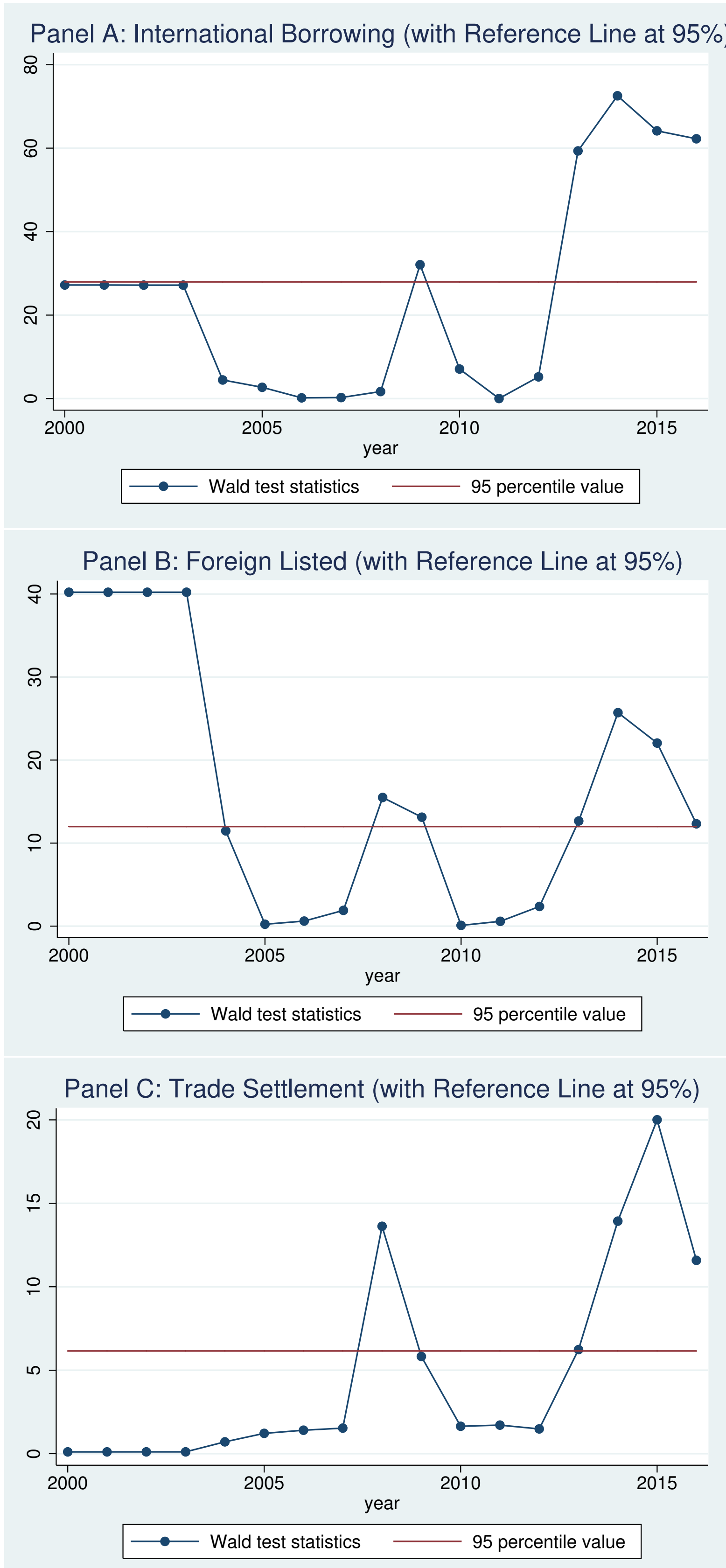


Figure 1.4 Parallel Trend

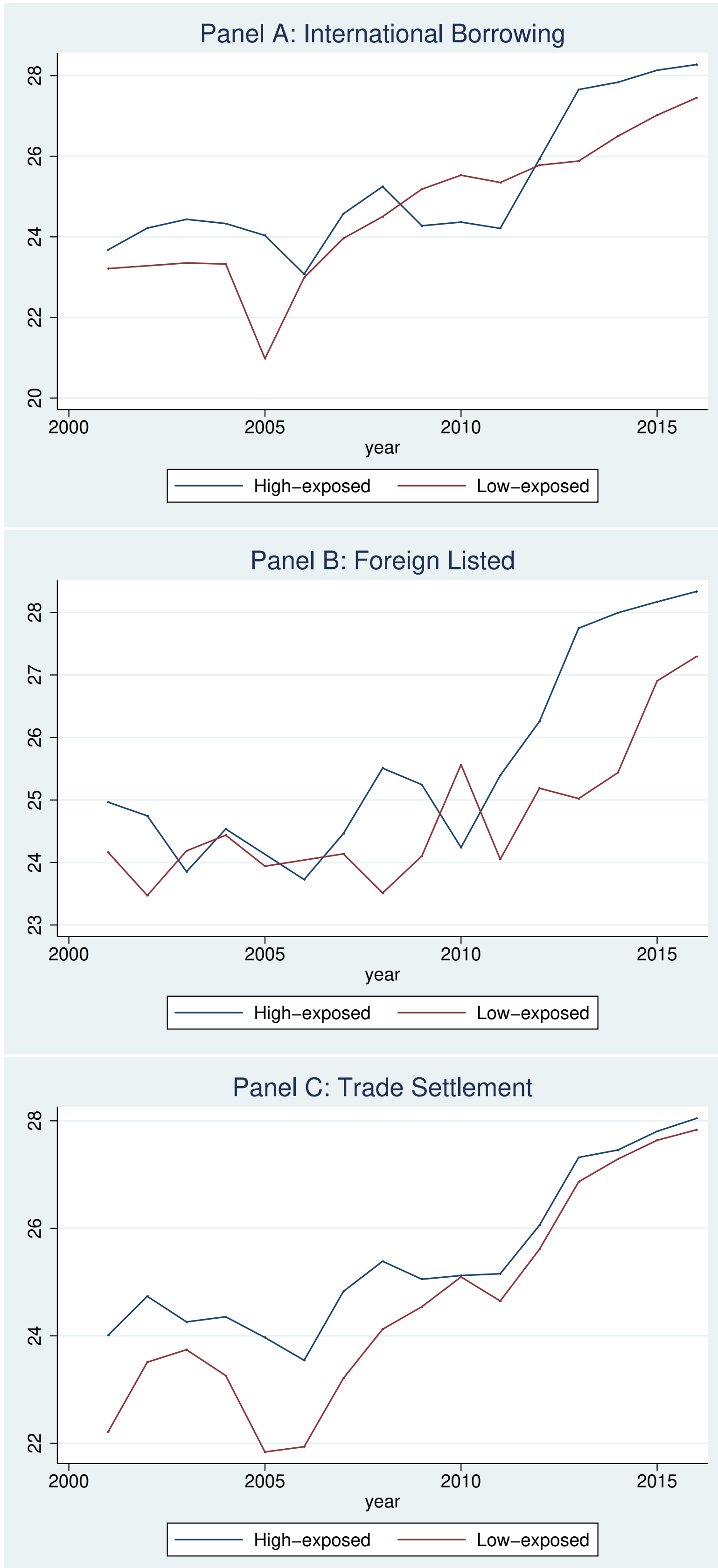


Figure 1.5 Year-specific Effects (Coefficients β_1)

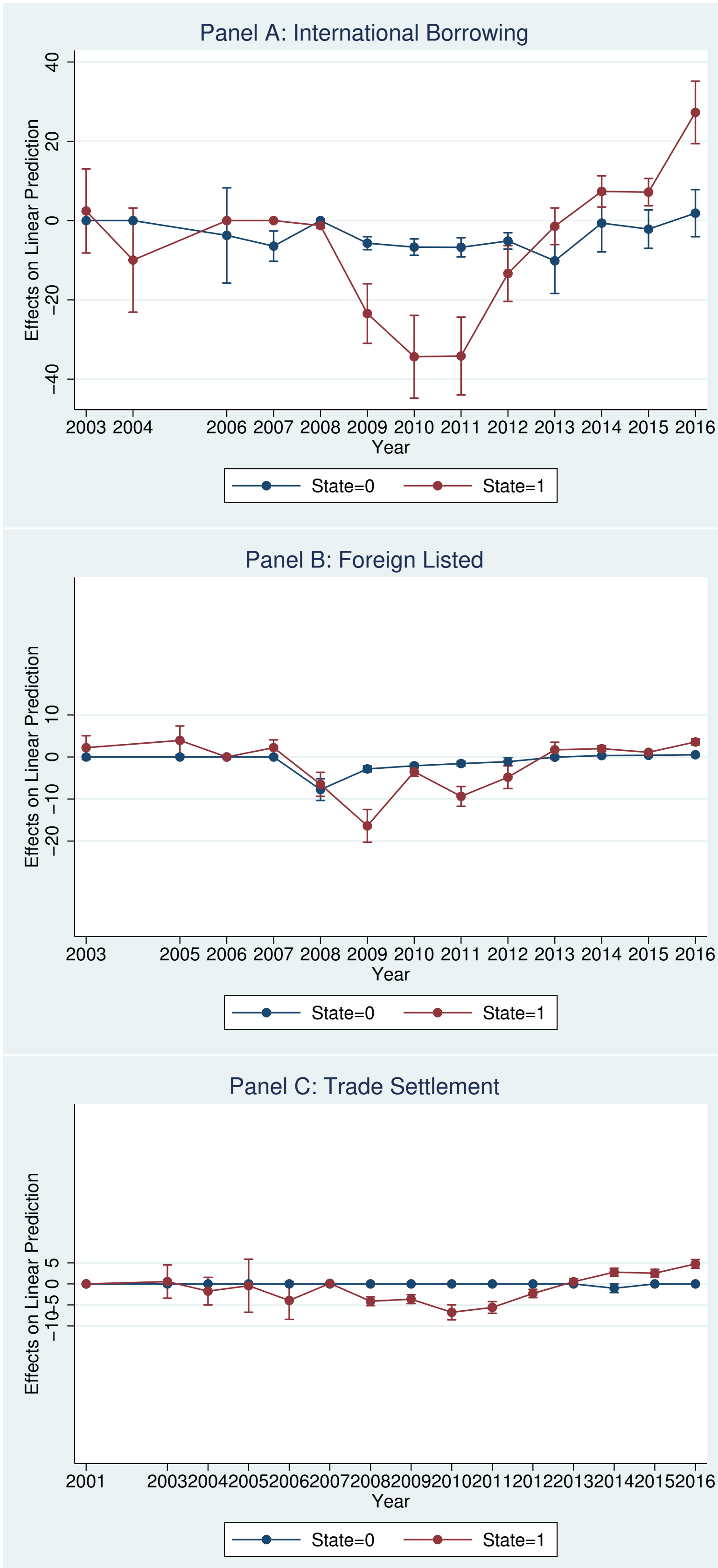


Figure 1.6 Marginal Effects

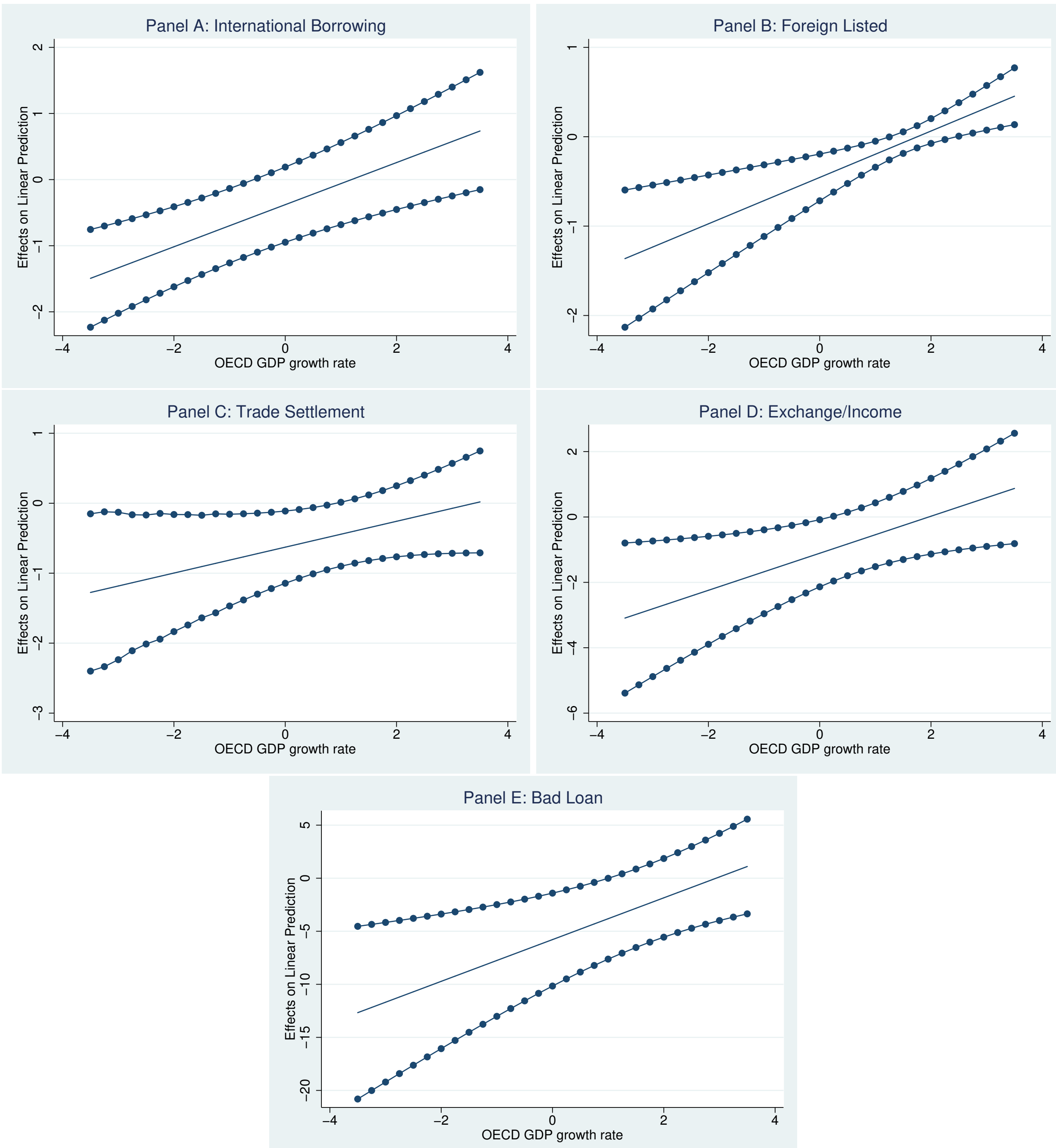


Figure 1.7 Firm Projects

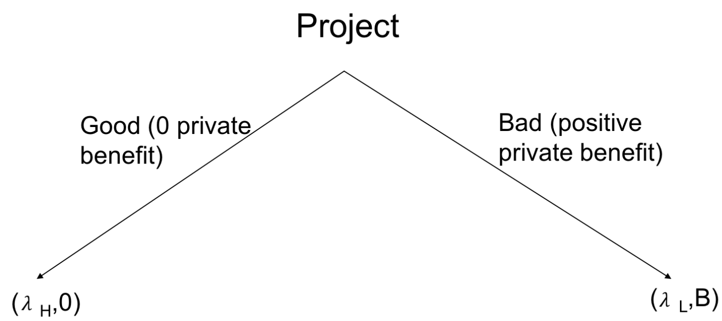
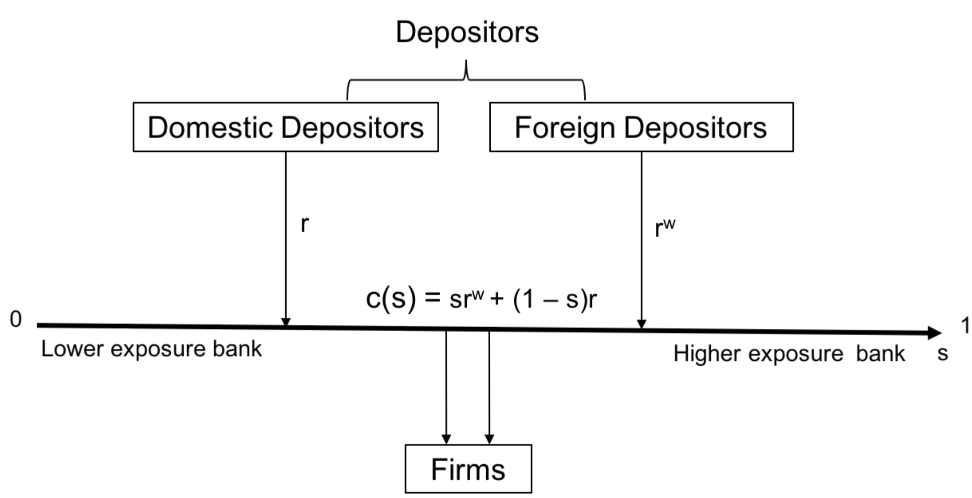


Figure 1.8 Timeline



1.12 Appendix

1.12.1 Techniques for Structural Breaks

Denote the sample period as $t = 1, ..., n$, the break date (date of the change) as T_1, the full break model could be written as

$$\begin{aligned} Y_1 &= X_1\beta_1 + e_1 \\ Y_2 &= X_2\beta_2 + e_2 \end{aligned} \tag{1.38}$$

or

$$y_t = \beta_1' x_t 1(t \leq T_1) + \beta_2' x_t 1(t > T_1) + e_t \tag{1.39}$$

where $Y_1 = (y_1, ..., y_{T_1})'$, $Y_2 = (y_{T_1+1}, ..., y_n)'$, $y_t = Log(loan_t)$, $X_1 = (x_1, ..., x_{T_1})'$, $X_2 = (x_{T_1+1}, ..., x_n)'$, $x_t =$ (1, exposure measurement, size, policy, rural, roa, bad loan)$'_{t-1}$, $e_1 = (\epsilon_1, ..., \epsilon_{T_1})'$, $e_2 = (\epsilon_{T_1+1}, ..., \epsilon_n)'$.

Thus we have

$$\begin{aligned} \hat{\beta}_1 &= (X_1'X_1)^{-1}(X_1'Y_1) \\ \hat{\beta}_2 &= (X_2'X_2)^{-1}(X_2'Y_2) \end{aligned} \tag{1.40}$$

Assume break dates are unknown; the null hypothesis is $\beta_1 = \beta_2$, I use the standard linear hypothesis test (Wald test). The Wald test statistic is

$$W(T_1) = n(\hat{\beta}_1 - \hat{\beta}_2)'(\hat{V}_1 \frac{n}{T_1} + \hat{V}_2 \frac{n}{n - T_1})^{-1}(\hat{\beta}_1 - \hat{\beta}_2) \tag{1.41}$$

where $\hat{V}_1$ and $\hat{V}_2$ are standard asymptotic variance estimators for $\hat{\beta}_1$ and $\hat{\beta}_2$ (on the split samples):

$$\begin{aligned}\hat{V}_1 &= \hat{Q}_1^{-1}\hat{\Omega}_1\hat{Q}_1^{-1} \\ \hat{V}_2 &= \hat{Q}_2^{-1}\hat{\Omega}_2\hat{Q}_2^{-1}\end{aligned} \tag{1.42}$$

And

$$\begin{aligned}\hat{Q}_1 &= \frac{1}{T_1}X_1'X_1 \\ \hat{Q}_2 &= \frac{1}{n-T_1}X_2'X_2\end{aligned} \tag{1.43}$$

We assume that e_t is independent identical distributed, thus

$$\begin{aligned}\hat{\Omega}_1 &= \frac{1}{T_1-k}(\hat{e}_1'\hat{e}_1)\hat{Q}_1 \\ \hat{\Omega}_2 &= \frac{1}{n-T_1-k}(\hat{e}_2'\hat{e}_2)\hat{Q}_2\end{aligned} \tag{1.44}$$

Under H_0, if the number of observations pre- and post-break are large, then under homoskedasticity, and in general

$$W(T_1) \longrightarrow_d \chi_k^2 \tag{1.45}$$

where k represents the number of the independent variables, we have $k = 7$.

We can reject H_0 in favor of H_1 if the test exceeds the critical value, thus "find a break" if the test rejects.

1.12.2 Proof of Proposition 1

Given $r^w > \frac{\beta(1+\rho)}{1+\rho-\delta}$ and $r > r^w$, we have $r > \frac{\beta(1+\rho)}{1+\rho-\delta}$.

$$r^w > \frac{\beta(1+\rho)}{1+\rho-\delta} \Leftrightarrow \frac{1+\rho-\delta}{\delta} - \frac{\beta}{r^w-\beta} > 0$$

$$r > \frac{\beta(1+\rho)}{1+\rho-\delta} \Leftrightarrow \frac{1+\rho-\delta}{\delta} - \frac{\beta}{r-\beta} > 0$$

$$r > r^w \Leftrightarrow \frac{\beta}{r-\beta} - \frac{\beta}{r^w-\beta} < 0$$

$$\Rightarrow \frac{dr}{dr^w} = \frac{\frac{1+\rho-\delta}{\delta} - \frac{\beta}{r^w-\beta}}{\frac{1+\rho-\delta}{\delta} - \frac{\beta}{r-\beta}} > 0; \quad \frac{dr}{dr^w} - 1 = \frac{\frac{\beta}{r-\beta} - \frac{\beta}{r^w-\beta}}{\frac{1+\rho-\delta}{\delta} - \frac{\beta}{r-\beta}} < 0 \Rightarrow \frac{dr}{dr^w} < 1$$

1.12.3 Proof of Proposition 2

Given $r^w > \frac{\beta(1+\rho)}{1+\rho-\delta}$ and $r > r^w$, we have $r > \frac{\beta(1+\rho)}{1+\rho-\delta}$.

$$r > \frac{\beta(1+\rho)}{1+\rho-\delta} \Leftrightarrow \frac{1+\rho-\delta}{\delta} - \frac{\beta}{r-\beta} > 0$$

$$r > r^w \Leftrightarrow \frac{1+\rho-\delta}{\beta\delta}(r-r^w); \quad \frac{\beta}{r^w-\beta} - \frac{\beta}{r-\beta} > 0 \Rightarrow \frac{1+\rho-\delta}{\beta\delta}(r-r^w) + (\frac{\beta}{r^w-\beta} - \frac{\beta}{r-\beta}) > 0$$

$$r^w > \frac{\beta(1+\rho)}{1+\rho-\delta} \quad and \quad r^w \approx r \Leftrightarrow \frac{\beta}{r^w-\beta} + \frac{1+\rho-\delta}{\delta}(\frac{r-r^w}{\beta} - 1) < 0$$

$$\Rightarrow \frac{dr}{dR} = \frac{\frac{1+\rho-\delta}{\beta\delta}(r-r^w) + (\frac{\beta}{r^w-\beta} - \frac{\beta}{r-\beta})}{\frac{1+\rho-\delta}{\delta} - \frac{\beta}{r-\beta}} > 0; \quad \frac{dr}{dR} - 1 = \frac{\frac{\beta}{r^w-\beta} + \frac{1+\rho-\delta}{\delta}(\frac{r-r^w}{\beta} - 1)}{\frac{1+\rho-\delta}{\delta} - \frac{\beta}{r-\beta}} < 0 \Rightarrow \frac{dr}{dR} < 1$$

1.12.4 Proof of Proposition 3

Given $c(s) > \beta = R - \gamma/\Delta\lambda > 0$, $0 < s < 1$, and $0 < \frac{dr}{dr^w} < 1$, we have

$$-\frac{\beta s}{(c-\beta)^2} < 0\,; \; -\frac{\beta(1-s)}{(c-\beta)^2}\frac{dr}{dr^w} < 0$$

$$\Rightarrow \frac{\partial I}{\partial r^w} = -\frac{\beta s}{(c-\beta)^2} - \frac{\beta(1-s)}{(c-\beta)^2}\frac{dr}{dr^w} < 0$$

$$(1 - dr/dr^w)(c - \beta) + 2(r - r^w)(s + (1-s)dr/dr^w) > 0\ ;\ (c-\beta)^3 > 0$$

$$\Rightarrow \frac{\partial I^2}{\partial r^w \partial s} = -\beta \frac{(1 - dr/dr^w)(c - \beta) + 2(r - r^w)(s + (1-s)dr/dr^w)}{(c-\beta)^3} < 0$$

1.12.5 Proof of Proposition 4

Given $c(s) > \beta = R - \gamma/\Delta\lambda > 0$, $0 < s < 1$, and $0 < \frac{dr}{dR} < 1$, we have

$$c - \beta(1-s)\frac{dr}{dR} > 0\ ;\ (c-\beta)^2 > 0$$

$$\Rightarrow \frac{\partial I}{\partial R} = \frac{c}{(c-\beta)^2} - \frac{\beta(1-s)}{(c-\beta)^2}\frac{dr}{dR} > 0$$

$$\beta dr/dR(c-\beta) > 0\ ;\ (r - r^w)(c + \beta - 2\beta(1-s)dr/dR) > 0\ ;\ (c-\beta)^3 > 0$$

$$\Rightarrow \frac{\partial I^2}{\partial R \partial s} = \frac{\beta dr/dR(c-\beta) + (r - r^w)(c + \beta - 2\beta(1-s)dr/dR)}{(c-\beta)^3} > 0$$

Chapter Two

Financial Corruption and Bank Loan Allocation: Evidence from China

2.1 Introduction

The research question is, what is the local effects of the anti-corruption campaign in 2012 in China on the credit supply of Chinese banks?

The identification strategy is the exogenous shock of the anti-corruption campaign in China. A far-reaching campaign against corruption began in China following the conclusion of the 18th National Congress of the Communist Party of China in 2012. The campaign, carried out under the aegis of Xi Jinping, General Secretary of the Communist Party of China (paramount leader), was the most significant organized anti-graft effort in the history of Communist rule in China. Figure 2.1 illustrates that the anti-corruption campaign is unanticipated in China.

Why do we choose the anti-corruption campaign in 2012 as the supply side shock? Figure 2.2 shows the bank assets and ROA of Chinese banks, we could find an obvious turning point in the year 2012, from 2000 to 2012, the ROA of Chinese banks increased dramatically (except for the financial crisis in 2008), however, after 2012 the ROA decreased significantly. Thus, something happened in Chinese bank system in the year 2012, after reading a large amount of the literature, I find that this domestic supply-side shock should be the anti-corruption campaign.

There is a growing literature studying the effects of the corruption, some literature study the negative effects of the corruption and the rent-seeking activities (Shleifer and Vishny, 1993, 1994; Mauro, 1995; Fisman, 2001; Fisman and Svensson, 2007; Butler et al., 2009; Bertrand et al., 2018). On the other side, some papers analyze the positive impacts of the corruption and the political connections. (Faccio, 2006; Goldman et al., 2008; Amore and Bennedsen, 2013; Dreher and Gassebner, 2013). Some other literature argue that the relationship between political connections and bank financing decisions is quite complex, especially across countries (Johnson and Mitton, 2003; Sapienza, 2004; Khwaja and Mian, 2005; Leuz and Oberholzer-Gee, 2006; Claessens et al., 2008; Zeume, 2017).

Recent studies related to the anti-corruption campaign in China mainly focus on the stock market price reactions (Griffin et al., 2016; Lin et al., 2016; Liu et al., 2017). However, bank financing still constitutes a dominant source of corporate financing (about 85 percent), while equity financing is only a tiny portion in China (1.3 percent) (Wang et al., 2019). Table 2.2 compares the financial development of China and the United States in 2016, column 1 shows that the absolute value of the bank credit in China is even more significant than the bank credit in the United States (15.45 vs. 12.44). When we consider the amount of bank credit as a fraction of the GDP, the difference between China and the United States is significantly amplified (137.95% vs. 67.00%). Moreover, the amount of bank credit is the largest among all financial instruments (stock, fixed income, insurance, and investment funds). Therefore I focus on Chinese bank credit in this paper.

Li et al. (2018) provides a novel empirical finding that the recent anti-corruption investigations in China are associated with credit reallocation from less productive, state-owned enterprises (SOEs) to more productive, non-SOEs. There are several differences between Li et al. (2018) and my paper. First of all, I use a completely different confidential data set to measure corruption in China. This paper employs the confidential data set compiled by the Central Commission for Discipline Inspection (CCDI) of the Communist Party of China. I got the data from the Center for Anti-corruption and Governance at Tsinghua University. Secondly, Li et al. (2018) emphasize on the loan rebate, which is the entry fee for the enterprise to obtain credit funds. My paper focus on the bank's monopoly power, which is the entry fee for the bank to get the monopoly power as the fund provider in the specific prefecture. Thirdly, this paper focus on the local effects. I use difference-in-difference estimation to emphasize the different levels of corruption between different prefectures in China. Thus, this paper analyzes the behavior of local officials instead of firm managers in Li et al. (2018). Moreover, I employ the misreporting extent of the specific prefecture as the robustness check. Finally, I use the bank-firm matched loan-level data, which means I could use Khwaja and Mian (2008) technique (year-firm fixed effect) to focus on the variation in loans due to banks primarily.

Khwaja and Mian (2008) analyze how supply-side bank liquidity shocks get transmitted to the rest of the economy. They examine the impact of liquidity shocks by exploiting cross-bank liquidity variation induced by unanticipated nuclear tests in Pakistan. When Pakistan tested nuclear devices in 1998, the IMF suspended the exchange rate liquidity support. Banks experienced the deposit run with larger dollar deposit accounts. The liquidity shock varied substantially across banks.Khwaja and Mian (2008) estimate the bank lending channel and firm borrowing channel simultaneously, and show that for the same firm borrowing from two different banks, its loan from the bank experiencing a 1 percent larger decline in liquidity drops by an additional 0.6 percent. This paper also uses the fixed effect technique in Khwaja and Mian (2008) to estimate the bank lending channel and firm borrowing channel simultaneously. Regarding to the ownership, Cong et al. (2019) find that the stimulus-driven credit expansion disproportionately favored state-owned firms and firms with a lower average product of capital, reversing the process of capital reallocation toward private firms that characterized China's high growth before 2008.

Most of the literature in the field of Political Science states that China is a highly centralized country, and neither provinces nor prefectures have any specific law, regulations and policies. Thus we could consider the enforcement is determined by the central government and is almost the same across regions. However, the enforcement differs over time. In this case, I conduct a robustness check by using the index in the initial year as well as the average of all the pre-shock years of my sample to measure the corruption.

The article unfolds as follows. Section 2.2 illustrates the background of anti-corruption campaign and banking system in China. Section 2.3 describes the data and key indexes used in this study. Section 2.4 discusses our empirical strategy and identification strategy. Section 2.5 reports our main finding regarding bank loan effects of the corruption as well as the results from a number of robustness checks. Section 2.6 investigates the mechanisms through which corruption may affect bank loans. Section 2.7 summarizes and concludes with a discussion of policy implications.

2.2 Background

After China's economic system reform in 1978, the planned economic system has broken, introduced market as well as the price system. In the transition process, the change of the property rights protection law has also created a group of beneficiaries, and banks are also one of the beneficiaries formed during the reform process. In the process of reform, state-owned commercial banks replaced government finance as the primary source of corporate funds (Ting, 1997; Wedeman, 2004). The provider uses the funds and powers granted by the government to obtain a monopoly position in the credit market. Although the central bank has reduced monopoly profits through interest rate controls, banks can circumvent restrictions through various other means. The banks are trying to acquire economic benefits beyond the law, which will undoubtedly increase the financing costs of the firms, both private and state-owned.

The anti-corruption investigation started in the year 2012. Most of the officials investigated were removed from office and faced accusations of bribery and abuse of power, although the range of alleged abuses varied widely. As of 2016, the campaign has 'netted' over 120 high-ranking officials, including about a dozen high-ranking military officers, several senior executives of state-owned companies, and five national leaders. More than 100,000 people have been indicted for corruption. Executed mainly under the direction of the Central Commission for Discipline Inspection and its Secretary from 2012 to 2019 Qishan Wang along with corresponding military and judicial organs, the campaign was notable in implicating both incumbent and former leaders. The anti-corruption campaign continues up to now. The second plenary meeting of the 19th Central Commission for Discipline Inspection of the Communist Party of China was held in Beijing from January 11 to 13, 2018. Xi Jinping emphasized that it is necessary to increase the anti-corruption efforts in the financial sector. The communique of the meeting also showed that it is essential to focus on resource-rich areas and critical positions and strengthen supervision to decrease power concentration.

In China, the corruption for bank credit could be classified into two forms: one is the loan rebate. The deducted benefit fee, which is the entry fee for the enterprise to obtain credit funds. This kind of rental income could be referred as the "entry fee". Most of the existing literature focus on the first form (Lu, 2000; Fan and Grossman, 2001; Ping and Lei, 2003; Yong and An-gang, 2003; Xie and Lu, 2005; Kwong, 2015); the other is the bank's monopoly power as the fund provider. The abuse of the pricing power of the fund to seek high markups in interest rates. The additional high-interest rate rental income from the monopoly power is called "price rent". This paper focuses on the second form. Traditionally, the traditional banking industry has gained unseen profits under its monopoly advantage. Among them, spread income is the primary source of income for the traditional banking industry in the past. It is reported that the income from interest in the listed banking industry accounted for more than 80% of the annual income.

Xie and Lu (2005) investigate the central bank, commercial bank, and policy bank in 29 cities in China. There are around 80.5% (975/1211) bank staffs admit that the rent-seeking by the right of financial resource allocating is very common to see or common to see in their work. Besides, approximately 61.5% (652/1061) staffs in firms consider they need to pay high markup to get the bank credit. Moreover, when they decrease the sample size to only staffs in private firms, the percentage increases from 61.5% (652/1061) to around 73.7% (350/475). Qian et al. (2015) states that PBOC limits the movements of interest rates on both deposits and loans by setting base rates along with upper and lower bounds. These rates and bounds vary over business cycles and with loan maturities. Local commercial banks could adjust the interest rate within the boundaries according to their specific issues.

2.3 Data and Descriptive Statistics

Below we describe our data, the definition and evolution of key indexes, and summary statistics.

2.3.1 Data Source

Our analysis uses several data sources, which we can link through unique identifiers for each individual and every prefecture. Our data set combines four sources of information existing in China: (i) bank loan data; (ii) corruption data; (iii) luminosity data; and, (iv) economic data.

(i) Bank Loan Data. This paper uses a novel data set that contains bank-firm relationships in China, along with detailed firm and bank-specific information. Our sample period is from 2001 to 2016, such that we have a symmetric time window surrounding the beginning of the financial crisis in the United States. Chinese data in this paper comes

three primary data sets: Wind Datafeed Service (referred to as WDS), GTA The China Stock Market and Accounting Research (referred to as GTA CSMAR) Database as well as Almanac of China's Finance and Banking (2001-2016).

Information about bank-firm relationships is from the bank loan data in GTA CSMAR Database. GTA CSMAR access data on the China stock markets and the financial statements of China's listed companies. GTA CSMAR is a unique, comprehensive database of China stock returns, covering all companies listed on the Shanghai Stock Exchange and the Shenzhen Stock Exchange. I collect information on bank loans to all of the listed firms from China.

I augment the data on bank-firm relationships with bank-level and firm-level data taken from WDS. WDS provides historical reference data, real-time market data, and historical intraday market data, covering stocks, bonds, futures, foreign exchanges, funds, indices, warrants, and macro market data as well as descriptions, real-time market data, financial data, dividend data, corporate actions, and historical intraday data.

In addition, I combine the data set with bank-level information from Almanac of China's Finance and Banking (2001-2016). Almanac of China's Finance and Banking is a high informative yearbook that is supervised by the People's Bank of China, sponsored by China society for finance and banking. Established in 1986, the yearbook has been consecutively published for 29 volumes at the publication frequency of one volume a year.

(ii) Corruption Data. Our corruption-related data set combines two sources of corruption information existing in China: 1. CCDI dataset; 2. Inside dataset at Tsinghua University.

1. CCDI dataset. This is the dataset compiled by the Central Commission for Discipline Inspection (CCDI). The data is provided by the Center for Anti-corruption and Governance at Tsinghua University. This is a cross-sectional dataset for the government officials been investigated and removed from 2012 to 2018 (after the starting of the anti-corruption campaign). Every person in the dataset is known to be investigated and removed, solidly found to be corrupted, and turned to the procuratorate (prosecutors of the Chinese government). In other words, if a government official is investigated but found to be "clean", s/he would not be in this dataset. I have their basic information: province, prefecture, county, administrative level, name, position, violation, and penalty. But I don't have the information regarding when the government officials are investigated and removed. The data contains 282 prefectures with 18453 officials been investigated and removed during the anti-corruption investigation.

2. Inside dataset at Tsinghua University. This is a panel dataset for the officials been investigated and removed from 1994 to 2018. In this dataset, I have the information regarding when the government officials are investigated and removed. Only less than 10% of the government officials were investigated and removed from 1994 to 2012 (before the starting of the anti-corruption campaign). But this dataset is not complete compared with the CCDI dataset, in other words, this dataset is a subset of the CCDI dataset from the year 2012 to 2018. Professors and graduate students at Tsinghua University collect all the data in this dataset by hand (news searching). I use this dataset to supplement the CCDI dataset for the government officials been investigated and removed between 1994 to 2012.

3. CV dataset. I also have the CV panel data for all the government officials regarding their name, age, educational background, and working experience from 1994 to 2018. I merged them using the name, position, prefecture, and year. The assumption is that being corrupt is a persistent characteristic of the individual.

(iii) Luminosity Data. The night-lights data are gathered by Air Force satellites that have been circling the earth 14 times a day since the 1970s, which measure the light intensity emanating from specific geographic pixels. Henderson et al. (2012) argued that the night-lights data are a good proxy for economic activity because the consumption of goods in the evening requires light. The luminosity data have latitudes and longitudes.

Use QGIS to locate them in a China GIS map. Then the area (prefecture) codes will be automatically assigned by QGIS. The luminosity data contains total, average, median, minimum, and maximum levels of light measurement. I also have the information on how may light observations are aggregated into each prefecture.

(iv) Economic Data. I also combine the data set with economic information from GTA CSMAR Database. The economic data set contains all the basic prefecture-level economic characteristics including GDP, government revenue, government spending, employment, FDI, population, financial asset, consumption, loan, deposit, and other economic variables.

2.3.2 Measure and Evolution of Corruption

There are plenty of different methods to measure the corruption, given the data we have, in this study, the Corruption Index is defined as the probability of being investigated and removed by the central government at the prefecture-level, and it is measured as

$$Corruption\ Index_{p,t} = \sum_{o'=1}^{O} W_{o',p,t} \times \mathbb{1}(Corrupted_{o',p,t} = 1) \tag{2.1}$$

where o represents government official for prefecture, p represents prefecture, and t represents year. $Corrupted_{o,p,t}$ equals 1 if official o was investigated and removed, 0 otherwise. $W_{o,p,t}$ is the weights given to each official regarding the characteristics of the position (1/m to leaders, 0.5/m to vice officials for Economics, 0.2/m to others), where $\sum_{o'=1}^{O} W_{o',p,t} = 1$.

To further understand the variations and evolution of corruption in China over time, this paper plots several figures to illustrate. Figure 2.4 plots the corruption index from 2001 to 2016. The corruption index increased from 2001 to 2011, then it peaked in the year around 2012, and finally decreased after 2012. I acknowledge that the trend of corruption index itself may suffer from the construction of index bias. However, I only focus on the variations among different prefectures rather than the average trend.

Figure 2.6 reports the percentiles of the corruption index from 2001 through 2016. I find that the dispersion of the corruption index increased significantly with the percentiles of the index. In other words, the variation of corruption index is highest for the 90% of the corruption index. Moreover, the indexes are peaked around the year 2012 for 60, 70, 80, and 90 percentiles.

Figure 2.7 plots the average variation in our measure of corruption across prefectures from 2001 to 2016. It shows that the northern areas are more corrupted compared with southern areas, which is consistent with the existing literature that southern areas are more developed, market-oriented, transparent, and independent from the central government compared with the northern areas in China.

Since the dispersion among different prefectures over time is what I focus on, I plot the variation in corruption index across prefectures over time (every four years) in Figure 2.8. First and foremost, from the period 2001-2004 to 2005-2008, the dispersion increased substantially. Then, in the next four years, the dispersion increased even more. This uptrend in variations has coincided with the rapid economic growth as well as the rising trend of the openness in China from 2001 to 2012. Lastly, from 2009-2012 to 2013-2016, the dispersion reduced significantly because the anti-corruption investigation started in 2012.

2.3.3 Measure of Misreporting

The misreporting index measures the difference between the reported data and the luminosity data at the prefecture-level, which could be standardized and represented by

$$Misreporting\ Index_{p,t} = ln(\frac{GDP_{p,t}}{sd(GDP_{p,t})}) - ln(\frac{light(sum)_{p,t}}{sd(light(sum)_{p,t})}) \tag{2.2}$$

where p represents prefecture and t represents year. By construction, a higher misreporting index implies the more exaggerated prefecture-level GDP reported by the local government.

2.3.4 Summary Statistics

To reduce the measurement error, we winsorize all variables at the 1% and 99% levels to lessen the influence of outliers.

Table 2.3 presents summary statistics for the loan level variables in our primary data set. Since our data covers the universe of all business loans of listed firms, there is considerable variation in loan sizes. For example, the average loan size is about 358.828 million yuan, and the standard deviation is 2418.99 million yuan. Given the considerable size variation, I choose to use the log loan volume instead of the loan volume. Table 2.3 also shows that the average of the log loan volume of state-owned banks is similar to that of private banks (18.842 vs. 18.548), and the average of the change of the log loan volume of state-owned banks are very similar to that of private banks (11.4% vs. 12.3%).

Table 2.4 presents summary statistics for the bank-level and prefecture-level variables in our data set. It shows that the average corruption index is relatively higher for state-owned banks, which means that state-owned banks are located in more corrupted areas compared with the regions of private banks located. Moreover, the average misreporting index is relatively higher for state-owned banks, which implies that state-owned banks are located in prefectures with more misreporting of GDP compared with the prefectures private banks located.

2.4 Empirical Strategy

2.4.1 Bank Loan Effects

For the level effect of the bank-firm matched loans, I employ the following specification for bank i and firm j in the prefecture p and year t:

$$Y_{ijpt} = \beta_1 Index_{p,t-1} + \beta_2 Index_{p,t-1} \times Post_t + \gamma X_{i,t-1} + \mu_i + \lambda_{jt} + \epsilon_{ijpt} \tag{2.3}$$

where i represents bank, j represents firm,p represents prefecture, t represents year. Y_{ijct} could be the log of the loan amount, the log of the interest rate, the log of the maturity, and the collateral. $Post_t$ is an indicator variable equals to one after 2012, $X_{i,t}$ is the bank level control variables, μ_i is the bank fixed effects, λ_{jt} is the year-firm fixed effects. $Index_{p,t}$ could be $Corruption\ Index_{p,t}$ and $Misreporting\ Index_{p,t}$ respectively. All the bank branches located in the same prefecture share the unique $Corruption\ Index_{p,t}$, and the prefecture level $Corruption\ Index_{p,t}$ is the probability of been investigated and removed because of corruption. All the bank branches located in the same prefecture share the unique $Misreporting\ Index_{p,t}$, and the prefecture level $Misreporting\ Index_{p,t}$ is the standardized differences between GDP and the luminosity data. Since we focus on the banks' behavior over time, the standard errors are clustered at the bank level and robust to heteroskedasticity.

2.4.2 Year-specific Effects

To further explore the relationship between the bank loan amount and corruption, I employ the following specification for bank i and firm j in year t for the year-specific effects.

$$L_{ijpt} = \beta_1 Index_{p,t-1} \times Year\ Dummy_t + \gamma X_{i,t-1} + \lambda_{jt} + \epsilon_{ijpt} \tag{2.4}$$

where i represents bank, j represents firm, t represents year, L_{ijpt} is the log loan size, Year Dummy equals to one for each specific year, otherwise it equals to zero, $X_{i,t-1}$ is the bank-level control variables, λ_{jt} is the year-firm fixed effects.

2.4.3 Loan Allocation Effects

For the level effect of the bank-firm matched loans, I employ the following specification for bank i and firm j in the prefecture p and year t:

$$\begin{aligned} L_{ijpt} &= \beta_1 Index_{p,t-1} + \beta_2 Index_{p,t-1} \times Post_t \\ &\quad + \beta_3 Index_{p,t-1} \times Post_t \times State\text{-}owned(Firm)_{jt} + \mu_i + \lambda_{jt} + \epsilon_{ijpt} \end{aligned} \tag{2.5}$$

where i represents bank, j represents firm,p represents prefecture, t represents year. Y_{ijct} could be the log of the loan amount, the log of the interest rate, the log of the maturity, and the collateral. $Post_t$ is an indicator variable equals to one after 2012, $X_{i,t}$ is the bank level control variables, μ_i is the bank fixed effects, λ_{jt} is the year-firm fixed effects. $Index_{p,t}$ could be $Corruption\ Index_{p,t}$ and $Misreporting\ Index_{p,t}$ respectively. All the bank branches located in the same prefecture share the unique $Corruption\ Index_{p,t}$, and the prefecture level $Corruption\ Index_{p,t}$ is the probability of been investigated and removed because of corruption. All the bank branches located in the same prefecture share the unique $Misreporting\ Index_{p,t}$, and the prefecture level $Misreporting\ Index_{p,t}$ is the standardized differences between GDP and the luminosity data. $State\text{-}owned(Firm)_{jt}$ equals one if the firm is a state-owned firm. Since we focus on the banks' behavior over time, the standard errors are clustered at the bank level and robust to heteroskedasticity.

2.4.4 Identification Strategy

The identification strategy of this paper could be summarized in four dimensions: omitted variable, reverse causality, measurement error, and sample selection.

I use the Khwaja and Mian (2008) technique to simultaneously estimate the bank lending and firm borrowing channels stems from identification concerns. Identification concerns arise because events that trigger changes in liquidity supply, such as monetary policy innovations or financial shocks are often accompanied by changes in investment returns and consequently, credit demand. Changes in firm borrowing, therefore, reflect both changes in credit supply as well as credit demand. This paper uses year-firm fixed effects to control credit shocks from the demand side.

Another concern about the omitted variable problem is the heterogeneity in bank response to political shocks. Could the lending channel coefficient be driven by inherent differences in how banks respond to the shocks induced by the anti-corruption campaign in China? It is possible if there is such response heterogeneity, and it is systematically correlated with a bank's liquidity shock. For example, perhaps the lending channel estimate is picking up differences in how state-owned and private banks react to political shocks since we know that the Chinese government has more control power on state-owned banks.

We also test for such concerns by including various bank characteristics that proxy for such differential lending sensitivity as controls, such as the bank size, the bank's return on assets, the bad loan rate, the amount lend from central bank, tangibility, cash flow, and dummies state-owned, policy, rural and listed banks. These bank-level controls are likely to capture a banks' sensitivity to political shocks. In particular, I use lagged value to avoid concern about the endogeneity problem.

Although firm fixed effects address the main identification concerns expressed in the literature, there may remain some additional questions. While the fixed effects strategy does not require or make any assumptions about the correlation between liquidity supply and demand shocks, the concern about the reverse causality problem is that if the liquidity supply shocks are anticipated, banks may adjust their lending or firms adjust their borrowing prior to the shock. This would lead to either an under or overestimate of the bank lending channel depending on the direction of the pre-shock loan adjustments. However, in this paper, the natural experiment anti-corruption campaign is unanticipated, the anti-corruption is happened out of the financial market. Therefore, it is difficult for Chinese banks and firms to anticipate this kind of liquidity supply shock. In particular, the identifying assumption for all the level effect regressions in this paper is that year-before financial positions are not positively correlated with unobserved within-bank changes in loans lending following the onset of the campaign.

To reduce the measurement Error, I winsorize all variables at the 1% and 99% level to lessen the influence of outliers. For the sample selection problem, my data set provides more comprehensive coverage of small, micro, and rural banks than other data sets. I include all of the banks, both listed and non-listed banks. A primary concern about the sample selection problem in my paper is that my data set only provides the information of the listed firms, there are also many small and micro firms (non-listed) in China. Gertler and Gilchrist (1994) suggest that since size may proxy for financial constraints, a higher sensitivity of small firms would provide evidence in favor of the "financial accelerator", the view that financial frictions can amplify downturns. Mehrotra et al. (2017) use new and confidential data on income statements and balance sheets of US manufacturing firms to bear on this idea. Thus, my analysis of the impact of the anti-corruption campaign on Chinese bank lending channel could be regarded as the "lower bound" impact, since we only consider the listed firms, if listed firms are affected by the anti-corruption campaign, small and micro firms should be affected more.

2.5 Bank Loan Effects of the Corruption

2.5.1 Baseline Results

Table 2.5 reports results from the regression model given by equation (2.3). Column (3) in Table 2.5 shows that one percentage increased of the leaders' probability of being investigated and removed is associated with 0.161% decline in the bank loan amount, and this reduction effect changes the direction and becomes significantly positive after the anti-corruption campaign in 2012. Column (6) in Table 2.5 presents that one additional increased of the difference between the log of the economic and luminosity data is associated with 0.122% decline in the bank loan amount, and this reduction effect changes the direction and becomes significantly positive after the anti-corruption campaign in 2012.

The estimated effects of interest rate is reported in Table 2.9. Column (3) in Table 2.9 shows that one percentage increased of the leaders' probability of being investigated and removed is associated with 0.900% increase in the interest rate, and this increment becomes significantly negative after the anti-corruption campaign in 2012. Column (6) in Table 2.5 presents that one additional increased of the difference between the log of the economic and luminosity data is associated with 0.103% increase in the interest rate, and this increment becomes significantly negative after the anti-corruption campaign in 2012.

Table 2.12 focus on the maturity of the loan. Column (3) in Table 2.12 shows that one percentage increased of the leaders' probability of being investigated and removed is associated with 0.637% increase in the maturity, and this increment significantly changes the direction after the anti-corruption campaign in 2012. Column (6) in Table 2.12 presents that one additional increased of the difference between the log of the economic and luminosity data is associated with 0.487% increase in the maturity, and this increment significantly changes the direction after the anti-corruption campaign in 2012.

Table 2.13 discusses the influence on having the collateral or not. Column (3) in Table 2.13 shows that one percentage increased of the leaders' probability of being investigated and removed is associated with 0.106% increase in the collateral, and this increment could be offset significantly after the anti-corruption campaign in 2012. Column (6) in Table 2.13 presents that one additional increased of the difference between the log of the economic and luminosity data is associated with 0.120% increase in the collateral, and this increment could be offset significantly after the anti-corruption campaign in 2012.

2.5.2 Year-specific Effects

Figure 2.10 plots the estimations of the year-specific effects from the equation (2.4). We find that the coefficient before the interaction term corruption index and year dummy, β_1, increased substantially around the year 2012. Therefore, we find that banks located in a more corrupted prefecture will increase their credit supply significantly in the starting year of the anti-corruption campaign.

2.5.3 Loan Allocation Effects

In part 2.5.1, we find that the banks located in the more corrupted areas increase lending more after the anti-corruption investigation. In this part, we discuss where the increase in the credit supply goes? Does it go to state-owned firms or private firms? Table 2.14 reports results from the regression model given by equation (2.5). The estimated results in Table 2.14 shows that banks located in more corrupted areas are more likely to increase lending after the anti-corruption campaign. Moreover, they have a preference to reallocate toward private firms, but not significantly so. In addition, the results in table 2.14 also indicates that banks located in areas with more misreporting of GDP will increase lending more after the anti-corruption campaign, and they are more likely to reallocate toward private firms, but not significantly so. my results in Table 2.14 are consistent with the main findings in Li et al. (2018).

2.5.4 Robustness Analysis

Dealing with the Endogeneity Problem of Indices

We just saw that our estimates are robust to the inclusion of a wide range of controls and fixed effects. We now report results from additional specification checks to further increase our confidence in the estimates. To further address the concern of the possible endogeneity problem of the corruption and misreporting index, we use the index in the initial year as well as the average of the entire pre-shock years of my sample to measure the corruption and misreporting. Table 2.6 reports estimates from the regression model given by equation (1.7) using the corruption or misreporting index of the initial year instead of lagged one year, the results for the loan amount are nearly identical to our baseline estimates. Table 2.7 shows results from equation (1.7) using the corruption or misreporting index of the average of all the pre-shock years, again we find that these are close to our baseline estimates.

Moreover, Table 2.10 reports the interest rate effects given by equation (1.7) using the initial year corruption or misreporting index, we find similar results as the baseline. Table 2.11 reports the interest rate effects also given by equation (1.7) using the corruption or misreporting index of the average of all the pre-shock years. It is reassuring to find no evidence that our results are driven by the endogeneity bias of the corruption or misreporting index.

Dealing with the Tight Fixed Effects

I employ the firm-year interacted fixed effects, λ_{jt}. Such tight firm-level fixed effects could help us focus primarily on the variation in loans due to banks. However, there is the concern that a given Chinese manufacturing firm receives only one credit from only one prefecture per year. The tight fixed effects hurt my fixed results estimation quite a bit as substantial variation is eliminated (there are no cross-bank loans; just one loan from one prefecture in a given year). To alleviate such concerns, first of all, I calculate the percentage of the firms in my sample take multiple loans per year from different prefectures. This number is 61.85%, which is higher than 60%, and the average number of years for those firms is 3.13. Thus the tight fixed effects in this paper should be warranted. However, in order to avoid my fixed firm-year interacted effects model could do more harm than bring improvements to my causal identification.

Therefore, I employ the estimation of the firm borrowing channel to make sure that the whole sample has been considered.

The exposure of firm j to local corruption index is measured by

$$\overline{Exposure}_{j,t} = \sum_{p=1}^{P} (Index_{p,t} \times \frac{L_{pjt}}{\sum_{p'=1}^{P} L_{p'jt}}) \tag{2.6}$$

Let Y_{jt} be a firm-level attribute of interest in period t (such as a firm's net debt, log cash, log sales, log capital investment, and log employment.) The reduced form firm borrowing channel can be determined by estimating the following equation:

$$\begin{aligned} Y_{jt} \quad = \quad & \beta_1^F \overline{Exposure}_{j,t-1} + \beta_2^F \overline{Exposure}_{j,t-1} \times Post_t \\ & + \beta_3^F \frac{CF_{j,t-1}}{Assets_{j,t-2}} + \beta_4^F \frac{Sales_{j,t-1}}{Assets_{j,t-2}} + \gamma X_{j,t-1} + \lambda_j + \mu_t + \eta_{jt} \end{aligned} \tag{2.7}$$

where j represents firm, p represents prefecture, and t represents year. Y_{jt} is a firm-level attribute of interest in period t. $Post_t$ equals to one after 2012. $X_{j,t-1}$ is the firm level control variables. While λ_j is the firm fixed effects, μ_t is the year fixed effects.

2.6 Mechanism

The corruption index could disclose the extent of the market competition or marketization. The corrupted banks bribe the government of the local prefecture to obtain market power. Banks with enough market power could have more ability to set prices and control quantity. Thus the loan amounts and interest rates could be affected by the extent of the market competition. Banks in the corrupted areas may have higher markup and lower credit supply.

Moreover, firms who want to acquire credits with banks with relatively higher market power need to pay the higher fixed cost (entry cost) and higher variable cost (interest rate). Thus they are more likely to use collateral and set longer maturity. In addition, banks in the corrupted areas may have lower efficiency and higher default risk. Thus they may suffer from higher bad loans. Therefore, this section explores the mechanisms behind the effect of bank loans relating to the bank concentration effects and bank health effects.

2.6.1 Bank Concentration Effects

To justify this mechanism, we examine the effects of corruption and misreporting on both the market share of the top three banks and the HHI index in the specific prefecture as follows.

Bank Share

I employ the bank share to measure the bargaining power of the banks in a specific prefecture in China. The Bank share is a variable used to measure the share of bank i in year t and prefecture p, and it could be written as

$$Bank\ Share_{it} = \frac{Size_{it}}{Total\ size_{pt}} \tag{2.8}$$

Size could be bank assets or loans. Here, we consider not only assets but loans because different banks in China have certain various operations, some banks may be small regarding the assets, but they may provide a considerable amount of the credit supply. Since we focus on bank lending, we need to consider the size of the bank loans.

For the bank-level information, I employ the following specification for bank i in the prefecture p and year t:

$$Bank\ Share_{ipt} = \beta_1 Index_{p,t-1} + \beta_2 Index_{p,t-1} \times Post_t + \gamma X_{i,t-1} + \mu_p + \lambda_t + \epsilon_{ipt} \tag{2.9}$$

where i represents bank, p represents prefecture, t represents year, Bad Loan$_{i,t}$ is (Subprime loan$_{i,t}$+Doubt loan$_{i,t}$+Loss loan$_{i,t}$)/Asset$_{i,t}$ or (Subprime loan$_{i,t}$+Doubt loan$_{i,t}$+Loss loan$_{i,t}$)/Total Loan$_{i,t}$, $Post_t$ is an indicator variable equals to one after 2012, $X_{i,t}$ is the bank level control variables, μ_p is the prefecture fixed effects, λ_t is the year fixed effects. $Index_{p,t}$ could be $Corruption\ Index_{p,t}$ and $Misreporting\ Index_{p,t}$ respectively. All the bank branches located in the same prefecture share the unique $Corruption\ Index_{p,t}$, and the prefecture level $Corruption\ Index_{p,t}$ is the probability of been investigated and removed because of corruption. All the bank branches located in the same prefecture share the unique $Misreporting\ Index_{p,t}$, and the prefecture level $Misreporting\ Index_{p,t}$ is the standardized differences between GDP and the luminosity data. Since our main variable of interest varies at the prefecture level over time, the standard errors are clustered at the prefecture level and robust to heteroskedasticity.

When calculating the bank share with the bank asset, column (1) in Table 2.15 shows that one percent increased of the leaders' probability of being investigated and removed is associated with 0.307% increase in the bank share, and this incremental effect significantly changes the direction after the anti-corruption campaign in 2012. Column (3) in Table 2.15 presents that one additional increase of the difference between the log of the economic and luminosity data is associated with a 0.360% increase in the bank share, and this incremental effect significantly changes the direction after the anti-corruption campaign in 2012.

In addition, when measuring the bank share with the total loans, column (1) in Table 2.15 shows that one percent increased of the leaders' probability of being investigated and removed is associated with 0.258% increase in the bank share, and this incremental effect could be offset significantly after the anti-corruption campaign in 2012. Column (3) in Table 2.15 presents that one additional increase of the difference between the log of the economic and luminosity data is associated with a 0.358% increase in the bank share, and this incremental effect could be offset significantly after the anti-corruption campaign in 2012.

HHI Index

Since the analysis of the bank share in part 2.6.1 only consider the change of the share for the top three banks, we need to discuss the effects to all the banks to be complete. Therefore, I further use the HHI index to indicate the monopoly power of the banks in a specific prefecture in China. The Herfindahl-Hirschman Index of the banks at the prefecture-level could be calculated as

$$HHI_{pt} = (s_1^2 + s_2^2 + s_3^2 + ... + s_n^2)/10000 \tag{2.10}$$

where s_n is the market share percentage of bank n. For the prefecture-level information, I employ the following specification for prefecture p and year t:

$$HHI_{pt} = \beta_1 Index_{p,t-1} + \beta_2 Index_{p,t-1} \times Post_t + \mu_p + \lambda_t + \epsilon_{pt} \tag{2.11}$$

where p represents prefecture and t represents year. $Post_t$ is an indicator variable equals to one after 2012, $X_{i,t}$ is the bank level control variables, μ_p is the prefecture fixed effects, λ_t is the year fixed effects. $Index_{p,t}$ could be $Corruption\ Index_{p,t}$ and $Misreporting\ Index_{p,t}$ respectively. All the bank branches located in the same prefecture share the unique $Corruption\ Index_{p,t}$, and the prefecture level $Corruption\ Index_{p,t}$ is the probability of been investigated and removed because of corruption. All the bank branches located in the same prefecture share the unique $Misreporting\ Index_{p,t}$, and the prefecture level $Misreporting\ Index_{p,t}$ is the standardized differences between GDP and the luminosity data. SInce our main variable of interest varies at the prefecture level over time, the standard errors are clustered at the prefecture level and robust to heteroskedasticity.

Table 2.16 presents the bank concentration effects estimated using the Herfindahl-Hirschman index. Column (1) calculates the HHI index using the bank assets, and we find that the HHI index was higher in the more corrupted areas before the anti-corruption investigation. The HHI index decreased substantially in the more corrupted areas after the anti-corruption campaign. Column (2) recalculates the HHI index via the bank loans, and the results barely move, if anything, the estimates became more substantial. Moreover, column (3) shows that the HHI index measured by asset was higher in the prefectures with more misreporting of GDP before the anti-corruption campaign, and the HHI index reduced significantly in the prefectures with a high level of misreporting after the anti-corruption investigation. Column (4) remeasured the HHI index using the bank loans, and the results are quite consistent with column (3); indeed, the results are more significant.

To sum up, I find that banks located in more corrupted prefectures or prefectures with more misreporting of GDP have a more substantial monopoly or bargaining power, higher markups as well as lower efficiencies. As we have seen in section 2.5, these monopoly effects lead to the higher interest rate and lower bank loan amount in more corrupted areas before the anti-corruption investigation. After 2012 the phenomena of higher price and lower quantity disappears because of the collapse of the monopoly or bargaining power.

2.6.2 Bank Health Effects

Does corruption affect bank health? We use the bad loan to disclose the effect of corruption and misreporting on the bank's balance sheet. The bad credit is a variable used to measure the health of a specific bank i in year t, and it is constructed as

$$Bad\ Loan_{it} = \frac{Subprime\ loan_{it} + Doubt\ loan_{it} + Loss\ loan_{it}}{Size_{it}} \tag{2.12}$$

where size could be measured by assets or total loans. Similarly, we consider both the assets and credits to rule out the condition of a bank with small assets and a massive amount of credit supply. In this case, even if the bad loans regarding the assets is high, it may be acceptable when considering the credit amounts.

For the bank-level information, I employ the following specification for bank i in the prefecture p and year t:

$$Bad\ Loan_{ipt} = \beta_1 Index_{p,t-1} + \beta_2 Index_{p,t-1} \times Post_t + \gamma X_{i,t-1} + \mu_p + \lambda_t + \epsilon_{ipt} \tag{2.13}$$

where i represents bank, p represents prefecture, t represents year, Bad Loan_{ipt} is (Subprime loan_{ipt}+Doubt loan_{ipt}+Loss loan_{ipt})/Asset_{ipt} or (Subprime loan_{ipt}+Doubt loan_{ipt}+Loss loan_{ipt})/Total Loan_{ipt}, Bank Share_{ipt} is $\text{Asset}_{ipt}/\sum_{i=1}^{I}\text{Asset}_{ipt}$ or $\text{Loan}_{ipt}/\sum_{i=1}^{I}\text{Loan}_{ipt}$, $Post_t$ is an indicator variable equals to one after 2012, $X_{i,t}$ is the bank level control variables, μ_p is the prefecture fixed effects, λ_t is the year fixed effects. $Index_{p,t}$ could be $Corruption\ Index_{p,t}$ and $Misreporting\ Index_{p,t}$ respectively. All the bank branches located in the same prefecture share the unique $Corruption\ Index_{p,t}$, and the prefecture level $Corruption\ Index_{p,t}$ is the probability of been investigated and removed because of corruption. All the bank branches located in the same prefecture share the unique $Misreporting\ Index_{p,t}$, and the prefecture level $Misreporting\ Index_{p,t}$ is the standardized differences between GDP and the luminosity data. Since our main variable of interest varies at the prefecture level over time, the standard errors are clustered at the prefecture level and robust to heteroskedasticity.

When re-scaling the bad loan with the bank asset, column (1) in Table 2.17 shows that one percent increase of the leaders' probability of being investigated and removed is associated with 0.486% increase in the bad loan as a fraction of the asset, and this incremental effect could be offset significantly after the anti-corruption campaign in 2012. Column (3) in Table 2.17 presents that one additional increased of the difference between the log of the economic and luminosity data is associated with 0.542% increase in the bad loan as a fraction of the asset, and this incremental effect could be offset significantly after the anti-corruption campaign in 2012.

Furthermore, when re-scaling the bad loan with the total credit, column (2) in Table 2.17 shows that one percent increased of the leaders' probability of being investigated and removed is associated with 0.216% increase in the bad loan as a fraction of the total loan, and this incremental effect could be offset significantly after the anti-corruption campaign in 2012. Column (4) in Table 2.17 presents that one additional increased of the difference between the log of the economic and luminosity data is associated with 1.128% increase in the bad loan as a fraction of the total loan, and this incremental effect could be offset significantly after the anti-corruption campaign in 2012.

Therefore, we find that the anti-corruption investigation indeed improves the banks' health reflected in the balance sheet. This finding is consistent with the results in part 2.5.3. Banks located in the more corrupted prefectures choose to reallocate more credits to the private firms, which have higher profitability and debt paying ability compared with the state-owned firms.

2.7 Conclusion

In this paper, I showed that banks located in more corrupted prefectures offer significantly less credits before the anti-corruption investigation, and this effect changes the direction after the investigation. Moreover, banks located in more corrupted prefectures tend to use higher interest rates, longer maturity, more collateral, higher bad loans before the campaign; all of these effects change the direction after the campaign.

The corruption index indicates the extent of the market competition (marketization). The corrupted banks bribe the local government to obtain market power. First of all, higher corruption leads to less loan supply and higher interest rates through the channel of less competition and higher markup. Secondly, higher corruption leads to higher interest rates through the channel of less competition, lower efficiency, and higher variable cost. Thirdly, higher corruption leads to longer maturity through the channel of less competition, lower efficiency, and higher fixed cost. This monopoly effect could be proved by that the bank concentration ratio is higher in the more corrupted areas, and this effect disappears after the campaign.

In my analysis, I took advantage of the confidential and novel data sets, covering a large number of small, privately owned, and rural banks, with information on firm-bank relationships and prefecture-level corruption index. This paper is the first paper to analyze the impact of the anti-corruption investigation on the Chinese bank lending channel using a prefecture-level corruption index.

My findings foster the understanding of the unfolding of the impact of the anti-corruption investigation on the Chinese bank lending channel. My results also indicate that both markets oriented and government-oriented effects are critical in the financial market in China. Moreover, my contribution provides evidence on the crucial role of monopoly power as a determinant of bank loan allocation. Furthermore, the analysis of which industry is affected most by the anti-corruption campaign in the future would be very meaningful and interesting.

2.8 Tables

Prefecture List

Anshan
Anyang
Baotou
Beijing (Districts)
Cangzhou
Changchun
Changsha
Changzhou
Chaoyang
Chengde
Chengdu
Chongqing (Districts)
Dalian
Datong
Dazhou
Deyang
Dezhou
Dongguan
Dongying
Foshan
Fushun
Fuxin
Fuzhou
Guangzhou
Guilin
Guiyang
Haerbin
Handan
Hangzhou
Hefei
Huhehaote
Huludao
Huzhou
Jiangmen
Jiaozuo
Jiaxing
Jinan
Jincheng
Jinhua
Jining
Jinzhou
Jiujiang
Kaifeng
Kelamayi
Kunming
Laiwu
Lanzhou
Leshan
Linyi
Liuzhou
Longyan
Luoyang
Luzhou
Mianyang
Nanchang
Nanchong
Nanjing
Nanning
Ningbo
Ningde
Panzhihua
Pingdingshan
Qingdao
Qinhuangdao
Quanzhou
Qujing
Rizhao
Shanghai (Districts)
Shantou
Shaoxing
Shenyang
Shenzhen
Shijiazhuang
Shizuishan
Suining
Suzhou
Taian
Taiyuan
Taizhou
Tangshan
Tianjin (Districts)
Tongling
Weifang
Weihai
Wenzhou
Wuhan
Wuhu
Wulumuqi
Wuxi
Xiamen
Xining
Yancheng
Yangzhou
Yantai
Yibin
Yinchuan
Yingkou
Yuxi
Zaozhuang
Zhanjiang
Zhengzhou
Zhuhai
Zibo
Zigong

Table 2.1 Variable Definitions

Variable	Definition
Bank-level	
Dependent Variables (winsorized at the 1% level)	
L	Natural logarithm of bank loans
ΔL	ln(bank loan$_{t+1}$−bank loan$_t$)
$Bad\ Loan_{i,t}$ (Asset)	(Subprime loan$_{i,t}$+Doubt loan$_{i,t}$+Loss loan$_{i,t}$)/Asset$_{i,t}$
$Bad\ Loan_{i,t}$ (Total loan)	(Subprime loan$_{i,t}$+Doubt loan$_{i,t}$+Loss loan$_{i,t}$)/Total loan$_{i,t}$
$Bank\ Share_{i,t}$ (Asset)	Asset$_{i,t}$/Total asset$_{p,t}$
$Bank\ Share_{i,t}$ (Loan)	Loan$_{i,t}$/Total loan$_{p,t}$
$HHI_{p,t}$ (Asset)	$(s_{1A}^2 + s_{2A}^2 + s_{3A}^2 + ... + s_{nA}^2)/10000$
$HHI_{p,t}$ (Total loan)	$(s_{1TL}^2 + s_{2TL}^2 + s_{3TL}^2 + ... + s_{nTL}^2)/10000$
Key Explanatory Variables (winsorized at the 1% level)	
$Post_t$	Dummy equal to one if the year is after 2012.
$Bad\ Loan_{i,t}$	(Subprime loan$_{i,t}$+Doubt loan$_{i,t}$+Loss loan$_{i,t}$)/Asset$_{i,t}$
$Corruption\ Index_{i,t}$	$\sum_{o'=1}^{O} W_{o',p,t} \times \mathbb{1}(Corrupted_{o',p,t} = 1)$
$Misreporting\ Index_{i,t}$	$ln(\frac{GDP_{p,t}}{sd(GDP_{p,t})}) - ln(\frac{light(sum)_{p,t}}{sd(light(sum)_{p,t})})$
$Trade\ Settlement_{i,t}$	Trade settlements$_{i,t}$/ Total loan$_{i,t}$
$(Exchange/Income)_{i,t}$	Exchange gains$_{i,t}$/Total income$_{i,t}$
Control Variables (winsorized at the 1% level)	
$State\text{-}owned_i$	Dummy equals to one if the bank is a state-owned bank.
$Policy_i$	Dummy equals to one if the bank is a policy bank.
$Rural_i$	Dummy equals to one if the bank is a rural bank.
$List_{i,t}$	Dummy equals to one if the bank is a listed bank.
$Size_{i,t}$	ln(Assets$_{i,t}$)
$Roa_{i,t}$	Net income$_{i,t}$/((Assets$_{i,t-1}$+Assets$_{i,t}$)/2)
$Tangibility_{i,t}$	Fixed assets$_{i,t}$/Assets$_{i,t}$
$Cash\ Flow_{i,t}$	Operating Income Before Depreciation$_{i,t}$/((Assets$_{i,t-1}$+Assets$_{i,t}$)/2)
$Lend\ Central\ Bank_{i,t}$	Debt lent from central bank$_{i,t}$/Assets$_{i,t}$
Firm-level	
Dependent Variables (winsorized at the 1% level)	
$Net\ Debt$	(Current+Non-Current Liabilities−Cash)/(Total Assets)
$\Delta Cash$	(Cash$_{t+1}$−Cash$_t$)/(Total Asset$_t$)
$Employment\ Growth$	ln(Employment$_t$)−ln(Employment$_{t-1}$)
$CAPX$	(Fixed Assets$_{t+1}$−Fixed Assets$_t$+Depreciation$_t$)/Assets$_t$, set to 0 if negative
Key Explanatory Variables (winsorized at the 1% level)	
$\overline{TL}_{j,t-1}$	$\frac{1}{I}\sum_{i=1}^{I}$Total Loan$_{ijt}$
$State\text{-}owned_j$	Dummy equals to one if the firm is a state-owned firm.
Control Variables (winsorized at the 1% level)	
$Sales\ Growth$	ln(Sales$_t$)−ln(Sales$_{t-1}$)
$Cash\ Flow_{i,t}$	Operating Income Before Depreciation$_{i,t}$/Assets$_{i,t-1}$

Table 2.2 Financial Development: China and US (USD Trillion) (2016)

Sector	Bank Credit	Stock	Fixed Income	Insurance	Investment Funds
Size (China)	15.45	7.32	10.43	2.19	1.32
Size (US)	12.44	27.35	39.36	8.46	19.20
% GDP (China)	137.95%	64.30%	93.12%	19.55%	11.76%
% GDP (US)	67.00%	147.40%	211.95%	45.56%	103.39%

Table 2.3 Summary Statistics (Firm-bank Pairwise)

	State-owned Banks			Private Banks			All Banks		
	mean	sd	count	mean	sd	count	mean	sd	count
loan	502.619	3,717.35	9,840	264.614	780.71	15,018	358.828	2,418.99	24,858
lnloan	18.842	1.38	9,840	18.548	1.22	15,018	18.664	1.29	24,858
dlnloan	0.114	0.91	4,822	0.123	0.77	6,655	0.119	0.83	11,477
Observations	9840			15018			24858		

Notes. This table presents descriptive statistics of firm-bank pairwise dependent variables split into state-owned and private banks. State-owned is an indicator variable equal to one if the bank is a state-owned bank. The sample consists of all firms that are listed in the A-share, B-share, H-share, and oversea stocks market.

Table 2.4 Summary Statistics (Bank and Prefecture Level)

	State-owned Banks			Private Banks			All Banks		
	mean	sd	count	mean	sd	count	mean	sd	count
Corruption Index	0.105	0.03	122	0.084	0.17	912	0.087	0.16	1,034
Misreporting Index	0.048	0.42	106	-0.173	0.60	658	-0.142	0.58	764
Size	28.367	1.83	103	25.401	1.32	756	25.757	1.69	859
State-owned (Bank)	1.000	0.00	122	0.000	0.00	912	0.118	0.32	1,034
Policy	0.344	0.48	122	0.000	0.00	912	0.041	0.20	1,034
Rural	0.000	0.00	122	0.154	0.36	912	0.135	0.34	1,034
Observations	122			912			1034		

Notes. This table presents descriptive statistics of bank-level explanatory variables split into state-owned and private banks. State-owned is an indicator variable equal to one if the bank is a state-owned bank. The sample consists of all banks that are located in China.

Table 2.5 Bank Lending Channel (Level Effect)

Log(Loan)	(1) Index (Corruption)	(2) Index (Corruption)	(3) Index (Cor.)	(4) Index (Misreporting)	(5) Index (Misreporting)	(6) Index (Misr.)
State-owned	0.160***	0.108*		0.0970**	0.0160	
	(0.0530)	(0.0600)		(0.0442)	(0.0788)	
Policy	0.512***	0.794***		0.493***	0.704***	
	(0.123)	(0.0804)		(0.0101)	(0.0309)	
Rural	0.161*	0.0340		0.143	-0.00184	
	(0.0888)	(0.147)		(0.116)	(0.211)	
Size	0.0809***	0.0982***		0.0932***	0.130***	
	(0.0160)	(0.0214)		(0.0189)	(0.0314)	
Index	-0.210*	-0.199	-0.161*	-0.135***	-0.173***	-0.122***
	(0.120)	(0.143)	(0.0931)	(0.0319)	(0.0643)	(0.0280)
Index x Post	0.317**	0.468***	0.314**	0.0721**	0.0615*	0.0638*
	(0.124)	(0.133)	(0.158)	(0.0363)	(0.0362)	(0.0358)
Observations	23630	23630	23296	13485	13485	12999
R^2	0.576	0.112	0.579	0.576	0.119	0.577
Bank Fixed Effect	NO	NO	YES	NO	NO	YES
Firm Fixed Effect	NO	NO	NO	NO	NO	NO
Year Fixed Effect	NO	YES	NO	NO	YES	NO
Firm-Year Fixed Effect	YES	NO	YES	YES	NO	YES
Clusters at Bank Level	162	162	160	144	144	141

Notes. This table presents the results of a Chinese bank lending channel regression. The unit of observation is a bank-firm-year. The dependent variable is the level of log loan volume. All regressions include fixed effects. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Robust standard errors given below coefficient estimates are clustered at the bank level.

Table 2.6 Bank Lending Channel (Level Effect) (Initial Year)

Log(Loan)	(1) Index (Corruption)	(2) Index (Corruption)	(3) Index (Cor.)	(4) Index (Misreporting)	(5) Index (Misreporting)	(6) Index (Misr.)
State-owned	0.160***	0.0931***		0.115***	0.0390	
	(0.0208)	(0.0268)		(0.0285)	(0.0344)	
Policy	0.509***	0.807***		0.471***	0.670***	
	(0.0455)	(0.0615)		(0.0479)	(0.0662)	
Rural	0.161***	0.0351		0.148*	0.0155	
	(0.0495)	(0.0679)		(0.0789)	(0.105)	
Size	0.0797***	0.101***		0.0703***	0.0969***	
	(0.00684)	(0.00847)		(0.0107)	(0.0124)	
Index	-0.120	-0.103		-0.173***	-0.269***	
	(0.0920)	(0.117)		(0.0427)	(0.0479)	
Index x Post	0.109**	0.275***	0.120*	0.0789**	0.0748**	0.0603*
	(0.0471)	(0.0642)	(0.0667)	(0.0308)	(0.0371)	(0.0324)
Observations	23630	23630	23296	13485	13485	12999
R^2	0.576	0.112	0.579	0.577	0.120	0.577
Bank Fixed Effect	NO	NO	YES	NO	NO	YES
Firm Fixed Effect	NO	NO	NO	NO	NO	NO
Year Fixed Effect	NO	YES	NO	NO	YES	NO
Firm-Year Fixed Effect	YES	NO	YES	YES	NO	YES
Clusters at Bank Level	162	162	160	144	144	141

Notes. This table presents the results of a Chinese bank lending channel regression. The unit of observation is a bank-firm-year. The dependent variable is the level of log loan volume. All regressions include fixed effects. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Robust standard errors given below coefficient estimates are clustered at the bank level.

Table 2.7 Bank Lending Channel (Level Effect) (Average of the pre-shock Years)

Log(Loan)	(1) Index (Corruption)	(2) Index (Corruption)	(3) Index (Cor.)	(4) Index (Misreporting)	(5) Index (Misreporting)	(6) Index (Misr.)
State-owned	0.159***	0.0802**		0.0957**	0.0132	
	(0.0313)	(0.0322)		(0.0460)	(0.0309)	
Policy	0.513***	0.813***		0.489***	0.699***	
	(0.0683)	(0.0560)		(0.133)	(0.110)	
Rural	0.161***	0.0336		0.150***	0.0115	
	(0.0312)	(0.0594)		(0.0477)	(0.124)	
Size	0.0808***	0.103***		0.0891***	0.125***	
	(0.00501)	(0.00902)		(0.0116)	(0.00994)	
Index	-0.108	-0.114		-0.193***	-0.255***	
	(0.0817)	(0.0746)		(0.0633)	(0.0685)	
Index x Post	0.102***	0.234***	0.109**	0.0670*	0.0629*	0.0603*
	(0.0334)	(0.0272)	(0.0547)	(0.0366)	(0.0381)	(0.0305)
Observations	23630	23630	23296	13485	13485	12999
R^2	0.576	0.112	0.579	0.577	0.119	0.577
Bank Fixed Effect	NO	NO	YES	NO	NO	YES
Firm Fixed Effect	NO	NO	NO	NO	NO	NO
Year Fixed Effect	NO	YES	NO	NO	YES	NO
Firm-Year Fixed Effect	YES	NO	YES	YES	NO	YES
Clusters at Bank Level	162	162	160	144	144	141

Notes. This table presents the results of a Chinese bank lending channel regression. The unit of observation is a bank-firm-year. The dependent variable is the level of log loan volume. All regressions include fixed effects. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Robust standard errors given below coefficient estimates are clustered at the bank level.

Table 2.8 Bank Lending Channel (Level Effect) (Coefficient of Variation)

Log(Loan)	(1) Index (Corruption)	(2) Index (Corruption)	(3) Index (Cor.)	(4) Index (Misreporting)	(5) Index (Misreporting)	(6) Index (Misr.)
State-owned	0.152***	0.0921		0.0970**	0.0160	
	(0.0330)	(0.0631)		(0.0442)	(0.0788)	
Policy	0.517***	0.804***		0.493***	0.704***	
	(0.0395)	(0.0820)		(0.0101)	(0.0309)	
Rural	0.161**	0.0329		0.143	-0.00184	
	(0.0771)	(0.145)		(0.116)	(0.211)	
Size	0.0821***	0.101***		0.0932***	0.130***	
	(0.0123)	(0.0217)		(0.0189)	(0.0314)	
Index	-0.197**	-0.132	-0.104	-0.135***	-0.173***	-0.122***
	(0.0977)	(0.166)	(0.120)	(0.0319)	(0.0643)	(0.0280)
Index x Post	0.196**	0.327***	0.159*	0.0721**	0.0615*	0.0638*
	(0.100)	(0.110)	(0.0962)	(0.0363)	(0.0362)	(0.0358)
Observations	23630	23630	23296	13485	13485	12999
R^2	0.575	0.112	0.579	0.576	0.119	0.577
Bank Fixed Effect	NO	NO	YES	NO	NO	YES
Firm Fixed Effect	NO	NO	NO	NO	NO	NO
Year Fixed Effect	NO	YES	NO	NO	YES	NO
Firm-Year Fixed Effect	YES	NO	YES	YES	NO	YES
Clusters at Bank Level	162	162	160	144	144	141

Notes. This table presents the results of a Chinese bank lending channel regression. The unit of observation is a bank-firm-year. The dependent variable is the level of log loan volume. All regressions include fixed effects. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Robust standard errors given below coefficient estimates are clustered at the bank level.

Table 2.9 Bank Lending Channel (Level Effect)

Log(Interest)	(1) Index (Corruption)	(2) Index (Corruption)	(3) Index (Cor.)	(4) Index (Misreporting)	(5) Index (Misreporting)	(6) Index (Misr.)
State-owned	-0.106***	-0.00990		-0.0943**	-0.0725*	
	(0.0370)	(0.0385)		(0.0371)	(0.0370)	
Policy	-0.934***	-1.236***		-0.135	-0.189***	
	(0.0999)	(0.0665)		(0.106)	(0.0579)	
Rural	-0.00535	0.00564		-0.0929	0.109	
	(0.0897)	(0.0958)		(0.125)	(0.122)	
Size	0.00304	0.0592***		0.0327**	0.00181	
	(0.0107)	(0.0105)		(0.0127)	(0.0123)	
Index	0.409***	0.617***	0.547***	0.0107	-0.0703	0.245**
	(0.129)	(0.151)	(0.141)	(0.0482)	(0.0443)	(0.111)
Index x Post	-0.501***	-0.640***	-0.498***	-0.141**	-0.156***	-0.185***
	(0.137)	(0.158)	(0.158)	(0.0583)	(0.0559)	(0.0636)
Observations	806	806	661	603	603	494
R^2	0.880	0.507	0.885	0.841	0.332	0.849
Bank Fixed Effect	NO	NO	YES	NO	NO	YES
Firm Fixed Effect	NO	NO	NO	NO	NO	NO
Year Fixed Effect	NO	YES	NO	NO	YES	NO
Firm-Year Fixed Effect	YES	NO	YES	YES	NO	YES
Clusters at Bank Level	73	73	65	61	61	56

Notes. This table presents the results of a Chinese bank lending channel regression. The unit of observation is a bank-firm-year. The dependent variable is the level of the log interest rate. All regressions include fixed effects. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Robust standard errors given below coefficient estimates are clustered at the bank level.

Table 2.10 Bank Lending Channel (Level Effect) (Initial Year)

Log(Interest)	(1) Index (Corruption)	(2) Index (Corruption)	(3) Index (Cor.)	(4) Index (Misreporting)	(5) Index (Misreporting)	(6) Index (Misr.)
State-owned	-0.0748**	0.0574		-0.0949***	-0.0633*	
	(0.0361)	(0.0373)		(0.0364)	(0.0363)	
Policy	-0.957***	-1.287***		-0.145	-0.206***	
	(0.101)	(0.0666)		(0.106)	(0.0579)	
Rural	0.00317	0.0214		-0.0643	0.121	
	(0.0906)	(0.0968)		(0.125)	(0.122)	
Size	0.00311	0.0619***		0.0266**	-0.00883	
	(0.0109)	(0.0109)		(0.0131)	(0.0124)	
Index	0.249	-0.102		-0.0748	0.0806*	
	(0.222)	(0.206)		(0.0485)	(0.0473)	
Index x Post	-0.176*	-0.0825	-0.165	-0.104**	-0.0463	-0.101**
	(0.104)	(0.112)	(0.181)	(0.0436)	(0.0433)	(0.0511)
Observations	806	806	661	603	603	494
R^2	0.877	0.496	0.877	0.842	0.332	0.846
Bank Fixed Effect	NO	NO	YES	NO	NO	YES
Firm Fixed Effect	NO	NO	NO	NO	NO	NO
Year Fixed Effect	NO	YES	NO	NO	YES	NO
Firm-Year Fixed Effect	YES	NO	YES	YES	NO	YES
Clusters at Bank Level	73	73	65	61	61	56

Notes. This table presents the results of a Chinese bank lending channel regression. The unit of observation is a bank-firm-year. The dependent variable is the level of the log interest rate. All regressions include fixed effects. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Robust standard errors given below coefficient estimates are clustered at the bank level.

Table 2.11 Bank Lending Channel (Level Effect) (Average of the pre-shock Years)

Log(Interest)	(1) Index (Corruption)	(2) Index (Corruption)	(3) Index (Cor.)	(4) Index (Misreporting)	(5) Index (Misreporting)	(6) Index (Misr.)
State-owned	-0.0700*	0.0525		-0.101***	-0.0767**	
	(0.0358)	(0.0368)		(0.0371)	(0.0371)	
Policy	-0.958***	-1.283***		-0.134	-0.193***	
	(0.101)	(0.0662)		(0.106)	(0.0576)	
Rural	-0.000397	0.0231		-0.0804	0.119	
	(0.0910)	(0.0971)		(0.125)	(0.122)	
Size	0.00157	0.0608***		0.0343***	-0.000492	
	(0.0108)	(0.0106)		(0.0124)	(0.0120)	
Index	0.118	0.0686		-0.0424	0.0967**	
	(0.176)	(0.171)		(0.0499)	(0.0480)	
Index x Post	-0.221**	-0.1733	-0.316**	-0.107**	-0.161***	-0.101**
	(0.109)	(0.117)	(0.161)	(0.0492)	(0.0499)	(0.0511)
Observations	806	806	661	603	603	494
R^2	0.877	0.496	0.877	0.841	0.333	0.846
Bank Fixed Effect	NO	NO	YES	NO	NO	YES
Firm Fixed Effect	NO	NO	NO	NO	NO	NO
Year Fixed Effect	NO	YES	NO	NO	YES	NO
Firm-Year Fixed Effect	YES	NO	YES	YES	NO	YES
Clusters at Bank Level	73	73	65	61	61	56

Notes. This table presents the results of a Chinese bank lending channel regression. The unit of observation is a bank-firm-year. The dependent variable is the level of the log interest rate. All regressions include fixed effects. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Robust standard errors given below coefficient estimates are clustered at the bank level.

Table 2.12 Bank Lending Channel (Level Effect)

Log(Maturity)	(1) Index (Corruption)	(2) Index (Corruption)	(3) Index (Cor.)	(4) Index (Misreporting)	(5) Index (Misreporting)	(6) Index (Misr.)
State-owned	0.116***	0.122***		0.200*	0.0959	
	(0.0430)	(0.0467)		(0.108)	(0.102)	
Policy	1.849***	2.388***		1.560***	2.112***	
	(0.0960)	(0.259)		(0.125)	(0.257)	
Rural	0.366***	0.452***		0.586*	0.782***	
	(0.107)	(0.0948)		(0.334)	(0.203)	
Size	0.0847***	0.106***		0.177***	0.239***	
	(0.0137)	(0.0152)		(0.0344)	(0.0318)	
Index	0.829***	0.931***	0.637**	-0.292**	-0.223***	0.487
	(0.268)	(0.329)	(0.274)	(0.113)	(0.0824)	(0.324)
Index x Post	-0.824***	-1.003***	-0.854***	-0.412**	-0.453***	-0.431**
	(0.280)	(0.333)	(0.286)	(0.210)	(0.152)	(0.180)
Observations	12379	12379	12410	4869	4869	4733
R^2	0.478	0.077	0.566	0.536	0.098	0.589
Bank Fixed Effect	NO	NO	YES	NO	NO	YES
Firm Fixed Effect	NO	NO	NO	NO	NO	NO
Year Fixed Effect	NO	YES	NO	NO	YES	NO
Firm-Year Fixed Effect	YES	NO	YES	YES	NO	YES
Clusters at Bank Level	141	141	142	125	125	122

Notes. This table presents the results of a Chinese bank lending channel regression. The unit of observation is a bank-firm-year. The dependent variable is the level of log maturity. All regressions include fixed effects. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Robust standard errors given below coefficient estimates are clustered at the bank level.

Table 2.13 Bank Lending Channel (Level Effect)

Collateral	(1) Index (Corruption)	(2) Index (Corruption)	(3) Index (Cor.)	(4) Index (Misreporting)	(5) Index (Misreporting)	(6) Index (Misr.)
State-owned	0.0206***	0.0358***		0.0352***	0.0667***	
	(0.00688)	(0.00989)		(0.00967)	(0.0135)	
Policy	0.00790	-0.0264		-0.00305	0.0164	
	(0.0137)	(0.0215)		(0.00669)	(0.0205)	
Rural	0.0167	-0.000975		-0.0253	-0.0266	
	(0.0215)	(0.0249)		(0.0304)	(0.0370)	
Size	-0.000713	-0.0240***		-0.00508	-0.0263***	
	(0.00357)	(0.00325)		(0.00479)	(0.00477)	
Index	0.0651*	0.139***	0.106**	0.0343**	0.0431**	0.120**
	(0.0366)	(0.0445)	(0.0490)	(0.0151)	(0.0178)	(0.0530)
Index x Post	-0.0877**	-0.139***	-0.0240	-0.0324*	-0.0408**	-0.0375**
	(0.0355)	(0.0492)	(0.0527)	(0.0170)	(0.0201)	(0.0190)
Observations	23630	23630	23296	13485	13485	12999
R^2	0.572	0.123	0.622	0.582	0.157	0.629
Bank Fixed Effect	NO	NO	YES	NO	NO	YES
Firm Fixed Effect	NO	NO	NO	NO	NO	NO
Year Fixed Effect	NO	YES	NO	NO	YES	NO
Firm-Year Fixed Effect	YES	NO	YES	YES	NO	YES
Clusters at Bank Level	162	162	160	144	144	141

Notes. This table presents the results of a Chinese bank lending channel regression. The unit of observation is a bank-firm-year. The dependent variable is an indicator variable equals one if the loan is a collateral loan. All regressions include fixed effects. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Robust standard errors given below coefficient estimates are clustered at the bank level.

Table 2.14 Bank Lending Channel (Level Effect) (Credit Allocation across Firms)

Log(Loan)	(1) Index (Corruption)	(2) Index (Corruption)	(3) Index (Cor.)	(4) Index (Misreporting)	(5) Index (Misreporting)	(6) Index (Misr.)
State-owned (Bank)	0.161***	0.110***		0.101**	0.00936	
	(0.0298)	(0.0331)		(0.0472)	(0.0302)	
Policy	0.508***	0.778***		0.479***	0.683***	
	(0.0669)	(0.0556)		(0.132)	(0.112)	
Rural	0.161***	0.0303		0.146***	-0.00804	
	(0.0310)	(0.0588)		(0.0443)	(0.115)	
Size	0.0803***	0.0970***		0.0934***	0.133***	
	(0.00570)	(0.00844)		(0.0125)	(0.00903)	
Index	-0.224*	-0.190	-0.154	-0.122***	-0.185***	-0.275***
	(0.131)	(0.122)	(0.156)	(0.0365)	(0.0386)	(0.101)
Index x Post	0.367***	1.224***	0.169*	0.0631*	0.352***	0.197*
	(0.142)	(0.127)	(0.0928)	(0.0380)	(0.0492)	(0.118)
Index x Post x State (Firm)	-0.0614	-1.291***	-0.0585	-0.0215	-0.428***	-0.0390
	(0.0880)	(0.168)	(0.0868)	(0.0307)	(0.0399)	(0.0407)
Observations	23175	23175	22845	13027	13027	12556
R^2	0.579	0.117	0.582	0.582	0.129	0.584
Bank Fixed Effect	NO	NO	YES	NO	NO	YES
Firm Fixed Effect	NO	NO	NO	NO	NO	NO
Year Fixed Effect	NO	YES	NO	NO	YES	NO
Firm-Year Fixed Effect	YES	NO	YES	YES	NO	YES
Clusters at Bank Level	162	162	160	144	144	141

Notes. This table presents the results of a Chinese bank lending channel regression. The unit of observation is a bank-firm-year. The dependent variable is the level of log loan volume. All regressions include fixed effects. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Robust standard errors given below coefficient estimates are clustered at the bank level.

Table 2.15 Bank Share (Top 3 Banks in the Specific Prefecture)

	(1) Bank Share (Asset) Index (Corruption)	(2) Bank Share (Total Loan) Index (Corruption)	(3) Bank Share (Asset) Index (Misreporting)	(4) Bank Share (Total Loan) Index (Misreporting)
State-owned	0.135***	0.111***	0.0182	0.000422
	(0.0420)	(0.0150)	(0.0310)	(0.0171)
Policy	-0.146***	-0.0851***	-0.156***	-0.0886***
	(0.0480)	(0.0163)	(0.0216)	(0.00509)
Rural	-0.220***	-0.143***	-0.234***	-0.135***
	(0.0195)	(0.0207)	(0.00807)	(0.0164)
Index	0.101*	0.0883*	0.655***	0.607***
	(0.0592)	(0.0519)	(0.0756)	(0.0696)
Index x Post	-0.139**	-0.197***	-0.0548*	-0.143**
	(0.0703)	(0.0596)	(0.0318)	(0.0548)
Constant	0.282***	0.0802	0.292***	0.107***
	(0.0964)	(0.0518)	(0.0128)	(0.0134)
Observations	2019	2358	1431	1695
R^2	0.633	0.612	0.669	0.646
Year Fixed Effect	YES	YES	YES	YES
Prefecture Fixed Effect	YES	YES	YES	YES
Clusters at Prefecture Level	113	123	109	123

Notes. This table presents the results of a Chinese bank lending channel regression. The unit of observation is a bank-year. The dependent variable is the bank share. All regressions include fixed effects. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Robust standard errors given below coefficient estimates are clustered at the prefecture level.

Table 2.16 Bank Concentration (Herfindahl-Hirschman Index)

	(1) HHI (Asset) Index (Corruption)	(2) HHI (Total Loan) Index (Corruption)	(3) HHI (Asset) Index (Misreporting)	(4) HHI (Total Loan) Index (Misreporting)
Index	0.0813*	0.0881**	0.245***	0.216***
	(0.0441)	(0.0405)	(0.0306)	(0.0336)
Index x Post	-0.0979*	-0.107**	-0.0342*	-0.0411**
	(0.0561)	(0.0517)	(0.0186)	(0.0169)
Constant	0.535***	0.126**	0.494***	0.0879***
	(0.0598)	(0.0537)	(0.0152)	(0.0150)
Observations	724	841	529	616
R^2	0.555	0.533	0.599	0.577
Year Fixed Effect	YES	YES	YES	YES
Prefecture Fixed Effect	YES	YES	YES	YES
Clusters at Prefecture Level	129	137	124	137

Notes. This table presents the results of a Chinese bank lending channel regression. The unit of observation is a bank-year. The dependent variable is the (Herfindahl-Hirschman Index/10000). All regressions include fixed effects. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Robust standard errors given below coefficient estimates are clustered at the prefecture level.

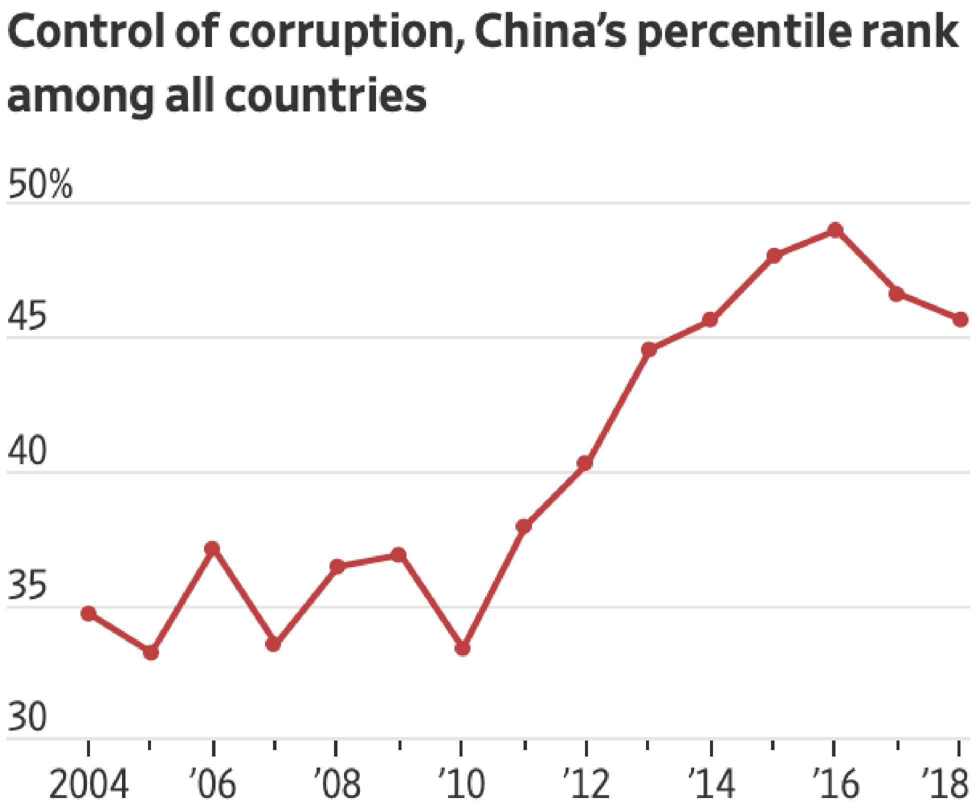
Table 2.17 Bank Health

	(1) Bad Loan/Asset Index (Corruption)	(2) Bad Loan/ Total Loan Index (Corruption)	(3) Bad Loan/Asset Index (Misreporting)	(4) Bad Loan/ Total Loan Index (Misreporting)
State-owned	1.697***	3.142***	1.846***	3.507***
	(0.469)	(0.707)	(0.545)	(0.834)
Policy	-1.428**	-4.301***	-1.557**	-4.548***
	(0.669)	(0.900)	(0.706)	(0.953)
Rural	1.164***	1.677***	1.462***	2.081***
	(0.121)	(0.147)	(0.185)	(0.230)
Index	0.218	0.359	0.346***	0.692***
	(0.233)	(0.372)	(0.110)	(0.174)
Index x Post	-0.212**	-0.827**	-0.308***	-0.710***
	(0.107)	(0.413)	(0.108)	(0.169)
Constant	4.956***	11.42***	5.029***	11.26***
	(1.324)	(2.314)	(1.319)	(2.287)
Observations	1273	1530	1099	1327
R^2	0.454	0.510	0.466	0.524
Year Fixed Effect	YES	YES	YES	YES
Prefecture Fixed Effect	YES	YES	YES	YES
Clusters at Prefecture Level	95	100	84	89

Notes. This table presents the results of a Chinese bank lending channel regression. The unit of observation is a bank-year. The dependent variable is the bad loan rate. All regressions include fixed effects. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Robust standard errors given below coefficient estimates are clustered at the prefecture level.

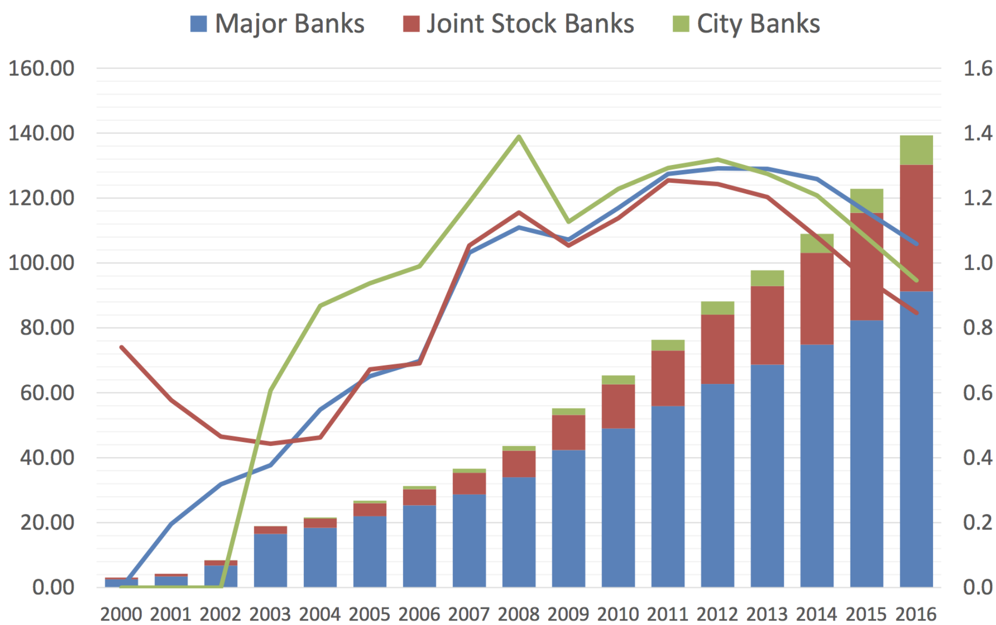
2.9 Figures

Figure 2.1 Cleanup Costs



Source: "China's Corruption Paradox", The Wall Street Journal.

Figure 2.2 Bank Assets (RMB Trillion) and ROA (%)



Source: Jiang Wang (MIT), "China's Financial System: Developments and Challenges", MIT Golub Center for Finance and Policy 4th Annual Conference.

Figure 2.3 Corruption Perception Index in China

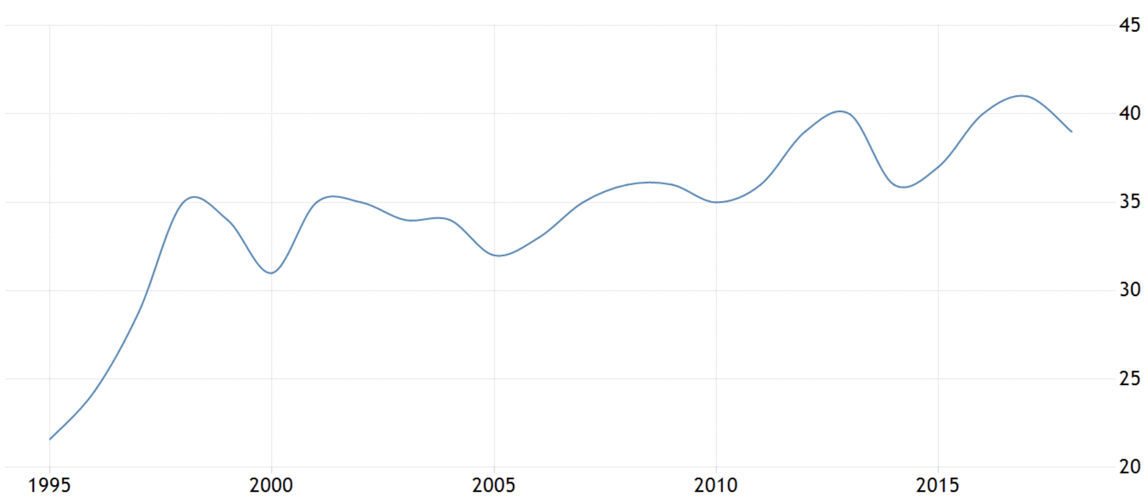


Figure 2.4 Corruption Index over Time

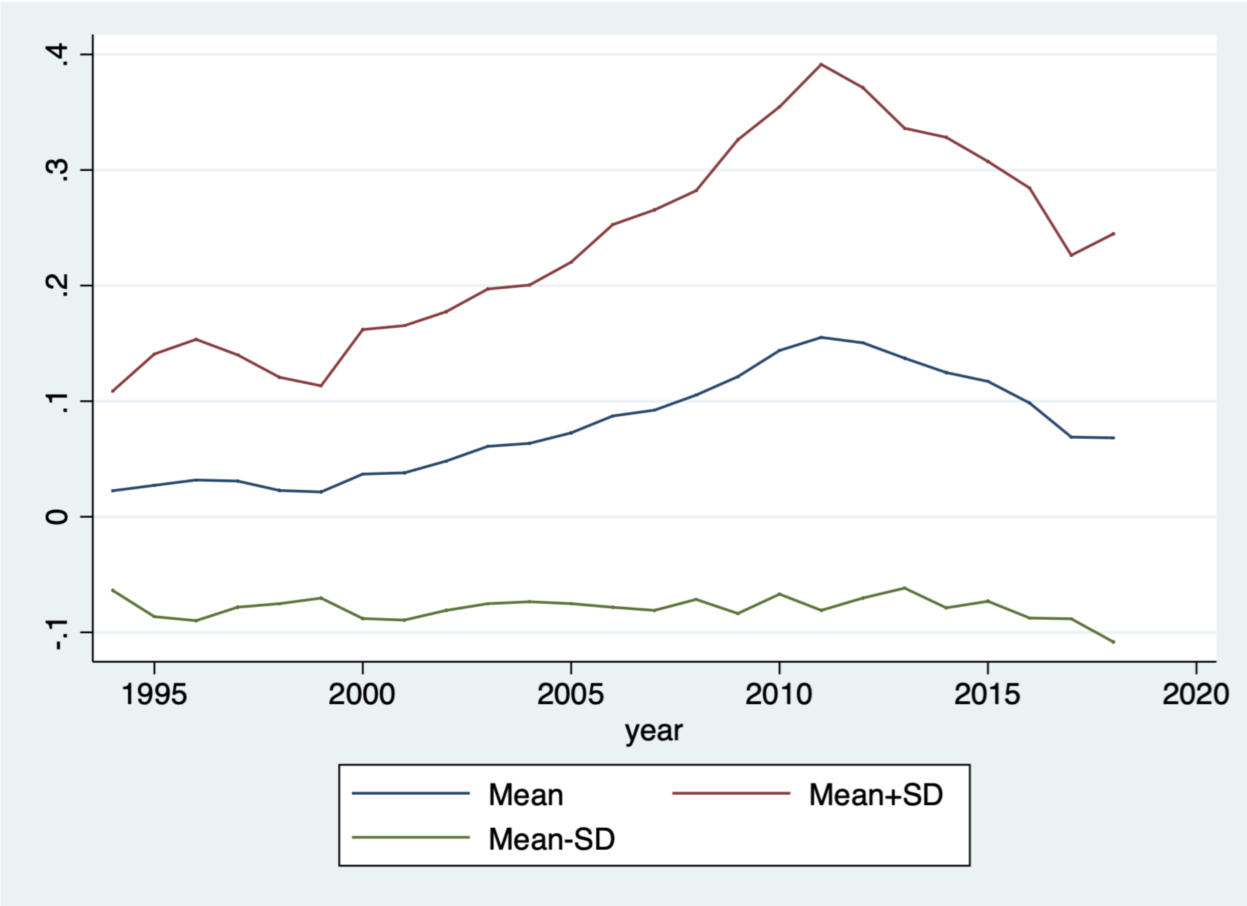


Figure 2.5 Corruption Index (Coefficient of Variation) over Time

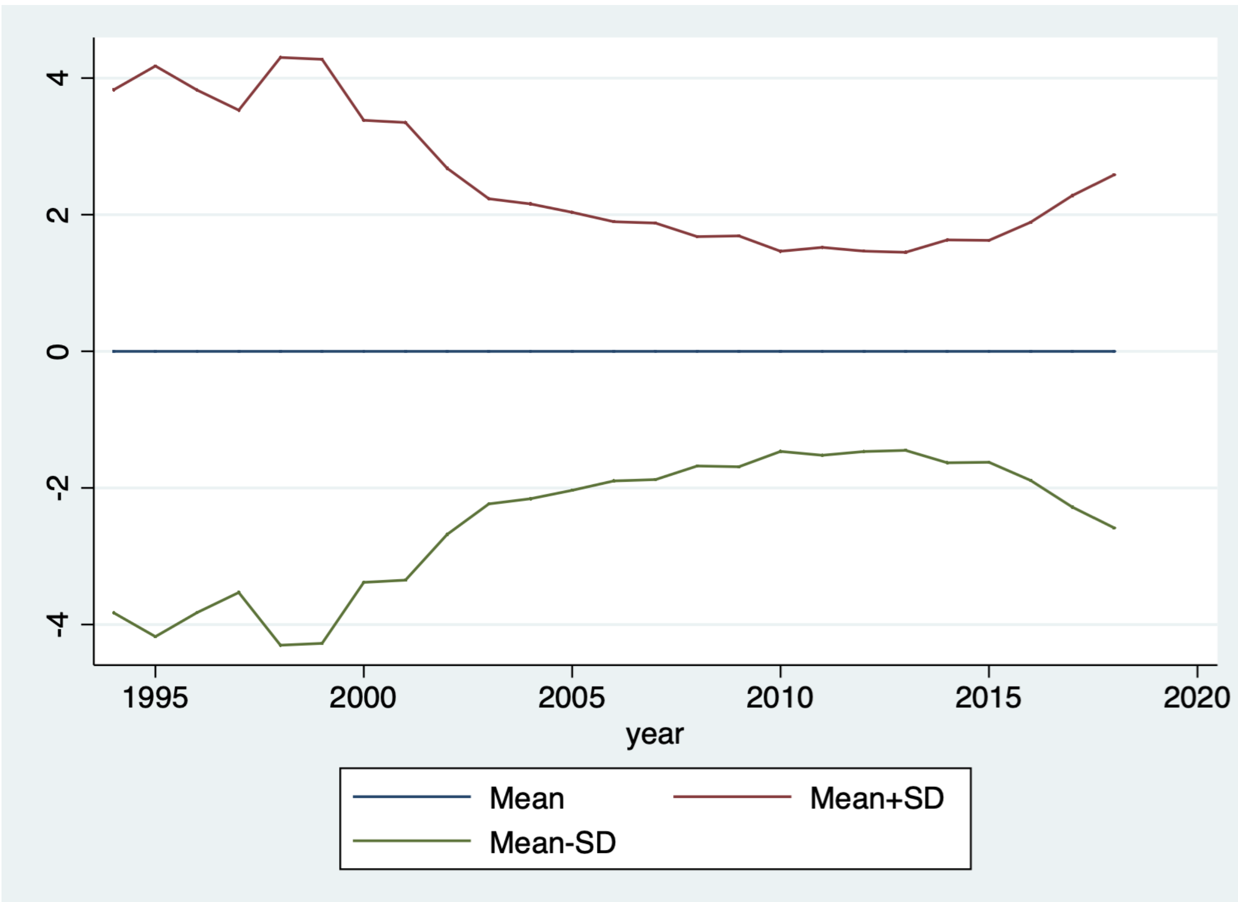


Figure 2.6 Corruption Index over Time

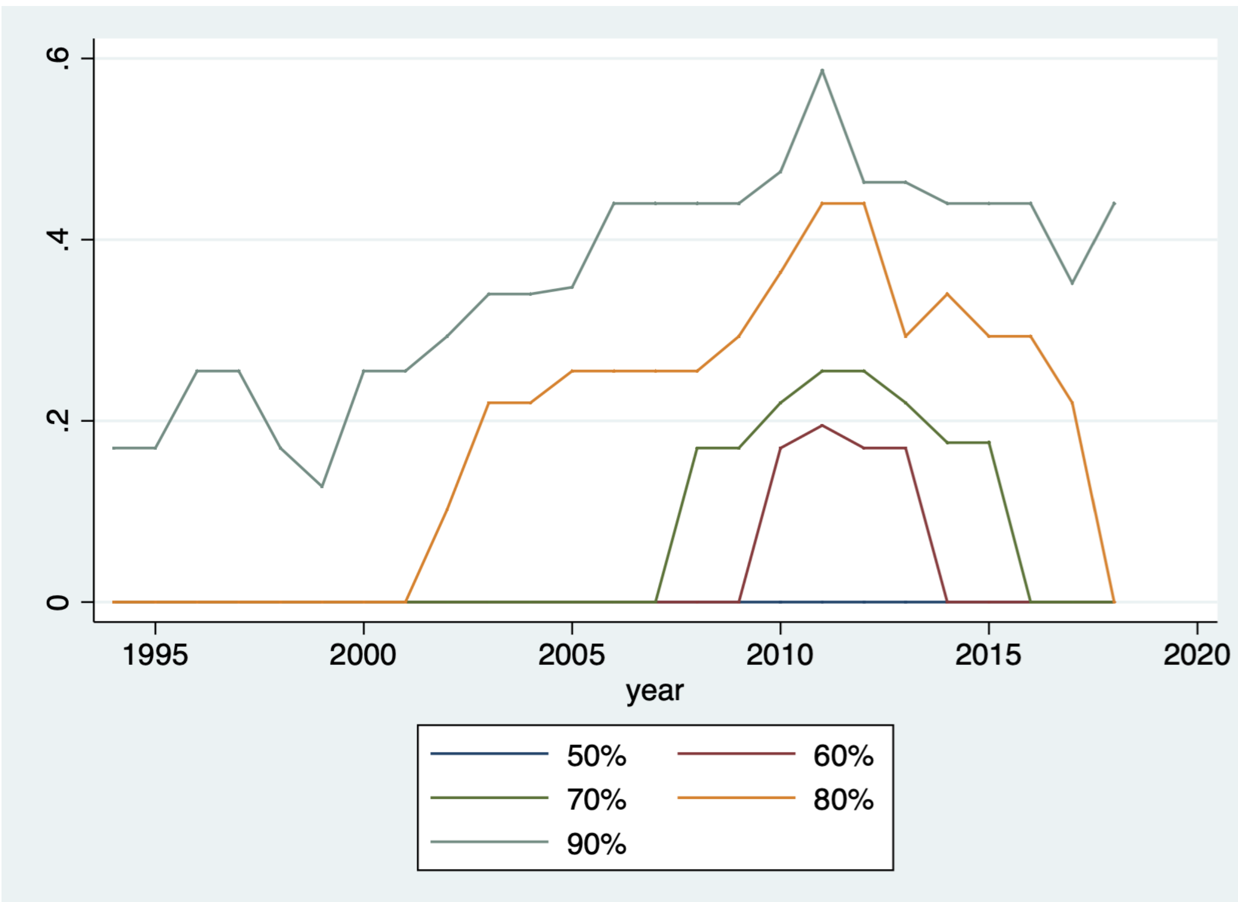


Figure 2.7 Corruption Index (DM) Map (2001-2016)

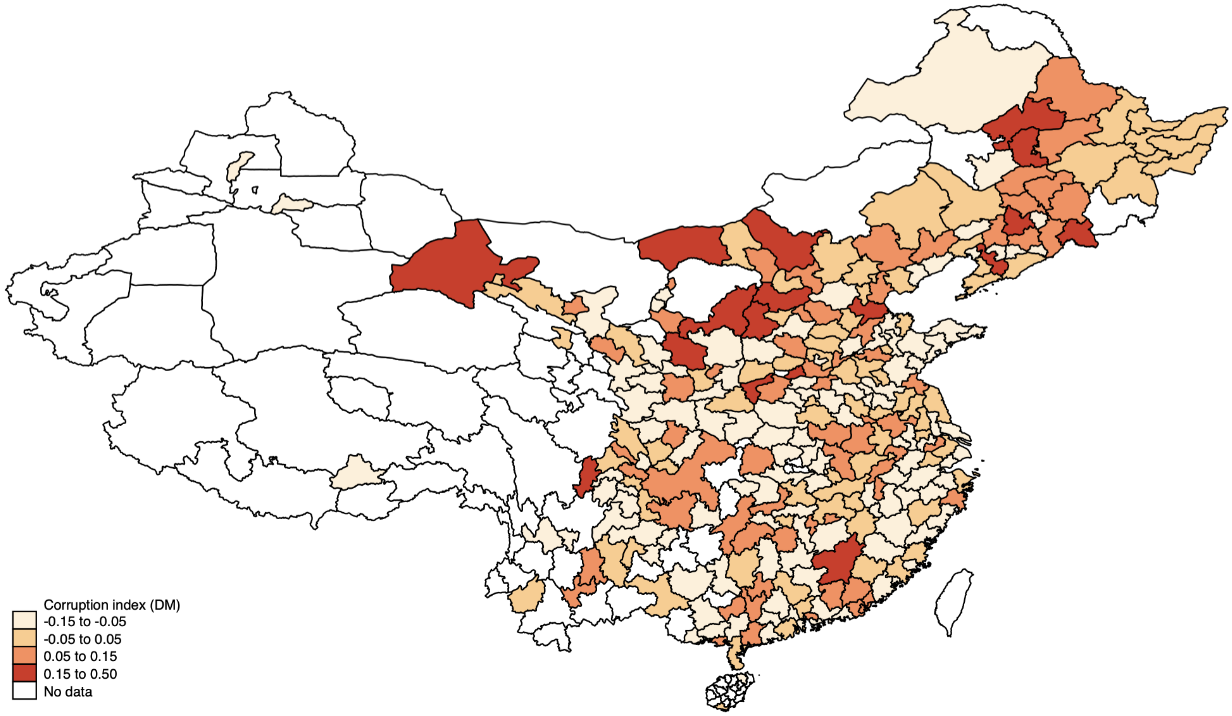


Figure 2.8 Evolution of Corruption Index (DM) (2001-2016)

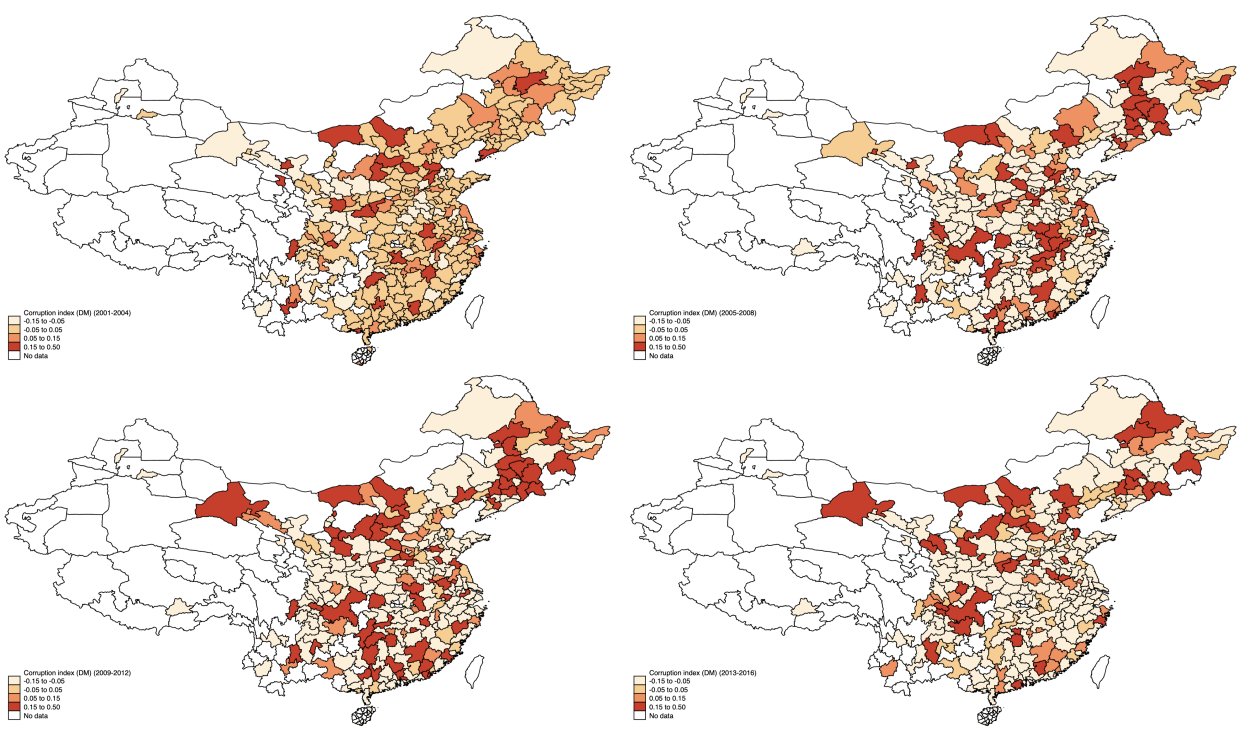


Figure 2.9 Parallel Trend (The Log Amount of the Loan)

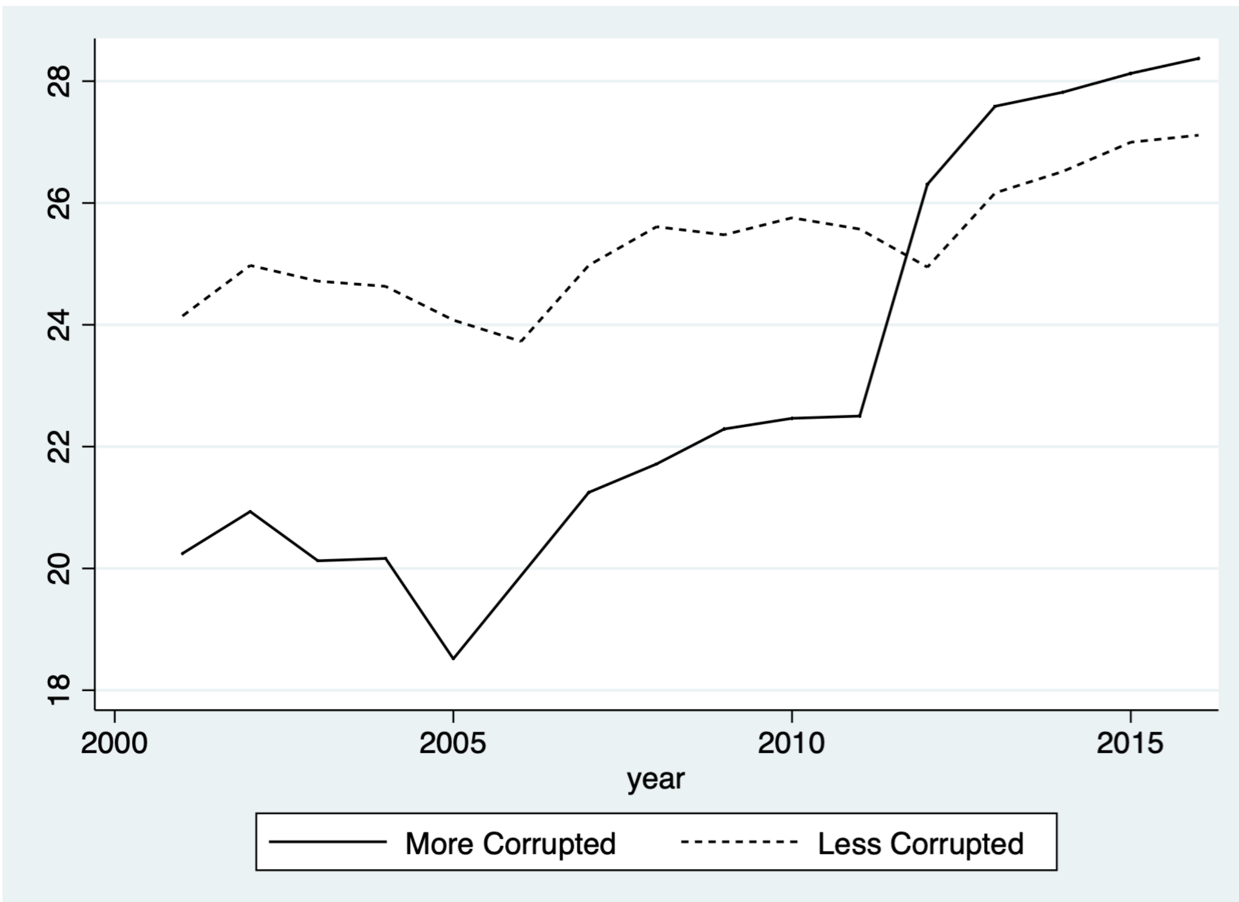
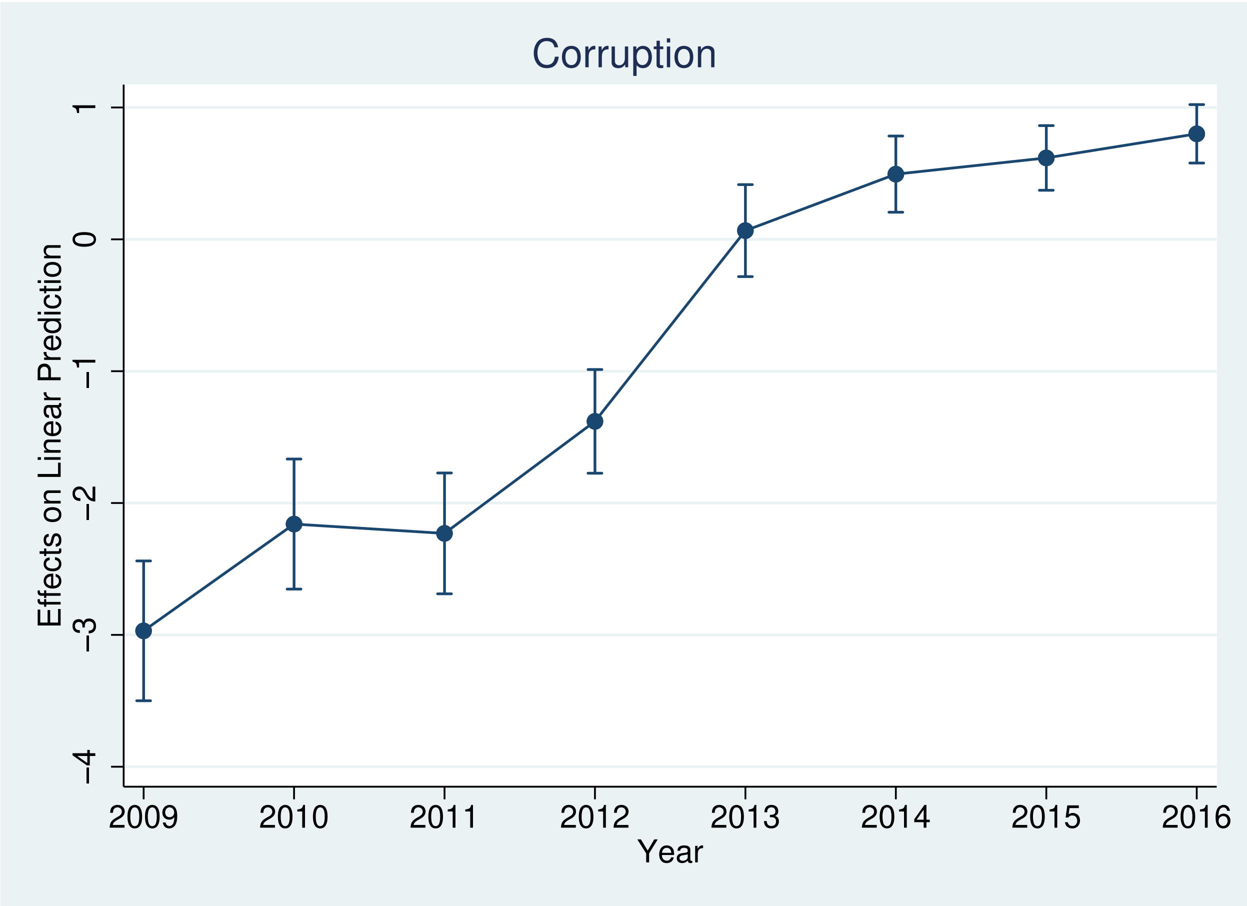


Figure 2.10 Year-specific Effects (Coefficients β_1)



Chapter Three

Cash Flow Forecasting: Dealing with Serial Correlation and Idiosyncratic Heterogeneity

with Hongtao Guo and Zhijie Xiao.

3.1 Introduction

Cash flow from operations (CFO) measures the amount of cash generated by a company's normal business operations. Prediction of CFO impacts a variety of economic decisions including valuation methodologies employing discounted cash flows, distress prediction, risk assessment, the accuracy of credit-rating predictions, and the provision of value- relevant information to security markets, among others (see, for example, Bowen et al., 1986, DeFond and Hung (2003)). Kim and Kross (2005) indicate that while practitioners use net earnings in security valuation and performance evaluations more widely, theoretical valuation models generally favor CFO as an input. Moreover, we observe that CFO prediction has been central to the deliberations of both the Financial Accounting Standards Board (FASB) and the International Accounting Standards Board (IASB). For example, the Statement of Financial Concepts No. 1 states: ". . . financial reporting should provide information to help investors, creditors, and others assess the amount, timing, and uncertainty of prospective net cash inflows to the related enterprise." (FASB 1978, par. 37). Statements issued jointly by IASB and FASB underscore the primacy of CFO prediction to investors and creditors. (IASB, FASB 2006, p. 18). Therefore, research similar to that contained herein is particularly relevant given the importance of the predictive ability of CFO data to the aforementioned topics.

Extant literature on statistically-based cash-flow prediction (e.g. Beaver (1970), Bernard and Stober (1989), Dechow et al. (1998), Barth et al. (2001), Dechow and Dichev (2002), Hribar and Collins (2002), Givoly et al. (2009), and Kim and Kross (2005), among others) has pursued cross-sectional versus time-series estimation procedures in a mutually exclusive fashion. Doing so has not allowed statistical models to realize their potential in terms of predictive performance. Specifically, the cross-section regression based approach has the advantage of minimal data requirements that serves to maximize sample size (n is usually large in practice). However, predictive models based on a cross-sectional regression restricts that the beta parameters to be the same for different firms. Cumulated empirical evidence indicates that the beta value varies across firms of different sizes, and the cross-sectional regression can not capture an idiosyncratic beta. The beta variability across firms suggests that it might be useful to consider estimating the coefficients based on a time series regression for each firm. Such a time series based predictive model has the advantage of allowing for firm-specific variability in beta, but requires a long enough time series data. With the passage of SFAS No. 95 requiring firms to present a statement of cash flow for fiscal year ending after July 15, 1988 (FASB, 1987), such time series of reported quarterly CFO can be obtained to conduct our analysis.

In this paper, we contribute to the CFO prediction literature by proposing a new prediction model that takes into account of both time series and cross sectional issues. In particular, our device uses firm specific time series quarterly CFO information, but also allows the coefficient estimators (beta parameters) to be varying with firm size. The proposed approach uses the idea of machine learning in making predictions. It allows the information about coefficients to be represented by a subset of local observations, thus the beta coefficients vary with these local observations. We are also interested in quarterly CFO predictions for the following reasons: 1) Both Dechow et al. (1998) and Barth et al. (2001) suggest that structural modeling of annual cash-flow prediction models ignore the complex seasonal patterns of CFO and also rely on less timely data. 2) Quarterly CFO series have the potential to exhibit both quarter-to-quarter (adjacent) as well as quarter-by-quarter (seasonal) autocorrelation, which is not present in annual CFO series. Time-series models have the capability to track such complex autocorrelation patterns in quarterly CFO series very efficiently. 3) The time-series literature in accounting provides empirical evidence that predictive ability of earnings numbers was dramatically enhanced when researchers employed quarterly time-series models versus annual models (Brown and Rozeff (1979)).

We provide empirical evidence that the prediction of cash flows from operations (CFO) is enhanced by jointly adopting features specific to both cross-sectional and time-series modeling simultaneously. In doing so, we extend the literature on statistically-based, annual cash-flow prediction models by introducing estimation procedures that combine the favorable attributes of both cross-sectional estimation via the use of "local" cross-sectional data for firms of similar size and time-series estimation via the capturing of firm-specific variability in the beta parameters for the independent variables.

The rest of the paper is organized as follows: In Section 3.2, we discuss the relevant accounting research to provide the background of our research. We introduce the research question and summarize existing forecasting models used in the CFO forecasting literature in Section 3.3. We also present preliminary empirical results that motivate our "learning" approach. Section 3.4 introduces our proposed model. Section 3.5 presents some Monte Carlo results. Empirical investigation on the proposed method and an extensive comparison with existing model are discussed in Section 3.6. In Section 3.7 we discuss the results from several robustness checks. Section 3.8 concludes.

3.2 Background

Research on annual CFO prediction models has been championed by Dechow et al. (1998) and Barth et al. (2001). These studies develop analytical linkages between aggregate and disaggregate earnings components and CFO and specify the functional forms of CFO prediction models derived from their analytical work. Two widely cited annual CFO prediction models are presented below:

$$CFO_{i,t+1} = \alpha + \sum_{\tau=0}^{k} \beta' EARN_{i,t-\tau} + u_{i,t+1} \tag{3.1}$$

$$CFO_{i,t+1} = c_0 + c_1 CFO_{it} + c_2 \Delta AR_{it} + c_3 \Delta INV_{it} + c_4 \Delta AP_{it} + c_5 DEPR_{it} + c_6 AMORT_{it} + c_7 OTHER_{it} + u_{it} \tag{3.2}$$

where i and t denote firm and year, τ ranges between 0 and 2, $EARN$ is net earnings before extraordinary items, ΔAR is a change in accounts receivable, ΔINV is a change in inventory, ΔAP is a change in accounts payable, $DEPR$ is depreciation expense, $AMORT$ is amortization expense, and $OTHER$ is the aggregate of remaining accruals not specifically detailed above. Barth et al. (2001) compared the predictive ability of the aggregate earnings prediction model in equation 3.1 versus the disaggregated earnings prediction model in equation 3.2 and determined that use of disaggregated accruals enhances predictive performance relative to the aggregate earnings model. Barth et al. (2001)'s prediction models were estimated cross-sectionally and the sole criterion used to assess predictive performance was in-sample descriptive fit (i.e., adjusted R2). Unfortunately, exclusive reliance on cross- sectional estimation procedures precludes capturing the information content in potentially idiosyncratic beta coefficients in the aforementioned models. Lorek and Willinger (2011) provide descriptive evidence that the beta parameter in equation

3.1 exhibits considerable firm-specific variability thus casting doubt upon the propriety of cross-sectional estimation procedures that suppress such variability. Moreover, sell-side analysts and users interested in firm valuations require accurate out-of-sample CFO predictions rather than simply relying upon descriptive goodness of fit measures.

3.3 The Research Question and Existing Methods

Given observation on CFO $\{CFO_{i,s}\}$, $s = 1, 2, \cdots\cdot, T; i = 1, \cdots, n$, and related information denoted by $\{x_{i,s}\}$, $s = 1, 2, \cdots\cdot, T; i = 1, \cdots, n,$, we want to predict the cash flow for the next period $CFO_{r,T+1}$ for each firm $r = 1, 2, \cdots, n$. The predictive regression models in the previous literature can be written in the following form:

$$CFO_{i,t+1} = \alpha + \beta' x_{i,t} + u_{i,t+1} \tag{3.3}$$

where i signifies the $i - th$ firm, and $x_{i,t}$ is a vector of informative variables that are available at time t. Popular models of CFO prediction includes:

1. Predictive model using past CFO information: $x_{i,t} = CFO_{i,t}$,

$$CFO_{i,t+1} = \alpha + \beta CFO_{i,t} + u_{i,t+1} \tag{3.4}$$

2. Predictive model using past values of net earnings: $x_{i,t} = E_{i,t}$,

$$CFO_{i,t+1} = \alpha + \beta E_{i,t} + u_{i,t+1} \tag{3.5}$$

3. Multiple regression model: $x_{i,t}$ contains past CFO and other important information, such as operating income.

We first discuss the two leading predictive regressions in the CFO forecasting literature, based on cross-sectional or time-series estimation procedures respectively.

3.3.1 Prediction Based on the Cross-sectional Regression

Given the large number of firms at each time period, much of the traditional literature estimate the parameters α and β based on a cross sectional regression at time t over i:

$$CFO_{i,t} = \alpha + \beta' x_{i,t-1} + u_{it}, i = 1, \cdots, n \tag{3.6}$$

Let

$$\theta = (\alpha, \beta')', W_{i,t} = (1, x'_{i,t-1})'. \tag{3.7}$$

and

$$\widehat{\theta} = \left(\widehat{\alpha}, \widehat{\beta}'\right)', \tag{3.8}$$

then the cross-sectional regression estimator of α and β is given by

$$\widehat{\theta} = \left(\sum_{i=1}^{n} W_{i,t} W'_{i,t}\right)^{-1} \sum_{i=1}^{n} W_{i,t} CFO_{i,t} \tag{3.9}$$

The one-step ahead forecast of CFO can be obtained by

$$\widehat{CFO}_{r,t+1} = \widehat{\alpha} + \widehat{\beta}' x_{r,t}. \tag{3.10}$$

The above cross-section regression-based approach has the advantage of minimal data requirements that serve to maximize sample size (n is usually large in practice). However, predictive models based on a cross-sectional regression restricts that the beta parameters to be the same for different firms. Cumulated empirical evidence indicates that the beta value varies across firms of various sizes, and the cross-sectional regression can not capture the idiosyncratic beta.

3.3.2 Prediction Based on the Time-series Regression

The beta variability across firms suggests that it might be useful to consider estimating the coefficients based on a *time series* regression for each given r

$$CFO_{r,s} = \check{\alpha} + \check{\beta}' x_{r,s-1} + \check{u}_{rs}, s = 2, \cdots, t. \tag{3.11}$$

i.e.

$$\check{\theta} = \left(\sum_{s=2}^{t} W_{r,s}W'_{r,s}\right)^{-1} \sum_{s=2}^{t} W_{r,s}CFO_{r,s} \tag{3.12}$$

where

$$\check{\theta} = (\check{\alpha}, \check{\beta}')'. \tag{3.13}$$

Then obtain one-step ahead forecast of CFO by

$$\widehat{CFO}_{r,t+1} = \check{\alpha} + \check{\beta}' x_{r,t}. \tag{3.14}$$

Such a time series based predictive model has the advantage of allowing for the firm-specific variability in beta, but requires a long enough time series data. This imposes strong data requirements since FASB standard N0. 95 didn't require firms to present a statement of cash flow until fiscal year ending after July 15, 1988.

3.3.3 Empirical Results Based on The Existing Methods: A Motivation

In this section, we present some preliminary empirical analysis revealing important features of CFO, which motivates the local learning methods we propose in later sections.

We obtained data from the quarterly Compustat from the first quarter, 2010, to the fourth quarter, 2018. Sample firms had calendar year-ends that met two sampling criteria: (i) They had a complete time series of quarterly CFO reported in accordance with SFAS No. 95 across the aforementioned interval, and (2) they had a complete time series of quarterly financial statement subcomponents necessary to operationalized multivariate time-series regression model (MULT). Lorek and Willinger (1996) developed a multivariate time-series regression model (MULT) that employs lagged values of CFO, operating income, receivables, payables, and inventory in a time series regression, which allows for firm-specific parameter estimation. The variables we used are Operating Activities - Net Cash Flow, Account Payable/Creditors, Inventories, Receivables, Operating Income Before Depreciation, Current Assets, Deferred Taxes, Income Taxes, Income Taxes, Cash, Current Liabilities, Assets, Standard Industry Classification Code, Debt in Current Liabilities, Long-Term Debt - Total as well as Interest and Related Expense.

For each firm i, we consider time series regression

$$CFO_{i,s} = \check{\alpha}_i + \check{\beta}_{i,1}CFO_{i,s-1} + \check{\beta}_{i,2}PAY_{i,s-1} + \check{\beta}_{i,3}INV_{i,s-1} + \check{\beta}_{i,4}OIBD_{i,s-1} + \check{\beta}_{i,5}REC_{i,s-1} + \check{u}_{is}, s = 2, \cdots, t. \tag{3.15}$$

where CFO is the cash flow from operations, PAY is the accounts payable, INV is the inventories, OIBD is the operating income before depreciation, and REC is the receivables, and obtain estimators $(\check{\alpha}_i, \check{\beta}_{i,1}, \cdots, \check{\beta}_{i,5})'$.

In order to get some insights of the idiosyncratic behavior of beta, we next consider regressions of $\check{\beta}_{i,j}$ on the size of firm i : size$_i$, (measured by assets):

$$\check{\beta}_{i,j} = \widehat{a} + \widehat{b} \cdot size_i + \widehat{e}_i, \tag{3.16}$$

Table 3.1 reports the regression results. These results shows that the beta values are closely related to the firm size.

Figure 3.1 depicts the coefficient heterogeneity in firm size for $\check{\beta}_{i,1}, \cdots, \check{\beta}_{i,5}$. The result of figure 3.1 further confirms the close relationship between $\check{\beta}_{i,j}$ and the size of firm i In addition, figure 3.1 also suggest a nonlinear relationship between $\check{\beta}_{i,j}$ and the size. For this reason, we next consider a "nonlinear" regression by introducing the quadratic term of size:

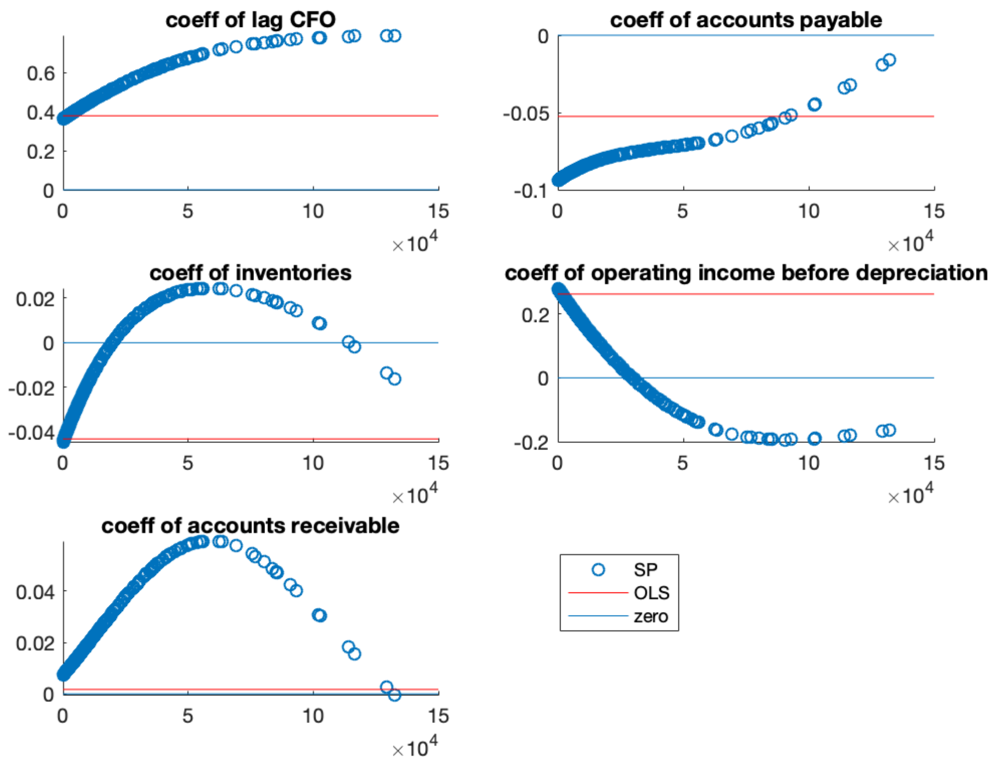
$$\check{\beta}_{i,j} = \widehat{a} + \widehat{b}_1 \cdot size_i + \widehat{b}_2 \cdot size_i^2 + \widehat{e}_i, \tag{3.17}$$

Table 3.2 reports the regression results. Table 3.2 re-confirms the close relationship between beta and size. In addition, after introducing the nonlinear term, the goodness of fit substantially increases, suggesting a nonlinear connection between beta and size.

As a robustness check, we also perform a regression of CFO on the interaction terms constructed as products of the previous regressors and assets.

Table 3.3 shows the coefficients of interaction terms of accounts payable/creditors, inventories, operating income before depreciation, receivables and cash flow from opera-

Figure 3.1 Coefficient Heterogeneity in Firm Size



Notes: This figure presents the relationship between the estimated coefficient of lag cash flow from operations, accounts payable, inventories, operating income before depreciation, receivable and firm size z.

Table 3.1 Regression of β to Assets

	(1) coeff of lag CFO	(2) coeff of PAY	(3) coeff of INV	(4) coeff of OIBD	(5) coeff of REC
Assets	0.0288***	0.00429***	0.00512***	-0.0384***	0.00361***
	(0.000946)	(0.000416)	(0.000351)	(0.00109)	(0.000230)
Constant	0.206***	-0.118***	-0.0698***	0.492***	-0.0120***
	(0.00708)	(0.00311)	(0.00263)	(0.00815)	(0.00172)
Observations	1029	1029	1029	1029	1029
R^2	0.475	0.094	0.172	0.548	0.193

Notes. This table presents the results of a beta to assets regression. The unit of observation is a firm-year. The dependent variable is the coefficient of lag cash flow from operations, account payable/creditors, inventories, operating income before depreciation and receivables. The independent variable is the assets. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Robust standard errors are given below the coefficient estimates.

Table 3.2 Regression of β to Assets and Assets-squared

	(1) coeff of lag CFO	(2) coeff of PAY	(3) coeff of INV	(4) coeff of OIBD	(5) coeff of REC
Assets	-0.0543***	-0.0116***	-0.00261*	0.0824***	-0.00679***
	(0.00327)	(0.00178)	(0.00154)	(0.00293)	(0.000970)
$(\text{Assets})^2$	0.00593***	0.00113***	0.000551***	-0.00862***	0.000741***
	(0.000227)	(0.000124)	(0.000107)	(0.000204)	(0.0000674)
Constant	0.468***	-0.0680***	-0.0454***	0.110***	0.0208***
	(0.0115)	(0.00625)	(0.00541)	(0.0103)	(0.00340)
Observations	1029	1029	1029	1029	1029
R^2	0.685	0.162	0.192	0.835	0.278

Notes. This table presents the results of a beta to assets regression. The unit of observation is a firm-year. The dependent variable is the coefficient of lag cash flow from operations, account payable/creditors, inventories, operating income before depreciation and receivables. The independent variable is the assets and assets-squared. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Robust standard errors are given below the coefficient estimates.

Table 3.3 Assumption Check

	(1)	(2)
	CFO_{t+1}	CFO_{t+1}
Account Payable/Creditors × Assets	0.00244***	0.00168***
	(0.000132)	(0.000115)
Inventories × Assets	-0.00181***	-0.00117***
	(0.0000835)	(0.0000724)
Operating Income Before Depreciation × Assets	0.166***	0.112***
	(0.000852)	(0.00101)
Receivables × Assets	-0.00208***	-0.00162***
	(0.000150)	(0.000131)
Cash Flow From Operations × Assets		0.0281***
		(0.000353)
Constant	22.93**	29.75***
	(10.71)	(9.106)
Observations	38123	38123
Prob > chi2	0.0000	0.0000
Overall R^2	0.5836	0.6430

Notes. This table presents the results of an assumption check regression. The unit of observation is a firm-year. The dependent variable is the cash flow from operations at the period t+1. The independent variable is the interaction between account payable/creditors, inventories, operating income before depreciation, receivables, cash flow from operations and the assets. ***, **, or * indicates that the coefficient estimate is significant at the 1%, 5%, or 10% level, respectively. Robust standard errors are given below the coefficient estimates.

tions with the firm sizes are statistically significant. Thus we could assume that the betas are varying with firm size.

3.4 The Proposed Method

The extensive empirical literature on CFO forecasting documented cross-sectional heterogeneity of beta, indicating the limitation of cross-sectional or panel data regression-based models, which assume the existence of a constant beta across all firms. Time series regression-based predictive models capture the heterogeneity across firms but require a long time series data which may not be available. A short time series will affect the efficiency of the prediction.

The preliminary empirical analysis in the previous section provides strong evidence that the idiosyncratic beta for each firm is closely related to the firm size. This finding provides important information to use local learning to find "local" cross-sectional information to improve upon the predictive power of single predictive time-series regressions.

There is an extensive literature on machine learning and nonparametric functional-coefficient estimation, see, among other things, Bontempi et al. (2012), Hastie et al. (2009), Zhang et al. (2002), Cai and Xiao (2012). The local learning approach assumes no a priori knowledge on the constancy of the beta coefficient. It allows that the information about coefficients is represented by only a subset of observations. This feature is particularly relevant in the CFO model, where the beta values are only related to cross-sectional data information that is "local" to its size.

In this paper, we propose a new predictive model that takes into account both local information and time-series information. In particular, we look through the data set for the "neighbor" of the size of the current firm and predicting that this firm will evolve in the same manner as the neighbor did.

According to the empirical evidence, the betas vary with firm size. We consider the following model:

$$CFO_{i,t+1} = \alpha + \beta(z_i)' x_{i,t} + u_{i,t+1} \tag{3.18}$$

where z_i is the size of firm i. In this model, we allow β to be dependent on the firm size, in particular, we assume that beta is a smooth function of the firm size z_i.

We next consider how to estimate the model (3.18) using panel data. Suppose that we want to forecast, $CFO_{r,t+1}$, the Cash Flow of a particular firm, say firm r, at time $t+1$, using all available information at time t. The basic procedure could be stated as follows:

We first estimate the coefficients α and $\beta(\cdot)$ using existing information upto time t, i.e. $\left\{\{CFO_{i,s}, z_{i,s}, x_{i,s}\}_{i=1}^{n}\right\}_{s=1}^{t}$. We use both time series information upto t and also "local" cross-sectional information for all firms that are similar to the specified firm. Such a device allows the beta to be varying with firm size. However, by utilizing the "local" cross-sectional information, the sample sizes in the above regression are many times than the simple time series regression model. Denote the estimators by $\widehat{\alpha}$ and $\widehat{\beta}(\cdot)$.

Then we move to the prediction of CFO for Firm r at time $t+1$. To predict the cash flow of firm r at time $t+1$, i.e. $CFO_{r,t+1}$. Notice that

$$CFO_{r,t+1} = \alpha + \beta(z_r)' x_{r,t} + u_{r,t+1} \tag{3.19}$$

our predictor for $CFO_{r,t+1}$ is given by

$$\widehat{CFO}_{r,t+1} = \widehat{\alpha} + \widehat{\beta}(z_r)' x_{r,t}. \tag{3.20}$$

We need to obtain estimates $\widehat{\alpha}$ and $\widehat{\beta}(z_r)$ first. For convenience, in our discussion below, we denote $z = z_r$, thus we need to estimate α and $\beta(z)$.

Step 1. we first estimate α and $\beta(z)$ based on the following local weighted regression:

$$\sum_{s=1}^{t-1} \sum_{i=1}^{n} \rho \left(CFO_{i,s+1} - \alpha - \sum_{j=0}^{m} \beta_j^T x_{i,s} \left(z_{is} - z\right)^j \right) K\left(\frac{z_{is} - z}{h}\right) \tag{3.21}$$

where $\rho(\cdot)$ is an appropriate criterion function (we consider the widely used $\rho(\cdot) = |\cdot|$ or $\rho(\cdot) = (\cdot)^2$ in this paper), $K(\cdot)$ is a kernel function, and $h = h(n)$ is a sequence of positive numbers tending to zero and it controls the amount of smoothing used in estimation. We denote the above local polynomial estimator of α as $\widehat{\alpha}(z)$, and construct

$$\widetilde{\alpha} = \frac{1}{n} \sum_{i=1}^{n} \widehat{\alpha}(z_i). \tag{3.22}$$

Step 2. Let $\underline{CFO}_{i,s} = CFO_{i,s} - \widetilde{\alpha}$, we next consider the following estimation of $\beta(z)$

$$\left(\widetilde{\theta}_0^T(z), \widetilde{\theta}_1^T(z)\right) = \arg\min \sum_{s=2}^{t} \sum_{i=1}^{n} \rho \left(\underline{CFO}_{i,s} - \sum_{j=0}^{m} \theta_j^T x_{i,s-1} \left(z_i - z\right) \right) K \left(\frac{z_i - z}{h} \right). \tag{3.23}$$

Let the corresponding estimator from the above regression to be $\widetilde{\theta}_j(z)$, then

$$\widetilde{\beta}(z) = \widetilde{\theta}_0(z) \tag{3.24}$$

Step 3. Finally we predict the cash flow of firm r at time $t+1$, i.e. $CFO_{r,t+1}$, by

$$\widetilde{CFO}_{r,t+1} = \widetilde{\alpha} + \widetilde{\beta}(z_r)' x_{r,t} \tag{3.25}$$

In practice, a popular and useful choice of m and $\rho(\cdot)$ is "$m = 1$, and $\rho(\cdot) = (\cdot)^2$".

Consider a semivarying coefficient model, which is an extension of the varying coefficient model, which is an extension of the varying coefficient model, which is called the semivarying-coefficient model. Procedures for estimation of the linear part and the non-parametric part are developed, and their associated statistical properties are studied. The proposed methods are illustrated by some simulation studies and a real example.

Suggest that a partially varying coefficient model allows appreciable flexibility on the structure of fitted models. The proposed model allows for linearity in coefficients in some variables and nonlinearity in other variables. In such a way, the model has the ability to capture the individual variations and easing the so-called “curse of dimensionality".

3.5 Monte Carlo Simulations

We conduct a Monte Carlo experiment to examine the finite sample performance of the proposed estimation procedures. The data generating process is given by:

Generate $\{\{CFO_{i,t}, z_i, x_{i,t}\}_{i=1}^{n}\}_{t=1}^{T}$ according to the following description where $n = 100$. $i = 1, \cdots, 100$ and $T = 200$. $t = 1, \cdots, 200$.

Where we use both $u_{i,t} =$ iid N$(0,1)$ and $u_{i,t} =$ iid t(5), $z_i \sim U[0,1]$, $x_{i,t} =$ iid N$(0,1)$.

Then, we generate $CFO_{i,t}$ based on

$$CFO_{i,t} = \alpha + (1 + \beta_1 z_i)\, x_{i,t-1} + u_{i,t} \tag{3.26}$$

using $\alpha = 1$. We consider several values of β_1 as $\beta_1 = 0$, $\beta_1 = 0.5$, $\beta_1 = 1$, $\beta_1 = 2$, and $\beta_1 = 5$. After generating the data, we start from $T_0 = 100$, predict $CFO_{i,t}$ at time periods $t = 101, \cdots, 200$, using information upto time $t-1$, i.e. using $\{\{CFO_{i,s}, z_i, x_{i,s}\}_{i=1}^{n}\}_{s=1}^{t-1}$, using the three prediction methods (cross-sectional regression, time-series regression, and the proposed). Then, we compare the three prediction methods based on their MAPE and RMSE:

$$\text{RMSE} \equiv \left[\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{T}\sum_{t=T_0+1}^{T-1}\left(\widehat{CFO}_{i,t} - CFO_{i,t}\right)^2\right)\right]^{1/2} \tag{3.27}$$

$$\text{MAPE} \equiv \frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{T}\sum_{t=T_0+1}^{T-1}\left|1 - \frac{\widehat{CFO}_{i,t}}{CFO_{i,t}}\right|\right) \tag{3.28}$$

The advantages, as well as disadvantages of the forecast accuracy measures, are: 1) RMSE measures the mean of the deviations, and MAPE measures the median of the deviations. 2) MAPE gives the same weight to all errors, the RMSE penalizes variance as it gives errors with larger absolute values more weight than errors with smaller absolute values. 3) MAPE has the disadvantage of being infinite or undefined if $CFO_{i,t} = 0$, and it also puts a heavier penalty on negative errors than on positive errors.

We choose the optimal bandwidth h by minimizing estimated AMISE over the bandwidth h. We did two out-of-sample groups of simulations: The first group is that $n = 100$, $T = 200$, starting from $T_0 = 100$, predict $CFO_{i,t}$ at time periods $t = 101, ..., 200$, using information up to time $t-1$. And the second group is that $n = 100$, $T = 100$, starting from $T_0 = 3$, predict $CFO_{i,t}$ at time periods $t = 4, ..., 100$, using information up to time $t-1$.

Table 3.4 Monte Carlo Comparison (MAPE) of Three Methods, n = 100 and T = 200

	Method 1	Method 2	Method 3	Efficency 1/3	Efficiency 2/3
Parameters			MAPE		
			ui,t = iid N(0, 1)		
$\beta_1 = 0$	2.4701	3.1865	2.4618	1.0034	1.2944
$\beta_1 = 0.5$	2.6400	3.7770	2.6299	1.0038	1.4362
$\beta_1 = 1$	2.5869	4.0587	2.3223	1.1139	1.7477
$\beta_1 = 2$	2.8912	3.5767	2.7008	1.0705	1.3243
$\beta_1 = 5$	3.6477	3.1097	2.3934	1.5241	1.2993
			ui,t = iid t(5)		
$\beta_1 = 0$	2.9523	3.8040	2.9402	1.0041	1.2938
$\beta_1 = 0.5$	5.3324	9.6390	5.3322	1.0000	1.8077
$\beta_1 = 1$	3.1553	3.8691	3.0533	1.0334	1.2672
$\beta_1 = 2$	3.0172	3.5155	2.7451	1.0991	1.2806
$\beta_1 = 5$	3.7159	2.7651	2.1466	1.7311	1.2881

Notes. This table presents the MAPE results of a Monte Carlo simulations. The unit of observation is a firm-year. Method 1 is the cross-sectional method. Method 2 is the time-series method. Method 3 is the proposed method. Efficiency 1/3 represents for the relative effeciencies of method 1 on method 3. Similarly, Efficiency 2/3 represents for the relative effeciencies of method 2 on method 3. The number of firms and time periods is designed as 100 and 200 respectively.

Table 3.5 Monte Carlo Comparison (RMSE) of Three Methods, n = 100 and T = 200

	Method1	Method 2	Method 3	Efficiency 1/3	Efficiency 2/3
Parameters			RMSE		
			ui,t = iid N(0, 1)		
$\beta_1 = 0$	1.0261	1.4252	1.0264	0.9997	1.3885
$\beta_1 = 0.5$	1.0374	1.6067	1.0301	1.0071	1.5598
$\beta_1 = 1$	1.0671	1.8122	1.0347	1.0313	1.7514
$\beta_1 = 2$	1.1754	2.2685	1.0464	1.1233	2.1679
$\beta_1 = 5$	1.7469	3.7932	1.1003	1.5877	3.4474
			ui,t = iid t(5)		
$\beta_1 = 0$	1.3075	1.6457	1.3077	0.9998	1.2585
$\beta_1 = 0.5$	1.3214	1.8136	1.3101	1.0086	1.3843
$\beta_1 = 1$	1.3521	2.0083	1.3132	1.0296	1.5293
$\beta_1 = 2$	1.4590	2.4519	1.3215	1.1040	1.8554
$\beta_1 = 5$	2.0307	3.9838	1.3621	1.4909	2.9247

Notes. This table presents the RMSE results of a Monte Carlo simulations. The unit of observation is a firm-year. Method 1 is the cross-sectional method. Method 2 is the time-series method. Method 3 is the proposed method. Efficiency 1/3 represents for the relative effeciencies of method 1 on method 3. Similarly, Efficiency 2/3 represents for the relative effeciencies of method 2 on method 3. The number of firms and time periods is designed as 100 and 200 respectively.

Table 3.6 Monte Carlo Comparison (MAPE) of Three Methods, n = 100 and T = 100

	Method 1	Method 2	Method 3	Efficiency 1/3	Efficiency 2/3
Parameters			MAPE		
			ui,t = iid N(0, 1)		
$\beta_1 = 0$	3.5214	5.9915	3.5312	0.9972	1.6967
$\beta_1 = 0.5$	3.6324	4.7366	3.6119	1.0057	1.3114
$\beta_1 = 1$	2.7979	3.8689	2.6514	1.0553	1.4592
$\beta_1 = 2$	8.5294	8.8414	3.1696	2.6910	2.7894
$\beta_1 = 5$	2.5774	3.0436	1.9812	1.3009	1.5362
			ui,t = iid t(5)		
$\beta_1 = 0$	2.9507	4.1142	2.9079	1.0147	1.4148
$\beta_1 = 0.5$	12.1027	18.0153	11.8537	1.0210	1.5198
$\beta_1 = 1$	8.9982	13.4050	8.7984	1.0227	1.5236
$\beta_1 = 2$	2.0346	2.9388	1.7395	1.1696	1.6895
$\beta_1 = 5$	2.2551	2.9586	1.7798	1.2671	1.6623

Notes. This table presents the MAPE results of a Monte Carlo simulations. The unit of observation is a firm-year. Method 1 is the cross-sectional method. Method 2 is the time-series method. Method 3 is the proposed method. Efficiency 1/3 represents for the relative effeciencies of method 1 on method 3. Similarly, Efficiency 2/3 represents for the relative effeciencies of method 2 on method 3. The number of firms and time periods is designed as 100 and 100 respectively.

Table 3.7 Monte Carlo Comparison (RMSE) of Three Methods, n = 100 and T = 100

	Method 1	Method 2	Method 3	Efficiency 1/3	Efficiency 2/3
Parameters			RMSE		
			ui,t = iid N(0, 1)		
$\beta_1 = 0$	0.9975	1.4074	0.9994	0.9981	1.4082
$\beta_1 = 0.5$	1.0127	1.6135	1.0031	1.0096	1.6085
$\beta_1 = 1$	1.0515	1.8478	1.0077	1.0435	1.8337
$\beta_1 = 2$	1.1890	2.3666	1.0201	1.1656	2.3200
$\beta_1 = 5$	1.8772	4.0834	1.0795	1.7390	3.7827
			ui,t = iid t(5)		
$\beta_1 = 0$	1.3173	1.6707	1.3211	0.9971	1.2646
$\beta_1 = 0.5$	1.3308	1.8585	1.3236	1.0054	1.4041
$\beta_1 = 1$	1.3629	2.0770	1.3266	1.0274	1.5657
$\beta_1 = 2$	1.4780	2.5736	1.3343	1.1077	1.9288
$\beta_1 = 5$	2.0957	4.2752	1.3693	1.5305	3.1222

Notes. This table presents the RMSE results of a Monte Carlo simulations. The unit of observation is a firm-year. Method 1 is the cross-sectional method. Method 2 is the time-series method. Method 3 is the proposed method. Efficiency 1/3 represents for the relative effeciencies of method 1 on method 3. Similarly, Efficiency 2/3 represents for the relative effeciencies of method 2 on method 3. The number of firms and time periods is designed as 100 and 100 respectively.

The values of both MAPE and RMSE for method 1, method 2 and method 3 are summarized in Table 3.4 and Table 3.5 when $T = 200$ and Table 3.6 and Table 3.7 when $T = 100$. To gauge the efficiency gain for the proposed method, we compute the ratios of both MAPE and RMSE of method 1 and 2 over the MAPE and RMSE of the proposed method, and the ratios are given in the right part of Table 3.4, 3.5, 3.6 and 3.7. From Table 3.4 and Table 3.5, we find that when true $\beta_1 = 0$, there is no heterogeneity in the model. The cross-sectional prediction method approximately equals the proposed method. Our method is significantly better than other methods (the efficiency gain for the proposed method is huge) when the heterogeneity is significant (or true β_1 is large).

From Table 3.6 and Table 3.7, we could also find that when true $\beta_1 = 0$, there is no heterogeneity in the model, cross-sectional prediction method approximately equals the proposed method. Our method is significantly better than other methods (the efficiency gain for the proposed method is huge) when the heterogeneity is significant (or true β_1 is large).

Compare Table 3.4, 3.5 and Table 3.6, 3.7, the proposed method is significantly better than other methods when the time periods T is small. It is evident that the semiparametric predictor provides a much better estimator for the operating cash flow prediction than the fully parametric models in the presence of heterogeneity. Therefore, one can conclude that the proposed method performs very well compared to the misspecified linear model.

3.6 Empirical Applications

3.6.1 Data

We obtained data from the Compustat (North American). We constructed a time series of quarterly observations for each cash flow series beginning in the first quarter of 2010 and ending in the fourth quarter of 2018. Sample firms had calendar year-ends that met two sampling criteria: (i) They had a complete time series of quarterly CFO reported in accordance with SFAS No. 95 across the aforementioned interval, and (2) they had a complete time series of quarterly financial statement subcomponents necessary to operationalized multivariate time-series regression model (MULT).

3.6.2 Key Variables

The dependent variable is the net cash flow from operating activities.

The first set of independent variables consisted of lagged values of the dependent variable cash flow from operations. Intuitively, the selection of lagged cash flow variables is consistent with the modeling procedures that rely on past values of the dependent variable to predict future values.

The second set of independent variables included lagged values of accrual-based earnings. We selected operating income before depreciation (OIBD) as the proxy for accrual-based earnings since Wilson (1986) and Rayburn (1986) have determined that long-term accruals possess little information content in a capital market setting.

The employment of the final set of independent variables (accounts receivable (REC, accounts payable (PAY), and inventory (INV)) is consistent with Wilson's use of current accruals in his regression model. We disaggregate the current accruals variable into REC, PAY, and INV to allow firm-specific parameter estimation for each subcomponent. We stress that the selection of independent variables is based on our intuition regarding possible ways to improve existing cash flow prediction models.

3.6.3 Summary Statistics

Table 3.8 shows the summary statistics of all the variables we used. Table 3.9 shows the correlations between these variables. There are 38124 observations and 36 quarters in our sample.

Table 3.8 Summary Statistics

Variable	Mean	Std. Dev.	Min.	Max.
Account Payable/Creditors	3175.735	37748.814	0.008	999486
Assets	13454.349	83796.051	0.560	2002213
Inventories	1559.013	14503.343	0.004	415293
Cash Flow From Operations	563.5	2883.891	-52280.758	102896.148
Operating Income Before Depreciation	307.171	1070.59	-5431.191	29019
Receivables	2915.15	32727.155	0.003	831427
N		38124		

Notes. This table presents descriptive statistics of firm-year dependent and independent variables. The sample consists of all firms in the North American.

Table 3.9 Correlations

	Assets	PAY	INV	CFO	OIBD	REC
Assets	1.0000					
PAY	0.9239	1.0000				
INV	0.8317	0.7777	1.0000			
CFO	0.4461	0.3136	0.2453	1.0000		
OIBD	0.5365	0.3288	0.3391	0.7430	1.0000	
REC	0.9317	0.9870	0.7700	0.3168	0.3407	1.0000

Notes. Assets represent the average of the assets across the periods. PAY is the accounts payable, INV is the inventories, CFO is the cash flow from operations, OIBD is the operating income before depreciation, and REC is the receivables.

3.6.4 Quarterly CFO Prediction Models

We compare the proposed model with a wide range of predictive models that have been used in the CFO forecasting literature. We also examine the performance of these models for different choices of covariates and error structures. For the convenience of comparison, we discuss below each of the models used in our empirical investigation.

Method 1 (Cross-section) and 2 (Panel):

The first model that we consider is a widely used prediction based on the following cross-sectional regression.

$$CFO_{i,t} = a + b_1 OIBD_{i,t-1} + b_2 REC_{i,t-1} + b_3 INV_{i,t-1} + b_4 PAY_{i,t-1} + e_{i,t},\ i = 1, ..., n. \quad (3.29)$$

where CFO_{it} is firm i's cash flow from operations at time t, $OIBD_{i,t-1}$ is firm i's operating income before depreciation at time $t-1$, $REC_{i,t-1}$ is firm i's receivables at time $t-1$, $INV_{i,t-1}$ is firm i's inventories at time $t-1$, $PAY_{i,t-1}$ is firm i's accounts payable at time $t-1$, and $e_{i,t}$ is the error term.

The traditional approach **run this regression for over firms** $i = 1, ..., n$**, at time period** t**.** Let $\widehat{a}$, $\widehat{b}_1$, $\widehat{b}_2$, $\widehat{b}_3$, $\widehat{b}_4$ be the estimators, then, for any firm r, $CFO_{r,t+1}$ is predicted by

$$\widehat{CFO}_{r,t+1} = \widehat{a} + \widehat{b}_1 OIBD_{r,t} + \widehat{b}_2 REC_{r,t} + \widehat{b}_3 INV_{r,t} + \widehat{b}_4 PAY_{r,t}. \quad (3.30)$$

In our paper, we refer this predictor to "Predictor 1".

We also consider a similar predictor using panel data regression estimator of the above model. i.e. the model (3.29) is estimated based on a panel regression over observations $(s = 2, \cdots, t, i = 1, ..., n)$, and then the prediction for $CFO_{r,t+1}$ is constructed based on these panel regression estimators of (a, b_1, b_2, b_3, b_4). In our paper, we refer this predictor to "Predictor 2".

Method 3 (Time-series) and 4 (Panel):

The second model is the traditional time series predictor based on the time-series regression

$$CFO_{r,s} = \alpha + \Phi CFO_{r,s-1} + e_{r,s}, \; s = 2, ..., t. \tag{3.31}$$

for each firm. Since the estimator of Φ is based on firm r information only, it captures the idiosyncratic firm-specific effect. Let $\widehat{\Phi}$ be the time series regression estimator, then, $CFO_{r,t+1}$ is predicted by

$$\widehat{CFO}_{r,t+1} = \widehat{\alpha} + \widehat{\Phi} CFO_{r,t}. \tag{3.32}$$

In our paper, we refer this predictor to "Predictor 3".

Similar to our treatment on model 1, we also consider a similar predictor using panel data regression estimator of the above model. i.e., the model (3.31) is estimated based on a panel regression over observations $(s = 2, \cdots, t, i = 1, ..., n)$, and then the prediction for $CFO_{r,t+1}$ is constructed based on these panel regression estimators of Φ. Notice that such an estimator is no longer firm-specific. In our paper, we refer this predictor to "Predictor 4".

Method 5 (Times-series) and 6 (Panel):

Lorek and Willinger (1996) developed a multivariate time-series regression model (MULT) that employs lagged values of CFO, operating income, receivables, payables, and inventory in a time series regression, which allows for firm-specific parameter estimation. The time series regression model can be written as

$$\begin{aligned} CFO_{it} \;\; = \;\; & a + b_1 CFO_{i,t-1} + b_2 CFO_{i,t-4} + b_3 OIBD_{i,t-1} + b_4 OIBD_{i,t-4} \\ & + b_5 REC_{i,t-1} + b_6 INV_{i,t-1} + b_7 PAY_{i,t-1} + e_{it}, \end{aligned} \tag{3.33}$$

Where CFO_{it} is operating cash flows for firm i at time t, $OIBD_{i,t-j}$ is operating income before depreciation for firm i at time $t-j$, $REC_{i,t-1}$ is accounts receivable for firm i at time $t-1$, $INV_{i,t-1}$ is inventory for firm i at time $t-1$, $PAY_{i,t-1}$ is accounts payable for firm i at time $t-1$, and e_{it} is a current disturbance term.

For each firm, they run a time series regression over observations $s = 2, \cdots, t$. Since the estimators are based on specified firm information only, they capture the idiosyncratic firm-specific effect. $CFO_{i,t+1}$ is predicted using the parameter estimators based on the above time series regression. In our paper, we refer this predictor to "Predictor 5".

Again, as our treatment on models 1 and 2, we consider the predictor using panel data regression estimator of the above model. i.e. the model (3.33) is estimated based on a panel regression over observations $(s = 2, \cdots, t, i = 1, ..., n)$, and then the prediction for $CFO_{r,t+1}$ is constructed based on these panel regression estimators of $(a, b_1, \cdots, b_7)$. In our paper, we refer this predictor to "Predictor 6".

Method 7:

This is the simple version of our proposed method that uses local learning. However, compared to the more advanced versions of our model (Models 7 and 8), we only consider a simple error structure. In particular, we estimate the coefficients based on

$$CFO_{r,s+1} = \alpha + \beta_1(z_r)CFO_{r,s} + \beta_2(z_r)OIBD_{r,s} + \beta_3(z_r)REC_{r,s} + \beta_4(z_r)INV_{r,s} + \beta_5(z_r)PAY_{r,s} + u_{r,t+1} \tag{3.34}$$

where z_i is the average assets for the firm i.

Our predictor for $CFO_{r,t+1}$ is given by

$$\widehat{CFO}_{r,t+1} = \widehat{\alpha} + \widehat{\beta}_1(z_r)CFO_{r,t} + \widehat{\beta}_2(z_r)OIBD_{r,t} + \widehat{\beta}_3(z_r)REC_{r,t} + \widehat{\beta}_4(z_r)INV_{r,t} + \widehat{\beta}_5(z_r)PAY_{r,t}. \tag{3.35}$$

In our paper, we refer this predictor to "Predictor 7".

The next few models consider more complicate error structures. For the convenience of explanation, the next 4 models employ the Box-Jenkins notation ((pdq)×(PDQ) notation) where: p(P) is the number of regular (seasonal) autoregressive parameters, d(D) is the number of consecutive (seasonal) differences, and q(Q) is the number of regular (seasonal) moving-average parameters.

Method 8 (Time-series) and 9 (Panel):

The SAR ARMA Model (000)×(100) [1]. Lorek et al. (1993) developed a univariate seasonal autoregressive (SAR) ARMA model, (000)×(100).

$$CFO_{it} = a + \Phi CFO_{i,t-4} + e_{it} - \Theta e_{i,t-4} \tag{3.36}$$

where CFO_{it} is operating cash flows for firm i at time t, Φ is an autoregressive parameter, Θ_1 is a seasonal moving-average parameter, e_{it} is a current disturbance term.

For each firm, Lorek et al. (1993) run a time series regression over observations over $s = 2, \cdots, t$. The estimators are based on specified firm information only, thus capture the idiosyncratic firm-specific effect. $CFO_{i,t+1}$ is predicted using the parameter estimators based on the above time series regression. In our paper, we refer this predictor to "Predictor 8".

Again, we consider the predictor using panel data regression estimator of the above model. i.e. the model (3.36) is estimated based on a panel regression over observations $(s = 2, \cdots, t, i = 1, ..., n)$, and then the prediction for $CFO_{r,t+1}$ is constructed based on these panel regression estimators. In our paper, we refer this predictor to "Predictor 9 ".

Method 10 (Time-series) and 11 (Panel):

The BROWN-ROZEFF ARIMA Model (100)×(011). This model was originally proposed by Brown and Rozeff (1979) several decades ago as a premier, a statistically-based predictive model for quarterly EPS (earnings per share). Then, Lorek and Willinger (2011) first document the descriptive fit of the (100) × (011) ARIMA model on quarterly CFO data. The regression model is:

[1]Using customary (pdq)× (PDQ) notation, quarterly earnings time-series models popularized by Brown and Rozeff (1979) and Griffin (1977). The (p, P) variables represent the number of autoregressive or seasonal autoregressive parameters; (d, D) represent the levels of consecutive or seasonal differencing and (q, Q) represents the number of moving-average or seasonal moving-average parameters.

$$CFO_{it} = a + CFO_{i,t-4} + \Phi_1(CFO_{i,t-1} - CFO_{i,t-5}) + e_{it} - \Theta_1 e_{i,t-4} \tag{3.37}$$

where CFO_{it} is operating cash flows for firm i at time t, Φ_1 is an autoregressive parameter, Θ_1 is a seasonal moving-average parameter, e_{it} is a current disturbance term.

For each firm, Lorek and Willinger (2011) run a time series regression over observations over $s = 2, \cdots, t$. $CFO_{i,t+1}$ is predicted using the parameter estimators based on the above time series regression. In our paper, we refer this predictor to "Predictor 10".

Again, we consider the predictor using panel data regression estimator of the above model. i.e. the model (3.37) is estimated based on a panel regression over observations $(s = 2, \cdots, t, i = 1, ..., n)$, and then the prediction for $CFO_{r,t+1}$ is constructed based on these panel regression estimators. In our paper, we refer this predictor to "Predictor 11 ".

Method 12:

This is our proposed method with quarterly-seasonal adjustment (ARMA). We consider the following model:

$$CFO_{r,t+1} = \alpha + \beta(z_r)CFO_{r,t} + \gamma(z_r)X_{r,t} + u_{r,t+1} - \lambda(z_r)u_{r,t-3} \tag{3.38}$$

which can be estimated using information in a neighborhood of z_r. Here $X_{i,t}$ represents all the other explanatory variables including $CFO_{i,t}$ (cash flow from operations), $OIBD_{i,t}$ (operating income before depreciation), $REC_{i,t}$ (receivables), $INV_{i,t}$ (inventories) and $PAY_{i,t}$ (accounts payable) for firm i at time t.

Given z_r, we can select an appropriate neighborhood of z_r, denote it by $I(z_r)$, then we estimate a MULT-ARMA model

$$CFO_{i,s} = \alpha + \beta CFO_{i,s-1} + \gamma X_{i,s} + u_{i,s} - \lambda u_{i,s-4} \tag{3.39}$$

using observations over $i \in I(z_r)$, and $s = 2, ..., t$.

Thus, our predictor for $CFO_{r,t+1}$ can then be obtained as

$$\widetilde{CFO}_{r,t+1} = \widetilde{\alpha} + \widetilde{\beta}(z_r)CFO_{r,t} + \widetilde{\gamma}(z_r)X_{r,t} - \widetilde{\lambda}(z_r)u_{r,t-3} \tag{3.40}$$

We denote this predictor by "Predictor 12".

Method 13:

The Proposed Method with Quarterly Seasonal Adjustment (ARIMA). We consider

$$CFO_{r,t+1} = CFO_{r,t-3} + \alpha + \beta(z_r)[CFO_{r,t} - CFO_{r,t-4}] + \gamma(z_r)X_{r,t} + u_{r,t+1} - \lambda(z_r)u_{r,t-3} \tag{3.41}$$

which can be estimated using information in a neighborhood of z_r. Here $X_{i,t}$ represents the same group of explanatory variables as Model 7.

Parallel to the approach in Predictor 12, given z_r, we can select an appropriate neighborhood of z_r, denote it by $I(z_r)$, the we estimate a MULT-ARIMA model

$$CFO_{i,t} = CFO_{i,t-4} + \alpha + \beta[CFO_{i,t-1} - CFO_{i,t-5}] + \gamma X_{i,t} + u_{i,t} - \lambda u_{i,t-4}, \tag{3.42}$$

using observations over $i \in I(z_r)$, and $s = 2, ..., t$.

Our predictor for $CFO_{r,t+1}$ can then be obtained as

$$\widetilde{CFO}_{r,t+1} = CFO_{r,t-3} + \widetilde{\alpha} + \widetilde{\beta}(z_r)[CFO_{r,t} - CFO_{r,t-4}] + \widetilde{\gamma}(z_r)X_{r,t} - \widetilde{\lambda}(z_r)u_{r,t-3}. \tag{3.43}$$

We denote this predictor by "Predictor 13".

Table 3.10 summarizes all of the methods.

3.6.5 Dealing with the Latent Variables

Model

Predicting quarterly cash flow for operations.

$$CFO_{i,t} = \alpha + \beta(z_i)CFO_{i,t-1} + \gamma(z_i)X_{i,t} + u_{i,t} + \lambda(z_i)u_{i,t-4} \tag{3.44}$$

Given data $\{CFO_{i,t}, X_{i,t}, z_{i,t}, i = 1, \cdots, n; t = 1, \cdots, T\}$, we want to predict $CFO_{i_0,T+1}$.

Table 3.10 Method Summarization

Method	Dimension	Dependent	Independent	Note
1	Cross-sectional	Y	X	previous one period
2	Panel	Y	X	all previous periods
3	Time-series	Y	lag Y	firm-by-firm
4	Panel	Y	lag Y	homogeneous β
5	Time-series	Y	lag Y + X	firm-by-firm
6	Panel	Y	lag Y + X	homogeneous β
7	**Panel**	**Y**	**lag Y + X**	**2 + 4 (heterogeneous β)**
8	Time-series	Y	lag Y + lag U	firm-by-firm
9	Panel	Y	lag Y + lag U	homogeneous beta
10	Time-series	ΔY	Δlag Y + lag U	firm-by-firm
11	Panel	ΔY	Δlag Y + lag U	homogeneous beta
12	**Panel**	**Y**	**lag Y + lag U**	**7 + 9 (heterogeneous β)**
13	**Panel**	**ΔY**	**Δlag Y + lag U**	**7 + 11 (heterogeneous β)**

Notes. Method 1 is the simple cross-sectional method. Method 2 is the cross-sectional method with panel estimation. Method 3 is the simple time-series method. Method 4 is the time-series method with panel estimation. Method 5 is the multivariate time-sries regression model (MULT). Method 6 is MULT with panel estimation. Method 7 is the simple version of our proposed method that uses local learning. Method 8 is the seasonal autoregressive ARMA model. Method 9 is the seasonal autoregressive ARMA model with panel estimation. Method 10 is the Brown-Rozeff ARIMA model. Method 11 is the Brown-Rozeff ARIMA model with panel estimation. Method 12 is our proposed method with quarterly-seasonal adjustment (ARMA). Method 13 is our proposed method with quarterly-seasonal adjustment (ARIMA). lag U represents for moving average term.

Two-step Estimation

For each given firm i,

$$\left[1+\lambda(z_i)L^4\right]^{-1}(CFO_{i,t}-\alpha-\beta(z_i)CFO_{i,t-1}-\gamma(z_i)X_{i,t})=u_{i,t} \tag{3.45}$$

Under the assumption that

$$\left[1+\lambda(z_i)L^4\right]^{-1}=1+b_1(z_i)L+b_2(z_i)L^2+\cdots=\sum_{j=0}^{\infty}b_j(z_i)L^j \tag{3.46}$$

$$CFO_{i,t}-\theta_0(z_i)-\sum_{j=1}^{\infty}\rho_j(z_i)CFO_{i,t-j}-\sum_{l=0}^{\infty}\theta_j(z_i)X_{i,t-l}=u_{i,t} \tag{3.47}$$

We consider the following Two-Step Estimation

1. Step 1. For each individual i, We conduct a preliminary long autoregression

$$CFO_{i,t}=\rho_0(z_i)+\sum_{j=1}^{K}\rho_j(z_i)CFO_{i,t-j}+\sum_{j=0}^{K}\theta_j(z_i)X_{i,t-j}=u_{i,t} \tag{3.48}$$

based on minimizing

$\min\sum_{t=1}^{T}\sum_{l=1}^{n}K\left(\frac{z_{lt}-z_i}{h}\right)[CFO_{l,t}-\rho_{01}-\rho_{02}(z_{lt}-z_i)-\sum_{j=1}^{K}\rho_{j1}CFO_{l,t-j}-\sum_{j=1}^{K}\rho_{j2}(z_{lt}-z_i)CFO_{l,t-j}-\sum_{j=0}^{K}\theta_{j1}X_{l,t-j}-\sum_{j=0}^{K}\theta_{j2}(z_{lt}-z_i)X_{l,t-j}]^2$

Let the solution of the above estimation be

$$\left(\widehat{\rho}_{01},\widehat{\rho}_{02},\left(\widehat{\rho}_{j1},\widehat{\rho}_{j2},\widehat{\theta}_{j1},\widehat{\theta}_{j2}\right),j=1,\cdots,K,\right) \tag{3.49}$$

we obtain

$$\widehat{u}_{i,t}=CFO_{i,t}-\widehat{\rho}_{01}-\sum_{j=1}^{K}\widehat{\rho}_{j1}CFO_{i,t-j}-\sum_{j=0}^{K}\widehat{\theta}_{j1}X_{i,t-j} \tag{3.50}$$

Obtain estimate $\widehat{u}_{i,t}$ for $i=1,\cdots,n$.

2. Step 2. For given firm i, Estimate the ARMA model by minimizing

$\min\sum_{t=1}^{T}\sum_{l=1}^{n}K\left(\frac{z_{lt}-z_i}{h}\right)[CFO_{l,t}-\alpha-\beta_0CFO_{l,t-1}-\beta_1(z_{lt}-z_i)CFO_{l,t-1}-\gamma_0X_{l,t-1}-\gamma_1(z_{lt}-z_i)X_{l,t-1}-\lambda_0\widehat{u}_{l,t-4}-\lambda_1(z_{l,t}-z)\widehat{u}_{l,t-4}]^2$

and let the

Grid-Searching

Under smoothness condition of coefficient functions $\beta(z)$, for any given firm size $z = z_r$, $\beta(z_i))$ can be approximated by a polynomial function as

$$\beta(z_i) \approx \beta(z) + \frac{d\beta(z)}{dz}(z_i - z) + \cdots \tag{3.51}$$

Taking first order expansion:

$$CFO_{i,t} \approx \alpha + \beta_0 CFO_{i,t-1} + \beta_1(z_{i,t} - z)CFO_{i,t-1} + \gamma_0 X_{i,t} + \gamma_1(z_{i,t} - z)X_{i,t} + u_{i,t} + \lambda_0 u_{i,t-4} + \lambda_1(z_{i,t} - z)u_{i,t-4} \tag{3.52}$$

Thus,

$$u_{i,t} = CFO_{i,t} - \alpha - \beta_0 CFO_{i,t-1} - \beta_1(z_{i,t} - z)CFO_{i,t-1} - \gamma_0 X_{i,t} - \gamma_1(z_{i,t} - z)X_{i,t} - \lambda_0 u_{i,t-4} - \lambda_1(z_{i,t} - z)u_{i,t-4} \tag{3.53}$$

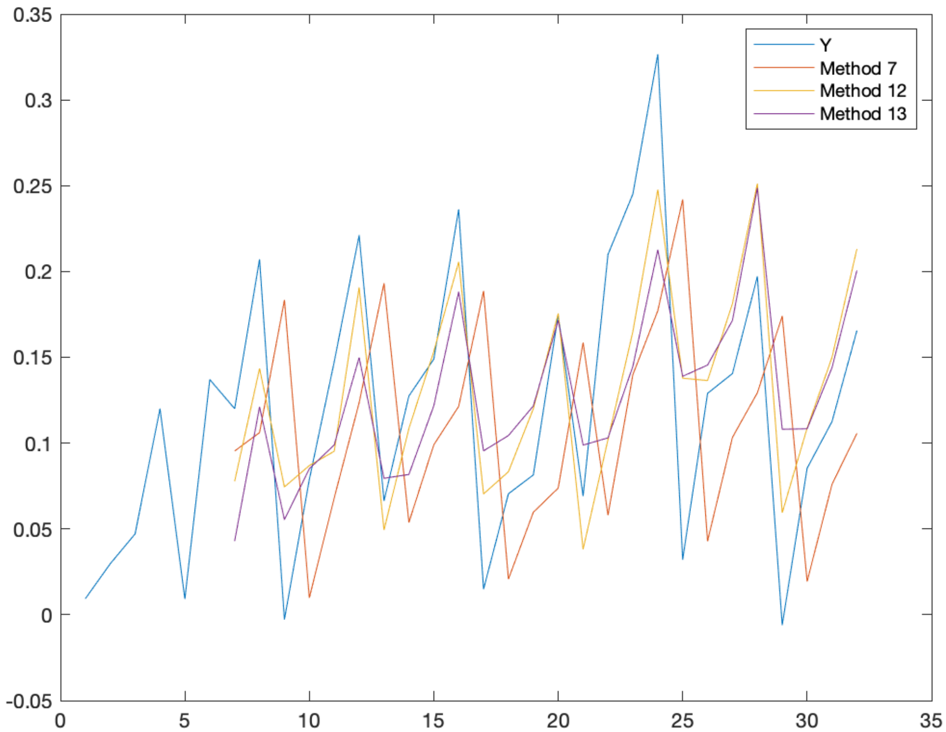
Initialization: let $\{u_{i,0}, u_{i,-1}, \cdots, u_{i,-4}\} = \{0, \cdots, 0\}$, and

$$CFO_{i,0} = \frac{1}{T}\sum_{s=1}^{T} CFO_{i,s} \tag{3.54}$$

Conditional on $\{u_{i,0}, u_{i,-1}, \cdots, u_{i,-4}\}$ and $CFO_{i,0}$, then, given $\theta = (\alpha, \beta_0, \beta_1, \gamma_0, \gamma_1, \lambda_0, \lambda_1)$,

$$\begin{aligned}
u_{i,1} &= CFO_{i,1} - \alpha - \beta_0 CFO_{i,0} - \beta_1(z_{i,1} - z)CFO_{i,0} - \gamma_0 X_{i,1} - \gamma_1(z_{i,1} - z)X_{i,1}, \\
u_{i,2} &= CFO_{i,2} - \alpha - \beta_0 CFO_{i,1} - \beta_1(z_{i,2} - z)CFO_{i,1} - \gamma_0 X_{i,2} - \gamma_1(z_{i,2} - z)X_{i,2}, \\
u_{i,3} &= CFO_{i,3} - \alpha - \beta_0 CFO_{i,2} - \beta_1(z_{i,3} - z)CFO_{i,2} - \gamma_0 X_{i,3} - \gamma_1(z_{i,3} - z)X_{i,3}, \\
u_{i,4} &= CFO_{i,4} - \alpha - \beta_0 CFO_{i,3} - \beta_1(z_{i,4} - z)CFO_{i,3} - \gamma_0 X_{i,4} - \gamma_1(z_{i,4} - z)X_{i,4}, \\
u_{i,5} &= CFO_{i,5} - \alpha - \beta_0 CFO_{i,4} - \beta_1(z_{i,5} - z)CFO_{i,4} - \gamma_0 X_{i,5} - \gamma_1(z_{i,5} - z)X_{i,5} - \lambda_0 u_{i,1} - \lambda_1(z_{i,5} - z)u_{i,1} \\
& \cdots\cdots \\
u_{i,t} &= CFO_{i,t} - \alpha - \beta_0 CFO_{i,t-1} - \beta_1(z_{i,t} - z)CFO_{i,t-1} - \gamma_0 X_{i,t} - \gamma_1(z_{i,t} - z)X_{i,t} - \lambda_0 u_{i,t-4} - \lambda_1(z_{i,t} - z)u_{i,t-4} \\
& \text{, for any } t \geq 5
\end{aligned}$$

Figure 3.2 Empirical Application: Forecasting



Notes: This figure represents the forecasting results of a specific firm in the sample. Method 7 is the proposed method. Method 12 is the proposed method with quarterly-seasonal adjustment (ARMA). Method 13 is the proposed method with quarterly seasonal adjustment (ARIMA).

Notice that the above $u_{i,t}$ are depdent on $(\alpha, \beta_0, \beta_1, \gamma_0, \gamma_1, \lambda_0, \lambda_1)$, $u_{i,t} = u_{i,t}(\alpha, \beta_0, \beta_1, \gamma_0, \gamma_1, \lambda_0, \lambda_1)$, we estimate the parameters by

$$\widehat{\theta}(z) = \arg\min \sum_{t=1}^{T} \sum_{i=1}^{n} K\left(\frac{z_{it} - z}{h}\right) u_{i,t}(\alpha, \beta_0, \beta_1, \gamma_0, \gamma_1, \lambda_0, \lambda_1)^2 \tag{3.55}$$

3.6.6 Empirical Results

Figure 3.2 compares the forecasting results of the operating cash flow using the proposed method 7, 12 and 13 and the real value for a specific firm in our sample. Figure 3.2 shows that our proposed model could be regarded as a good predictor for the trend and fluctuations of the operating cash flow.

Table 3.11 Empirical Application (MAPE and RE Based on MAPE), n = 1029 and T = 36

Method	MAPE	Relative Efficiencies on Method 13
1 (Cross-sectional)	6.9999	2.9903
2 (Panel)	3.8100	1.6276
3 (Time-series)	3.7885	1.6184
4 (Panel)	3.9468	1.6860
5 (Time-series)	4.1638	1.7788
6 (Panel)	2.7964	1.1946
7 (The proposed)	**3.8072**	**1.6264**
8 (Time-series)	3.6660	1.5661
9 (Panel)	2.3666	1.0110
10 (Time-series)	6.5354	2.7919
11 (Panel)	2.5984	1.1100
12 (ARMA)	**2.4965**	**1.0665**
13 (ARIMA)	**2.3409**	**1.0000**

Notes. The observations begining in the first quarter of 2010 and ending in the fourth quarter of 2018. There are 1029 firms in our sample. Method 1 is the simple cross-sectional method. Method 2 is the cross-sectional method with panel estimation. Method 3 is the simple time-series method. Method 4 is the time-series method with panel estimation. Method 5 is the multivariate time-sries regression model (MULT). Method 6 is MULT with panel estimation. Method 7 is the simple version of our proposed method that uses local learning. Method 8 is the seasonal autoregressive ARMA model. Method 9 is the seasonal autoregressive ARMA model with panel estimation. Method 10 is the Brown-Rozeff ARIMA model. Method 11 is the Brown-Rozeff ARIMA model with panel estim-ation. Method 12 is our proposed method with quarterly-seasonal adjustment (ARMA). Method 13 is our proposed method with quarterly-seasonal adjustment (ARIMA).

Table 3.12 Empirical Application (RMSE and RE Based on RMSE), n = 1029 and T = 36

Method	RMSE	Relative Efficiencies on Method 13
1 (Cross-sectional)	0.1434	1.6377
2 (Panel)	0.1100	1.2565
3 (Time-series)	0.1149	1.3123
4 (Panel)	0.1105	1.2621
5 (Time-series)	0.1419	1.6203
6 (Panel)	0.0884	1.0102
7 (The proposed)	**0.1088**	**1.2431**
8 (Time-series)	0.9033	10.3162
9 (Panel)	0.0971	1.1094
10 (Time-series)	2.9178	33.3242
11 (Panel)	0.0944	1.0784
12 (ARMA)	**0.0894**	**1.0206**
13 (ARIMA)	**0.0876**	**1.0000**

Notes. The observations begining in the first quarter of 2010 and ending in the fourth quarter of 2018. There are 1029 firms in our sample. Method 1 is the simple cross-sectional method. Method 2 is the cross-sectional method with panel estimation. Method 3 is the simple time-series method. Method 4 is the time-series method with panel estimation. Method 5 is the multivariate time-sries regression model (MULT). Method 6 is MULT with panel estimation. Method 7 is the simple version of our proposed method that uses local learning. Method 8 is the seasonal autoregressive ARMA model. Method 9 is the seasonal autoregressive ARMA model with panel estimation. Method 10 is the Brown-Rozeff ARIMA model. Method 11 is the Brown-Rozeff ARIMA model with panel estimation. Method 12 is our proposed method with quarterly-seasonal adjustment (ARMA). Method 13 is our proposed method with quarterly-seasonal adjustment (ARIMA).

The values of MAPE and RMSE for methods 1-13 are summarized in Table 3.11 and Table 3.12. To gauge the efficiency gain for the proposed method, we compute the relative effects of MAPE for methods 1-13 over MAPE of the proposed method 13 in Table 3.11. Moreover, we also compute the relative effects of RMSE for methods 1-13 over RMSE of the proposed method 13 in Table 3.12. From Table 3.11 and Table 3.12, we can find that the pro- posed method could beat all of the existing methods in terms of MAPE and RMSE in both the baseline framework and the quarterly model with seasonal adjustment framework.

3.7 Robustness Check

3.7.1 Using Initial Value of Firm Size

Due to the endogeneity concern, we changed the firm size from the mean value across the periods to the initial value. The results are similar to the baseline results, which means the proposed method is robust.

The value of MAPE and RMSE for method 1-13 are summarized in Table 3.13 and Table 3.14. To gauge the efficiency gain for the proposed method, we compute the relative effects of MAPE for method 1-13 over MAPE of the proposed method 13 in Table 3.13. Moreover, we also compute the relative effects of RMSE for method 1-13 over RMSE of the proposed method 13 in Table 3.14. From Table 3.13 and Table 3.14, we could find that the proposed method could beat all of the existing methods, consider MAPE and RMSE in both the baseline framework and the quarterly model with seasonal adjustment framework.

Table 3.13 Initial Value (MAPE and RE Based on MAPE), n = 1029 and T = 36

Method	MAPE	Relative Efficiencies on Method 13
1 (Cross-sectional)	6.9999	2.9735
2 (Panel)	3.8100	1.6185
3 (Time-series)	3.7885	1.6093
4 (Panel)	3.9468	1.6766
5 (Time-series)	4.1638	1.7688
6 (Panel)	2.7964	1.1879
7 (The proposed)	**3.8066**	**1.6170**
8 (Time-series)	3.6660	1.5573
9 (Panel)	2.3666	1.0053
10 (Time-series)	6.5354	2.7762
11 (Panel)	2.5984	1.1038
12 (ARMA)	**2.4961**	**1.0603**
13 (ARIMA)	**2.3541**	**1.0000**

Notes. The observations begining in the first quarter of 2010 and ending in the fourth quarter of 2018. There are 1029 firms in our sample. Method 1 is the simple cross-sectional method. Method 2 is the cross-sectional method with panel estimation. Method 3 is the simple time-series method. Method 4 is the time-series method with panel estimation. Method 5 is the multivariate time-sries regression model (MULT). Method 6 is MULT with panel estimation. Method 7 is the simple version of our proposed method that uses local learning. Method 8 is the seasonal autoregressive ARMA model. Method 9 is the seasonal autoregressive ARMA model with panel estimation. Method 10 is the Brown-Rozeff ARIMA model. Method 11 is the Brown-Rozeff ARIMA model with panel estimation. Method 12 is our proposed method with quarterly-seasonal adjustment (ARMA). Method 13 is our proposed method with quarterly-seasonal adjustment (ARIMA).

Table 3.14 Initial Value (RMSE and RE Based on RMSE), n = 1029 and T = 36

Method	RMSE	Relative Efficiencies on Method 13
1 (Cross-sectional)	0.1434	1.6948
2 (Panel)	0.1100	1.3003
3 (Time-series)	0.1149	1.3581
4 (Panel)	0.1105	1.3060
5 (Time-series)	0.1419	1.6767
6 (Panel)	0.0884	1.0454
7 (The proposed)	**0.1094**	**1.2932**
8 (Time-series)	0.9033	10.6757
9 (Panel)	0.0971	1.1481
10 (Time-series)	2.9178	34.4854
11 (Panel)	0.0944	1.1160
12 (ARMA)	**0.0887**	**1.0489**
13 (ARIMA)	**0.0846**	**1.0000**

Notes. The observations begining in the first quarter of 2010 and ending in the fourth quarter of 2018. There are 1029 firms in our sample. Method 1 is the simple cross-sectional method. Method 2 is the cross-sectional method with panel estimation. Method 3 is the simple time-series method. Method 4 is the time-series method with panel estimation. Method 5 is the multivariate time-sries regression model (MULT). Method 6 is MULT with panel estimation. Method 7 is the simple version of our proposed method that uses local learning. Method 8 is the seasonal autoregressive ARMA model. Method 9 is the seasonal autoregressive ARMA model with panel estimation. Method 10 is the Brown-Rozeff ARIMA model. Method 11 is the Brown-Rozeff ARIMA model with panel estimation. Method 12 is our proposed method with quarterly-seasonal adjustment (ARMA). Method 13 is our proposed method with quarterly-seasonal adjustment (ARIMA).

3.7.2 Including Firm Size in Each Prediction Method

Due to the number of variables concern, we add the additional variable firm size in every method. The results are also very similar to the baseline results, which means the proposed method is robust.

The value of MAPE and RMSE for method 1-13 are summarized in Table 3.15 and Table 3.16. To gauge the efficiency gain for the proposed method, we compute the relative effects of MAPE for method 1-13 over MAPE of the proposed method 13 in Table 3.15. Moreover, we also compute the relative effects of RMSE for method 1-13 over RMSE of the proposed method 13 in Table 3.16. From Table 3.15 and Table 3.16, we could find that the proposed method could beat all of the existing methods, consider MAPE and RMSE in both the baseline framework and the quarterly model with seasonal adjustment framework.

3.8 Concluding Remarks

We provide empirical results supportive of a quarterly CFO prediction model that uses firm specific time series quarterly CFO information, but also allows the coefficient estimators (beta parameters) to be varying with firm size. Specifically, our estimation method utilizes "local" cross-sectional information to improve upon the predictive power of single predictive time series regressions. We compare our proposed method with the prediction models in existing accounting literature and find our method provides better quarterly CFO predictions in terms of both MAPE and RMSE. This methodological improvement in quarterly CFO prediction should be of interests to investors and creditors in their firm valuation methodology, and to accounting and finance researchers who are developing statistical proxies for market expectation on future quarterly CFOs. The proposed forecasting model can also be easily applied to a number of settings that involve making predictions, for example, the prediction of stock price, future revenues and future earnings.

Table 3.15 Including Firm Size (MAPE and RE Based on MAPE), n = 1029 and T = 36

Method	MAPE	Relative Efficiencies on Method 13
1 (Cross-sectional)	6.9453	2.9670
2 (Panel)	3.8208	1.6322
3 (Time-series)	3.9583	1.6909
4 (Panel)	3.9471	1.6862
5 (Time-series)	3.7002	1.5807
6 (Panel)	2.8017	1.1969
7 (The proposed)	**3.8072**	**1.6264**
8 (Time-series)	4.4947	1.9201
9 (Panel)	2.3684	1.0118
10 (Time-series)	4.2232	1.8041
11 (Panel)	2.5981	1.1099
12 (ARMA)	**2.5065**	**1.0707**
13 (ARIMA)	**2.3409**	**1.0000**

Notes. The observations begining in the first quarter of 2010 and ending in the fourth quarter of 2018. There are 1029 firms in our sample. Method 1 is the simple cross-sectional method. Method 2 is the cross-sectional method with panel estimation. Method 3 is the simple time-series method. Method 4 is the time-series method with panel estimation. Method 5 is the multivariate time-sries regression model (MULT). Method 6 is MULT with panel estimation. Method 7 is the simple version of our proposed method that uses local learning. Method 8 is the seasonal autoregressive ARMA model. Method 9 is the seasonal autoregressive ARMA model with panel estimation. Method 10 is the Brown-Rozeff ARIMA model. Method 11 is the Brown-Rozeff ARIMA model with panel estimation. Method 12 is our proposed method with quarterly-seasonal adjustment (ARMA). Method 13 is our proposed method with quarterly-seasonal adjustment (ARIMA).

Table 3.16 Including Firm Size (RMSE and RE Based on RMSE), n = 1029 and T = 36

Method	RMSE	Relative Efficiencies on Method 13
1 (Cross-sectional)	0.1415	1.6158
2 (Panel)	0.1100	1.2563
3 (Time-series)	0.1224	1.3982
4 (Panel)	0.1105	1.2621
5 (Time-series)	0.1501	1.7144
6 (Panel)	0.0884	1.0100
7 (The proposed)	**0.1088**	**1.2431**
8 (Time-series)	0.4170	4.7622
9 (Panel)	0.0971	1.1094
10 (Time-series)	0.1517	1.7323
11 (Panel)	0.0944	1.0784
12 (ARMA)	**0.0897**	**1.0246**
13 (ARIMA)	**0.0876**	**1.0000**

Notes. The observations begining in the first quarter of 2010 and ending in the fourth quarter of 2018. There are 1029 firms in our sample. Method 1 is the simple cross-sectional method. Method 2 is the cross-sectional method with panel estimation. Method 3 is the simple time-series method. Method 4 is the time-series method with panel estimation. Method 5 is the multivariate time-sries regression model (MULT). Method 6 is MULT with panel estimation. Method 7 is the simple version of our proposed method that uses local learning. Method 8 is the seasonal autoregressive ARMA model. Method 9 is the seasonal autoregressive ARMA model with panel estimation. Method 10 is the Brown-Rozeff ARIMA model. Method 11 is the Brown-Rozeff ARIMA model with panel estimation. Method 12 is our proposed method with quarterly-seasonal adjustment (ARMA). Method 13 is our proposed method with quarterly-seasonal adjustment (ARIMA).

3.9 Appendix

3.9.1 Technical Details

Under smoothness condition of coefficient functions $\beta(z)$, for any given firm size z, $\beta(z_i))$ can be approximated by a polynomial function as

$$\beta(z_i) \approx \beta(z) + \frac{d\beta(z)}{dz}(z_i - z) + \cdots + \frac{d^m\beta(z)}{dz^m}(z_i - z)^m / m\,!,$$

where $\approx$ denotes the approximation by ignoring the higher orders, thus, $\beta(z_i)' x_{i,t}$ can be approximated by

$$\beta(z_i)' x_{i,t} \approx \sum_{j=0}^{m} \theta_j^T x_{i,t}\,(z_i - z)^j,$$

where $\theta_j = \frac{d^j\beta(z)}{dz^j} / j!$ for $0 \leq j \leq m$. Based on this motivation, we estimate α and $\beta(z)$ using the following local weighted regression:

$$\sum_{s=2}^{t} \sum_{i=1}^{n} \rho \left(CFO_{i,s} - \alpha - \sum_{j=0}^{m} \beta_j^T x_{i,s-1}\,(z_i - z)^j \right) K \left(\frac{z_i - z}{h} \right)$$

where $\rho(\cdot)$ is an appropriate criterion function (we may use either $\rho(\cdot) = |\cdot|$ or $\rho(\cdot) = (\cdot)^2$ in this paper), $K(\cdot)$ is a kernel function, and $h = h(n)$ is a sequence of positive numbers tending to zero and it controls the amount of smoothing used in estimation.

The we move to the prediction of CFO for Firm r at time $t+1$. To predict the cash flow of firm r at time $t+1$, i.e. $CFO_{r,t+1}$. Notice that

$$CFO_{r,t+1} = \alpha + \beta(z_r)' x_{r,t} + u_{r,t+1}$$

our predictor is given by

$$\widehat{CFO}_{r,t+1} = \widehat{\alpha} + \widehat{\beta}(z_r)' x_{r,t}.$$

When we choose $m = 1$, and $\rho(\cdot) = (\cdot)^2$, we have

$$\left(\widehat{\alpha}(z), \widehat{\theta}_0^T(z), \widehat{\theta}_1^T(z)\right) = \arg\min \sum_{s=2}^{t} \sum_{i=1}^{n} K \left(\frac{z_{is} - z}{h} \right) \left(CFO_{i,s} - \alpha - \beta_0^T x_{i,s-1} - \beta_1^T x_{i,s-1}\ (z_{is} - z) \right)^2.$$

If we re-write

$$\begin{aligned}
& CFO_{is} - \alpha - \beta_0^T x_{i,s-1} - \beta_1^T x_{i,s-1}\ (z_{is} - z) \\
= \ & CFO_{is} - \left(\alpha, \beta_0^T, \beta_1^T\right) \begin{bmatrix} 1 \\ x_{i,s-1} \\ x_{i,s-1}\ (z_{is} - z) \end{bmatrix} = CFO_{is} - \Theta^T \mathbb{Z}_{is} \\
= \ & Y_t - \left(\theta_0^T, \theta_1^T, \cdots, \theta_p^T\right) \begin{bmatrix} Z_t \\ Z_t\left(X_t - x_0\right) \\ \cdots \\ Z_t\left(X_t - x_0\right)^p \end{bmatrix} = Y_t - \theta^T \mathbb{Z}_t
\end{aligned}$$

where

$$\Theta^T = \left(\alpha, \beta_0^T, \beta_1^T\right), \text{and } \mathbb{Z}_{is} = \begin{bmatrix} 1 \\ x_{i,s-1} \\ x_{i,s-1}\ (z_{is} - z) \end{bmatrix}.$$

Let $K_{is} = K((z_{i,s} - z)/h)$, then

$$\begin{aligned}
Q_n\left(z; \boldsymbol{\Theta}\right) &= \sum_{s=2}^{t} \sum_{i=1}^{n} K\left(\frac{z_{is} - z}{h}\right) \left(CFO_{i,s} - \alpha - \beta_0^T x_{i,s-1} - \beta_1^T x_{i,s-1}\ (z_{is} - z)\right)^2 \\
&= \sum_{s=2}^{t} \sum_{i=1}^{n} K_{is} \left(CFO_{is} - \Theta^T \mathbb{Z}_{is}\right)^2 \\
&= \sum_{t=1}^{T} K_t \left(Y_t - \boldsymbol{\Theta}^T \mathbb{Z}_t\right)^2
\end{aligned}$$

Thus, the first order condition corresponding to the local optimization with a quadratic loss function is given by

$$\left(\sum_{s=2}^{t} \sum_{i=1}^{n} K_{is} \mathbb{Z}_{is} \mathbb{Z}_{is}^T\right) \widehat{\Theta} = \sum_{s=2}^{t} \sum_{i=1}^{n} K_{is} \mathbb{Z}_{is} CFO_{is}$$

Thus,

$$\widehat{\Theta} = \left(\sum_{s=2}^{t} \sum_{i=1}^{n} K_{is} \mathbb{Z}_{is} \mathbb{Z}_{is}^T\right)^{-1} \sum_{s=2}^{t} \sum_{i=1}^{n} K_{is} \mathbb{Z}_{is} CFO_{is}$$

References

ACHARYA, V. V., T. EISERT, C. EUFINGER, AND C. HIRSCH (2018): "Real effects of the sovereign debt crisis in Europe: Evidence from syndicated loans," *The Review of Financial Studies*, 31, 2855–2896.

ALLEN, F., J. QIAN, AND M. QIAN (2005): "Law, finance, and economic growth in China," *Journal of financial economics*, 77, 57–116.

AMORE, M. D. AND M. BENNEDSEN (2013): "The value of local political connections in a low-corruption environment," *Journal of Financial Economics*, 110, 387–402.

ARIFF, M. AND C. LUC (2008): "Cost and profit efficiency of Chinese banks: A non-parametric analysis," *China Economic Review*, 19, 260–273.

ASMILD, M. AND K. MATTHEWS (2012): "Multi-directional efficiency analysis of efficiency patterns in Chinese banks 1997–2008," *European Journal of Operational Research*, 219, 434–441.

BAILEY, W., W. HUANG, AND Z. YANG (2011): "Bank loans with Chinese characteristics: some evidence on inside debt in a state-controlled banking system," *Journal of Financial and Quantitative Analysis*, 46, 1795–1830.

BALDUZZI, P., E. BRANCATI, AND F. SCHIANTARELLI (2017): "Financial markets, banks' cost of funding, and firms' decisions: Lessons from two crises," *Journal of Financial Intermediation*.

BARTH, M. E., D. P. CRAM, AND K. K. NELSON (2001): "Accruals and the prediction of future cash flows," *The accounting review*, 76, 27–58.

BEAVER, W. H. (1970): "The time series behavior of earnings," *Journal of Accounting Research*, 62–99.

BERGER, A. N., I. HASAN, AND M. ZHOU (2009): "Bank ownership and efficiency in China: What will happen in the world's largest nation?" *Journal of Banking & Finance*, 33, 113–130.

BERNARD, V. L. AND T. L. STOBER (1989): "The nature and amount of information in cash flows and accruals," *Accounting review*, 624–652.

BERTRAND, M., F. KRAMARZ, A. SCHOAR, AND D. THESMAR (2018): "The Cost of Political Connections," *Review of Finance*, 22, 849–876.

BONTEMPI, G., S. B. TAIEB, AND Y.-A. LE BORGNE (2012): "Machine learning strategies for time series forecasting," in *European business intelligence summer school*, Springer, 62–77.

BROWN, L. D. AND M. S. ROZEFF (1979): "Univariate time-series models of quarterly accounting earnings per share: A proposed model," *Journal of Accounting Research*, 179–189.

BUTLER, A. W., L. FAUVER, AND S. MORTAL (2009): "Corruption, political connections, and municipal finance," *The review of financial studies*, 22, 2873–2905.

CAI, Z. AND Z. XIAO (2012): "Semiparametric quantile regression estimation in dynamic models with partially varying coefficients," *Journal of Econometrics*, 167, 413–425.

CHEN, Y., Y. SHI, X. WEI, AND L. ZHANG (2014): "Domestic systemically important banks: a quantitative analysis for the Chinese banking system," *Mathematical Problems in Engineering*, 2014.

CHODOROW-REICH, G. (2013): "The employment effects of credit market disruptions: Firm-level evidence from the 2008–9 financial crisis," *The Quarterly Journal of Economics*, 129, 1–59.

CLAESSENS, S., E. FEIJEN, AND L. LAEVEN (2008): "Political connections and preferential access to finance: The role of campaign contributions," *Journal of financial economics*, 88, 554–580.

CONG, L. W., H. GAO, J. PONTICELLI, AND X. YANG (2019): "Credit allocation under economic stimulus: Evidence from China," *The Review of Financial Studies*, 32, 3412–3460.

DECHOW, P. M. AND I. D. DICHEV (2002): "The quality of accruals and earnings: The role of accrual estimation errors," *The accounting review*, 77, 35–59.

DECHOW, P. M., S. P. KOTHARI, AND R. L. WATTS (1998): "The relation between earnings and cash flows," *Journal of accounting and Economics*, 25, 133–168.

DEFOND, M. L. AND M. HUNG (2003): "An empirical analysis of analysts' cash flow forecasts," *Journal of accounting and economics*, 35, 73–100.

DONG, Y., M. FIRTH, W. HOU, AND W. YANG (2016): "Evaluating the performance of Chinese commercial banks: A comparative analysis of different types of banks," *European Journal of Operational Research*, 252, 280–295.

DREHER, A. AND M. GASSEBNER (2013): "Greasing the wheels? The impact of regulations and corruption on firm entry," *Public Choice*, 155, 413–432.

DUCHIN, R., O. OZBAS, AND B. A. SENSOY (2010): "Costly external finance, corporate investment, and the subprime mortgage credit crisis," *Journal of Financial Economics*, 97, 418–435.

FACCIO, M. (2006): "Politically connected firms," *American economic review*, 96, 369–386.

FAN, C. S. AND H. I. GROSSMAN (2001): "Incentives and corruption in Chinese economic reform," *Journal of Economic Policy Reform*, 4, 195–206.

FENECH, J.-P., Y. K. YAP, AND S. SHAFIK (2014): "Can the Chinese banking system continue to grow without sacrificing loan quality?" *Journal of International Financial Markets, Institutions and Money*, 31, 315–330.

FISMAN, R. (2001): "Estimating the value of political connections," *American economic review*, 91, 1095–1102.

FISMAN, R. AND J. SVENSSON (2007): "Are corruption and taxation really harmful to growth? Firm level evidence," *Journal of development economics*, 83, 63–75.

FU, X. M. AND S. HEFFERNAN (2009): "The effects of reform on China's bank structure and performance," *Journal of Banking & Finance*, 33, 39–52.

FUNGÁČOVÁ, Z., R. HERRALA, AND L. WEILL (2013): "The influence of bank ownership on credit supply: Evidence from the recent financial crisis," *Emerging Markets Review*, 15, 136–147.

GAO, H., H. RU, AND D. Y. TANG (2019): "Subnational debt of China: The politics-finance nexus," *Available at SSRN 2801878*.

GARCÍA-HERRERO, A., S. GAVILÁ, AND D. SANTABÁRBARA (2006): "China's banking reform: an assessment of its evolution and possible impact," *CESifo Economic Studies*, 52, 304–363.

GERTLER, M. AND S. GILCHRIST (1994): "Monetary policy, business cycles, and the behavior of small manufacturing firms," *The Quarterly Journal of Economics*, 109, 309–340.

GERTLER, M. AND P. KARADI (2011): "A model of unconventional monetary policy," *Journal of monetary Economics*, 58, 17–34.

GERTLER, M. AND N. KIYOTAKI (2010): "Financial intermediation and credit policy in business cycle analysis," in *Handbook of monetary economics*, Elsevier, vol. 3, 547–599.

GIVOLY, D., C. HAYN, AND R. LEHAVY (2009): "The quality of analysts' cash flow forecasts," *The Accounting Review*, 84, 1877–1911.

GOLDMAN, E., J. ROCHOLL, AND J. SO (2008): "Do politically connected boards affect firm value?" *The Review of Financial Studies*, 22, 2331–2360.

GRIFFIN, J., C. LIU, AND T. SHU (2016): "Is the Chinese anti-corruption campaign effective," *University of Texas, Austin*.

HASTIE, T., R. TIBSHIRANI, AND J. FRIEDMAN (2009): *The elements of statistical learning: data mining, inference, and prediction*, Springer Science & Business Media.

HENDERSON, J. V., A. STOREYGARD, AND D. N. WEIL (2012): "Measuring economic growth from outer space," *American economic review*, 102, 994–1028.

HOLMSTROM, B. AND J. TIROLE (1997): "Financial intermediation, loanable funds, and the real sector," *the Quarterly Journal of economics*, 112, 663–691.

HRIBAR, P. AND D. W. COLLINS (2002): "Errors in estimating accruals: Implications for empirical research," *Journal of Accounting research*, 40, 105–134.

HUANG, Q., J. DE HAAN, AND B. SCHOLTENS (2015): "Analysing systemic risk in the Chinese banking system," *Pacific Economic Review*.

JOHNSON, S. AND T. MITTON (2003): "Cronyism and capital controls: evidence from Malaysia," *Journal of financial economics*, 67, 351–382.

KHWAJA, A. I. AND A. MIAN (2005): "Do lenders favor politically connected firms? Rent provision in an emerging financial market," *The Quarterly Journal of Economics*, 120, 1371–1411.

——— (2008): "Tracing the impact of bank liquidity shocks: Evidence from an emerging market," *American Economic Review*, 98, 1413–42.

KIM, M. AND W. KROSS (2005): "The ability of earnings to predict future operating cash flows has been increasing—not decreasing," *Journal of Accounting research*, 43, 753–780.

KIYOTAKI, N. AND J. MOORE (2012): "Liquidity, business cycles, and monetary policy," Tech. rep., National Bureau of Economic Research.

KWONG, J. (2015): *The political economy of corruption in China*, Routledge.

LEUZ, C. AND F. OBERHOLZER-GEE (2006): "Political relationships, global financing, and corporate transparency: Evidence from Indonesia," *Journal of financial economics*, 81, 411–439.

LI, B., Z. WANG, AND H. ZHOU (2018): "China's anti-corruption campaign and credit reallocation from SOEs to non-SOEs," *PBCSF-NIFR Research Paper*.

LIN, C., R. MORCK, B. YEUNG, AND X. ZHAO (2016): "Anti-corruption reforms and shareholder valuations: Event study evidence from China," Tech. rep., National Bureau of Economic Research.

LIN, X. AND Y. ZHANG (2009): "Bank ownership reform and bank performance in China," *Journal of Banking & Finance*, 33, 20–29.

LIU, L. X., H. SHU, AND K. J. WEI (2017): "The impacts of political uncertainty on asset prices: Evidence from the Bo scandal in China," *Journal of financial economics*, 125, 286–310.

LOREK, K. S., T. F. SCHAEFER, AND G. L. WILLINGER (1993): "Time-series properties and predictive ability of funds flow variables," *Accounting Review*, 151–163.

LOREK, K. S. AND G. L. WILLINGER (1996): "A multivariate time-series prediction model for cash-flow data," *Accounting Review*, 81–102.

——— (2011): "Multi-step-ahead quarterly cash-flow prediction models," *Accounting Horizons*, 25, 71–86.

LU, X. (2000): "Booty socialism, bureau-preneurs, and the state in transition: Organizational corruption in China," *Comparative Politics*, 273–294.

MAURO, P. (1995): "Corruption and growth," *The quarterly journal of economics*, 110, 681–712.

MEHROTRA, N., N. CROUZET, ET AL. (2017): "Small and Large Firms over the Business Cycle," in *2017 Meeting Papers*, Society for Economic Dynamics, 600.

PING, X. AND L. LEI (2003): "Financial Corruption in Transition: the Transaction Characteristics and Institutional Incentives for Gray Financial Activities [J]," *Economic Research Journal*, 6.

QIAN, J. Q., P. E. STRAHAN, AND Z. YANG (2015): "The impact of incentives and communication costs on information production and use: Evidence from bank lending," *The Journal of Finance*, 70, 1457–1493.

SAPIENZA, P. (2004): "The effects of government ownership on bank lending," *Journal of financial economics*, 72, 357–384.

SHLEIFER, A. AND R. W. VISHNY (1993): "Corruption," *The quarterly journal of economics*, 108, 599–617.

——— (1994): "Politicians and firms," *The Quarterly Journal of Economics*, 109, 995–1025.

SONG, Z., K. STORESLETTEN, AND F. ZILIBOTTI (2011): "Growing like china," *American Economic Review*, 101, 196–233.

TING, G. (1997): "Forms and Characteristics of China's Corruption in the 1990s: Change with Continuity," *Communist and Post-Communist Studies*, 30, 277–288.

WANG, H., H. WANG, L. WANG, AND H. ZHOU (2019): "Shadow banking: China's dual-track interest rate liberalization," *Available at SSRN 2606081*.

WANG, Y., X. SHAN, AND J. GENG (2015): "Estimating the Systemic Risk of China's Banking Industries Based on Merton Model," *Applied Mathematics & Information Sciences*, 9, 957.

WEDEMAN, A. (2004): "Great disorder under heaven: Endemic corruption and rapid growth in contemporary China," *China Review*, 1–32.

XIE, P. AND L. LU (2005): "Economic analysis of China's financial corruption," .

YONG, G. AND H. AN-GANG (2003): "Administrative Monopoly, Rent Seeking and Corruption: An Analysis of the Corruption Texture in the Transition Economy [J]," *Comparative Economic and Social Systems*, 2.

ZEUME, S. (2017): "Bribes and firm value," *The Review of Financial Studies*, 30, 1457–1489.

ZHANG, W., S.-Y. LEE, AND X. SONG (2002): "Local polynomial fitting in semivarying coefficient model," *Journal of Multivariate Analysis*, 82, 166–188.