Spring 5-15-2016

Essays in Macroeconomics, Economic Development, and Growth

Lin Shao
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS
Department of Economics

Dissertation Examination Committee:

B. Ravikumar, Chair
Gaetano Antinolfi
Costas Azariadis
Yongseok Shin
Stephen D. Williamson

Essays in Macroeconomics, Economic Development, and Growth

by
Lin Shao

A dissertation presented to the
Graduate School of Arts & Sciences
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

May 2016
St. Louis, Missouri

Contents

List of Figures

List of Tables

Acknowledgments

Almost ten years ago, I made a decision to switch from computer science to economics, for I had never been as interested in machines as in people. I had, of course, envisioned many possible scenarios for the next 10 years. But life, as complex and mysterious as it can be, always beats one's widest imaginations, and the journey to the finalization of this thesis is anything but boring.

I am grateful for all the great minds that I have met. I am indebted to my advisor B. Ravikumar for molding my mind into thinking like an economist and like a serious researcher. This thesis can never come into this shape without the encouragement and advices from Yongseok Shin. My former advisor Zheng Song at Fudan introduced economics to me, and I forever owe him for bringing me to the frontier of the macroeconomic research. I consider myself extremely lucky to have studied under Costas Azariadis, Srihari Govindan, Rody Manuelli, Stephen Williamson, and Fabrizio Zilibotti. They have made me realize how good an economist one can aspire to be. I do not know how to or whether I will ever be able to repay this huge intellectual debt; I can only wish to be given an opportunity to pass it along in the future.

I am thankful for the people who kindly shared a part of their lives with me, and allowed me to do the same with them–Simon Alder, Jan Auerbach, Brian Bergfeld, Wan-jung Cheng, Silvio Contessi, Lihe Guo, Mushegh Harutyunyan, Elisa Keller, Jan-Aaron Klaasson, Andreas Müller, Luciene Torres de Mello Pereira, Faisal Sohail, Chien-chiang Wang, Yikai Wang, Wen Yao, Emircan Yurdagul, and Zhuoyao Zhang–for all the good and the bad, the happiness and the sadness, the goofiness and the insights, and everything else.

Lastly, and more importantly, there is my brother, who generously helped me through difficult times. And there are my parents. Their endless and sometimes aimless love always reminds me of Rilke's words: "...believe in a love that is being stored up for you like an inheritance, and have faith that in this love there is a strength and a blessing so large that you can travel as far as you wish without having to step out of it." I dedicate this thesis to them.

To my parents.

ABSTRACT OF THE DISSERTATION

Essays in Macroeconomics, Economic Development, and Growth

for Arts & Sciences Graduate Students

by

Lin Shao

Doctor of Philosophy in Economics

Washington University in St. Louis, 2016

Professor B. Ravikumar, Chair

Chapter 1: Trade credit in production chains In an economy where production takes place in multiple stages and is subject to financial frictions, how firms finance intermediate inputs matters for aggregate outcomes. This paper focuses on trade credit – the lending and borrowing of input goods between firms in a production chain – and quantifies its aggregate impact. Motivated by empirical evidence, our model shows how trade credit alleviates financial frictions through a process of credit redistribution and creation, thus leading to a higher out- put level in the steady state. However, the flow of trade credit is prone to disruptions when financial crises hit the economy. The decline in economic activities following crises is in turn amplified by disruptions in trade credit. The model simulation suggests that the drop in trade credit during the Great Recession can account for almost one-fourth of the observed decline in output.

Chapter 2: Financial development beyond the formal financial market This paper argues that to understand the quantitative importance of finance in economic development, it is important to look beyond the formal financial market. We document that informal financing is more accessible in countries with a highly developed formal financial market. The volume of informal financing, as well as the substitutability of informal financing for formal financing, are both positively correlated with the development of the formal financial market.

We build a quantitative model in which a fundamental contract enforcement problem delivers the documented empirical patterns. The model is then disciplined to match aggregate and distributional moments of bank credit and trade credit – an important informal financial institution – of the U.S. and Chinese manufacturing firms, respectively. Our quantitative analysis suggests that by focusing on bank credit only, we understate the importance of finance in explaining the income differences between these two countries.

Chapter 3: Economic reforms and industrial policy in a panel of Chinese cities
We study the effect of place-based industrial policy on economic development, focusing on the establishment of Special Economic Zones (SEZ) in China. We use data from a panel of Chinese (prefecture-level) cities from 1988 to 2010. Our difference-in-difference estimation exploits the variation in the establishment of SEZ across time and space. We find that the establishment of a state-level SEZ is associated with an increase in the level of GDP of about 20%. This finding is confirmed with alternative specifications and in a sub-sample of inland provinces, where the selection of cities to host the zones was based on administrative criteria. The main channel is a positive effect on physical capital accumulation, although SEZ also have a positive effect on total factor productivity and human capital investments. We also investigate whether there are spillover effects of SEZ on neighboring regions or cities further away. We find positive and often significant spillover effects.

Chapter 1

Trade Credit in Production Chains

1.1 Introduction

Most studies on the macroeconomic effect of finance focus on financial activities between nonfinancial and financial sectors, or those within financial sectors, such as the inter-bank lending market. In this paper, we look into the nonfinancial sector and show how trade credit – a short-term loan in the form of input goods between nonfinancial firms – affects macroeconomic outcomes.

In practice, trade credit is an important financial resource for U.S. firms. Take the U.S. nonfinancial corporate sector as an example. Accounts payable (trade credit liability) are about one-third the size of the quarterly GDP in 2006. Trade credit is pro-cyclical and very volatile. The standard deviation of trade credit is about twice as high as that of quarterly GDP. During the 2007–09 financial crisis, short-term liabilities decrease for about 457 billion dollars from 2007Q4 to 2009Q2. More than 70 percent of the drop could be accounted for by the drop in trade credit.

Despite the size of trade credit, its macroeconomic implication is largely unexplored. We try to fill in the void by providing a quantitative assessment of how trade credit affects the level and volatility of the aggregate economy.

We find that the existence of trade credit significantly increases the aggregate output of the U.S. economy during normal times as well as its volatility over credit cycles.

During normal times, that is, the steady state of the economy without aggregate shocks, trade credit channels financial resources to flow into firms whose production is constrained by their capacity to obtain bank loans. By borrowing input goods from their suppliers, constrained firms achieve a larger production scale that is closer to the optimal level. Resources are better allocated in the economy, and aggregate output is higher.

During a financial crisis, bank lending, including short-term lending, is tightened. Short of liquidity, firms cut back trade credit extension and demand that sales be made on the spot. Constrained firms who are able to use trade credit before the crisis are no longer able to do so. In this case, the financial crisis affects the real economy through its impact on trade credit. This indirect effect of financial crisis, *vis-à-vis* the direct effect through the tightening of bank credit, brings additional damages to the economy and amplifies the magnitude of financial shocks.

We begin our exploration by documenting two sets of empirical facts using firm level data. First, in a pooled sample of firms from Compustat and the Survey of Small Business Finances (SSBF), we show that in net terms, trade credit flows from unconstrained to constrained firms. Second, using a panel of Compustat firms that borrow from the syndicated loan market, we document that firms respond to a disruption in their access to the financial market by cutting back their lending of trade credit.

We then build trade credit into a dynamics general equilibrium model with heterogeneous entrepreneurs to deliver the documented empirical patterns at the aggregate and firm levels.

In the model, production takes place in two stages: intermediate goods production stage and final goods production stage. Each stage is populated by a measure one of heterogeneous entrepreneurs operating a decreasing return to scale technology. Entrepreneurs differ from each other by productivity and wealth.

There is a competitive banking sector. The production scale of entrepreneurs is bounded by a working capital constraint *à la* Jermann and Quadrini (2012). Due to the moral hazard problem of entrepreneurs, bank loans are limited by the amount of collateral that they have to offer. As a result, only entrepreneurs with enough wealth can obtain enough bank loans to finance working capital and achieve their optimal production scale.

Trade credit exists because intermediate input entrepreneurs have a certain comparative advantage over banks in lending input goods. In particular, following Burkart and Ellingsen (2004), we assume that unlike bank loans, intermediate input goods can not be diverted. Under this assumption, it is secure for intermediate goods entrepreneurs to lend input goods. Banks internalize this comparative advantage by lending to intermediate goods entrepreneurs against accounts receivable (accounts receivable financing).

Although intermediate goods entrepreneurs have a comparative advantage in lending input goods, trade credit can be expensive. For constrained intermediate goods entrepreneurs, cash flow is valuable since it can be used to finance their own working capital. Extending trade credit means postponing cash flow, and thus it needs to be compensated for by an interest rate on trade credit. The more constrained intermediate goods entrepreneurs are, the higher the trade credit interest rate is in equilibrium.

We capture the demand and supply forces of trade credit by assuming a competitive trade credit market. Very constrained final goods entrepreneurs borrow from this market while relatively unconstrained intermediate goods entrepreneurs lend to it. This creates a flow of trade credit from unconstrained to constrained entrepreneurs.

Trade credit changes entrepreneurs' access to financial resources in two ways: First, trade credit increases available financial resources for final goods entrepreneurs by reallocating unused bank credit from intermediate goods entrepreneurs to their constrained customers. This is labeled as *credit redistribution* channel. Second, there is the *credit creation* channel.

Credit is created when banks lend against accounts receivable. Trade credit extension, combined with accounts receivable financing, increases the collective access to credit for both final and intermediate goods entrepreneurs.

We calibrate the steady state of the model to match aggregate and distributional data moments for the U.S. nonfinancial corporate sector. In particular, the steady state is calibrated to match the ratio of trade credit to gross value added, in order to capture the importance of trade credit in financing production and the distribution in firms size and revenue.

With the calibrated model as benchmark, we proceed to quantify the aggregate effects of trade credit by comparing key economic outcomes generated by the benchmark and a counterfactual economy, in which all the transactions of input goods have to be made on the spot.

In the steady state analysis, the parameters of the counterfactual economy without trade credit are set to be the same as the benchmark economy. With the same set of parameters, the aggregate output of the counterfactual economy is 24 percent lower than that of the benchmark. This can be decomposed into 15 percent lower capital stock, 24 percent lower labor, and 8 percent lower aggregate TFP. Trade credit greatly alleviates the borrowing constraint of final goods entrepreneurs. This leads to a higher aggregate TFP and output of the final goods sector. Although the impact of trade credit on the aggregate TFP of the intermediate goods sector is negligible; trade credit increases the output of the intermediate goods sector through a general equilibrium effect – a higher demand for intermediate input goods from the final goods sector.

To analyze the role of trade credit during the 2007-09 financial crisis, we simulate the benchmark model with an unexpected shock on the collateral constraint. The magnitude of the shock is calibrated to match the observed drop in both the ratio of credit market liability and trade credit to nonfinancial assets during this period.

Following the shock, the aggregate output drops by 6 percent in the benchmark, closely matching the data. Constrained entrepreneurs are affected by the shock. They are forced to scale down their production. Unconstrained entrepreneurs increase their production scale in response to lower prices for input goods. Because unconstrained entrepreneurs are on average less productive than constrained ones, this shift in production scales leads to a lower aggregate TFP and a drop in the aggregate output.

The aggregate effect of the financial crisis can be decomposed into two parts: The first one operates directly through the tightening of bank credit, which is the standard mechanism in financial crisis research. The second part operates indirectly through changing firms' ability to use trade credit. Following the shock, both intermediate and final goods entrepreneurs become more constrained. Final goods entrepreneurs want to borrow more trade credit, while intermediate goods entrepreneurs want to cut back their trade credit lending. This leads to a spike in the trade credit interest rate from a pre-crisis level of 2.7 percent to as high as 8.3 percent. Modestly constrained final goods entrepreneurs find it no longer profitable to use trade credit to finance production. Only very constrained ones use trade credit, but at a smaller scale. The aggregate volume of trade credit experiences a huge drop–the percentage drop in trade credit is almost twice as high as that in output. The drop in trade credit volume and the increase in trade credit interest rate both indicate that the financial shock hinders the role of trade credit in improving allocation efficiency.

To quantify the indirect effect of financial crisis through trade credit, we first recalibrate the counterfactual economy without trade credit such that it is comparable with the benchmark in steady state. We then feed into the counterfactual economy the same financial shock that generates the 2007–09 financial crisis in the benchmark. Following the shock, aggregate output only drops by 4.6 percent, which is a 1.4 percentage points smaller than the 6 percent drop of output in the benchmark. This means that, without trade credit, the output drop during the 2007–09 financial crisis would have been 23 percent smaller.

Related Literature This paper belongs to several strands of literature.

First, our model builds on previous research that studies the financial aspects of trade credit.[1] Empirically, starting from Meltzer (1960) and more recently in Cunningham (2004), authors document that small firms rely more on trade credit than large firms do. Theoretically, previous papers such as those of Cunat (2007), Biais and Gollier (1997), and Burkart and Ellingsen (2004) show that the existence of trade credit is a result of a certain comparative advantage that intermediate goods producers have over financial intermediaries in lending to their customers. Our model builds up on the above papers but differs from them by jointly analyzing the demand and supply of trade credit.

Second, this paper is also related to papers that study how trade credit propagates financial shocks. Kiyotaki and Moore (1997a) build a model in which firms are linked through an input-output channel as well as a trade credit channel. Shocks to one firm are propagated to their suppliers through a chain of defaults in trade credit. Raddatz (2010) constructs a cross-country dataset on trade credit linkage between sectors. The paper shows trade credit contributes to the comovement between sectors along the business cycle. In Jacobson and von Schedvin (2015), using Swedish firm-level data, the authors show how bankruptcy propagates through loss of trade credit. In these papers, the propagation mechanism operates through trade credit default while in our paper, the mechanism operates through changes in trade credit on the intensive margin.

Third, this paper also contributes to a relatively new strand of literature that studies how financial frictions and financial shocks affect real economy when taking account of the input-output linkage. Among them, Zetlin-Jones and Shourideh (2016) show that a financial shock that affects only a subset of firms can be transmitted to other firms through an input-output channel and amplified by the complementarity of input goods. Bigio and La'O (2014) shows

[1] See Petersen and Rajan (1997) and Cunat and Garcia-Appendini (2012) for excellent surveys of the literature.

that the input-output production network leads to a "liquidity multiplier" that amplifies liquidity shocks. They assume that only a fixed fraction of the input goods can be purchased using trade credit. This economy is more fragile in the face of liquidity shocks compared to an economy without the input-output linkage, because it requires a higher level of liquidity to sustain production. Kalemli-Ozcan et al. (2014) builds up on Kim and Shin (2012) and develops a model of production chain in which firms hold interlocking claims and obligations of trade credit. A financial shock is amplified in this economy because long production chains are less viable during a financial crisis. We think that our paper is complementary to Bigio and La'O (2014) because we take as given the production structure of the economy while explicitly modeling firms' choice of trade credit. The endogenous changes in trade credit, rather than the input-output structure itself, amplifies the financial shocks.

More broadly speaking, this paper belongs to a long strand of literature that studies how shocks originating from the financial sector are transmitted and amplified to affect the real economy (see Bernanke et al., 1989; Kiyotaki and Moore, 1997b; Jermann and Quadrini, 2012). More recently, researchers further the understanding of this question by studying the 2007–09 financial crisis. Among them, Bigio (forthcoming) and Gertler and Kiyotaki (2010) study how changes in banks' net worth affect their ability of financial intermediation. Guerrieri and Lorenzoni (2012) study the impact of credit crunch on consumption. Mian and Sufi (2011) and Midrigan and Philippon (2011) show how financial crises affect the economy through the consumer demand channel. Buera et al. (2015a), Siemer (2014), and Arellano et al. (2012) focus on explaining how the 2007–09 financial crisis affects the labor market. We contribute to this literature by discovering an amplification mechanism outside the financial sector and quantifying its size during the 2007–09 financial crisis.

Finally, this paper is reminiscent of a strand of literature that studies the efficiency and stability in the interbank lending market. According to Lee (2015), the circulation of repo contract improves the allocation efficiency, but endogenous changes in repo spread exacerbate the financial crisis through a positive feedback loop. Zhang (2014) illustrates how a shock to the collateral risk of the repo is amplified through a chain of repo defaults. According to Boissay and Cooper (2014), through the process of lending to firms, banks create "inside collateral," which can be used to borrow in the interbank lending market. The creation of collateral gives rise to multiple equilibria in the interbank lending market and makes it more fragile. Our paper shows that similar to interbank lending, trade credit helps channel financial resources to their most productive use but makes the economy more vulnerable to financial shocks.

The rest of the paper is organized as follows: Section 1.2 presents empirical evidence on trade credit that motivates our modeling choices, section 1.3 introduces the benchmark model and section 1.4 defines its recursive equilibrium. We calibrate the steady state of the benchmark model to match the U.S. economy in section 1.5 and study the model dynamics during the 2007-09 financial crisis in section 1.6. Section 1.7 introduces a counterfactual economy without trade credit and studies the aggregate implication of trade credit by comparing the benchmark and the counterfactual economy. Section 1.8 concludes.

1.2 Empirical motivation

In this section, we document two sets of empirical facts. In the first part, we use firm level data to establish the linkage between firms' choice of trade credit and financial constraints that they face. In the second part, we examine the impact of financial market disruptions on trade credit by studying how disruptions in the syndicated loan market during the 2007-09 financial crisis impacts the extension of trade credit for firms who borrow from that market.

The measurement of trade credit deserves some discussions before we present the empirical evidence. Ideally, a comprehensive measure of trade credit should consist of transaction records between firms, including for example value of sales, value of sales made through trade credit, interest rate of trade credit, and its maturity. Unfortunately such transaction records are not available at large scales, forcing us to take an indirect approach to measure trade credit. We infer trade credit extension and usage from firms' balance sheet. Trade credit extended to other firms appears on the asset side of balance sheet as accounts receivable (AR). Trade credit borrowed from other firms appears on the liability side of the balance sheet as accounts payable (AP).[2] Net accounts receivable (Net AR), which is defined as AR net AP, measures the firm's net lending through trade credit. Dividing AR, AP, and net AR by sales gives three variables that measure trade credit in this paper: 1) AR to sales ratio (AR/Sales), measuring the extension of trade credit, 2) AP to sales ratio (AP/Sales), measuring the using of trade credit, and 3) Net AR to sales ratio (Net AR/Sales), measuring the net lending through trade credit.

1.2.1 Trade credit and being financially constrained

In this section, we show how being financially constrained affects firms' choice of trade credit. Section 1.2.1.1 discusses the data and the definition of financially constrained firms. Section 1.2.1.2 presents and discusses the results.

1.2.1.1 Data

To construct our sample of firms, we combine Compustat North America annual and SSBF data for the years when SSBF data is available (1987, 1993, 1998 and 2003). Firms in the financial sector (SIC 60-69) and wholesale and retail sector (SIC 50-59) are dropped. Since we focus on firms that engage in production, it is clear why financial firms are excluded. The decision to exclude wholesale and retail sector firms are due to two reasons. First, previous

[2]Accounts receivable and payable include receivable and payable that are not related to trade credit. Whenever the data is available, we use "trade receivable" and "trade payable" on the balance sheet to measure trade credit. Using the Compustat data, we find that on average approximately 90 percent of the account receivable is trade receivables. Moreover, the ratio of trade receivable over accounts receivable is rather consistent across different industries and firm sizes.

research shows that the choice of trade credit between retailers and their suppliers is affected by the monopolistic power of large retail stores such as Walmart. Second, accounts receivable of retails might contain consumer credit, which is not the object of this study.

We first consider a sample consisting of only Compustat firms. Following Almeida and Campello (2007), we create three separate dummy variables for financially constrained firms in this sample. The first one is based on payout ratio – firm i with 0 payout in year t is identifies as being financially constraint in that year ($I\_constrained_{it} = 1$). The second one is based on firms' access to bond market, a firm is identified as financially constrained if it has neither a long-term nor a short-term bond rating from the Standard & Poor. The third one is based on asset size of firms. A firm is financially constrained if it is among bottom 30 percentile in the asset size distribution.

Second, we augment the above sample of Compustat firms with SSBF data, which contains small and private firms. The combined Compustat-SSBF sample offers a more comprehensive view of the whole population of U.S. firms. For this sample, we again define a firm to be financially constrained ($I\_constrained_{it} = 1$) if it is among bottom 30 percentile in the asset size distribution. Compustat firms consist of 22 percent of the financially constrained firms in this Compustat-SSBF sample.

1.2.1.2 Being constrained and the choice of trade credit

We apply the following specification to estimate the effect of being financially constrained on the choice of trade credit,

$$y_{ist} = \phi_{st} + \alpha I\_constrained_{it} + \chi_i + \epsilon_{ist}, \tag{1.1}$$

where y_{ist} is one of the three measurement of trade credit – AR/Sales, AP/Sales, and Net AR/Sales – of firm i in sector s of year t, ϕ_{st} is the sector-year fixed effect, and χ_i is a vector of other non time-varying fixed characteristics of the firm including whether it is a Compustat firms and whether it is a corporation.[3]

The coefficient on the dummy variable $I\_constrained_{it}$, α, is the object of interests, for it estimates the effect of being financially constrained on trade credit choices. The results are presented in Table 1.1. In Panel (A), the dependent variable is Net AR/Sales. Compared to unconstrained Compustat firms, constrained ones maintain a significantly lower net AR/Sales: 6.2 percentage point lower for the firms that have 0 payout (column 1), 5.8 percentage point lower for the firms that do not have a S&P rating (column 2), 11.5 percentage point lower for the firms that are among the bottom 30 percentile of the asset size distribution (column 3). A similar result holds for the Compusta-SSBF sample. As shown in column (4), the net AR/Sales is 17.1 percentage points lower for constrained firms in this sample and highly significant. Since Net AR/Sales measures the relative position of firms in the network of trade credit, results in Panel A suggest that constrained firms are the net borrower of trade credit while unconstrained firms are net lenders.

We have established so far that net lending through trade credit is significantly higher for constrained firms. We continue to explore whether the difference in net lending is due to the difference in the trade credit extension or usage. In Panel (B) and (C), we run specification 1.1 while using AP/Sales and AR/Sales as dependent variable, respectively. As shown in Panel (B), constrained firms use significantly more trade credit than their unconstrained counterparts: 6.6 percentage point more for 0 payout firms, 7 percentage point more for firms without public bond rating, and 10.1 and 12.3 percentage point more for small firms in the two samples respectively. On the contrary, the effect of being financial constrained on AR/Sales is much weaker and more ambiguous (see Panel C). Firms with 0 payout ratio (column 1) and firms without public bond rating (column 2) extend slightly more trade

[3]Unfortunately, due to the limited information on firm age in the Compustat data, we can not control for firm age in these regression, which admittedly is an potentially important factor in trade credit choices.

credit compared to the unconstrained firms. Small firms in both samples extend less trade credit, but the point estimates are relatively low – 1.4 percentage point for Compustat sample (column 3) and 4.8 percentage points for the Compustat-SSBF sample (column 4).

In summary, we find that constrained firms borrow more trade credit than unconstrained ones. However, they do not seem to extend less trade credit to other firms.

Accounts receivable financing Accounts receivable financing is a practice in which firms use accounts receivable as collateral to obtain bank loans.[4]. It is an important and often neglected part of trade credit in the previous literature. Due to the lack of data, it is difficult for us to evaluate the aggregate volume of accounts receivable financing in the U.S.. We can infer, however, the importance of accounts receivable financing by looking at the syndicated loan market using Thomson-Reuters Dealscan data.[5] We document that, among all secured credit line facilities that were opened during 2004-06, 46.3 percent of which require accounts receivable as collateral. The rest of the facility requires other assets including inventory, property, plants, or equipment. In addition, accounts receivable is a type of rather liquid asset compared to the rest. Among these facilities, average advance rate (borrowing base percentage) for "accounts receivable" is 87 percent, which is much higher compared to a 59 percent for "inventory of all kinds", and a 29 percent for "property, plant and equipment".

We emphasize the importance of accounts receivable financing because it could potentially explain the weak correlation between being financially constrained and trade credit extension. Without accounts receivable financing, extending one unit of trade credit means one unit less cash flow for the firm. With accounts receivable financing, extending one unit of trade credit decreases cash flow by one unit. However, at the same time, it increases the firm's access to bank credit by creating one unit of collateral. In this way, the marginal cost of

[4]The accounts receivable financing in the U.S. was a financial innovation that had started in the early 1900s (see Murphy, 1992)

[5]More information about this dataset can be found in section 1.2.2.1.

trade credit extension is greatly reduced for the constrained firms. In fact, if the advance rate of AR is 100 percent, the cost of trade credit extension is 0.

Furthermore, the existence of accounts receivable financing changes the nature of trade credit. Without accounts receivable financing, trade credit serves merely as a redistribution channel that directs credit from unconstrained firms to constrained ones. With accounts receivable financing, in addition to the credit redistribution channel, collateral is created when firms extend trade credit to customers. Through a process of credit creation, accounts receivable financing increases collective access to bank credit for both trade credit lenders and borrowers.

1.2.2 Trade credit and the disruption in financial market

This section contains the second set of the empirical evidence that studies how the access to financial market affect firms' choice of extending trade credit. To this end, we adopt a similar strategy as in Chodorow-Reich (2014). Using the performance of firms' relationship bank during the 2007-09 financial crisis as an exogenous source of variation in the availability of credit, we estimate the impact of financial market disruption on firms' choice of trade credit extension.

This section uses data on the syndicated loan market and Compustat firms who borrow from that market. We introduce data on the syndicated loan market in section 1.2.2.1. Section 1.2.2.2 presents evidence on disruptions in the syndicated loan market during the 2007-09 financial crisis. In section 1.2.2.3, we discuss the empirical strategy and present the results.

1.2.2.1 Data

A syndicated loan is a type of loan whereby two or more lenders jointly issue fund to a borrowing firm. We use Thomson-Reuters Dealscan to obtain information of syndicated loans. This data set contains information on syndicated loans issued globally and in the United States. Its coverage of the U.S. syndicated loan market is rather comprehensive, especially in the post-1995 era. Each observation is a facility (loan). Detailed information of the loan, such as loan type, size, and maturity, are gathered from SEC filings, including 13-Ds, 14-Ds, 13-Es, 10-Ks, 10-Qs, 8-Ks, and S-series.

1.2.2.2 Disruption in the syndicated loan market during the 2007-09 financial crisis

We use characteristics of credit line facilities in the syndicated loan market to show its disruption during the 2007-09 financial crisis. Credit line facilities can be categorized into three types: 1) unsecured credit line, 2) secured credit line, with accounts receivable as collateral asset, and 3) secured credit line, with other types of asset as collateral. In Figure 1.8, we plot changes in several key characteristics of newly opened credit line facilities during 2006-10.

Across different types of credit lines, we see similar patterns of decreasing number of newly opened facilities, total facility size, and maturity. Take the accounts receivable backed credit line as an example. Compared to the 2006 pre-crisis level, in year 2009, total number of new facilities dropped by around 60 percent, total size of these facilities dropped by around 60 percent, and average maturity dropped by about 20 percent.

Interestingly, borrowing base percentage of the secured credit line facilities does not change by much during the same period. In fact, borrowing base percentage for "other assets backed facilities" even increased slightly compared to the pre-crisis level.

1.2.2.3 Effect of disruptions in the syndicated loan market on trade credit extension

The multiple lenders in a syndicated loan can be categorized into two types: lead lender and participants. As discussed in Sufi (2007) and Chodorow-Reich (2014), a leader lender differs from participants in various dimensions, the most important difference being that the lead lender has accumulated superior information about the borrower. It is costly to switch into a new lender, because the accumulated information of the current lead lender would be lost during the process. Therefore a firm's access to credit is hindered if its lead lender becomes unhealthy.

Following the strategy in Chodorow-Reich (2014), we construct a sample of firms and their lead lenders in the syndicated loan market. Using performance of lead lenders during the crisis as an exogenous variation of access to credit market, we estimate the effect of disruptions in the syndicated loan market on the trade credit extension.

We first construct the Dealscan-Compustat sample (DC sample) with firms and their relationship banks. We first use the link table provided by Chava and Roberts (2008) to match facilities to their corresponding borrowers in Compustat. We drop the observation if 1) the firm is in financial, insurance, retail, and wholesale sector, 2) the facility has multiple lead lenders, 3) the facility is not open during the period of Jan.1st 2004 to Dec.31 2006, and 4) the lead lender is not among the top 43 lenders as defined in Chodorow-Reich (2014). If a firm has only one open facilities during the period of Jan.1st 2004 to Dec.31 2006, we define the lead lender of that facility to be its pre-crisis relationship bank. If a firm has multiple open facilities during that period, we define the lead lender of the newest open facility as its relationship bank.

This process yields a panel of 1219 firm-bank pair over the period 2007Q1 to 2009Q4. This sample is a good representation of the whole universe of Compustat firms in terms of sectoral composition. However, comparing to the rest of Compustat firm, firms in the DC sample are much larger. On average, asset of DC firms is 8 times the size of the rest of Compustat firms. Among the 1291 DC firms, 393 have third party credit rating. In short, the DC sample consists of very large and financially advantageous firms.

For each firm-bank pair in this sample, we define a dummy indicator $Unhealthy_i$, which takes value 1 if the bank's percentage drop in new loan issuance during the financial crisis is higher than that of the median bank.[6]

We define a crisis indicator $Crisis_t$, which takes value 1 during the period of crisis times (2007Q4 to 2009Q4). The dependent variable is AR/Sales of firm i and time t. Our baseline regression is a fixed effect regression of the following form,

$$\begin{aligned} AR/Sales_{it} &= \beta_1 AP/Sales_{it} + \beta_2 Crisis_t + \beta_3 Crisis_t * Unhealthy_i \\ &\quad + \beta_4 Crisis_t * Rating_i + Crisis_t * \gamma_s + Crisis_t * \psi_i \\ &\quad + \chi_i + \epsilon_{it}, \end{aligned}$$

where χ_i is a set of firm-level fixed effect, which absorb time-invariant differences in terms of trade credit extension. We include accounts payable to sales ratio ($AP/Sales_{it}$), controlling for firms' borrowing from other firms through trade credit. The crisis indicator $Crisis_t$ captures the average change in accounts receivable to sales ratio during the crisis. The interaction term of $Crisis_t * Unhealthy_i$ thus captures the additional change in accounts receivable to sales ratio of firms with unhealthy banks, compared to the other firms. Other control variables include the interaction of crisis indicator with sectoral fixed effects (γ_s),3rd party bond rating indicator ($Rating_i$), and firm size fixed effects (ψ_i), capturing respectively the sectoral differential trend during the crisis, differential effects of crisis on firms with

[6] The information of banks' loan issuance is taken from Chodorow-Reich (2014). Unhealthy banks in the DC sample include BMO Capital Markets Financing, Banco Santander, Bank of New York Mellon, Bear Stearns, CIT Group, CIBC, Citi, Comerica, Credit Suisse, Deutsche Banks, GE Capital, Goldman Sachs, JP Morgan, KeyBank, Lehman Brothers, M&T Bank, Merrill Lych, Mitsubishi UFJ Financing Group, Morgan Stanley, National City, Rabobank, Scotiabank, Societe General, UBS, and Wachovia.

alternative financial resources, and effects on firms with different asset size.

The coefficient on the interaction term $Crisis_t * Unhealthy_i$, β_3, is the object of interests. We expect $\hat{\beta}_3$ to be negative and significant, indicating that having an unhealthy relationship bank during crisis time reduces firms' extension of trade credit more than firms with healthy banks.

The result is displayed in Table 1.2. Since firms in the DC sample are very large and financially integrated, not surprisingly, the effect of being in crisis time does not seem have a significant impact on the decision of trade credit extension. The coefficient on $Crisis_t$ is insignificantly and positive. The estimated coefficients on the interaction term $Crisis_t * Unhealthy_i$ show that having an unhealthy bank during crisis, on the other hand, significantly reduces the firm's trade credit extension. Firms' extension of trade credit, measured by AR/Sales, is around 1.3 to 1.8 percentage point lower, if their relationship bank is unhealthy. The estimates are consistently negative and significant with different sets of control variables.

1.3 Model

Time is discrete with infinite horizon. There are two types of goods in the economy: intermediate goods and final goods. Final goods are used for both consumption and investment. Intermediate goods are used exclusively as input in producing final goods.

The production of final goods takes place in two stages: intermediate goods and final goods production stage. Each stage is operated by a continuum of entrepreneurs with measure 1. Entrepreneurs differ from each other by wealth and productivity. The productivity process z is stochastic. It is parameterized by a poisson process with death rate π and a new draw from distribution $G(z)$ after death.

There is a continuum of workers with measure N. Workers provide labor and consume. They do not have access to asset markets, i.e. they are "hand-to-mouth".

The banking sector is competitive with one representative bank making zero expected profit.

1.3.1 Preference, endowment and production technology

The preference of workers is time separable, with instantaneous utility function $u(c_t^h, h_t)$, such that,

$$U^h(c^h, h) = \sum_t \beta^t u(c_t^h, h_t),\ u(c_t, h_t) = c_t^h - \psi \frac{h_t^{1+\theta}}{1+\theta},$$

where β is the discounting factor, ψ represents disutility from working, and θ is the inverse of Frisch elasticity.

The preference of entrepreneurs is time separable with instantaneous utility function of $log(c_t)$. The expected utility of the entrepreneur can be written as,

$$U^e(c) = \mathbb{E} \sum_t \beta^t log(c_t),$$

where the expectation is taken over the stochastic process of productivity z and wealth a.

Intermediate goods entrepreneurs operate a decreasing return to scale production technology ($\mu_1 < 1$) that transforms capital and labor into intermediate goods, such that

$$y_1 = A_1 z F_1(k, l) = A_1 z (k^\alpha l^{1-\alpha})^{\mu_1}.$$

Final goods entrepreneurs operate a decreasing return to scale production technology ($\mu_2 < 1$) that transforms capital, labor and intermediate goods into final goods, such that

$$y_2 = A_2 z F_2(k, l, x_1) = A_2 z ((k^\alpha l^{1-\alpha})^{1-\chi} x_1^\chi)^{\mu_2}.$$

1.3.2 Financing production without trade credit

Entrepreneurs enter each period with wealth a. After the productivity z is realized, the entrepreneurs make choices of production k, l, x_1, consumption c, and next period wealth a'. To achieve this production scale, entrepreneurs need to raise an inter-temporal debt $d = k - a$, with interest rate r.

Following Jermann and Quadrini (2012), we assume that due to mismatch between payment and revenue realization, entrepreneurs need to finance working capital using intra-temporal loan. In particular, they need the loan to cover 1) investment into next period $i = a' - a$, 2) consumption c, 3) interest payment $r(k - a)$, and 4) input goods cost: $\delta k + wl$ for the intermediate goods entrepreneurs and $\delta k + wl + p_1 x_1$ for the final goods entrepreneurs. The interest rate of the intra-temporal loan is assumed to be 0.

The amount of intra-temporal bank loan needed for the entrepreneurs are,

$$\begin{aligned} intermediate: \quad & m_1 = a' - a + c + r(k - a) + \delta k + wl, \\ final: \quad & m_2 = a' - a + c + r(k - a) + \delta k + wl + p_1 x_1. \end{aligned}$$

Using budget constraints of the entrepreneurs,

$$\begin{aligned} intermediate: \quad & c + a' = (1 + r)a + p_1 A_1 z F_1(k, l) - (r + \delta)k - wl, \; and \\ final: \quad & c + a' = (1 + r)a + A_2 z F_2(k, l, x_1) - (r + \delta)k - wl - p_1 x_1, \end{aligned}$$

we derive that intra-temporal bank loan needed to finance production is equal to the revenue of production. That is, $m_1 = p_1 A_1 z F_1(k, l)$ and $m_2 = A_2 z F_2(k, l, x_1)$.

Working capital constraint without trade credit The ability of entrepreneurs to borrow from the bank is constrained by limited enforceability of bank debt obligations. In particular, we assume that at the end of each period, after revenue is realized, entrepreneurs can default on their liabilities and divert the revenue. After default, the bank has an option to liquidate the wealth of entrepreneurs. Once the liquidation process starts, with probability γ_1, the bank can successfully liquidate the whole value of wealth a' and with probability $1-\gamma_1$ they get nothing. However, the bank and entrepreneurs can renegotiate on their debt before the liquidation option is exercised. In particular, entrepreneurs can make a take-it-or-leave-it offer such that the bank is indifferent between liquidation and non-liquidation. In this case, entrepreneurs will offer to repay only the expected liquidation value $\gamma_1 a'$. Hence the default value for intermediate goods entrepreneurs are $y-\gamma_1 a'$ and non-default value is $y-m_1$. The incentive compatibility constraint leads to a constraint on the intra-temporal loan, such that $m_1 \leq \gamma_1 a'$. Similarly, the incentive compatibility constraint for final goods entrepreneurs gives constraint $m_2 \leq \gamma_1 a'$. Therefore the limited enforceability yields a set of working capital constraints for entrepreneurs in an environment without trade credit,

$$intermediate: \quad p_1 A_1 z F_1(k,l) \leq \gamma_1 a', \ and \tag{1.2}$$

$$final: \quad A_2 z F_2(k,l,x_1) \leq \gamma_1 a'. \tag{1.3}$$

1.3.3 Trade credit

Following previous literature on the existence of trade credit such as Burkart and Ellingsen (2004) and Biais and Gollier (1997), we postulate that the existence of trade credit is due to a certain comparative advantage of intermediate entrepreneur in lending to their customers. Here we assume that the comparative advantage lies in the fact that intermediate goods entrepreneurs have better control over their customers than the bank does. As an example, suppliers can threat to stop supplying to the customer if trade credit is not repaid (see Petersen and Rajan, 1997). In addition, we assume that intermediate goods entrepreneurs can perfectly enforce repayment of trade credit. In other words, lending of trade credit to customers is always secure.

Modeling trade credit We assume that there is a competitive market for intermediate goods and trade credit. Two prices – price of intermediate goods p_1 and interest rate r^{tc} of trade credit – clear the market.

An intermediate goods entrepreneur enters the market with a contract $(y_1, AR \leq p_1 y_1)$ to offer. The contract provides intermediate goods of value $p_1 x_1$ and a loan of size AR. Once the contract is accepted by the market, the intermediate goods entrepreneur proceeds to produce and expects to collect a payment of size $p_1 x_1 + (1 + r^{tc})AR$ from the market by the end of this period.

A final goods entrepreneur enters the market to purchase a contract $(x_1, AP \leq p_1 x_1)$. By signing the contract, the final goods entrepreneur receives a loan of size AP and commits to purchase intermediate goods of value $p_1 x_1$. A payment of size $p_1 x_1 + (1 + r^{tc})AP$ is expected to be made to the market by the end of this period.

Assuming that intermediate goods are identical and infinitely divisible, both the supply and demand of contracts can be divided infinitely. Hence there exists a market clearing algorithm to allocate the contracts, as long as aggregate demand for intermediate goods and trade credit equate aggregate supply of the two.

Our modeling of trade credit merits some discussions. Trade credit in reality is not an explicit loan as is in our model. Instead, extending trade credit essentially means delaying receipt of cash flows that can be used to finance production, investment and etc.. However, modeling trade credit in this more realistic way requires an alternative timing and adds an additional state variable. As an example, in Appendix A.1, we provide an alternative model of trade credit extension as delayed cash flow. In it, we assume that trade credit is carried over into the next period, instead of repaying at the end of this period. Hence this specification adds trade credit from previous period as another state variable.

Clearly, the modeling of trade credit in this paper makes computation less intensive. At the same time, we believe that it very well captures the essence of trade credit – a reallocation of financial resources, whether it is cash flow as in Appendix A.1, or explicit loan as in here.

Working capital constraint with trade credit Offering trade credit AR increases intermediate goods entrepreneurs' need for intra-temporal bank loan by the amount AR. By using trade credit AP, final goods entrepreneurs' need for intra-temporal loan decreases by AP. Therefore, their needs for intra-temporal loan with trade credit are,

$$\begin{aligned} intermediate: &\quad \hat{m}_1 = m_1 + AR = a' - a + c + r(k-a) + \delta k + wl + AR, \\ final: &\quad \hat{m}_2 = m_2 - AP = a' - a + c + r(k-a) + \delta k + wl + p_1 x_1 - AP. \end{aligned}$$

Using budget constraints of the entrepreneurs,

$$\begin{aligned} intermediate: &\quad c + a' = (1+r)a + p_1 A_1 F_1(k,l) - (r+\delta)k - wl + r^{tc} AR, \ and \\ final: &\quad c + a' = (1+r)a + A_2 F_2(k,l,x_1) - (r+\delta)k - wl - p_1 x_1 - r^{tc} AP, \end{aligned}$$

we derive the need for intra-temporal bank loan for intermediate goods entrepreneurs as $\hat{m}_1 = p_1 A_1 F_(k,l) + (1 + r^{tc})AR$ and for final goods entrepreneurs it is $\hat{m}_2 = A_2 F_2(k,l,x_1) - (1 + r^{tc})AP$.

Upon default, a renegotiation process begins. Intermediate goods entrepreneurs would propose a take-it-or-leave-it offer to repay only $\gamma_1 a' + \gamma_2 AR$, where $\gamma_2 AR$ is the expected liquidation value of accounts receivable for the bank. The value of default for intermediate goods entrepreneurs is therefore $y + AR - (\gamma_1 a' + \gamma_2 AR)$ and the value of non-default is $y + AR - \hat{m}_1$. The incentive compatibility constraint gives $\hat{m}_1 \leq \gamma_1 a' + \gamma_2 AR$. Similarly, for final goods entrepreneurs, the incentive compatibility constraint leads to a constraint on intra-temporal bank loan $\hat{m}_2 \leq \gamma_1 a'$.

In the end, we can write the working capital constraints with trade credit as follows,

$$intermediate: \quad p_1 A_1 F_1(k,l) + (1+r^{tc})AR \leq \gamma_1 a' + \gamma_2 AR, \tag{1.4}$$

$$final: \quad A_2 F_2(k,l,x_1) - (1+r^{tc})AP \leq \gamma_1 a'. \tag{1.5}$$

A comparison between working capital constraint 1.2 and 1.4 shows that, for intermediate goods entrepreneurs, by extending one unit of trade credit, the need for intra-temporal bank loan increases by $1+r^{tc}$, but it is partially compensated by an increased intra-temporal bank loan limit with accounts receivable as collateral. A comparison between working capital constraint 1.3 and 1.5, on the other hand, indicates that by using one unit of trade credit, final goods entrepreneurs reduce the need for intra-temporal bank loan by $1+r^{tc}$.

Figure 1.8 summarizes the flow of goods and credit after trade credit is introduced in the economy. Intermediate goods entrepreneurs supply input goods of value $p_1 y_1$ (black arrow) as well as trade credit of size AR (green arrow) to final goods entrepreneurs. The bank provide intra-temporal loan of size $\gamma_1 a' + \gamma_2 AR$ and $\gamma_1 a'$ to intermediate and final goods entrepreneurs, respectively (blue arrows).

1.4 Recursive competitive equilibrium

In this section, we present the problem of households and entrepreneurs, define recursive competitive equilibrium, and analyze entrepreneurs' choice of trade credit.

The problem of households is stationary. It can be written simply as follows,

$$\max_{c_t^h, h_t} \frac{(c_t^h - \psi h_t^{1+\theta})^{1-\nu}}{1-\nu}, \ s.t. \ c_t^h = w h_t. \tag{1.6}$$

Given current state variables (a, z), intermediate goods entrepreneurs choose input goods k, l, trade credit extension AR, consumption c, and next period wealth a'. The choices are subject to an inter-temporal budget constraint 1.8 and a working capital constraint 1.9. Two additional constraints on accounts receivable require that it is nonnegative and does not exceed the value of output. We also require that entrepreneurs always hold positive wealth. The problem of intermediate goods entrepreneurs can be written recursively as follows,

$$\begin{aligned} V_1(a,z) \quad &= \quad \max_{c,k,l,AR,a'} log(c) + \beta \mathbb{E}_{z'} V_1(a',z'), && (1.7) \\ s.t. \quad & \quad c + a' = (1+r)a + p_1 A_1 z F_1(k,l) - (r+\delta)k - wl + r^{tc} AR, && (1.8) \\ & \quad p_1 A_1 z F_1(k,l) + (1+r^{tc})AR \leq \gamma_1 a' + \gamma_2 AR, && (1.9) \\ & \quad 0 \leq AR \leq p_1 A_1 z F_1(k,l), a' \geq 0. \end{aligned}$$

Similarly, we can write the problem of final goods entrepreneurs as follows,

$$\begin{aligned} V_2(a,z) \quad &= \quad \max_{c,k,l,x_1,AP,a'} log(c) + \beta \mathbb{E}_{z'} V_2(a',z'), && (1.10) \\ s.t. \quad & \quad c + a' = (1+r)a + A_2 z F_2(k,l,x_1) - (r+\delta)k - wl - p_1 x_1 - r^{tc} AP, && (1.11) \\ & \quad A_2 z F_2(k,l,x_1) - (1+r^{tc})AP \leq \gamma_1 a', && (1.12) \\ & \quad 0 \leq AP \leq p_1 x_1, a' \geq 0. \end{aligned}$$

where equation 1.11 is the inter-temporal budget constraint and inequality 1.12 is the working capital constraint.

We are now ready to define the recursive competitive equilibrium.

Definition 1 *The recursive competitive equilibrium consists of interest rate R, wage rate w, intermediate goods price p_1 and interest rate of trade credit r^{tc}, value functions of entrepreneurs $V_1(a,z)$ and $V_2(a,z)$, policy functions of entrepreneurs $c_1(a,z)$, $c_2(a,z)$, $k_1(a,z)$, $k_2(a,z)$, $a'_1(a,z)$, $a'_2(a,z)$, $l_1(a,z)$, $l_2(a,z)$, $x_1(a,z)$, $AR(a,z)$, $AP(a,z)$, consumption and labor supply of households $\{c^h, h\}$ and distributions of entrepreneurs $\Phi_1(a,z)$ and $\Phi_2(a,z)$,* *such that,*

1. *Given prices, value functions and policy functions solve the optimization problems of entrepreneurs 1.7 and 1.10.*

2. *Given prices, consumption and labor supply solve the households optimization problem 1.6.*

3. *Labor market clears*

$$\int l_1(a,z)d\Phi_1(a,z) + \int l_2(a,z)d\Phi_2(a,z) = N \cdot h.$$

4. *Inter-temporal debt market clears,*

$$\int (k_1(a,z) - a) \cdot d\Phi_1(a,z) + \int (k_2(a,z) - a) \cdot d\Phi_2(a,z) = 0.$$

5. *Intermediate goods market and trade credit market clear,*

$$\begin{aligned} \int A_1 z F_1(k,l) d\Phi_1(a,z) &= \int x_1(a,z) d\Phi_2(a,z), \\ \int AR(a,z) d\Phi_1(a,z) &= \int AP(a,z) d\Phi_2(a,z). \end{aligned}$$

6. *The stationary distributions evolve according to the following law of motion,*

$$\begin{aligned} \Phi_1(a',z') &= \int \mathbb{I}_{a'=a_1'(a,z)} \pi(z'|z) d\Phi_1(a,z), \\ \Phi_2(a',z') &= \int \mathbb{I}_{a'=a_2'(a,z)} \pi(z'|z) d\Phi_2(a,z). \end{aligned}$$

1.4.1 Some properties of the equilibrium

In the following propositions, we characterize 1) the state of being financially constrained, 2) the choice of extending and using trade credit, and 3) the relationship between the state of being constrained and the choice of trade credit.

The first proposition characterizes the state of being financially constrained.

Proposition 1 *There exist functions $g_1(z)$ and $g_2(z)$ such that,*

1. *For intermediate goods entrepreneurs with wealth a and productivity z, working capital constraint is not binding if $a > g_1(z)$ and it is binding if $a \leq g_1(z)$.*

2. *For final goods entrepreneurs with wealth a and productivity z, working capital constraint is not binding if $a > g_2(z)$ and it is binding if $a \leq g_2(z)$.*

The proof of this proposition can be found in Appendix A.2.1. It says that the state of being constrained follows a cut-off rule. For an entrepreneur with a given productivity z, she moves from being constrained to being unconstrained when a increases. This is an intuitive result. An increase in wealth a leads to a larger bank loan limit and thus relaxes the working capital constraint. For any z, there exists an optimal production scale because the production function is decreasing return to scale. The entrepreneur becomes unconstrained when a is large enough so that the optimal scale of production can be achieved with the available bank loan.

In the second proposition, we analysis the decision of trade credit extension and usage.

Proposition 2 *There exist functions $h_1(z)$ and $h_2(z)$ such that,*

1. *For any intermediate entrepreneurs with wealth a and productivity z, $AR = p_1 A_1 z F_1(k, l)$ if $a > h_1(z)$, $AR = 0$ if $a < h_1(z)$, and $AR \in (0, p_1 A_1 z F_1(k, l))$ if $a = h_1(z)$.*

2. *For any final goods entrepreneurs with wealth a and productivity z, $AP = 0$ if $a > h_2(z)$, $AP = p_1 x_1$ if $a < h_2(z)$, and $AP \in (0, p_1 x_1)$ if $a = h_2(z)$.*

The proof of this proposition can be found in Appendix A.2.2. To understand the choice of trade credit extension for intermediate goods entrepreneurs, consider the first order condition with respect to AR,

$$r^{tc} = \lambda(1 - \gamma_2) + \tau_1 - \tau_2. \tag{1.13}$$

In it, λ is the Lagrangian multiplier on working capital constraint representing the shadow price of bank loans, and τ_1 and τ_2 are the Lagrangian multipliers of two accounts receivable constraints ($AR \geq 0$ and $AR \leq p_1 A_1 z F_1(k, l)$ respectively). The Lagrangian multiplier of working capital constraint, λ, represents the shadow price of bank loan.

The marginal cost of extending trade credit is $\lambda(1 - \gamma_2)$, which is the shadow price of bank loan multiplied by loss of bank loan $1 - \gamma_2$. When wealth of the entrepreneur is very low and thus shadow price of bank loans is very high, marginal cost is higher than marginal benefit of extending trade credit, which is the interest rate of trade credit r^{tc}. In this case the entrepreneur is unwilling to extend trade credit. This can be seen in equation 1.13 – when λ is high, we have $\tau_1 > 0$, hence $AR = 0$ following the complementary slackness condition. The entrepreneur will extend trade credit if and only if shadow price of bank loan is low enough, i.e. her wealth level is high enough.

A similar analysis can be applied to the use of trade credit for final goods entrepreneurs.

The third proposition describes the relationship between choices of trade credit and being financial constrained.

Proposition 3 *The following properties hold if $r^{tc} > 0$,*

1. *If $\gamma_2 \in [0, 1]$, for any z, $h_1(z) < g_1(z)$.*

2. *For any z, $h_2(z) < g_2(z)$.*

The proofs of the proposition can be found in Appendix A.2.3. The first part of the above proposition says that, all unconstrained intermediate goods entrepreneurs will extend trade credit. If accounts receivable can be used as collateral to obtain bank loan, some of the constrained entrepreneurs will extend trade credit. The second part of the propositions says that, for final goods entrepreneurs, if they are financially unconstrained, they will not borrow from their supplier through trade credit. For constrained final goods entrepreneurs, they will only use trade credit if marginal benefit exceeds the interest rate of trade credit. Hence only the very productive final goods entrepreneurs use trade credit.

1.5 Calibration

In this section, we calibrate the stationary equilibrium to match key moments of the U.S. data. One period in the model corresponds to one quarter in the data.

1.5.1 Calibration strategy and result

The household utility function is of GHH form (see Greenwood et al., 1988). We pick $\theta = 0.5$, which gives a Frisch elasticity of 2. This value is well within standard macro estimation (see Chetty et al., 2011; Keane and Rogerson, 2012). Another parameter in the utility function is ψ, representing disutility from providing one unit of labor. We calibrate ψ such that 30 percent of households time is spent on working, i.e. $h = 0.3$. Entrepreneurs' instantaneous utility function is log-form. We calibrate the discount rate β of entrepreneurs to match an annual interest rate of 4 percent. Since the share of entrepreneurs in the U.S. data is around 10 percent, we pick $N = 18$ so that the share of entrepreneurs in the model matches that number.

These are two sectoral production function in the model. We assume that capital to labor ratio in the two sectors are the same. We then fix the relative capital share α to be 1/3. Consequently, the relative labor share is 2/3. Following Yi (2003), the intermediate goods share χ is fixed to be $2/3$. Depreciation rate δ is chosen to be 0.025 so that the annual depreciation rate of capital is equal to 10 percent. The Poisson death rate π is fixed at 10 percent, following Buera et al. (2011).

We assume that scale parameters in the two sectors are the same, i.e. $\mu_1 = \mu_2$. The productivity distribution $G(z)$ is assumed to be Pareto with scale parameter 1 and tail parameter ν. Following Buera et al. (2011), we calibrate $\mu_1(\mu_2)$ and ν to match the top 5 percentile earnings share and top 10 percentile employment share, respectively. Lastly, we pick γ_1 and γ_2, the collateral constraint on wealth a' and accounts receivable AR, such that the model delivers the ratio of credit market liability to nonfinancial assets and the ratio of accounts receivable to gross value added in the data. See Table 1.3 for a summary of the calibrated parameters, their value, targets, and calibration results. The algorithm to solve the stationary equilibrium can be found in Appendix A.3.1.

How we compute the model moment of credit market liability deserves some discussion here. In our model, credit market liability consists of both inter-temporal and intra-temporal debt. It is clear that the aggregate inter-temporal debt is

$$\int_0^1 \max(k_1(a,z) - a, 0)d\Phi_1(a,z) + \int_0^1 \max(k_2(a,z) - a, 0)d\Phi_2(a,z). \tag{1.14}$$

The size of intra-temporal debt of intermediate goods entrepreneurs is $p_1 y_1(a,z) + AR(a,z)$, where $y_1(a,z) = A_1 z F_1(k_1(a,z), l_1(a,z))$. It covers output and the amount of trade credit extended. The intra-temporal debt of final goods entrepreneurs is output minus trade credit used $y_2(a,z) - AP(a,z)$.

Given that in equilibrium, $\int_0^1 AR(a,z)d\Phi_1(a,z) = \int_0^1 AP(a,z)d\Phi_2(a,z)$, it follows that the aggregate intra-temporal debt is

$$\int_0^1 p_1 y_1(a,z)d\Phi_1(a,z) + \int_0^1 y_2(a,z)d\Phi_2(a,z). \tag{1.15}$$

Credit market liability in the model is therefore the sum of inter-temporal debt 1.14 and intra-temporal debt 1.15.

1.5.2 Discussions of the calibrated model

Heterogeneous entrepreneurs and their choice of trade credit As discussed in section 1.4.1, choices of extending and using trade credit depend on wealth and productivity of entrepreneurs. In particular, we know that given productivity, the usage of trade credit decreases with wealth while the extension of trade credit increases with wealth.

We take the calibrated model and compare trade credit choices of entrepreneurs with different level of wealth and productivity. We divide entrepreneurs into four groups based on their productivity and wealth: low wealth low productivity, low wealth high productivity, high wealth low productivity, and high wealth high productivity.

Trade credit choices for these four groups are presented in Table 1.4. Choice of trade credit extension of intermediate goods entrepreneurs is presented in the first row of the table. Two low productivity groups extend all of their output as trade credit. On the contrary, two high productivity groups extend only a fraction of their output as trade credit. Furthermore, conditional on being high productivity, low wealth and high wealth groups extend approximately the same fraction of output as trade credit. This is a result of the high collateral value of accounts receivable in the calibration ($\gamma_2 = 0.95$).

The four groups of entrepreneurs show very different patterns when it comes to the using of trade credit. For the group with high wealth and low productivity, there is no need for trade credit, therefore its accounts payable to output ratio is 0. The accounts payable to output ratio increases to 29.8 percent for the low productivity and low wealth group, an indication that due to low wealth level, some of the low productivity entrepreneurs need additional financing from trade credit. The accounts payable to output ratio increases even more for the two groups with high productivity – 47.9 percent if wealth level is low and 53.0

percent if wealth level is high.

Interest rate of trade credit One prominent empirical characteristics of trade credit is its high interest rate. As an example, Petersen and Rajan (1997) documents that the effective annual interest rate is around 43 percent for one of the most commonly used trade credit contracts in retail businesses. Costello (2014) calculates that the annual interest rate of trade credit is between 12 percent to 16 percent by comparing gross profit margin of firms before and after using trade credit. In our calibrated model, interest rate of trade credit is 2.7 percent quarterly, yielding an annual interest rate of 11.8 percent, which is very close to the calculation in Costello (2014).

The fact that our model generates a high trade credit interest rate is not surprising. According to Proposition 3 and Table 1.4, only very productive and constrained entrepreneurs use trade credit. The interest rate of trade credit therefore reflects marginal productivity of the most productive group of final goods entrepreneurs.

Nature of trade credit: Redistribution and creation of credit Trade credit in nature is a financial institution that alleviates the efficiency loss from financial frictions. It plays this role through two channels: 1) it redistributes unused credit from unconstrained intermediate goods entrepreneurs to constrained final goods entrepreneurs, and 2) it creates credit through accounts receivable financing and increases collective access to trade for all entrepreneurs.

With the calibrated model, we decompose total trade credit, $\bar{AR} = \int AR(a,z)d\Phi_1(a,z)$, into two parts by their nature: credit redistribution ($\bar{AR}^r$) and credit creation ($\bar{AR}^c$). The credit creation part $\bar{AR}^c$ is the amount of trade credit that are used by intermediate goods entrepreneurs as collateral to obtain bank loans. More formally, it is defined as

$$\bar{AR}^c = \frac{1}{\gamma_2}\int max(0, y_1(a,z) + (1+r^{tc})AR(a,z) - \gamma_1 a')d\Phi_1(a,z). \tag{1.16}$$

In it, $max(0, y_1(a, z) + (1 + r^{tc})AR(a, z) - \gamma_1 a')$ is the amount of bank loan that intermediate goods entrepreneurs receive by using accounts receivable as collateral. Dividing that by γ_2, collateral constraint of accounts receivable, yields the size of trade credit that is used as collateral. Consequently, the rest of the accounts receivable is a pure credit redistribution, i.e. $\bar{AR}^r = \bar{AR} - \bar{AR}^c$.

The decomposition exercise using our calibrated model shows that the credit creation part consists of 87 percent of the aggregate trade credit in the economy while the rest 13 percent is pure credit redistribution.

1.6 Engineering the 2007-09 financial crisis

In this section, we use the calibrated model to study the 2007-09 financial crisis. To this end, we engineer a financial shock by reducing the collateral constraint parameters γ_1 and γ_2 such that the model delivers the drop in the ratio of credit market liability and trade credit to nonfinancial assets of the U.S. nonfinancial corporate sector during the 2007-09 financial crisis.

To simulate the model, we assume that the collateral parameters $\gamma_{1,t}$ and $\gamma_{2,t}$ are hit by a common shock ρ_t. That is, $\gamma_{i,t} = \rho_t \gamma_i$ for $i = 1, 2$, in which γ_i takes the parameter value at steady state. Our simulation results suggest that the following ρ_t process yields the best fit of data: $\{\rho_1, \rho_2, \rho_3, \rho_4\} = \{0.975, 0.95, 0.925, 0.9\}$, $\rho_t = \rho_{t-1} + 0.014$ for $t = 5, ..., 10$, and $\rho_t = 1$ for $t \geq 11$.

We plot the simulation result in Figure 1.8. In the left panel, we plot credit market liability to capital stock ratio in the data (dotted line) and in the model (solid line). In the right panel, we plot trade credit to capital stock ratio in the data (dotted line) and in the model (solid line). From peak to trough following the 2007-09 financial crisis, credit market liability to capital stock ratio dropped by around 10 percent and trade credit to capital stock ratio dropped by around 13 percent in the data. As shown in the figure, the simulated model matches the magnitude of the drop in the data quite well. It delivers a 11 percent drop in credit market liability to capital stock ratio and a 12.5 percent drop in trade credit to capital stock ratio.

1.6.1 Aggregate dynamics after the financial crisis

How does our model perform in terms of capturing aggregate dynamics of the economy during the 2007-09 financial crisis? In this section, we presents aggregate dynamics of the model economy following the shock process ρ_t. The algorithm to solve the transitional dynamics can be found in Appendix A.3.2.

In the face of a tightening borrowing constraints during the financial crisis, constrained entrepreneurs are forced to reduced the scale of their production. This leads to a drop in hours, capital, and output. In addition, prices of input goods drop during crisis. As a result, unconstrained entrepreneurs, who are on average less productivity than constrained ones, increase their production scale. This reallocation of resources from productive entrepreneurs to less productive ones results in a lower aggregate TFP, which exacerbate the decline in aggregate output and input.

Figure 1.8 shows aggregate dynamics of the economy following the financial crisis. In the model, output drops by 6 percent, while in the data output shows a 6 percent deviation below trend (see the upper left panel of Figure 1.8). Aggregate dynamics in hours, capital stock, and TFP are presented in the other three panels. The model displays approximately a 8 percent drop in total hours, a 2 percent drop in aggregate TFP, and a 1 percent drop in total capital stock. Compared to data, the model generates a higher drop in hours and a lower drop in TFP. In general, it performs rather well to match the dynamics in the data.

1.7 The aggregate implications of trade credit

In this section, we construct a counterfactual economy in which bank credit is the only source of financing. By comparing the benchmark to the counterfactual economy in steady state and during financial shocks, we attempt to understand the aggregate implications of trade credit.

1.7.1 A counterfactual economy

The counterfactual economy has the same economic environment as the benchmark. The only difference is the missing of trade credit. In the counterfactual economy, limited enforcibility of bank loans leads to a set of working capital constraints 1.2 and 1.3 (see section 1.3.2).

The stationary equilibrium of the counterfactual economy is defined as follows,

Definition 2 *The recursive competitive equilibrium without trade credit consists of interest rate $\tilde{R}$, wage rate $\tilde{w}$, and intermediate goods price $\tilde{p}_1$, value functions of entrepreneurs $\tilde{V}_1(a,z)$ and $\tilde{V}_2(a,z)$, policy functions of entrepreneurs $\tilde{c}_1(a,z)$, $\tilde{c}_2(a,z)$, $\tilde{k}_1(a,z)$, $\tilde{k}_2(a,z)$, $\tilde{a}'_1(a,z)$, $\tilde{a}'_2(a,z)$, $\tilde{l}_1(a,z)$, $\tilde{l}_2(a,z)$, $\tilde{x}_1(a,z)$, consumption and labor supply of households $\{\tilde{c}^h, \tilde{h}\}$ and distributions of entrepreneurs $\tilde{\Phi}_1(a,z)$ and $\tilde{\Phi}_2(a,z)$, such that,*

1. *Given prices, value functions and policy functions solve the optimization problem of entrepreneurs.*

$$
\begin{aligned}
\tilde{V}_1(a,z) &= \max_{c,k,l,a'} log(c) + \beta \mathbb{E}_{z'} \tilde{V}_1(a',z') \\
s.t. \quad & c + a' = (1+r)a + p_1 A_1 z F_1(k,l) - (r+\delta)k - wl, \\
& p_1 A_1 z F_1(k,l) \leq \tilde{\gamma}_1 a', a' \geq 0.
\end{aligned}
$$

$$
\begin{aligned}
\tilde{V}_2(a,z) &= \max_{c,k,l,x_1,a'} log(c) + \beta \mathbb{E}_{z'} \tilde{V}_2(a',z'), \\
s.t. \quad & c + a' = (1+r)a + A_2 z F_2(k,l,x_1) - (r+\delta)k - wl - p_1 x_1, \\
& A_2 z F_2(k,l,x_1) \leq \tilde{\gamma}_1 a', a' \geq 0.
\end{aligned}
$$

2. *Given prices, consumption and labor supply solve households optimization problem 1.6.*

3. *Labor market clears,*

$$
\int \tilde{l}_1(a,z) d\tilde{\Phi}_1(a,z) + \int \tilde{l}_2(a,z) d\tilde{\Phi}_2(a,z) = N \cdot \tilde{h}.
$$

4. *Capital market clears,*

$$
\int (\tilde{k}_1(a,z) - a) \cdot d\tilde{\Phi}_1(a,z) + \int (\tilde{k}_2(a,z) - a) \cdot d\tilde{\Phi}_2(a,z) = 0.
$$

5. *Intermediate goods market clears,*

$$
\int A_1 z F_1(\tilde{k}_1(a,z), \tilde{l}_1(a,z)) d\tilde{\Phi}_1(a,z) = \int \tilde{x}_1(a,z) d\tilde{\Phi}_2(a,z).
$$

6. *The stationary distributions evolve according to,*

$$
\begin{aligned}
\tilde{\Phi}_1(a',z') &= \int \mathbb{I}_{a'=\tilde{a}'_1(a,z)} \pi(z'|z) d\tilde{\Phi}_1(a,z), \\
\tilde{\Phi}_2(a',z') &= \int \mathbb{I}_{a'=\tilde{a}'_2(a,z)} \pi(z'|z) d\tilde{\Phi}_2(a,z).
\end{aligned}
$$

1.7.2 The impact of trade credit in steady state

To see the role of trade credit in steady state, we take the calibrated parameters in the benchmark economy (see Table 1.3) and feed them into the counterfactual. In particular, we set $\tilde{\gamma}_1 = \gamma_1$, making collateral constraint of entrepreneurs' wealth to be the same across two economies.

We solve the above parameterized counterfactual economy numerically. In Table 1.5, we present the percentage difference of the counterfactual economy from the benchmark in aggregate and sectoral output, hours, capital, and TFP. As shown in the table, compared to the benchmark, aggregate output of the counterfactual economy is 23.9 percent lower, which can be decomposed into a 15.3 percent lower capital stock, a 24.4 percent lower labor, and a 8.4 percent lower aggregate TFP.

Why is output higher in the benchmark economy? In short, the existence of trade credit alleviates borrowing constraints of entrepreneurs. Therefore resources are allocated more efficiently in the benchmark, leading to a higher aggregate output level. Our computation shows that in the benchmark economy, when weighted by output, 82 percent of intermediate entrepreneurs are constrained and 79 percent of the final goods entrepreneurs are constrained. As a comparison, in the counterfactual economy, the fraction of constrained entrepreneurs is 85 percent and 86 percent in the intermediate goods and final goods sector, respectively.

A further examination of sectoral differences between the two economies provides a clearer understanding of the role of trade credit. As shown in the last column of Table 1.5, aggregate TFP of the counterfactual economy is 8.4 percent lower than that of the benchmark, indicating a higher level of resource misallocation. Furthermore, the difference in aggregate TFP is almost completely explained by the different in TFP of the final goods sector (7.5 percent). This is not surprising since trade credit relaxes borrowing constraint of final goods entrepreneurs, thus leading to a more efficient allocation of resources in that sector. On the contrary, difference in TFP of the intermediate goods sector is very small (0.9 percent). This is due to a smaller impact that trade credit has on the borrowing constraint of intermediate goods entrepreneurs. To see this, notice that in our calibration collateral constraint of accounts receivable is very high ($\gamma_2 = 0.95$). Therefore, although there is loss in bank credit when extending trade credit, it is almost offset completely by an increased access to bank loan through accounts receivable financing.

Although trade credit has a very small impact on TFP of intermediate goods sector, its impact on output of that sector is large. As shown in Table 1.5, intermediate goods output, measured by its quantity, of the counterfactual economy is 26.4 percent lower than that of the benchmark. This effect on output works mainly through general equilibrium forces. Higher demand of intermediate input goods from final goods sector increases the price of intermediate goods, which essentially makes all intermediate goods entrepreneurs more productive.

1.7.3 The impact of trade credit during 2007-09 financial crisis

So far we have established the role of trade credit in steady state: it relaxes the borrowing constraint and increases aggregate output. In this section, we ask: what is the role played by trade credit during the 2007-09 financial crisis? To answer this question, we again turn to the counterfactual economy without trade credit. More specifically, we introduce into the counterfactual economy the same financial shock as in the benchmark economy and study dynamics of that economy following the shock.

Before conducting the experiment, we first make sure that the benchmark and the counterfactual economy are comparable to each other. In particular, we want to make sure that the fraction of constrained entrepreneurs in these two economies are at the same level, because the effect of a financial shock crucially depends on the fraction of entrepreneurs whose production scale are constrained by their access to financial resources.

To this end, we recalibrate the counterfactual economy so that its aggregate output is at the same level as the benchmark. In particular, we need to increase $\tilde{\gamma}_1$, such that $\tilde{\gamma}_1 = 1.43\gamma_1 = 0.4$. Under this calibration, the rest of the model moments in the counterfactual economy are very close to those in the benchmark. In addition, the fraction of constrained entrepreneurs are at a very similar level as the benchmark. In steady state of the benchmark economy, around 82 percent of intermediate goods entrepreneurs and 79 percent of final goods entrepreneurs are constrained. The numbers are very similar in the recalibrated counterfactual economy: around 81 percent of entrepreneurs are constrained in each sector.

We then hit the recalibrated counterfactual economy with the same financial shock ρ_t as in the benchmark model (see section 1.6). More formally, we feed in the recalibrated counterfactual model a sequence of collateral constraint $\tilde{\gamma}_{1,t}$ such that $\tilde{\gamma}_{1,t} = \rho_t \tilde{\gamma}_1$ and compare the aggregate dynamics of the recalibrated counterfactual economy with the benchmark economy.

In Figure 1.8, we plot aggregate dynamics of the recalibrated counterfactual economy following the financial shock ρ_t in red dashed lines. In each figure, we also plot the corresponding aggregate dynamics of the benchmark economy following the same financial shock ρ_t in blue solid line. Notice that the blue solid lines in Figure 1.8 are identical to those in Figure 1.8. The recession is milder in the counterfactual economy – the drop in output, hours, TFP, and capital are all smaller in the counterfactual economy. In particular, the drop in output is around 1.4 percentage point smaller in the recalibrated counterfactual economy, which accounts for approximately 23 percent of the total decline in output in the benchmark. Based on these observations, we draw the conclusion that the existence of trade credit in the benchmark economy amplifies the financial shock.

1.7.4 Mechanisms behind the amplification effect

In this section, we explore mechanisms behind the amplification effect of trade credit in section 1.7.3. Since we have established that the two economies are comparable in terms of aggregate output and fraction of constrained entrepreneurs in steady state, a nature starting point of exploration is thus the dynamics of trade credit.

During crisis, entrepreneurs become more constrained and the collateral value of trade credit drop. As a result, on the one hand, intermediate goods entrepreneurs are less willing to extend trade credit. On the other hand, final goods entrepreneurs would like to use more trade credit to finance production. The shift in marginal willingness to extend and use trade credit leads to a spike in the trade credit interest rate. As shown in the left panel of Figure 1.8, trade credit interest rate more than triples during the crisis. It jumps from a pre-crisis level of 2.7 percent to as high as 8.3 percent. Accompanying the interest rate spike is a huge drop in the amount of trade credit relative to output (see the right panel of Figure 1.8).

These changes in trade credit indicate that financial shock brings more damage in the benchmark economy through the following channels.

First of all, on the aggregate level, a drop in trade credit leads to a drop in available collateral, thus exacerbates the drop in bank credit. We find that under the same ρ_t shock, the drop in bank credit is much larger in the benchmark economy.

Second, from the steady state analysis in section 1.7.2, we know the existence of trade credit leads to better allocation of resources. During crisis, this reallocation channel is hindered. This can be seen from a huge drop in the quantity of trade credit in the economy. In addition, the increase in trade credit interest rate indicates that productivity of the marginal final goods entrepreneur who uses trade credit increases. In other words, part of the entrepreneurs who use trade credit before crisis no longer use trade credit during crisis because it becomes more expensive.

Lastly, highly productive entrepreneurs are affected disproportionately more than unconstrained ones when bank credit shrinks, because they are the ones that are financially constrained. Compared to bank credit, the negative impact of a drop in trade credit is more disproportionately borne by the productive entrepreneurs. To see this, remember that as shown in Proposition 3, only the most productive and constrained final goods entrepreneurs use trade credit. Thus they are the ones that are directly affected by the drop in trade credit.

1.7.5 Trade credit and the U.S. business cycles

In this section, we extend our study beyond the 2007-09 financial crisis into U.S. business cycles in general. More specifically, we attempt to answer two questions using the model. First, what can we learn about the driving forces of U.S. business cycles? Second, does trade credit amplify business cycle fluctuations in general?

Figure 1.8 shows the cyclical property of trade credit. Trade credit exhibits a strong pro-cyclical pattern – the correlation between gross value added and trade credit is 0.61. It is also more volatile than gross value added – the standard deviation of trade credit is 0.05 compared to a standard deviation of 0.02 for gross value added.

Driving forces of the business cycles We consider two most widely used types of shock that generates the U.S. business cycle – TFP shock and financial shock. In order to understand the driving forces behind the U.S. business cycle, we ask whether these two types of shocks can generate model dynamics that are consistent with data.

The two upper panels in Figure 1.8 plot the dynamics of trade credit and output in the benchmark model following a negative and a positive financial shock to the collateral constraints. In essence, there are the same type of ρ_t shock that is used to generate the 2007-09 financial crisis in section 1.6. In both cases, as shown in the figure, after a financial shock, the percentage drop in trade credit is almost twice as high as the percentage drop in output.

The lower two panels of Figure 1.8 plot the dynamics of trade credit and output following a sectoral neutral TFP shock ρ_t, such that $A_{i,t} = \rho_t A_i$ for $i = 1, 2$. Interestingly, the percent drop in output and trade credit are of the similar magnitude following TFP shocks. Why is this the case? Take the negative TFP shock experiment as an example. Admittedly, in the face of a negative TFP shock, both intermediate goods and final goods entrepreneurs become less productive and therefore less constrained. This leads to a decreasing marginal willingness to borrow trade credit and an increasing marginal willingness to lend trade credit. However, these shifts do not seem to be quantitatively important. In fact, the trade credit interest rate under TFP shocks does not experience a significant change, which indicates that role played by trade credit in reallocating resources is not affected.

Therefore, given that the standard deviation of trade credit is more than twice as high as that of output, we infer that a sectoral neutral TFP can not be the sole source of the business cycles. Financial shocks must have contributed to the business cycle fluctuation.

Does trade credit amplify the business cycles? One important result of our analysis is that, in the face of a financial crisis, disruptions in trade credit amplify the magnitude of financial shocks. It is then nature to ask whether the existence of trade credit amplifies fluctuation of U.S. business cycles.

Our model shows that financial shocks are amplified in the benchmark model with trade credit (see the two upper panels of Figure 1.8). The mechanisms of the amplification effect are discussed previously in section 1.7.4. As shown in the lower panels of Figure 1.8, however, we find that following TFP shocks, dynamics in the benchmark and the counterfactual economy are almost indistinguishable. Further analysis shows that the percentage drop in trade credit is of a similar magnitude as that in output. In addition, the change in trade credit interest rate is very small. These facts show that the role of trade credit in the economy is not affected by TFP shocks. As a result, the amplification effect is missing in this case.

We have established that financial shocks are indispensable for the model to match the business cycle pattern of trade credit. The fact that trade credit amplifies financial shocks lead us to conclude that, through the lenses of our model, trade credit amplifies the U.S. business cycle fluctuations.

1.8 Conclusion

At the present time, production is becoming increasingly specialized. Generally, production of final goods takes place in several stages, with each stage operated by a different firm. Therefore, transactions of intermediate-input goods are carried across firm boundaries and need to be financed. This leads to potential misallocation of intermediate goods (see Jones, 2011). One way to restore this misallocation is to internalize transaction of intermediate inputs through vertically integration of firms. In this paper, we show that trade credit – lending input goods – serves a similar role as vertical integration to alleviate resource misallocation.

Following previous researches on the topic, we posit that the existence of trade credit is due to a certain comparative advantage of input goods suppliers over banks in lending to their customers. In this paper, we find that, the extent to which input goods suppliers can utilize this superior lending "technology" depends crucially on their own financial conditions. This technology is more efficiently used during credit booms and is hindered by financial crisis, which contributes to the aggregate volatility of the economy over credit cycles.

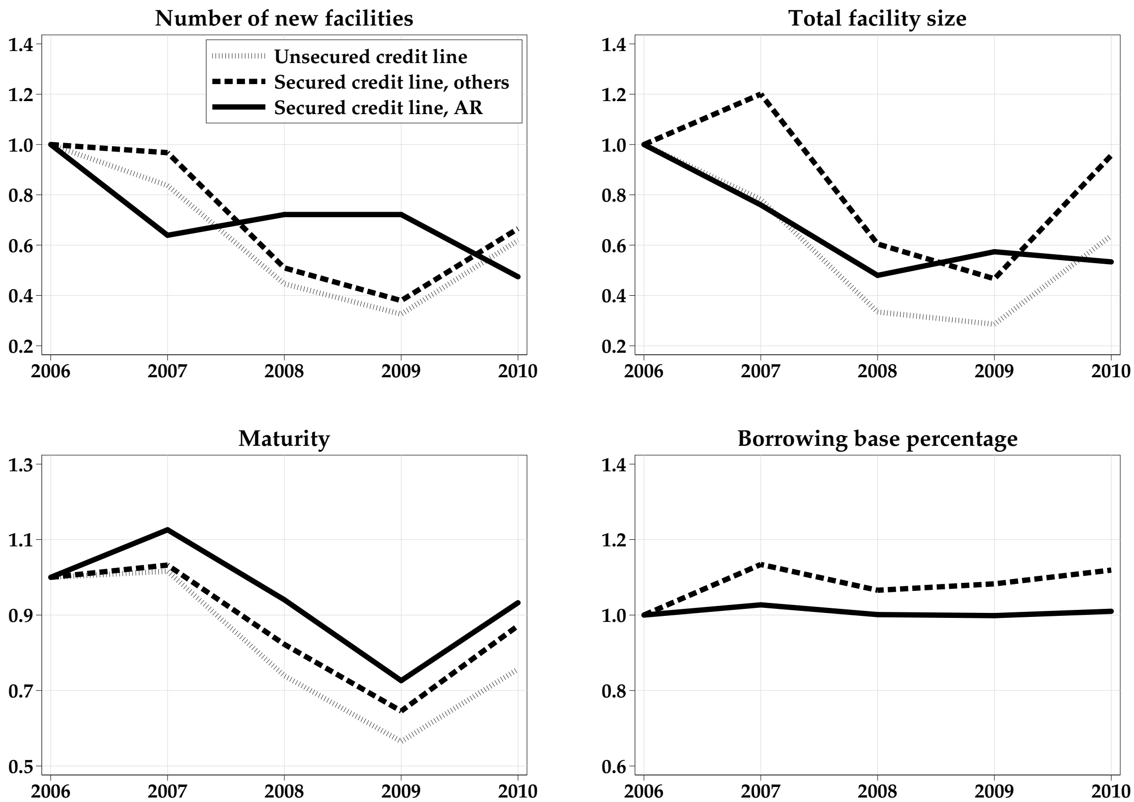


Figure 1.1: Changes in the characteristics of newly opened credit line facilities

Notes: We compute the characteristics of the newly opened credit line facilities of each year as the average of all credit line facilities that are opened in that year. The solid lines in these figures are credit line facilities that require accounts receivable as collateral. The dashed lines are credit line facilities that require other types of assets as collateral. The dotted lines are unsecured credit line facilities. The time series are normalized such that they are 1 in year 2006.

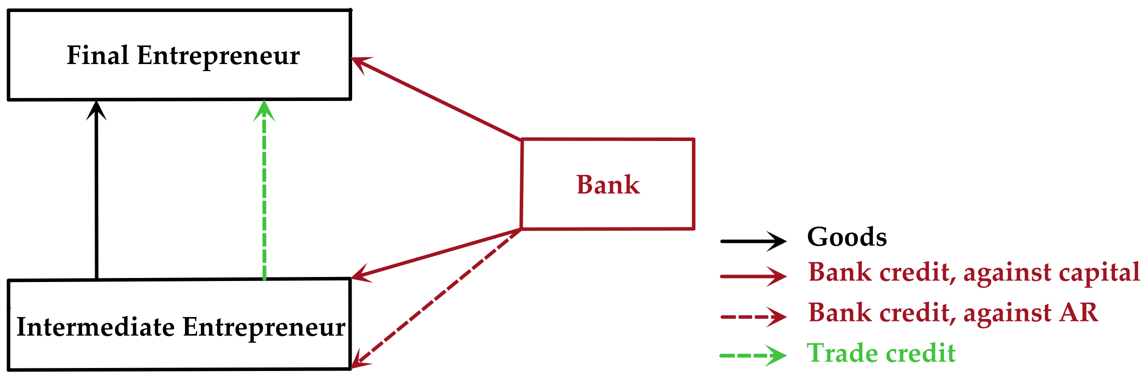


Figure 1.2: Flow of goods and credit

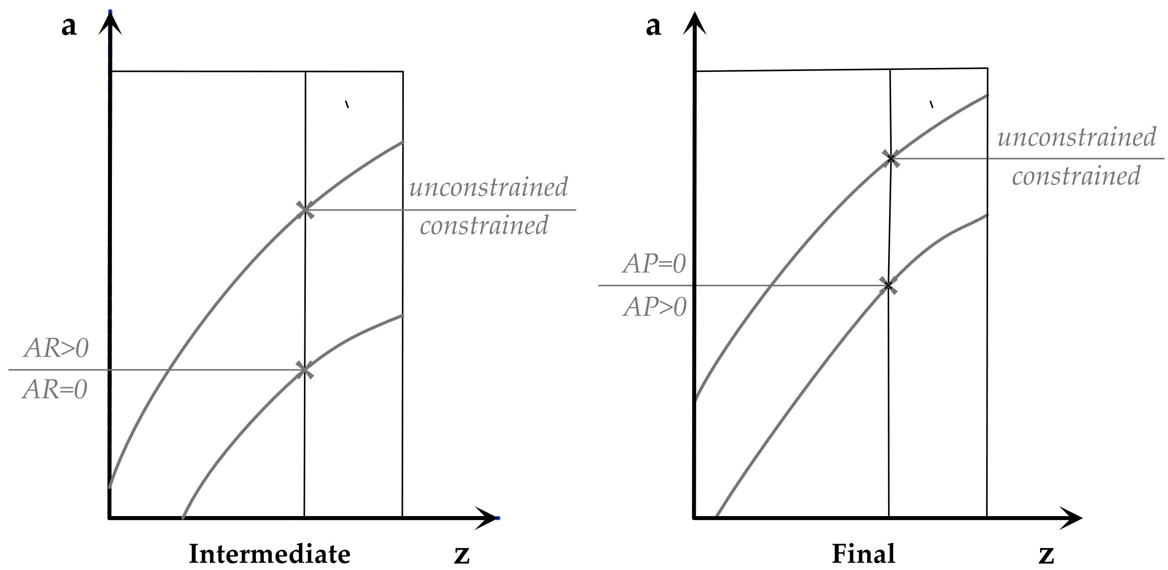


Figure 1.3: Trade credit and being financially constrained: A graphic illustration

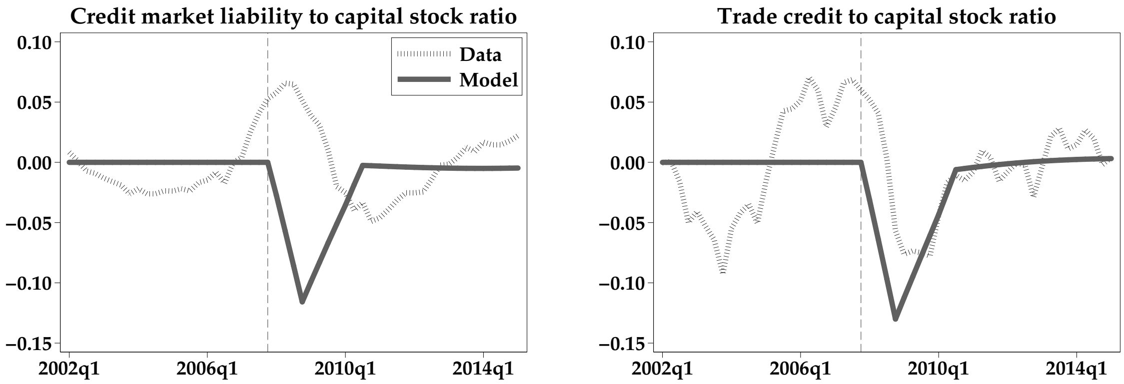


Figure 1.4: Changes in credit market liability and trade credit during the financial crisis

Notes: The data used in above figures are taken for the US nonfinancial corporate sector. Among them, credit market liability is taken from Flow of Funds Table L.103 line 23. Trade credit is calculated as the the average of accounts payable (line 30 of Flow of Funds Table L.103) and accounts receivable (line 15 of Flow of Funds Table L.103). Capital stock is constructed as the sum of equipment (line 46 of Flow of Funds Table B.103), IPP (line 47 of Flow of Funds Table B.103), and nonresidential structural capital (line 51 of Flow of Funds Table B.103), all valued at historical prices. Both credit market liability and trade credit to capital stock ratio are HP-filtered with a smoothing parameter of 1600 and the percentage derivation from trend is plotted in the figures. The corresponding model moments are normalized to be 0 at $t = 0$.

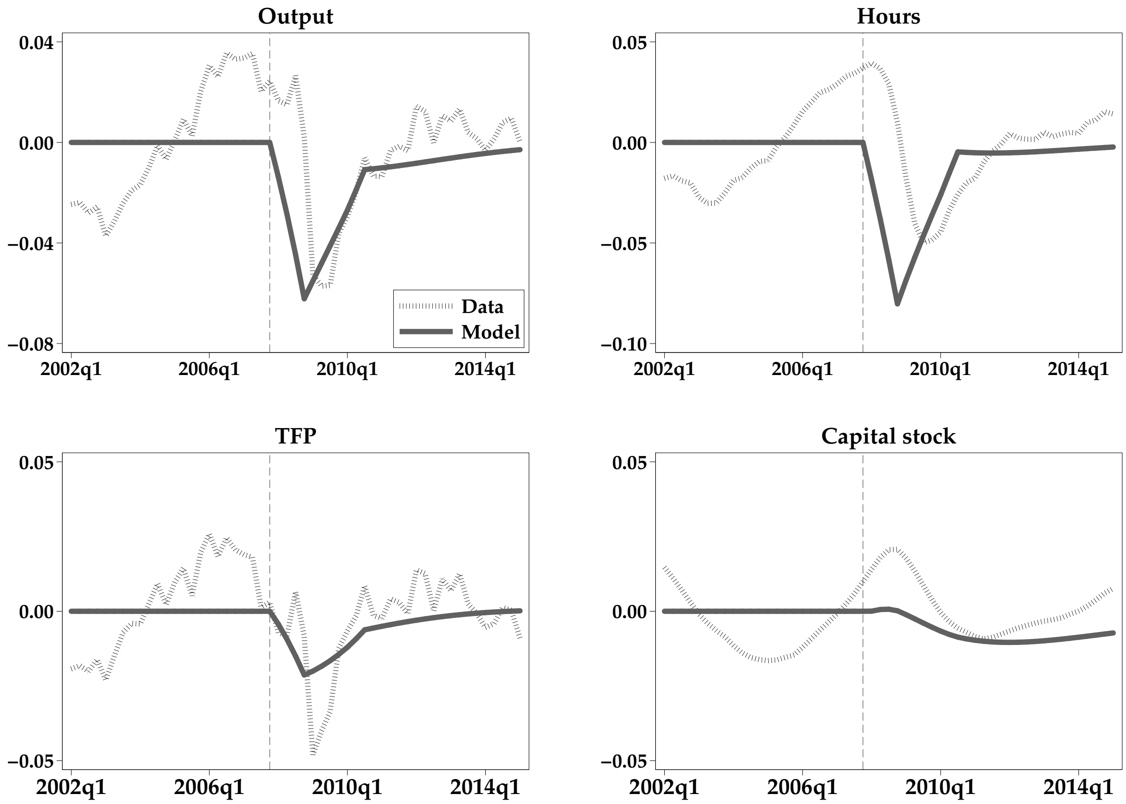


Figure 1.5: Aggregate dynamics of the economy during the financial crisis

Notes: The data used in above figures are taken for the US nonfinancial corporate sector. Among them, output (gross value added) is taken from NIPA Table 1.14 line 17. Data for hours worked is an index taken from Bureau of Labor Statistics Productivity and Cost database (BLS code PRS88003033). Data for capital stock is constructed in the same way as Figure 1.8. TFP is then constructed as a Solow-type residual using output, hours, and capital stock.

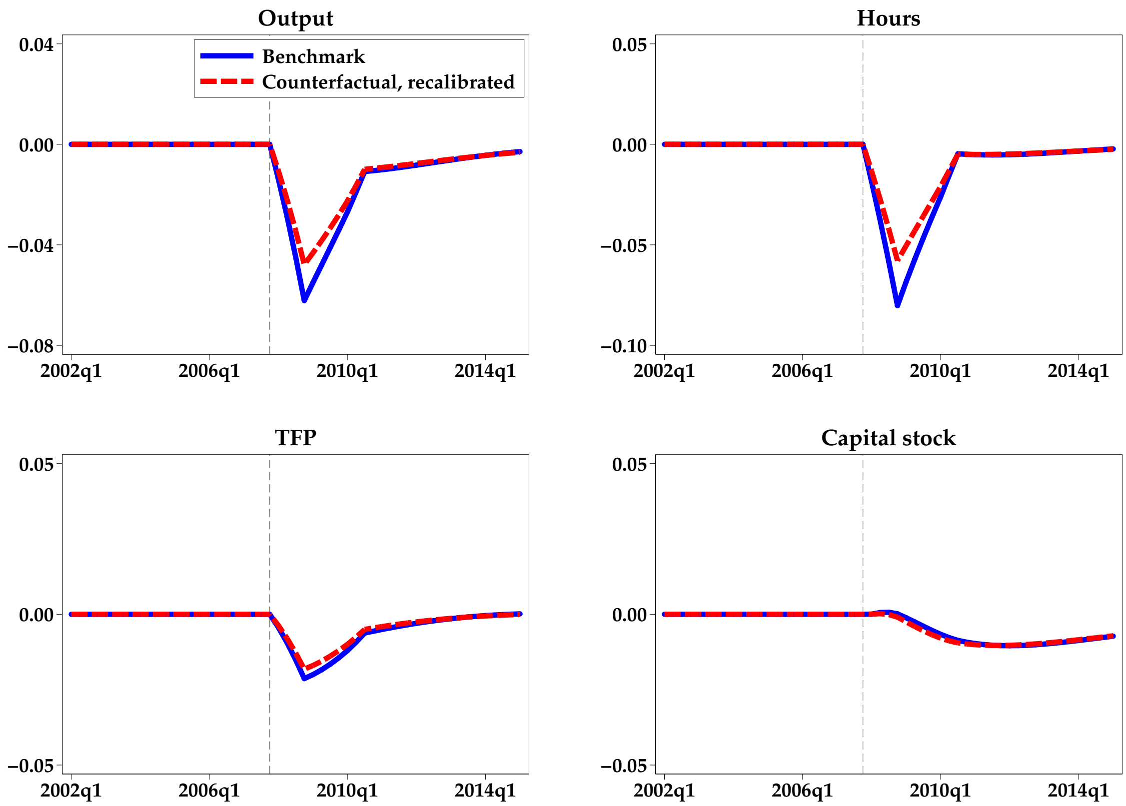


Figure 1.6: Aggregate dynamics after the financial crisis: Benchmark vs. counterfactual

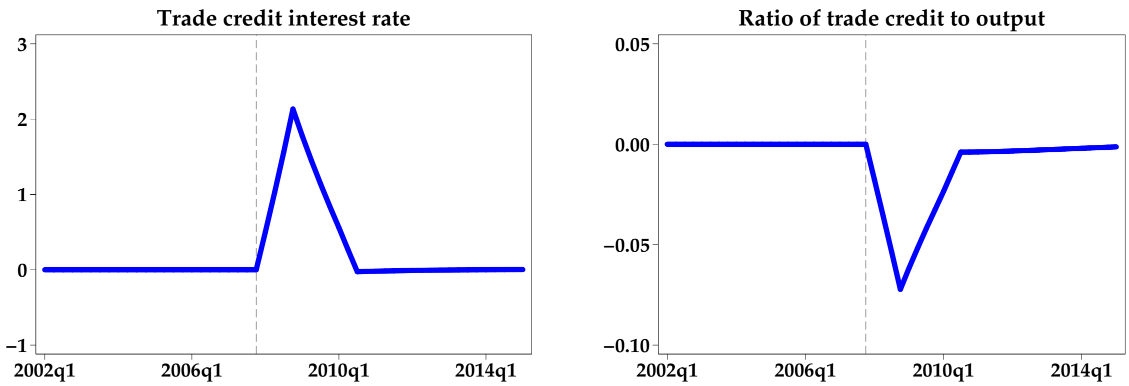


Figure 1.7: Dynamics of trade credit

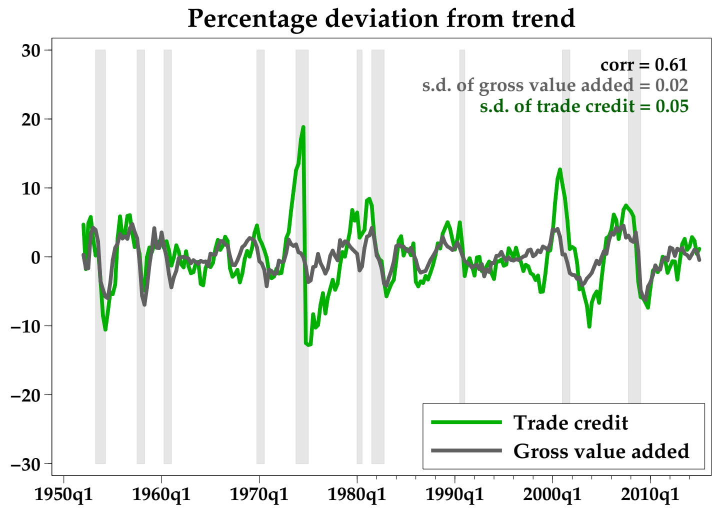


Figure 1.8: Cyclical property of trade credit

Notes: The data is for taken the nonfinancial corporate sector. Gross value added is taken from NIPA Table 1.14 line 17. Trade credit is computed as the average of accounts receivable (line 15 of Flow of Funds Table L.103) and accounts payable (line 30 of Flow of Funds Table L.103). Both time series are HP-filtered with a smoothing parameter of 1600.

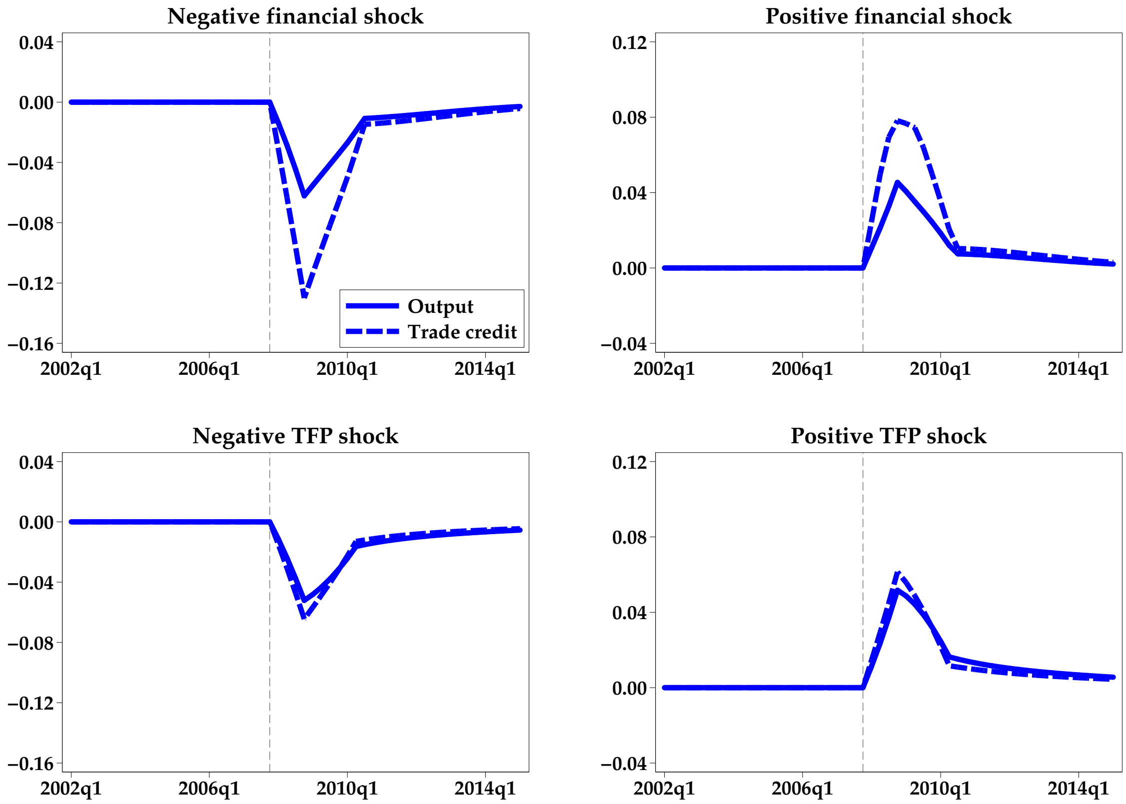


Figure 1.9: Dynamics of trade credit and output after a financial and a TFP shock

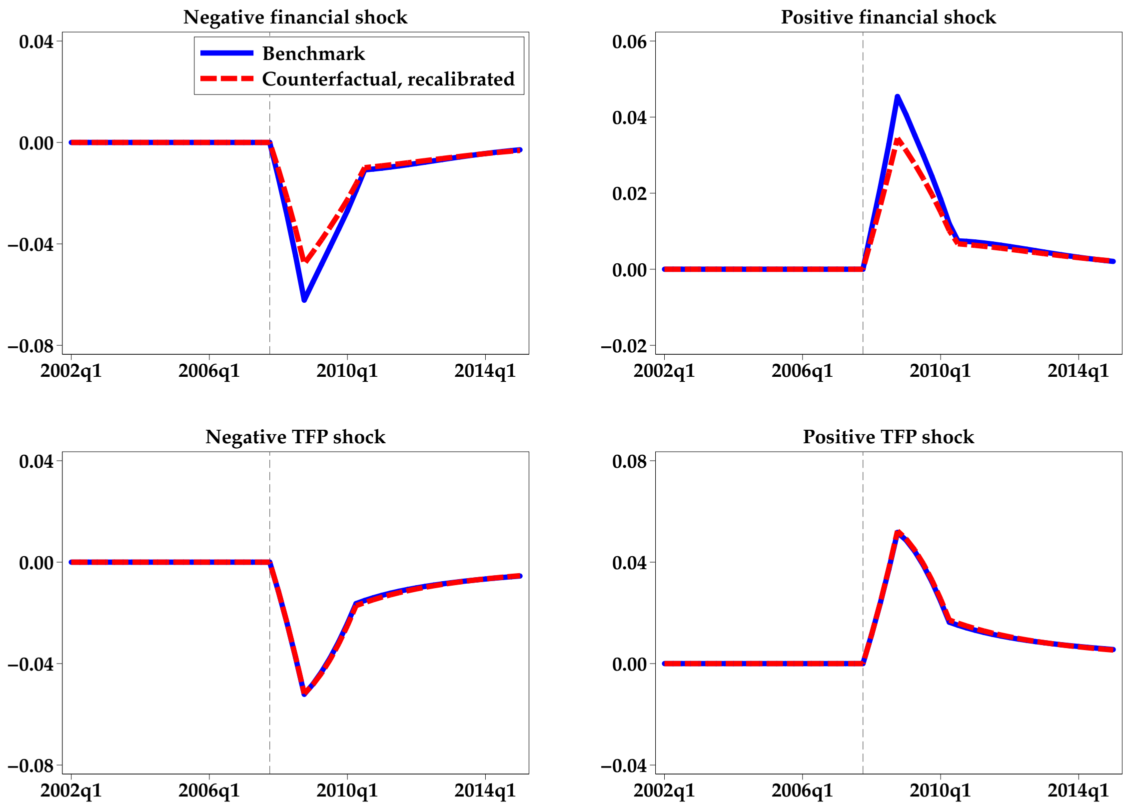


Figure 1.10: Dynamics of output after a financial and a TFP shock

Table 1.1: Trade credit and being financially constrained

Panel A: Net AR/Sales

	(1)	(2)	(3)	(4)
Financially Constrained based on Payout Ratio	-6.198***			
	(-29.86)			
Financially Constrained Based on S&P Rating		-5.766***		
		(-18.22)		
Financially Constrained Based on Size			-11.49***	
			(-40.85)	
Financially Constrained Based on Size				-17.07***
				(-38.33)
Dependent variable	Net AR/S	Net AR/S	Net AR/S	Net AR/S
Sample	Compustat	Compustat	Compustat	Compustat+SSBF
N	26036	26036	26036	34705
AR2	0.130	0.113	0.183	0.219

Panel B: AP/Sales

	(1)	(2)	(3)	(4)
Financially Constrained based on Payout Ratio	6.552***			
	(34.05)			
Financially Constrained Based on S&P Rating		6.964***		
		(23.65)		
Financially Constrained Based on Size			10.05***	
			(38.05)	
Financially Constrained Based on Size				12.30***
				(28.85)
Dependent variable	AP/S	AP/S	AP/S	AP/S
Sample	Compustat	Compustat	Compustat	Compustat+SSBF
N	26036	26036	26036	34705
AR2	0.137	0.120	0.173	0.161

Panel C: AR/Sales

	(1)	(2)	(3)	(4)
Financially Constrained based on Payout Ratio	0.354***			
	(3.00)			
Financially Constrained Based on S&P Rating		1.198***		
		(6.40)		
Financially Constrained Based on Size			-1.435***	
			(-9.92)	
Financially Constrained Based on Size				-4.765***
				(-21.26)
Dependent variable	AR/S	AR/S	AR/S	AR/S
Sample	Compustat	Compustat	Compustat	Compustat+SSBF
N	26036	26036	26036	34705
AR2	0.150	0.151	0.154	0.288

Notes: Our sample includes all but wholesale, retail, and financial and insurance firms in the Compustat and the SSBF data set for the fiscal year 1987, 1993, 1998, and 2003. All regressions include two-digit SIC industry-year fixed effects. Column (4) of every panel includes two dummy variables indicating whether the firms is a corporation and a Compustat firm, respectively. The dependent variables are winsorized at top and bottom 5% for each year. Standard errors are clustered at the firm level.

Table 1.2: Effects of bank health on trade credit extension

	(1)	(2)	(3)	(4)
Crisis X Unhealthy	-1.274*	-1.502**	-1.545**	-1.837**
	(0.681)	(0.696)	(0.714)	(0.718)
Crisis	0.446	0.0672	0.243	2.680
	(0.483)	(0.537)	(1.423)	(6.087)
AP to sales ratio	0.381***	0.382***	0.382***	0.382***
	(0.0294)	(0.0294)	(0.0293)	(0.0292)
Dependent variable	AR/S	AR/S	AR/S	AR/S
Crisis X Credit rating FE	N	Y	Y	Y
Crisis X Firm size bin FE	N	N	Y	Y
Crisis X SIC FE	N	N	N	Y
N	15275	15275	15275	15275
AR2	0.171	0.171	0.172	0.176

Notes: The dependent variables in these regressions are AR/Sales (percent). The sample include quarterly data of 1219 Compustat firms from 2007Q1 to 2009Q4. All regressions include a set of firm fixed effects. Standard errors are clustered at the firm level.

Table 1.3: Summary of calibration

Parameter		Value	Target/Source	Data	Model
θ	inverse of Frisch elasticity	1/2	standard	–	–
α	capital share in production function	1/3	capital share of 1/3	–	–
χ	intermediate goods share	2/3	Yi (2003)	–	–
π	Possion death rate	0.1	Buera et al. (2011)	–	–
N	measure of workers	18	share of entrepreneur	10%	10%
ψ	disutility from working	1.9	hours	0.3	0.3
δ	depreciation rate	0.025	annual 10% depreciation rate	10%	10%
β	discount rate	0.95	annual 4% interest rate	0.4	0.4
μ_1, μ_2	scale parameter	0.85	top 5 percentile earning share	0.3	0.3
ν	Pareto tail	4.0	top 10 percentile employment share	0.69	0.69
γ_1	collateral value of wealth	0.28	credit market liability to nonfinancial assets	0.36	0.36
γ_2	collateral value of AR	0.95	trade receivable to gross value added	0.31	0.31

Notes: The data moment for credit market liability to nonfinancial asset and accounts receivable to gross value added ratio is computed for the nonfinancial corporate sector, averaged over 4 quarters in year 2006. The credit market liability is taken from Flow of Funds Table L.103 line 23. The nonfinancial asset level data is taken from Flow of Funds Table B.103 line 2. The trade receivable data is taken from Flow of Funds Table L.103 line 15. Gross value added is taken from NIPA Table 1.14 line 17.

Table 1.4: Trade credit extension and usage by wealth and productivity of entrepreneurs

	low wealth low productivity	low wealth high productivity	high wealth low productivity	high wealth high productivity
AR to output ratio (%)	100.0	79.9	100.0	79.7
AP to output ratio (%)	29.8	47.9	0.0	53.0

Notes: An entrepreneur is defined to be low wealth (productivity) if she belongs to the bottom 50 percentile in the wealth (productivity) distribution of her own sector. The accounts receivable (payable) to output ratio for each group of entrepreneurs is defined as the sum of accounts receivable (payable) divided by the sum of output of all entrepreneurs in that group.

Table 1.5: Percent difference of the counterfactual relative to the benchmark economy

	output	capital	labor	input goods	TFP
Intermediate sector	-26.4	-23.8	-32.2	–	-0.9
Final sector	-23.9	-0.2	-10.6	-26.4	-7.5
Aggregate	-23.9	-15.3	-24.4	–	-8.4

Notes: This table displays the percent difference of the counterfactual economy relative to the benchmark economy. A negative number in the table suggest that aggregate statistics of the counterfactual economy is lower than that of the benchmark economy.

Chapter 2

Financial Development beyond the Formal Financial Market

2.1 Introduction

A well-developed financial market is crucial to the economic development (see Levine, 2005 for a review of the literature). It is difficult, however, to evaluate the quantitative importance of finance on economic development due to the existence of informal financial activities, which is large relative to the size of formal financial market, but because of its nature, very difficult to measure. Economists have long postulated that informal financial activities plays an important role in developing countries. More specifically, if informal lending, such as borrowing from relatives, friends, or suppliers, can substitute for the lack of a well-functioning formal financial sector, finance might not be an important obstacle to economic growth.

In this paper, we try to quantify the role of informal financial institutions in determining the huge differences in cross-country income.

Using a cross-country firm-level dataset, we document two facts about informal financing. First, on the aggregate level, there is a positive correlation between the size of formal and informal financing. In fact, firms in more developed countries borrow more from the formal financial market, as well as from informal channels. Second, financially constrained firms in countries with a better formal financial market can borrow more from informal channels compared to their constrained counterparts in countries with a less developed formal financial market. These facts indicate that finance might play a more important role in creating the huge cross-country income differences, once the informal financing activities are taken into account.

We build a quantitative model to provide an estimation of the importance of informal financing. Here we restrict our focus to one particular informal financial activity – trade credit. We explore the fact trade credit is relatively easier to measure, compared to other types of informal financing activities, since it appears on the balance sheet of firms according to the accounting rules in many countries. The result shows that, when only taking into account of the formal financial market, improving the quality of the formal financial market in China to the level of the U.S., would lead to a gain of aggregate output of 17%. However, when taking into account of both informal and formal financial market, an improvement of quality of the financial market in China to the level of the U.S. would lead to a gain of aggregate output of 23%.

In the model, this is a continuum of intermediate goods entrepreneurs who produces differentiable intermediate goods using capital, labor, and a composite of all intermediate input goods. Each intermediate goods entrepreneur is a supplier for all the other intermediate goods entrepreneurs and at the same time, demand the intermediate goods produced by the other entrepreneurs. Financial friction arises in the model due to a moral hazard problem of the entrepreneurs to default and divert bank loans. In order to introduce the existence of trade credit, we make two crucial assumptions. First, suppliers have a better technology to enforce the trade credit loan than banks over bank loans. Second, while banks have access to deep pockets, the lenders of trade credit can be financially constrained. In equilibrium, the substitutability of trade credit to the lack of access to bank credit crucially depends on how constrained the lenders of trade credit are. As the access to bank credit improves, the suppliers are more willing to lend trade credit to their customers. This will lead to an increase in the substitutability between trade credit and bank credit.

2.1.1 Literature review

This paper belongs to the following strands of literature.

First, it contributes to the understanding of the financial aspect of firm growth in countries with less developed formal financial sector. In Allen et al. (2005), the authors shows that despite that fact that the formal financial market favors state-owned firms in China, private firms outperform state-owned firms through lending in the informal financial sector. Similarly, Tsai (2001) points out the importance of the informal financial sector in help firm growth in China. Ayyagari et al. (2010) shows instead that firms with access to formal credit (bank loans) grow faster than firms that only utilize informal financing. Recently, Degryse et al. (2013) studies the relationship between informal and formal financing for firms with different sizes. Similar to Degryse et al. (2013) and Ayyagari et al. (2010), this paper focuses on the relationship between informal and formal financing and its impact on firm growth. In addition, we contribute to the literature by showing how the substitutability (or complementarities) between informal and formal financing correlates with the quality of the formal financial market. More specifically, we show that although financially constrained firm in countries with a less developed financial market benefit from lending through informal channels, their constrained counterpart in countries with a highly developed financial market benefit more.

Second, our paper is one of a few papers that model the coexistence of formal and informal financing. We build up on the insights from the previous literature that the efficiency of the financial market is determined fundamentally by the legal, regulatory, and tax system of the economy (see Beck and Levine, 2003 and Levine, 2005). Perhaps what is mostly related to our paper is Madestam (2014), in which the author captures the key trade-off between informal and formal financing as the following: First, informal lenders can better monitor their borrowers than formal lenders, and second, informal lenders might be financially constrained themselves, unlike formal lenders who have access to a deep pocket. In Madestam (2014), the substitutability between the informal and formal lending depends on the monopolistic power of the formal lender. Differing from Madestam (2014), in our paper, the degree of such substitutability is determined by the informal lenders' access to the formal financial market.

Third, this paper belongs to the strand of literature that quantifies the impact of financial friction on productivity loss. In a seminal paper, Hsieh and Klenow (2009) documents that the resource misallocation among firms can account for a large fraction of productivity difference between the U.S. and China. Other papers have shown that financial friction is an important reason behind the resource misallocation (see Buera et al., 2015b for a review of these papers). Our paper expands this literature by incorporating the formal financial market into the framework. Jones (2011) points out at the loss from misallocation can be amplified by the input-output structure, more specifically, by the misallocation of input goods. Following the insights of Jones (2011), we further elaborate in this paper that the choice of trade credit – how input goods are bought and sold – is crucial in understanding how input goods are misallocated among firms.

The rest of the paper is organized as the following: In section 2.2, we provide empirical evidence that motivates the paper. Section 2.3 explains the model. In section 2.4, we define the problem of the entrepreneurs and the recursive competitive equilibrium. In section 2.5, we calibrate the model and perform quantitative analysis. Section 2.6 concludes.

2.2 Empirical evidence

In this section, we provide empirical evidence on the cross-country difference in forma and informal financing. We organize the empirical findings in two parts. In section 2.2.2 we document cross-country patterns on informal and formal financing at the aggregate level; while in section 2.2.3, we present these patterns at the firm level.

2.2.1 Data

For the empirical findings in section 2.2.2, we take data from the World Bank Enterprise Survey (WBES) standardized dataset (2006-2014). We get the relevant information for 109 countries by taking an average over all the firms in each country. In section 2.2.3, we use firm-level data from four different sources. In section 2.2.3.1, we use the WBES firm-level data for 109 countries to document the cross-country difference in the substitutability of informal for formal financing. Then in section 2.2.3.2, we use firm level data on China and U.S. firms to look into the same issue. To this end, we use the Chinese Industrial Survey (2005–07) to study the firm behavior in China.[1] We use a combination of the Compustat and Survey of Small Business Finance (SSBF) to study firms in the U.S..[2]

2.2.2 Aggregate-level pattern

First we document the relationship between the volume of formal and informal financing. The aim of this exercise is to give an idea of how the aggregate use of informal financing changes with the development of the formal financial market. To measure the aggregate use of trade credit, we use the percent of working capital purchased using trade credit averaged over all firms in each country. Here we use the average percent of working capital that is

[1]The Chinese Industrial Survey covers a longer period but the information on trade credit is only available for 2005, 2006, and 2007.

[2]See Shao (2016) for a detailed discussion about how to use these two datasets.

financed by bank credit as a measurement for the development of formal financial market.

The first pattern we documented using aggregate-level data is presented in Figure 2.1. In Panel A, we show the correlation between the size of trade credit and bank credit, and in Panel B, we present the correlation between the size of other informal financing and bank credit. Both figures show the same pattern: The size of informal financing increases with the size of bank credit. In other words, firms in countries with a highly developed formal financial market on aggregate also borrow more from the informal financial market.

2.2.3 Firm-level pattern

In this section, we present evidence of the relationship between informal and formal financing at the firm level. First we utilize the WBES firm-level data and study two types of informal financing: trade credit and informal lending from relatives, friends, and moneylenders. Then we use U.S. and Chinese firm-level data and instead focus on studying only trade credit.

2.2.3.1 The enterprise survey

We run the following regression using enterprise survey data,

$$y_{istc} = \alpha + \beta_c I\_constrained_i + \chi_{st} + I\_young_i \times I\_small_i + \gamma_i + \varepsilon_{ist},$$

in which, y_{ist} is the percent of working capital of firm i in sector s of year t that is financed by trade credit or other informal financing channels, $I\_constrained$ is a dummy indicator of whether the firm i is financially constrained,[3] χ_{st} are a set of sector $\times$ year fixed effects, I_young_i is a dummy indicator of whether the firm is young (equal or less than 5 years old), I_small_i is a dummy indicator of whether the firms is small (equal less than 10 employees), and γ_i is a dummy indicator for the firm type: whether it is government-owned, private, or

[3] A firm is financially constrained if it reports that access to finance is its biggest obstacle of growth.

foreign.

Clearly, in country c, compared to financially unconstrained firms, constrained firms borrowers $\hat{\beta}_c$ percent more working capital in the form of trade credit (or other informal financing).

In Figure 2.2, we plot $\hat{\beta}_c$ agains the GDP per capita of country c.[4] In Panel A, the dependent variable is percent of working capital that is financed by trade credit and in Panel B it is the percent of working capital that is financed by relatives, friends, and moneylenders etc. In both cases, we see that $\hat{\beta}_c$ increases with the income level of that country. In other words, financially constrained firms in developed countries can rely more on informal channels to finance their production than their financially constrained counterparts in developing countries.

2.2.3.2 China versus the U.S.

We have established so far that the availability of informal financing to the financially disadvantaged firms increases with the development of economy and the formal financial market. In this session, we focus on 1) one particular type of informal financing – trade credit, and 2) firms in two countries China and U.S., in order to show the difference between the substitutability of trade credit over the lack of bank credit in these two countries.

We run a regression of the following form,

$$y_{ist} = \alpha + \beta_{c,1} I_{p50} + \beta_{c,2} I_{p75} + \beta_{c,3} I_{p100} + \chi_{st} + \gamma_i + \varepsilon_{ist},$$

in which $c \in \{China,\ U.S.\}$ denotes the two countries.

The dependent variable of this regression is the ratio of accounts receivable to sales, the ratio of accounts payable to sales, and the ratio of net accounts receivable to sales of firm i

[4] We only plot countries with a $\hat{\beta}_c$ that is significant at 10% level. The correlation in both figures is also computed for these countries.

in sector s and year t. We have three dummy variables I_{p50}, I_{p75}, and I_{p100}, indicating, in terms of total asset size, whether the firm belongs to the 25th to 50th percentile, 50th to 75th percentile, and 75th to 100th percentile, respectively. Clearly the control group in this regression are firms that belong to the bottom 25 percentile in terms of total asset, i.e. the smallest firms. Other control variables include a set of sector-year fixed effects χ_{st}, and a set of dummy variables γ_i that controls for firm types.[5]

The object of interests are the $\hat{\beta}_{c,1}$, $\hat{\beta}_{c,2}$, and $\hat{\beta}_{c,3}$. Since many empirical papers suggest that small firms are on average more financially constrained than large firms, if trade credit can substitutes for the lack of access to formal financing, we should see the larger firms borrow significantly less trade credit and lend significantly more. As shown in Table 2.1, this is indeed the case for the U.S. firms. Larger firms in the U.S. lend significantly more trade credit (column 1) and borrows significantly less (column 2). Not surprisingly, in net terms, large firms lend significantly more than their small counterparts (column 3).

However, this patter does not hold for the Chinese firms. As shown in column (4), smaller firms do borrow more trade credit, however, they also seem to lend more to their customers (column 5). In net terms, the difference between small and large firms in terms of the lending of trade credit is less than one percentage point (column 6).

2.3 Model

Time is discrete with infinite horizon. There is one type of final goods which is used to for consumption and investment. The final goods is produced by a representative firm using a continuum of intermediate goods. Each variety of intermediate goods i is produced by an entrepreneur with monopolistic power, using capital, labor, and a composition of

[5]In the U.S. data, we distinguish between the following firm types: Compustat firm or SSBF firm; and corporate or non-corporate firm. In the Chinese data, we control for firm types: State-owned, private, collectively-owned, etc..

intermediate goods. There is a measure N of households in the economy, who are the labor provider in the economy.

2.3.1 Preference

The household's preference is time-separable with instantaneous utility function $u^h(c_t^h, h_t) = log(c_t^h) - \psi h_t$, where ψ represents the disutility from working. The utility of households over a sequence of consumption $c^h = \{c_t^h\}_{t=0}^{\infty}$ and labor $h = \{h_t\}_{t=0}^{\infty}$ can be written as,

$$U^h(c^h, h) = \sum_{t=0}^{\infty} \beta^t u^h(c_t^h, h_t),$$

where β is the discounting factor.

The preference of intermediate goods entrepreneurs is time-separable with instantaneous utility function $u^e(c_t) = log(c_t)$. The expected utility of the entrepreneur over a sequence of consumption $c = \{c_t\}_{t=0}^{\infty}$ can be written as

$$U^e(c) = \mathbb{E} \sum_{t=0}^{\infty} \beta^t u^e(c_t),$$

where the expectation is taking over a stochastic process of productivity $\{z_t\}_{t=0}^{\infty}$, wealth $\{a_t\}_{t=0}^{\infty}$, and output $\{y_{t-1}\}_{t=0}^{\infty}$.

2.3.2 Production technology

The final goods firm produce final goods $\bar{y}$ using a composition of intermediate input goods $\{y_i\}_{i=0}^{1}$ with the following technology

$$\bar{y} = (\int_0^1 y_i^{\rho} di)^{1/\rho},$$

where $\rho \in (0,1)$ is the elasticity of substitution among the intermediate goods.

The intermediate goods entrepreneur i employs capital $k(i)$, labor $l(i)$, and intermediate goods $\{x_j(i)\}_{j=0}^1$ to produce the intermediate goods i, such that,

$$y_i(i) = AzF(k(i), l(i), \bar{x}(i)) = Az(k(i)^\alpha l(i)^{1-\alpha})^{1-\chi} \bar{x}(i)^\chi,$$

in which $\bar{x}(i) = (\int_0^1 x_j(i)^\rho dj)^{1/\rho}$, A is the aggregate TFP of the economy, and z the idiosyncratic productivity of the entrepreneur.[6]

2.3.3 Timing

An intermediate goods entrepreneur i enters each period with wealth a and last period output y. Her idiosyncratic productivity shock z is then realized. Notice that since that all intermediate goods are symmetric in the production of final goods and intermediate goods, we can simply use her state variables (a, y, z) to represent the identity of intermediate goods entrepreneur i. To simplify the notation, we drop the entrepreneur index i henceforth in this paper.

The entrepreneur makes the following choices: First, she provides two prices for her goods y. One is the price of the goods if it is sold on the spot, which is denoted by p^{ntc}. The other is the price of the goods if it is lent as trade credit, which is denoted by p^{tc}. Second, she decides the amount of capital k and labor l to employ. She also chooses the amount of intermediate input goods to use $\{x_j\}_{j=0}^1$, and whether or not to borrow them through trade credit $I^{tc} \in \{0, 1\}$. She then borrows from the bank to finance working capital and make a decision of whether or not to default and divert the bank loan. After the default decision is made, the entrepreneur then starts to produce.

At the beginning of the period, representative final goods producer purchases the inter-

[6]Note that $x_j(i)$ presents the amount of j^{th} goods used by entrepreneur i. Since entrepreneur i has the monopoly power in producing goods i, we have $y_i(j) = 0$ for all $j \neq i$.

mediate goods. The production of the final goods takes one period to finish. At the end of this period, intermediate goods entrepreneurs and workers consume. Intermediate goods entrepreneurs also save into the next period a'. Her current period output is carried out into the next period as y'.

2.3.4 Markets

The economy consists of the following markets: capital rental market, labor market, final goods market, and a continuum of intermediate goods market. The markets of capital, labor and final goods are perfectly competitive. The market-clearing prices for capital and labor are r and w, respectively. And the price for the final goods is normalized as 1. Each intermediate goods entrepreneur is the sole provider of its goods in that particular intermediate goods market. She posts two prices for her goods: p^{ntc} and p^{tc}. The former is the price when goods is paid on the spot, and the latter is the price when the goods is lent out as trade credit. The pricing schedule implies an implicit trade credit interest rate $\frac{p^{tc}}{p^{ntc}} - 1$.

2.3.5 Financial frictions

In this section, we discuss the moral hazard problem of the entrepreneurs as well as the co-existence of bank and trade credit as a result of it.

2.3.5.1 The existence of trade credit

Following Burkart and Ellingsen (2004), we assume that intermediate goods entrepreneurs can only default and divert bank loans before entering production stage. In other words, once the entrepreneurs start producing, the moral hazard problem ceases to exist.

Furthermore, we make the following assumptions regarding the comparative advantage of intermediate goods entrepreneurs in lending trade credit:

Assumption 1 *Intermediate goods entrepreneurs can observe and verify the input goods transaction in which they are part of.*

Assumption 2 *Entrepreneurs i has access to the market i as a seller. She can only enter the other market $j \neq i$ as a buyer.*

The above assumptions follow those made in Burkart and Ellingsen (2004) very closely. In fact, they are simplified, or more abstract versions of the assumptions made in that paper. The first assumption is that the supplier can observe the input goods purchase but the bank can not. This is a reasonable assumption because the supplier is one of the two parties that involve in the input goods transaction. The second assumption indicates that the entrepreneur does not receive any benefit by diverting the input goods, but they can fully enjoy the diverted bank loan. In order to benefit from the diversion of input goods, the entrepreneurs need to sell the input goods in exchange for final goods, but they can not do that. The underlying assumption is that they are not institutionalized or it is very costly for them to search for a potential buyer.

Proposition 4 *Under assumption 1 and 2, trade credit is essentially a trio-party financing arrangement, involving the intermediate goods entrepreneurs, the buyer of their goods, and the bank. In this arrangement, 1) it is secure for the intermediate goods entrepreneurs to lend their goods as trade credit, and 2) the bank lends against accounts receivable.*

The logic of proof goes as the following. The entrepreneur can not benefit from diverting input goods, therefore they will only purchase input goods from supplier if they would like to produce. Once they start producing, the moral hazard problem is resolved. Therefore, from the point of view of the entrepreneur, lending input good is always secured. After lending the input goods, the account receivable is a perfect collateral to obtain loans from the bank, because it knows that the account receivable serves as a signal that the input transaction is complete. At the same time, when the account receivable is posted as a collateral, the owner of this liability is transferred from the hand of the supplier to that of the bank. Therefore, the trade credit will always be paid back by the firm.[7]

One important feature of this model is that the account receivable serves as collateral to obtain bank loans. The collateral value of the accounts receivable depends on the fraction of the asset value that can be recovered by the bank upon the default of the entrepreneurs. From the perspective of the firm, part of the loss in cash flow due to the trade credit extension can be offset by the enlarged access to bank credit. In other words, the trade credit in this model is implicitly and partially financed by the banks.

2.3.5.2 Working capital constraint

Due to the mismatch between the production and the realization of revenue, firms need to raise loans to cover working capital. Work capital cost includes interest payment of capital, labor cost, as well as cost for intermediate input goods if entrepreneurs choose to purchase those on the spot.

To obtain bank loans, entrepreneurs can provide their wealth, which is in the form of physical capital, as well as accounts receivable as collateral. The bank can recover a fraction χ_1 of the wealth and χ_2 of the accounts receivable once the entrepreneur threats to default and renegotiate the loan. The Nash Bargaining problem can be written as the following,

$$\max_{x_b, x_e}(x_b - \chi_1 a - \chi_2 AR)^{\alpha}(x_e - b)^{1-\alpha}$$

$$s.t.$$

$$x_b + x_e = b + a + AR,$$

[7]In reality, account receivable is one of the most senior loan.

in which b is the size of the bank loan, and $\{x_b, x_e\}$ are the bargaining outcome for the bank and the entrepreneur. If the entrepreneur defaults, the payoff for the bank is $\chi_1 a + \chi_2 AR$. The payoff of the entrepreneur is b, the size of the bank loan that she can divert. However, if they reach an agreement, the sum of their payoffs will be $b + a + AR$. Solving the above bargaining problem gives a payoff schedule as following,

$$x_b = (\alpha + (1-\alpha)\chi_1)a + (\alpha + (1-\alpha)\chi_2)AR, \ and$$

$$x_e = b + (1-\alpha)(1-\chi_1)a + (1-\alpha)(1-\chi_2)AR.$$

The bank loan contact that takes away the incentive of the entrepreneurs to default in the first place has to satisfy the following condition:

$$a + AR \geq x_e.$$

This leads to the following constraint on the bank loan limit,

$$b \leq \gamma_1 a + \gamma_2 AR,$$

in which $\gamma_1 = 1 - (1-\alpha)(1-\chi_1)$ and $\gamma_2 = 1 - (1-\alpha)(1-\chi_2)$.

Now we are ready to write down the working capital constraint,

$$use\ trade\ credit: \ rk + wl \leq \gamma_1 a + \gamma_2 AR, \tag{2.1}$$

$$otherwise: \ rk + wl + \int p_j^{tc} x_j dj \leq \gamma_1 a + \gamma_2 AR. \tag{2.2}$$

2.3.5.3 Inverse demand function

Next we define the inverse demand function for the intermediate goods entrepreneurs. Notice that a demand for the intermediate input goods comes from three types of agents: 1) the final goods producer, 2) the intermediate goods entrepreneur who wants to use trade credit, and 3) the intermediate goods entrepreneurs who do not use trade credit. We derive the demand from them one by one.

Final goods producer First there is the demand from the final good producer. Since we assume that final goods producer always purchase the intermediate input goods on the spot, therefore the pricing schedule she faces is $\{p_j^{ntc}\}_{j=0}^1$. The optimization problem of the final goods producer reads,

$$\max \bar{p}\bar{y} - \int p_j^{ntc} dj \quad s.t.$$

$$\bar{y} = (\int_0^1 y_j^\rho dj)^{1/\rho}, \tag{2.3}$$

in which $\bar{y}$ is the final goods output, and $\bar{p}$ is the price of final goods.

The FOC w.r.t. y_j gives an inverse demand function for goods j from the final goods producer.

$$y_j = \bar{y}(\frac{p_j^{ntc}}{\bar{p}})^{\frac{1}{\rho-1}}. \tag{2.4}$$

Taking 2.4 back to 2.3,

$$\bar{y} = (\int_0^1 \bar{y}^\rho (\frac{p_i^{ntc}}{\bar{p}})^{\frac{\rho}{\rho-1}} di)^{\frac{1}{\rho}},$$

which leads to

$$\bar{p} = (\int_0^1 p_i^{ntc\frac{\rho}{\rho-1}} di)^{\frac{\rho-1}{\rho}}.$$

As discussed in section 2.3.4, we normalize the price of final goods to be 1. Therefore, it holds that $(\int_0^1 p_i^{ntc\frac{\rho}{\rho-1}} di)^{\frac{\rho-1}{\rho}} = 1$.

Intermediate goods entrepreneurs who do not use trade credit In order to derive the inverse demand function from this type of entrepreneurs, we need to first analysis their problem. Let $V^{ntc}(a, y, z)$ be the value function of the entrepreneur who decides to purchase the input goods on spot and whose state variables are (a, y, z).[8]

$$V^{ntc}(a, y, z) = \max_{p^{tc}, p^{ntc}, a', c, k, l, x_i} log(c) + \beta \mathbb{E} V(a', y', z') \quad s.t.$$

$$c + a' = (1+r)a + p^{tc} y^{tc} + p^{ntc} y^{ntc} - wl - (r+\delta)k - \int p_i^{ntc} x_i di,$$

$$rk + wl + \int p_i^{ntc} x_i \leq \gamma_1 a + \gamma_2 p^{tc} x^{tc} + p^{ntc} x^{ntc},$$

$$y' = AzF(k, l, (\int x_i^\rho di)^{1/\rho})).$$

We derive FOC w.r.t. x_i as the following,

$$x_i : \; -\frac{1}{c} p_i^{ntc} + \beta \mathbb{E} V_{y'} AzF_{\bar{x}} \bar{x}^{1-\rho} {x_i}^{\rho-1} - \mu p_i^{ntc} = 0,$$

where μ is the Lagrangian multiplier on the working capital constraint and $\bar{x} = (\int x_i^\rho di)^{1/\rho}$. Given any two goods $i \neq j$, the above equation leads to $\frac{p_i^{ntc}}{p_j^{ntc}} = (\frac{x_i}{x_j})^{\rho-1}$. Furthermore, we can show that entrepreneur's demand for goods x_i can be written as the following,

$$x_i = \bar{x}\left(\frac{p_i^{ntc}}{(\int_0^1 {p_j}^{ntc \frac{\rho}{\rho-1}} dj)^{\frac{\rho-1}{\rho}}}\right)^{\frac{1}{\rho-1}} = \bar{x} {p_i}^{ntc \frac{1}{\rho-1}},$$

in which the second equation holds since we normalize the price of the final goods to be 1, i.e. $(\int_0^1 {p_j}^{ntc \frac{\rho}{\rho-1}} dj)^{\frac{\rho-1}{\rho}} = 1$.

In addition, we can show that the cost of using input goods $\int p_i^{ntc} x_i di$ is equal to $\bar{x}$.

[8]Notice that this problem is missing three constraints: 1) the inverse demand function that link y^{tc} to p^{tc}, 2) the inverse demand function that links y^{ntc} to p^{ntc}, and 3) the constraint that $y^{tc} + y^{ntc} = y$. These three constraints are irrelevant for the analysis here. Therefore they are omitted.

Intermediate goods entrepreneurs who use trade credit Similar, for those entrepreneurs who use trade credit to finance their production, their problem can be written as,[9]

$$V^{tc}(a,y,z) = \max_{p^{tc},p^{ntc},a',c,k,l,x_i} log(c) + \beta \mathbb{E} V(a',y',z') \ \ s.t.$$

$$c + a' = (1+r)a + p^{tc}y^{tc} + p^{ntc}y^{ntc} - wl - (r+\delta)k - \int p_i^{tc} x_i di,$$

$$rk + wl \leq \gamma_1 a + \gamma_2 p^{tc}x^{tc} + p^{ntc}x^{ntc},$$

$$y' = AzF(k,l,(\int x_i^\rho di)^{1/\rho})).$$

From the FOC w.r.t. x_i, we can derive the inverse demand function as,

$$x_i = \bar{x}(\frac{p_i^{tc}}{(\int_0^1 p_j^{tc\frac{\rho}{\rho-1}} dj)^{\frac{\rho-1}{\rho}}})^{\frac{1}{\rho-1}} = \bar{x}(\frac{p_i^{tc}}{\bar{p}^{tc}})^{\frac{1}{\rho-1}},$$

where $\bar{p}^{tc} = (\int_0^1 p_j^{tc\frac{\rho}{\rho-1}} dj)^{\frac{\rho-1}{\rho}}$. And the cost of purchase input goods $\{x_j\}_{j=0}^1$ with trade credit can be written as

$$\int p_i^{tc} x_i di = \bar{p}^{tc}\bar{x}.$$

Taking stock By providing a pricing schedule $\{p^{ntc}, p^{tc}\}$, the intermediate goods entrepreneur is faced with the following inverse demand function,

$$y^{ntc} = Q^{ntc} p^{ntc\frac{1}{\rho-1}}, \ and$$

$$y^{tc} = Q^{tc}(\frac{p^{tc}}{\bar{p}^{tc}})^{\frac{1}{\rho-1}}.$$

In the first equation, y^{ntc} is the demand of her goods from the final goods producer and the intermediate goods producers who do not use trade credit, and Q^{ntc} is the sum of final

[9]Similarly, this problem is missing three constraints.

goods and of all $\bar{x}$, the composite of intermediate input goods used by these entrepreneurs. In the second equation, y^{ntc} is the demand for her goods from intermediate goods producers who use trade credit. Similarly, Q^{tc} is the sum of the composition composite intermediate input goods $\bar{x}$ used by these entrepreneurs who borrow through trade credit. A mathematical formula for Q^{ntc} and Q^{tc} can be found in Definition 3.

2.4 Recursive competitive equilibrium

In this section, we first present the problem of household and entrepreneurs in the recursive form, and then define the recursive competitive equilibrium.

2.4.1 Households

The state variable of household is her wealth a, and her problem can be written as,

$$V^h(a) = \max_{a'} u^h(c,h) + \beta V^h(a') \quad s.t. \tag{2.5}$$

$$c + a' = a(1+r) + wh.$$

2.4.2 Intermediate goods entrepreneurs

The intermediate goods entrepreneurs enter each period with wealth a, output carried over from the previous period y, and realized productivity shock z, they choose whether or not use trade credit $I^{tc} \in \{0,1\}$, the price of her goods p^{ntc} if they are sold on the spot, the price of her goods p^{tc} if they are sold through trade credit, the production scale k, l, and $\bar{x}$, this period consumption c, and next period wealth a'. Their problem can be written recursively as,

$$V^e(a,y,z) = \max_{c,k,l,I^{tc},a',\bar{x},y^{tc},y^{ntc}} u^e(c) + \beta \mathbb{E}_{z'} V^e(a',y',z'), \tag{2.6}$$

$$s.t.$$

$$c + a' = (1+r)a + p^{tc}y^{tc} + p^{ntc}y^{ntc} - wl - (r+\delta)k - (1-I^{tc})\bar{x} - I^{tc}\bar{p}^{tc}\bar{x},$$

$$y' = AzF(k,l,\bar{x}),$$

$$0 \leq y^{tc} = Q^{tc}(\frac{p^{tc}}{\bar{p}^{tc}})^{\frac{1}{\rho-1}} \leq y,$$

$$0 \leq y^{ntc} = Q^{ntc}p^{ntc\frac{1}{\rho-1}} \leq y,$$

$$y^{tc} + y^{ntc} = y,$$

$$rk + wl + (1-I^{tc})\bar{x} \leq \gamma_1 a + \gamma_2 p^{tc}y^{tc} + p^{ntc}y^{ntc}.$$

Now we are ready to define the recursive competitive equilibrium.

Definition 3 *The recursive competitive equilibrium consists of a set of prices* $\{\bar{p}^{tc}, w, r\}$*, aggregate demand* $\{Q^{ntc}, Q^{tc}\}$*, value function of the entrepreneurs* $V(a,y,z)$*, policy functions* $c(a,y,z)$*,* $k(a,y,z)$*,* $l(a,y,z)$*,* $I^{tc}(a,y,z)$*,* $\bar{x}(a,y,z)$*,* $y^{tc}(a,y,z)$*,* $y^{ntc}(a,y,z)$*,* $p^{tc}(a,y,z)$*,* $p^{ntc}(a,y,z)$ *and* $a'(a,y,z)$*, asset holding of the households* a^h*, value function of the households* $V^h(a)$*, policy functions* $c^h(a)$*,* $l^h(a)$ *and* $a^{h'}(a)$*, and CDF function for the distribution of state variables* $\Phi(a,y,z)$*, such that,*

1. Given r and w, the value function and policy functions solve the household's problem 2.5.

2. Given r, w, Q^{tc} and Q^{ntc}, the value function and policy functions solve the entrepreneur's problem 2.6.

3. Markets clear

(a) Capital market:

$$\int k(a,y,z)d\Phi(a,y,z) = \int ad\Phi(a,y,z) + N \cdot a^h.$$

(b) Labor market:

$$\int l(a,y,z)d\Phi(a,y,z) = N \cdot l^h(a^h).$$

(c) Aggregate demand $\{Q^{tc}, Q^{ntc}\}$ are:

$$Q^{tc} = \underbrace{\int I_{tc}(a,y,z)\bar{x}(a,y,z)d\Phi(a,y,z),}_{composite\ of\ intermediate\ input\ goods}$$

and

$$Q^{ntc} = \underbrace{N \cdot c^h(a^h) + \int c(a,y,z)d\Phi(a,y,z) + \delta K}_{final\ goods}$$
$$+ \underbrace{\int (1 - I_{tc}(a,y,z))\bar{x}(a,y,z)d\Phi(a,y,z),}_{composite\ of\ intermediate\ input\ goods}$$

where

$$K = N \cdot a^h + \int ad\Phi(a,y,z).$$

4. The CDF function $\Phi(a,y,z)$ is stationary.

2.5 Quantitative analysis

In this section, we first calibrate the model to match the data moments of the U.S. and China in section 2.5.1. In section 2.5.2, we use the calibrated model to examine the importance of trade credit in explaining the income difference between these two countries. Although in the empirical section 2.2, we study informal financing activities in general, in this section, due to data limitation, we will focus on studying one type of informal financing – trade credit. We notice two things from studying the WBES data: First, in terms of empirical pattern, trade credit is qualitatively no different than the other type of informal financing, and second, trade credit is about 4 times as large as the other informal financing.

2.5.1 Calibration strategy and results

Each period in the model corresponds to one quarter in the data. We first calibrate our model to match the U.S. aggregate and firm-level moments.

The U.S. The household utility function is $log(c_t^h)-\psi h_t$, in which ψ indicates the disutility from working. We calibrate ψ such that in equilibrium, household's working hour h is equal to 0.3. The discounting factor β is calibrated such that the annual risk-free interest rate is 0.04. Since the share of entrepreneurs in the U.S. data is around 10 percent, we pick $N = 9$ so that the share of entrepreneurs in the model matches that.

There are three crucial parameters in the production functions. In the intermediate goods production function, we pick $\alpha = 1/3$ to match the share of capital return. The share of intermediate input goods χ is set to be 0.43, following Zetlin-Jones and Shourideh (2016). The elasticity of substitution between intermediate input goods, which appears both in the intermediate goods production and final goods production function, is governed by parameter ρ. We set $\rho = \frac{3}{4}$, so that the elasticity of substitution between intermediate goods is 4, which is common in the literature and the same as in Zetlin-Jones and Shourideh (2016). The depreciation rate δ is chosen to be 0.025 so that the annual depreciation rate of capital is equal to 10 percent. The stochastic process of idiosyncratic shock z is parameterized following Buera et al. (2011): a Poisson death rate of $\pi = 0.1$ and a new draw of productivity z from a Pareto distribution with tail parameter μ, which is calibrated to match the top 10 percentile employment share. The two important parameters that govern the performance of financial market, γ_1 and γ_2, are calibrated to match the ratio of total debt to asset, and the ratio of trade credit to total asset.

China We then take the calibrated parameters as shown above, and recalibrate a subset of them to match data moments of Chinese manufacturing firms, which are computed using the Chinese industrial survey for the year 2006. The first data moment of target is the top 10 percentile share in employment. According to Buera et al. (2011), we know that in the U.S., the top 10 percentile firms employ about 69% of the total labor force in the manufacturing sector. In China, this number at 57%, which is significantly lower. This is consistent with finding in Hsieh and Klenow (2009). The second data moment is the total debt to asset ratio, which measures the indebtedness of the firms. And the third data moment is the ratio of trade credit to total asset, which measures the reliance of firms on trade credit in financing their production.[10]

A summary of the calibration strategy and results can be found in Table 2.2

2.5.2 The quantitative importance of trade credit

The question that we are after in this section is: How important is finance, and especially informal financing such as trade credit in explaining the cross-country productivity and income differences.

The benchmark To answer this question, we first look at the benchmark model (section 2.3) and its calibrated version (section 2.5.1). A comparison between the calibrated parameters of China and the U.S. suggests that both the collateral value of wealth (physical capital) as well as accounts receivable are lower in China than in the U.S.. In other words, through

[10]It is more nature to use the ratio of trade credit to total sales, or the ratio of trade credit over cost of goods sold to measure firms' reliance on trade credit. But here I use total asset instead because it is more comparable across the two datasets we have for the U.S. and China.

the lenses of our model, a lower γ_1 and γ_2 can be seen as a result of a worse contract enforcement environment in China (see section 2.3.5.1). It naturally leads to a lower leverage ratio in China and firms are more constrained in equilibrium. As a result, even if they have a better monitoring technology in lending input goods to their customers, they can not afford to do so in such an environment. Keeping other parameters fixed, an increase in γ_1 and γ_2 of the Chinese economy to that of the U.S. level leads to an increase in aggregate output of 23%.

The counterfactual We have established so far that the difference in the financial development between the U.S. and China contributes to the aggregate productivity difference between these two countries. The next question is then, how important is the role of trade credit (informal financing)?

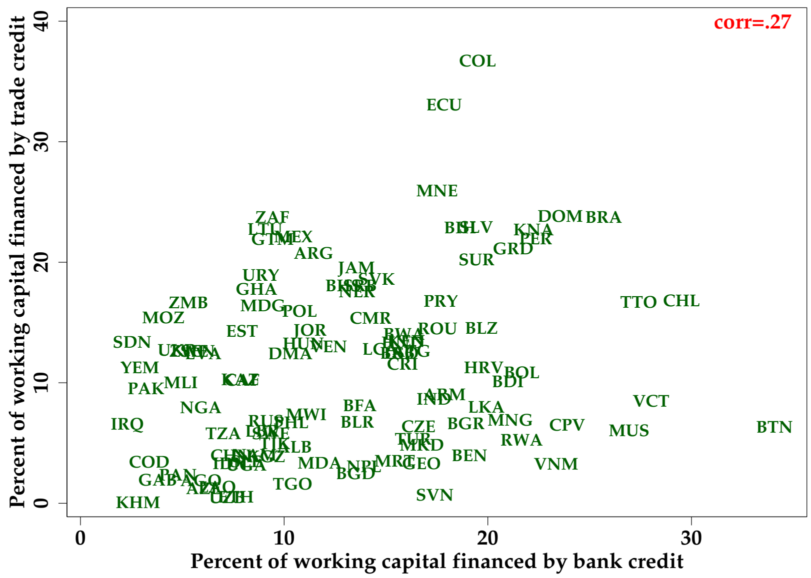
In order to answer this question, we consider a counterfactual economy, in which there is only the formal financial market. We recalibrate the model such that the counterfactual economy can match the data moments in the U.S. and China, except the those related to trade credit. In particular, to match the total debt to asset ratio in the economy, we would have to increase the parameter value of γ_1 to 0.72 for the U.S. and to 0.59 for China. We consider this calibrated version of the counterfactual economy as what is used by most previous researches. In this economy, we only consider what the financial friction presented in the formal financial market. If we improve the quality of China's formal financial market to that of the U.S. level, i.e. increase γ_1 to 0.72 for China, the total gain in aggregate productivity and output level is 17%. Taking into consideration that there is an 23% increase in productivity and output in the benchmark model if the quality of Chinese informal and formal financial market both improve to the level of the U.S., our exercise shows that 7% of that can be attributed to the informal market it self.

2.6 Conclusion

One puzzle about the three decades of stellar economic growth in China is the role of finance. While the financial institution in China has been widely regarded as one of the weakest link in the Chinese economy, the private sector firms, albeit suffering from financial repressions, have been growing rapidly, out-pacing their state-owned counterparts. Many suggests that one of the reasons to explain this puzzle is although private firms have limited access to the formal financial market, they can borrow through informal channels, such as from relatives, friends, and suppliers (see Allen et al. (2005)). This view coincides with findings from previous literature that informal financial institutions can somehow substitute the lack of formal financing in developing countries, therefore mitigating the loss from financial frictions.

The findings in our paper question this view. In our paper, the quality of both formal and informal financial market are determined by the quality of fundamental contract enforcement in an economy. Informal financial market channels fundings to firms that are on average more constrained by their access to the formal financial market. This particular role of informal financing is enhanced with the quality of the formal financial market. Informal financing might have contributed to growth of private firms in China, but its effects would have been much larger if those firms were in the United States.

Panel A: Trade credit



Panel B: Other informal financing

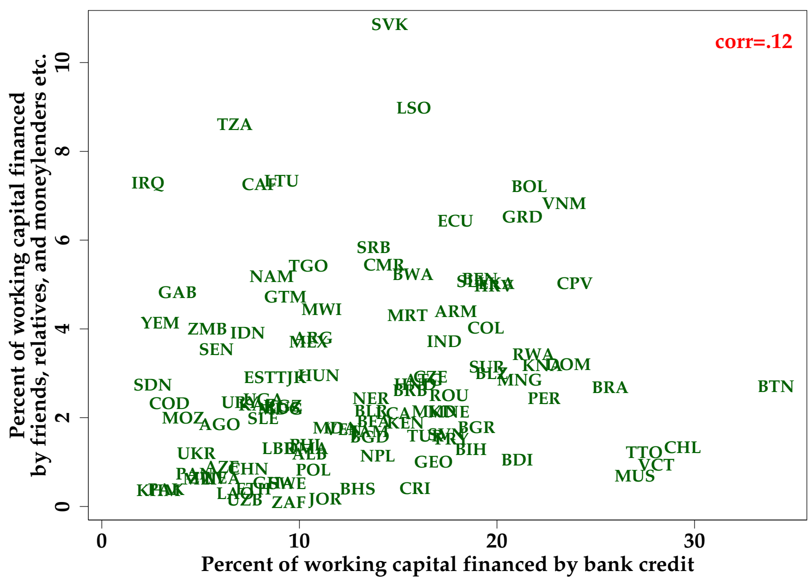
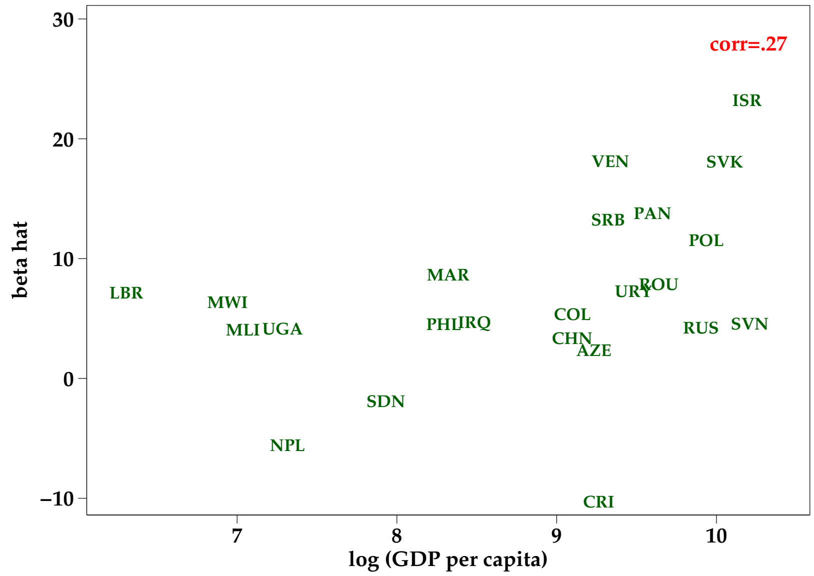


Figure 2.1: Informal financing and formal financing: Aggregate pattern

Notes: This figure shows the correlation between the percent of working capital that is financed by bank credit (x axis) and percent of working capital that is financed by trade credit (y axis) for 109 countries. Each point in the figure represents one country.

Panel A: Trade credit



Panel B: Other informal financing

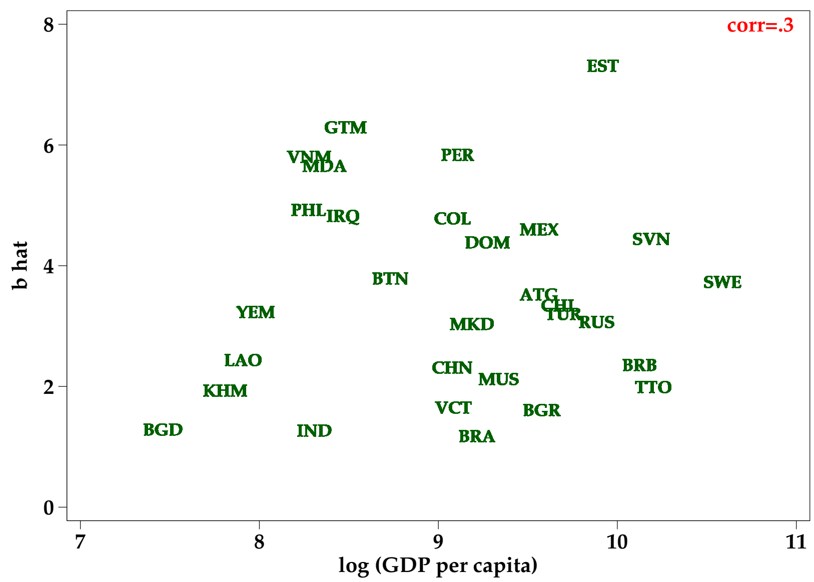


Figure 2.2: The availability of informal financing to financially constrained firms increases with economic development

Notes: This figure shows the correlation between the percent of working capital that is financed by bank credit (x axis) and percent of working capital that is financed by trade credit (y axis) for 109 countries. Each point in the figure represents one country.

Table 2.1: Trade credit and firm size: U.S. versus China

	(1)	(2)	(3)	(4)	(5)	(6)
25th to 50th percentile	6.827***	-8.051***	14.88***	3.267***	2.316***	0.951***
	(0.217)	(0.365)	(0.386)	(0.0478)	(0.0432)	(0.0493)
50th to 75th percentile	7.391***	-15.66***	23.05***	4.585***	3.799***	0.786***
	(0.237)	(0.399)	(0.422)	(0.0484)	(0.0438)	(0.0499)
75th to 100th percentile	6.525***	-17.85***	24.37***	5.086***	4.917***	0.169***
	(0.244)	(0.411)	(0.434)	(0.0508)	(0.0460)	(0.0524)
Dependent variable	AR/S	AP/S	Net AR/S	AR/S	AP/S	Net AR/S
Country	U.S.	U.S.	U.S.	China	China	China
N	34840	34840	34840	705312	705312	705312
AR2	0.287	0.177	0.238	0.113	0.0774	0.0245

Notes: The dependent variable for the regressions are the ratio of accounts receivable to sales in column (1) and (4), the ratio of accounts payable to sales in column (2) and (5), and the ratio of net accounts receivable to sales in column (3) and (6). Column (1)-(3) use data for Chinese firms and column (4)-(6) use data for the U.S. firms. All regressions include a set of sector times year fixed effects and a set of dummies of firm types.

Table 2.2: Summary of calibration

Parameter		Value	Target/Source	Data	Model
α	capital share in the production function	1/3	capital share of 1/3	–	–
ρ	elasticity of substitution between intermediate goods	3/4	Zetlin-Jones and Shourideh (2016)	–	–
χ	intermediate goods share	0.43	Zetlin-Jones and Shourideh (2016)	–	–
π	Poisson death rate	0.1	Buera et al. (2011)	–	–
N	measure of workers	9	share of entrepreneur	10%	10%
ψ	disutility from working	2.1	hours	0.3	0.3
δ	capital depreciation rate	0.025	annual 10% depreciation rate	10%	10%
β	discount rate	0.99	annual 4% risk-free interest rate	4%	4%
		United States			
μ	Pareto tail	4.1	top 10 percentile employment share	69%	67%
γ_1	collateral value of wealth	0.51	total debt to asset ratio	65%	63%
γ_2	collateral value of accounts receivable	0.85	trade credit to total asset	22%	24%
		China			
μ	Pareto tail	3.9	top 10 percentile employment share	57%	55%
γ_1	collateral value of wealth	0.42	total debt to asset ratio	57%	52%
γ_2	collateral value of accounts receivable	0.51	trade credit to total asset	9%	11%

Notes: For the U.S. data: The total debt to asset ratio is computed for the manufacturing sector using total asset and networth data from Table 14b of the corporate return statistics collection for the year 2006. The trade credit to total asset ratio is computed for the manufacturing sector using accounts receivable, accounts payable, and total asset from Table 6 of the corporate return statistics collection for the year 2006. For the China data: All the data moments are generated using the Chinese Industrial Survey for the year 2006.

Chapter 3

Economic Reforms and Industrial Policy in a Panel of Chinese Cities

3.1 Introduction

The process of economic reforms launched in 1978, and gradually extended until current days, has catapulted China into a stellar growth trajectory that has proven resilient. Because a variety of new policies and institutions were introduced simultaneously, even today it is difficult to pinpoint which of them were crucial. This paper aims at contributing to a better understanding of the policy roots of China's success by focusing on a major component of its industrial policy. It also provides new evidence in the debate about the effect of place-based policies.

We exploit the variation across cities and years in the establishment of different types of Special Economic Zones (SEZ) to estimate the effects of SEZ on economic development. SEZ are a salient component of the reform process for a variety of reasons. First, they have been a centerpiece of the gradualist Chinese development strategy based on the learning-through-experimentation principle. Second, they have fostered an uneven development across geographic areas and sectors. Last but not least important, their effects are easier to measure than those of other reforms, as they took the form of well-defined changes in the legal status staggered across different Chinese cities. The first SEZ were introduced as experiments in market allocation in geographically restricted areas along the coast. SEZ enjoyed special rules applying to labor markets, foreign direct investments, firms' ownership, and export controls. Another important difference from the rest of the country is that local political leaders were granted substantial autonomy and could shape key aspects of industrial policy. After the success of the early experiments, SEZ were extended first to other cities along the coast and then, starting in the early 1990s, to inland regions. The establishment of new zones has continued until today. For instance, in September 2013 the government of Li Keqiang has launched the Shanghai Pilot Free Trade Zone, which grants the Pudong area full liberalization of foreign trade and partial capital market liberalization.

We use a panel of 276 cities over the period 1988-2010.[1] Our econometric strategy is a difference-in-difference estimation controlling for time-invariant heterogeneity at the city level. We also control for province-specific shocks by using province×time fixed effects. We first regress (the logarithm of) GDP or GDP per capita on a reform indicator that switches on (i.e., takes the unit value) in the year after a city has received SEZ status, controlling for city characteristics such as land area. In our baseline specification, the introduction of a SEZ is associated with a permanent increase in the city's GDP level of about 12%. The effect on GDP per capita is about 9%. The result is robust to controlling for local government spending. To account for gradual effects of the reform, we also consider more flexible specifications where the effect of the reform is allowed to vary, both parametrically and non-parametrically, as a function of the time elapsed since the start of the treatment. We find an increasing cumulative effect of the policy treatment that flattens out after about

[1] More precisely, we use data on prefecture-level cities, which are administrative units below provinces and above counties. See Section 3.3 for details.

ten years; the long-term effect of a SEZ is estimated to be a differential increase of about 20% in the GDP level. We also study the channels through which GDP and GDP per capita increased as cities were granted SEZ status. SEZ attract larger populations, more investments in physical and human capital, and experience stronger increases in total factor productivity (TFP).

A common objection to place-based industrial policy is that it may induce a concentration of economic activity in some areas by drawing resources away from other locations. We find no evidence of such beggar-thy-neighbor effects on GDP. To detect potential cross-city spillovers, we investigate how the performance of cities varies with their distance from SEZ in other cities. The identifying assumption is that the spillover intensity decays with the distance from the SEZ. Distance is measured in three alternative ways: geodesic distance, driving time on the current roads network, and the computed shortest path through the physical geography. In all specifications, SEZ appear to generate positive (often highly significant) spillovers on nearby cities. We also compute measures of *exposure* to other zones by creating a sum over GDP in other cities with a SEZ, weighted by the inverse of distance to those cities. We again find some evidence of positive spillovers, especially strong in inland provinces, albeit often imprecisely estimated. We then investigate whether SEZ lead to a reallocation from areas that are further away from the zone to areas in the proximity of the zone. We consider various rings of up to 400 kilometer around the zone. Spillovers typically decline with distance. Interestingly, we find no negative spillovers even at these medium distances. These results are inconsistent with the hypothesis that the effect of SEZ is driven exclusively by direct transfers and political connections of the cities involved. This could not explain why cities *close* to SEZ benefited from the policy.

Our analysis is subject to two *caveats.* First, the assignment of cities to treatment and control groups may not be random. The Chinese government might have selected cities based on some prior knowledge that the conditions for industrial development might be especially favorable (picking winners), or to the opposite, in order to curb regional inequality. The narrative suggests that a picking-the-winner strategy may have been especially important in the first stage of the reforms, when all SEZ were chosen along the coast and close to potential trading partners and investors such as Hong Kong and Taiwan. Ideally, one would like to have instruments to isolate exogenous sources of variation in the reform treatment, but finding valid instruments is difficult in practice. We mitigate the concern with endogeneity through three complementary strategies. First, we restrict the sample to cities located in inland provinces where the selection of the zones was largely based on a rigid administrative criterion, i.e., being a provincial capital. Second, we augment the regressions with indicators for the immediate pre-reform years to capture differential trends. Third, we control for flexible differential trends depending on the initial conditions of the different cities. This is potentially important, since the cities hosting SEZ are on average more densely populated and more developed than those that did not host SEZ. The results are reassuring: the effect of SEZ is robust in the restricted sample, differentials before the actual establishment of the zone are insignificant, and allowing for differential trends based on the initial development or population density does not significantly affect the coefficients of interest.

The second *caveat* concerns data quality. One might worry that local statistics may be manipulated strategically by local officers in order to create the impression that an SEZ was successful so as to attract government support. In addition, while city-level nominal GDP data are available, city-level price deflators are more problematic (and only available for fewer cities/years). In our main specification, we use only nominal variables. The inclusion of city fixed effects removes any bias arising from time-invariant price level differences. Inflation differences across provinces are absorbed by the interaction between time and province fixed effects. Yet, this leaves open the possibility that different cities within the same province may experience different inflation rates. This would be a problem for our strategy if the SEZ status triggers systematically higher inflation rates, as in this case part of our estimated effect would be due to inflation. To address this concern, we first document that, in the more restricted sample for which we have data on prices at the city level, treated cities do not appear to have experienced higher inflation than did cities without SEZ. Next, we complement our analysis with alternative proxies of GDP that do not depend on prices: light intensity measured by satellites and electricity consumption. The results confirm the existence of robust significant effects of SEZ.

3.1.1 Related Literature

Our paper contributes to the large international literature studying the effect of place-based policy, comprehensively reviewed by the recent papers of Glaeser and Gottlieb (2008), Kline and Moretti (2014a), and Neumark and Simpson (2015). In developed countries, place-based policies often target the development of lagging regions. China's SEZ incorporate both efficiency and equity motivations together with the additional target of experimenting with market reforms. On the efficiency side, SEZ pursued the reduction of pre-existing distortions and the exploitation of agglomeration effects. On the equity side, the expansion since 1992 toward inland cities promoted the development of poorer Chinese regions.[2]

In line with the results of our paper, the literature finds positive effects of place-based policies in a number of instances. Criscuolo et al. (2012) use firm level data to study an investment subsidy program in the U.K. and find positive effects on employment, investment, and net entry. However, contrary to our study, they find no effect on TFP. Busso et al. (2013) compare locations selected for special treatment, such as tax-credits and subsidies for disadvantaged neighborhoods, with similar locations that were rejected or treated in a second round. They conclude that the policy had significant positive effects on employment and wages, while the efficiency costs were relatively small. Kline and Moretti (2014b) study

[2] Akinci and Crittle (2008) provide a cross-country comparison specifically focusing on different types of special economic zones and their role for development.

the long-run effects of place-based policies by focusing on a subsidy program in the U.S. that supported lagging regions. They find positive direct effects on productivity. Martin et al. (2011a) in contrast do not find positive effects from subsidies to Local Productive Systems in France.

Some papers try to assess, as we do, whether place-based policies generate spillovers – either positive or negative – to non-treated areas. The evidence is mixed. Criscuolo et al. (2012) aggregate their observations to larger geographical units that incorporate neighboring non-treated areas. They find that the positive treatment effect is not reduced by this aggregation, suggesting that there were no negative spillovers through reallocation from non-treated to treated firms within the same area. This is similar to our finding that SEZ had a positive effect on the prefecture area around the urban core. Furthermore, we also find some evidence of positive cross-city spillovers. Neumark and Kolko (2010) find insignificant employment spillovers of California's enterprise zones and Martin et al. (2011a) obtain a similar result for France. One economic rationale for place-based policy is to foster local agglomeration forces. Kline and Moretti (2014b) find no aggregate gains through agglomeration forces, because local gains are offset by losses elsewhere. Greenstone et al. (2010) estimate the effect of large plant openings on incumbent firms' TFP. They find that these agglomeration spillovers are positive but vary substantially across different cases. Briant et al. (2015) and Devereux et al. (2007) also find evidence of heterogeneous effects of place-based policies.

We are not the first to study the effects of China's SEZ. Most of the earlier studies, arguably due to data constraints, rely on comparisons of the cross-sectional variation in economic performance rather than on a difference-in-difference methodology. jin Wei (1993) uses city-level data for a sample of coastal cities where special policies were introduced in 1984, and documents that cities hosting SEZ have a significantly higher average growth rate during the early reform period, while other types of preferential policies do not produce the same effects.[3] Since his sample ends in 1990, when only a small subset of the cities had been granted the status of SEZ, his identification relies on the cross-sectional comparison between early reformers – a small and arguably selected group – and cities that were never granted the SEZ status at the time of his study. Wei's pioneer study is extended by Démurger et al. (2002) and Jones et al. (2003), who also document differences in growth rates between treated and non-treated cities. Different from these articles, our study exploits the staggered establishment of SEZ across cities. This allows us to estimate the treatment effect controlling for time-invariant heterogeneity (city fixed effects) and time-varying province-level shocks.

A recent study by Wang (2013) also uses a panel of Chinese cities and finds, using a difference-in-difference approach similar to ours, positive effects of SEZ on foreign direct investments (FDI), exports, and the output of foreign enterprises. The effects on other outcome variables (which do not comprise GDP) are smaller and less robust. Our findings are complementary to Wang (2013) insofar as we focus on GDP and GDP per capita, a comprehensive measure for the development of the local economy, while her study focuses on intermediate targets of the policy. An important difference for our analysis is that we distinguish between state-level and province-level SEZ (see below for a detailed motivation for this choice). Without drawing such a distinction, the introduction of SEZ would yield no statistically significant effect on GDP in our sample. Other studies focus on different economic outcomes. For example, Cheng and Kwan (2000) show that provinces hosting SEZ attract significantly more FDI than do other provinces. Head and Ries (1996) analyze the location decision of international firms in Chinese cities and find that SEZ have a positive effect that is amplified by agglomeration economies.

A number of studies look at firm-level data. Schminke and Biesebroeck (2013) estimate the effect of being located inside SEZ on firms' productivity and export behavior. They find

[3] jin Wei (1993) uses two samples: the first has 434 cities but only a limited time variation from 1988-1990. The second sample includes fewer cities (74) and covers the period 1980-1990.

that firms in SEZ export more, have higher output per worker and higher capital intensity, but no higher TFP once selection is controlled for. Their control group consists of firms outside of the SEZ in the same industry and in the same broadly defined regions (west, central and coastal). Our finding of positive effects of SEZ on TFP hinges on a comparison of the average performance of firms before and after the onset of a SEZ. Lu et al. (2015) compare firms that are located inside of SEZ with firms across the zone boundary and find positive effects. Similarly, Zheng et al. (2015) study local spillover effects of SEZ in eight Chinese cities using firm-level data for the period 1998–2007. They find positive spillover effects of SEZ on productivity and consumption in the area surrounding the SEZ. This result is consistent with our finding that there are positive effects of SEZ on the periphery around the urban core. Finally, Brooks et al. (2015) study the role of collusion in industrial clusters and find that collusion is particularly strong in SEZ.

Our study also relates more generally to a large literature on liberalization and industrial policy, including specific applications to the Chinese reform process.[4] Rodrik (2006) argues that government policies creating distortions in favor of more advanced industries played an important role in the success of Chinese reforms. Dewatripont and Roland (1995) and Rodrik (2004) argue that, through experimentation, the state can generate information about the potential of different sectors. Brandt and Zhu (2010) find that rising TFP in the private sector was an important driver of China's growth. Our findings are broadly consistent with these views. Finally, our study has some similarity in both the methodology and motivation with Aghion et al. studying the effect of industrial policy (the demise of the License Raj) in India. Similar to our study, they exploit the fact that the reforms were staggered across time and sectors. However, different from our study, they emphasize the interaction between the reform and state-level characteristics of the labor market. Moreover, they study an episode of pure liberalization (delicensing), while China's industrial policy also entails proactive policy

[4] See Perkins (1988), Naughton (2007), Rawski and Brandt (2008), and Xu (2011).

elements (tax credits, subsidies, etc.).

The rest of the paper is structured as follows. Section 3.2 provides the historical and institutional background of the Chinese industrial policy. Section 3.3 describes the data sources and the sample. Section 3.4 discusses the empirical strategy and the main results. Section 3.5 decomposes the effects of the SEZ into factor accumulation and total factor productivity. Section 3.6 discusses the spillover effects of the policy. Section 3.7 performs a variety of robustness checks. Section 3.8 concludes. The Online Appendix contains additional tables, figures and details on the data.

3.2 Institutional Features of SEZ

Since its establishment in 1949, the People's Republic of China relied on rigid economic planning. The two decades preceding Mao's death in 1976 were characterized by a volatile economic performance and by an intense social turmoil.[5] The reformist political leadership that won the battle for Mao's succession in 1978, led by Deng Xiaoping, faced the desperate need for measures to restore social cohesion and revitalize the economy. There were, however, no existing blueprints showing how to proceed. Learning-through-experimentation then became the guiding principle of economic reforms. As Deng put it: "one has to grope for stepping-stones as he crosses the river." The first policy breakthrough happened in rural areas, where the Household Responsibility System entitled farmers, after fulfilling their procurement quota, to the rest of their agricultural output. However, the leadership soon realized that reforms had to be extended to urban China, and that industrialization necessitated opening up China to foreign investments.

The idea of SEZ was *per se* no Chinese innovation. China's SEZ inherited some essential characteristics of the Export Processing Zones (EPZ), which had already been established

[5] For more detailed analyses of the economic growth of China before and after the start of economic reforms, see Cheremukhin et al. (2015), Song et al. (2011), and Storesletten and Zilibotti (2014).

in over 80 countries by 1980 (Naughton, 2007; Vogel, 2011). Like EPZ, SEZ were designed to circumvent the complex rules of import and export. China's SEZ were special in the sense that they also bore the responsibility of policy innovation and experimentation. They were the laboratories for the market economy (Vogel, 2011). The local officials of the zones were implicitly encouraged to be innovative in designing economic policies and institutions. Successful innovations were retained and extended to later waves of development zones (man Yeung et al., 2009).

3.2.1 The Timeline of SEZ

In the year 1980, four cities in the provinces of Fujian and Guangdong, Shenzhen, Zhuhai, Shantou, and Xiamen, were granted the SEZ status.[6] The success of the experiment was remarkable: between 1980 and 1984 Shenzhen grew at an annual rate of 54%, and in 1984 the four SEZ alone attracted 26% of China's total FDI. In addition, the zones had developed a set of well-functioning markets for labor, land, capital, transportation, and technology (Zeng, 2010).

The establishment of SEZ met the resistance of the conservative fraction of the Communist Party's (CCP) central committee that viewed renting China's land to foreign companies and allowing them to exploit China's cheap labor as unacceptable. However, the success of the experiment strengthened the reformist fraction in the CCP and softened the conservative opposition. In 1984, 14 coastal cities were granted the right to build Economic and Technological Development Zones (ETDZ). The ETDZ shared most of the policies and privileges granted earlier to the initial four SEZ.

During January and February of 1992, Deng made a celebrated tour to southern China,

[6]The SEZ status implied tax deductions, special tariffs for import and export, and exemptions from the regulations on foreign exchange and land use. Foreign firms that resided inside of the SEZ first enjoyed two years of tax holiday, then three years of a low tax rate of 7.5%, and after the initial five years a tax rate of 15%. Outside of the zones, the tax rate for foreign firms was 33% and for state-owned firms 55% (see jin Wei, 1993).

including stops at the SEZ of Shenzhen and Zhuhai, to mark the end of a period of political instability and to restate the commitment of the CCP to the reform process. Shortly afterwards, a new SEZ called Pudong New Area was established in Shanghai. In May, the CCP announced the plan to grant the five inland cities along the Yangtze River, nine border cities, and all thirty of the provincial capital the same privileges as the SEZ (Fewsmith, 2001). Following the instruction, several ETDZ and High-tech Industry Development Zones (HIDZ) were approved during 1992-1993 and 2000-2002, all located in inland provinces.

In the first decade of the XXIst Century, the introduction of SEZ spread quickly across China. By 2005, the system of state-level development zones comprised 54 ETDZ, 53 HIDZ, 15 Bonded Zones (BZ) and 60 Export Processing Zones (EPZ).[7] In the year 2005, the 54 ETDZ accounted for 4.49% of the national GDP and for 14.93% of national export (of Commerce, 2006). Establishing a development zone became a common strategy for the local government to attract FDI and foster local economic growth. Through shuffling local officials across different regions, the governments diffused the knowledge and experiences accumulated in the early zones to help develop new SEZ (Xu, 2011). Figure 3.8 shows that by 2010 SEZ had been established throughout the country.

3.2.2 Different Types of the Special Economic Zones

To summarize the discussion above, there exist five types of state-level SEZ: Comprehensive SEZ (CSEZ, a label we coin to distinguish the early zones from the general notion of SEZ), ETDZ, HIDZ, BZ, EPZ, and in addition Border Economic Cooperation Zones (BECZ). They all share preferential treatment in terms of tax deduction, custom duty deduction, reduced land-use price, flexibility in signing labor contract and financing. However, they are administered by different authorities: the CSEZ, ETDZ and HIDZ are directed by the State Council (the HIDZ being co-directed by the Ministry of Science and Technology); BZ and

[7]See section 3.2.2 for details on the difference between the zones.

EPZ are directed by customs; BECZ were directed by the State Council until 2007, and are now under the control of the Ministry of Commerce.

In addition, the zones differ in their stated mission. The goal of the CSEZ and of the ETDZ is to attract FDI and to boost export activity. They are also explicitly encouraged to design and experiment with new institutions and policies. The goal of HIDZ is to foster domestic high-tech industries. The BZ are free-trade zones located in coastal port cities or border cities where import and export can be expedited at a higher speed. The function of EPZ is to import raw materials from abroad, process them, and export the final goods without entering the real territory of China. Many of the EPZ are established within pre-existing ETDZ and HIDZ. The BECZ intend to take advantage of the location of the border cities to foster trade with other countries.

Aside from *de jure* changes, the central government is likely to have supported SEZ by assigning capable local leaders and providing administrative support. Because of data limitations (in particular, we have no data for transfers from the central government to cities), in the baseline regressions we simply regard any such complementary measure as part of the treatment. However, in the robustness analysis of Section 3.7.2 we attempt to separate the effects of government spending and road infrastructure, for which we have data.

Together with state-level SEZ, China saw the proliferation of a variety of development zones under the authority of provinces.[8] There are some important differences between state-level and province-level SEZ. The state council explicitly requests that "*the policies given to the province-level development zones should not be comparable to those given to the state-level ones,* " in order to prevent excessive competition between the zones and the waste of land resources (of Taxation, 2004).[9] The political autonomy of the province-level

[8]Online Appendix Table A1 lists the number of state-level and province-level development zones and their average share of industrial output in three coastal provinces hosting a large share of SEZ. The data are from WEFore (2010) for the year 2009. All three provinces have a larger number of province-level than of state-level zones (a ratio of 7:1). However, the state-level zones account for a far larger share of industrial output.

[9]Such competition is also a concern in other countries. See for example Ossa (2015) for a general equi-

zones is also much more limited. Finally, many province-level zones target specific industries whose selection depends on the capture of local interests. Overall, province-level SEZ are a patchwork of different policies rather than a coherent policy instrument. This causes a severe measurement error problem. In our analysis below, we find that province-level zones have an insignificant effect on economic development.

3.3 Data

The main data source is the National Bureau of Statistics of China (NBS), that publishes the *China City Statistical Yearbooks* including GDP, electricity consumption, population, education, investment, foreign direct investment, government spending, government income, and land area. In addition, we use the light intensity data from weather satellites as a proxy for GDP. More detailed information about the data sources is provided in the appendix.

The main unit of analysis is a *prefecture-level city*, an administrative division ranking below a province and above a county in China's administrative structure. A prefecture-level city comprises a core urban area and a surrounding periphery that may include rural areas, other smaller cities, towns and villages. The NBS reports separate statistics for the core and the periphery of each prefecture-level city. In our baseline we use the larger definition of the prefecture-level city that includes the core and the periphery, but we have also done the analysis when restricting to the urban core. One advantage of considering the larger area as opposed to focusing on the urban core is that border changes are less frequent for the former.[10] Henceforth, unless an ambiguity arises, we refer to a prefecture-level city as a *city*.

The sample period is 1988-2010. At instances, city borders were changed by administrative reforms. While this was less frequent for the borders for the broad definition of a

librium analysis of subsidy competition in the U.S.

[10] Although we can track border changes (of the core and the periphery) over time by controlling for land area as reported in the statistical yearbooks, they are less of a concern when considering the larger area.

city (including the periphery) than for the urban cores, it is important to take the changes of borders into account. This information on changes in the land area is reported in the *China City Statistical Yearbooks.* We focus on 276 cities, excluding from our analysis the four cities in which CSEZ were introduced before 1988, as well as Hainan, where the entire province received the status of SEZ in 1988. We drop two city-year observations where a county-level city was promoted to a prefecture-level city which implied that it incorporated the periphery, but the associated border change occurs with a one-year delay in some variables. Furthermore, we exclude Tibet, where we have data for only one city, and the province-level municipalities, including Beijing, Chongqing, Shanghai, and Tianjin, because our set of province-time fixed effects would absorb all variation in GDP.

3.3.1 Main Variables

We start by listing the outcome variables that are used as the dependent variables in the regression analysis. Unless stated otherwise, the variables from the yearbooks are for the city area that combines the urban core and the periphery.

- $\log GDP$ and $\log (GDP/L)$ are, respectively, the (logarithm of) GDP and of GDP per capita at the city level. Population measures are constructed based on the census and the statistical yearbooks.

- $\log Electricity$ is the electricity consumption and is available for the same set of cities as GDP but only for their urban cores. It measures the use of electricity for household consumption and industrial production.

- $\log Light$ is the average light intensity. In the data provided by the National Geographical Data Center, light intensity is measured on approximately each square km (pixel) on a discrete scale from 0-63. We use digital maps from 2010 to aggregate the light intensity of the pixels to administrative units. We use the maps of urban cores, which corresponds to the level at which the electricity data are available.[11] When using $\log Light$, we must restrict the sample to the period 1992-2010 for which the light data are available.

- $\log(K/L)$ is the physical capital per capita. The physical capital stock is constructed by applying the perpetual inventory method to the investment data for the period 1988-2010, assuming an annual depreciation rate of 8%. For some cities, we collect the investment data from the *New China in 60 Years Provincial Statistical Collection* for the earlier period 1978-1987. The province-specific investment deflator is from the *New China in 60 Years Statistical Collection*.

- $\log L$ is the population size (a proxy for the labor force). Population data is available from the census and, annually, from the *China City Statistical Yearbooks*. The census data is more comprehensive (in particular, it includes *non-hukou* population), but it is only available every ten years. Therefore, we construct the observations between two editions of the census based on the growth rate from the *China City Statistical Yearbooks*.[12]

- $\log h$ is the average human capital, constructed using average educational attainment of the population over the age of 6. The educational attainment data comes from the *China Population Census*.

- *log TFP* is total factor productivity, constructed with an estimated production function and physical capital, human capital, and population of each city.

Next, we discuss the construction of the explanatory variables. The main variables of interest are indicators for the presence of SEZ. For each of the different types of SEZ we

[11]Note that, unlike for GDP, we can hold the area of the urban core constant when measuring light intensity based on the 2010 maps. The concerns due to border changes therefore do not apply here.

[12]A detailed description of this process can be found in Online Appendix B.2.

construct a dummy, I_Reform_{it} (where i denotes the city, and t denotes the year), which switches on (i.e., takes the unit value) in the year *after* the establishment of a zone and retains the unit value in all following years. Formally, we define the reform indicator based on the establishment of a zone as

$$I\_Reform_{it} = \begin{cases} 1 & \text{if } ReformYear_i < t \\ 0 & \text{otherwise.} \end{cases},$$

where $ReformYear_i$ is the year in which the zone was established in city i and t is the current year. In our baseline specification we will focus on the *first* state-level zone that was established in city i. Note that for cities that never host a zone $I\_Reform_{it} = 0$ for all t. We also construct separate dummies for each lag from the reform year, as discussed in more detail in the empirical sections.

3.3.2 Control Variables

We use two main control variables from the *China City Statistical Yearbooks*. First, the geographic size of the city, to which we refer as land area measured in square kilometers. This variable is available annually and varies over time, reflecting changes in the legal city boundaries. Second, in some specifications, we control for population size.

In order to assess spillover effects that may depend on distance or transport costs between cities, we calculate a variety of different measures related to distance or driving time between cities. First, we calculate the geodesic distance in kilometers between all pairs of cities in our sample. Second, we calculate the driving time on the current road infrastructure using Google maps. Third, we use topographical features such as the slope of the terrain and use shortest path algorithms to construct transport cost measures.

3.3.3 Summary Statistics

Table 3.1 shows the summary statistics of the dependent variables and of the main control variables.[13] We have over 5000 observations for GDP from an unbalanced panel of 276 cities from 1988 to 2010. Our policy variable, the establishment of SEZ, is illustrated in Figure 3.8. This figure shows the time evolution of the shares of cities hosting the different types of zones in the balanced sample. The figure also shows the share of cities that have any state-level zone. The two most important types of zones are HIDZ and ETDZ with shares reaching 26% and 22% in 2010, respectively. Two types of zones existed before the start of our sample: the CSEZ, established in 1980, and a few early ETDZ, established in 1984. ETDZ and HIDZ are altogether the most frequent zone types. We also consider Export Processing Zones (EPZ) and other less frequent types of zones (e.g., BZ and BECZ), introduced in cities that already hosted either ETDZ or HIDZ.

We report the mean values of city characteristics separately for reformers and non-reformers in Table 3.2. We distinguish three broad categories, with breakdown by coastal and inland cities: cities that received the first SEZ before 1988, cities that received the first SEZ in 1988 or later, and cities that never hosted a SEZ in the sample. As the table shows, cities hosting a SEZ were larger in terms of population and richer in terms of GDP per capita. They also tended to have more universities relative to other cities. Government spending over GDP was instead higher in non-reformers. Our empirical specification controls for city fixed effects filtering out the effect of time-invariant heterogeneity. However, one might be concerned about pre-treatment differences having differential effects on growth or on the effectiveness of the policy treatment. Our strategies to address these challenges are explained in detail in Sections 3.4 and 3.7.

[13]For the dependent variable we show the statistics for real GDP based on provincial price deflators, but in the empirical analysis we use nominal GDP because the province-year fixed effects absorb price changes at the province level. See also the next section.

3.3.4 Price Deflators

The *China City Statistical Yearbooks* report nominal GDP for the period 1988-2010. Since Chinese price data are regarded as somewhat unreliable (see, for example, Young, 2003), we opt to use nominal data. Time-invariant differences in price levels across cities and time-varying inflation differences across provinces are absorbed, respectively, by city and province×time fixed effects. This approach would be problematic if inflation rates differed significantly across cities within each province. The main concern is that the SEZ treatment might systematically increase local inflation. We check if there are differences in inflation rates between treated and non-treated cities in those years for which real GDP data are available from the NBS. More precisely, we compute an implicit city-level deflator using the data on nominal and real GDP, and compare it between cities with and without a SEZ. We find that, within each province, cities with a SEZ did not have higher inflation.[14] As an alternative strategy that avoids relying on prices altogether, we use electricity consumption (in GWh) and light intensity as proxies for the level of economic activity.

3.4 Empirical Strategy and Results

3.4.1 Motivation

In this section, we discuss the econometric strategy and the main results. We use a difference-in-difference estimator exploiting the variation in economic policy across a panel of cities and over 23 years following the establishment of SEZ.

Although the focus of the paper is empirical, and we do not present a formal model, it is useful to motivate and interpret our analysis in the light of spatial equilibrium models

[14] The real GDP index of cities is available from the NBS for the period 1996-2010. For this period, cities with a SEZ had an average yearly inflation rate of 1.8%, while cities without a SEZ had an average of 2.3%. The difference is not statistically significant. We also run a panel regression of prices on the reform indicator and control for city and province-year fixed effects. The estimate is -0.008 and insignificant.

such as Greenstone et al. (2010) and Redding (2012). Greenstone et al. (2010) construct a model economy comprising many locations where firms produce using labor, capital, and land. Firms are perfectly mobile, and their profits are equalized in equilibrium. Workers are only partially mobile due to idiosyncratic preferences for certain locations, such that utility is equalized across location but wages are not. Local productivity spillovers imply that total factor productivity depends on the pool of labor that works and lives in a given location. Their framework can be applied to our environment by interpreting the onset of a SEZ as a policy shock that reduces firms' costs in the treated locations. This induces firms to relocate or expand their activity within the SEZ. Agglomeration externalities and technology transfer from foreign firms (or from more productive Chinese firms that relocate to the SEZ) may increase total factor productivity. The (possibly gradual) inflow of firms is limited by congestion externalities, as new firms bid up the prices for local factors such as land and labor. The higher costs offset the initial increase in profits, providing an equilibrating mechanism. The dynamic adjustment eventually comes to a halt when firms' profits and workers' utility are equalized across locations. In the new spatial equilibrium, total factor productivity, the stock of capital and labor, and ultimately the GDP are permanently higher in the SEZ.

Guided by this model, we investigate, first, if the onset of a SEZ triggers an increase in GDP and GDP per capita relative to other cities. In a world of perfect capital and labor mobility, we should expect a permanent increase in TFP, factor accumulation, and GDP while labor productivity (GDP per capita) should eventually be equalized across locations. To the opposite, in a world with no labor mobility GDP per capita would also be permanently higher in treated cities. In China, labor is not immobile but migration is subject to frictions such as the *hukou* system. Thus, we test whether the onset of a SEZ affects both GDP (and its components) *and* GDP per capita. We defer the analysis of the effect of the policy on factor accumulation and TFP to Section 3.5 below.

3.4.2 Baseline Specification

In this section, we run regressions whose dependent variables are the logarithms of either GDP or GDP per capita. When we run regressions for GDP, we do not control for changes in labor since these are part of the outcome variable. When we run regressions using GDP per capita as the dependent variable, we do control for population to account for decreasing returns to labor.[15]

The main explanatory variables are reform indicators switching on in the year after part of a city's territory is granted the status of a state-level SEZ.[16] All regressions control for city fixed effects and province-time interaction dummies. Standard errors are clustered at the city level. More formally, we run regressions of the form

$$y_{ipt} = \phi_i + \gamma_{pt} + \alpha\, I\_Reform_{it} + X_{it}\beta + \varepsilon_{it}, \tag{3.1}$$

where y_{ipt} is the logarithm of nominal GDP or nominal GDP per capita, ϕ_i is a city fixed effect, γ_{tp} is a province-time fixed effect, and I_Reform_{it} is an indicator switching on, for each city, in the year after a state-level SEZ is established. X_{it} is a vector of time-varying control variables and ε_{it} is a normal error term. City fixed effects absorb time-invariant heterogeneity in city characteristics like initial development or geographical location. Thus, the effects of reforms are identified across city-time within each province. Province-time fixed effects control for time-varying province-specific shocks that can play a confounding role. In particular, they absorb cross-province inflation differentials.

The econometric specification in (3.1) restricts the treatment effect to a shift in the after-reform GDP (GDP per capita) level path; namely, in reformed cities the GDP per capita *level* (or trend) is allowed to shift whenever the reform indicator switches on. Below, we

[15] In an earlier version, we also show results for GDP per capita if one does not control for population. The results are qualitatively similar to those shown in Table 3.3.

[16] We also construct a similar separate dummy variable for province-level reforms. Note that including the year of the reform in the dummy does not alter the baseline results significantly.

explore more flexible econometric specifications allowing for trend breaks and distributed lags.

The estimated coefficients are shown in Table 3.3. In column (1), we include no additional control variable except for the city fixed effects and province-time dummies. The coefficient of the state-level SEZ is positive and highly significant. Becoming the host of a SEZ increases the average GDP of the treated city by about 15.6% in post-reform years. In contrast, the effect of province-level reforms is small and insignificant. In column (2) we include the logarithm of the city's land area as a control. This variable controls for changes in city borders, which are relatively frequent in China and would change GDP mechanically.[17] Increases in land area appear to be positively associated with aggregate GDP. The estimated effect of the SEZ decreases to about 11.6% but remains highly significant.

In column (3) we show the results of regressions where GDP per capita is the outcome variable and where we control for the logarithm of population.[18] The estimated effect of the reform is 9.27%.[19] This suggests that part of the increase in GDP is due to labor reallocation (something we document more explicitly in Section 3.5 below). In columns (4)–(6) we repeat the analysis for the sub-sample of inland provinces.[20] This sub-sample involves a less discretionary selection of individual cities. To mitigate concerns about the selection further, we exclude cities that were granted the status of SEZ in spite of not being provincial capitals. Thus, the restricted inland sample only contains provincial capitals (treatment group) and cities that were never granted the SEZ status (control group). Columns (4)–(6)

[17]In the robustness section 3.7.5 we discuss the results when instead of controlling for land area we allow for structural breaks in the city fixed effects when there are border changes.

[18]The coefficient on population size is negative, suggesting that an increase in the population size due, e.g., to immigration, has a negative effect on labor productivity.

[19]This specification in column (3) is equivalent to controlling for the logarithm of population density and land area. In Section 3.7.5 we investigate the role of density in more detail, and we also discuss the concern that population and population density could be endogenous. The results are robust to using lagged variables and alternative ways of controlling for border changes.

[20]In the sub-sample of inland cities, 44 cities were granted SEZ status. Of these, 18 were provincial capitals.

in Table 3.3 show that the results are robust to restricting the sample to inland provinces.[21] The coefficient of interest is positive and significant, and even larger than in the full sample.

3.4.3 Pre-Reform Trends

A concern with the results of Table 3.3 is that cities hosting SEZ might already have been on a higher-growth trajectory – or might even have been selected precisely because of their promise of success. The focus on inland capitals alleviates such concerns. However, the year in which capitals were assigned to the treatment group may not be random. Moreover, provincial capitals may be a special group *per se*.

We address this point through a variety of strategies. First, we investigate whether the performance of treated cities was different from that of other cities in the same province in the years shortly pre-dating the reform. Table 3.4 is the analogue of Table 3.3, reporting the results of regressions where we add four pre-reform indicators taking on the unit value, respectively, in the year of reform and one, two and three years before the reform.[22] If cities were granted the status of SEZ due to their promising pre-reform trends, these coefficients ought to be positive and significant. In contrast, we find the estimated coefficient of the pre-reform dummies to be mostly negative and insignificant. In column (5) the indicators for the reform year and for one year before the reform are marginally significant but negative. The treatment effect in the full sample continues to be positive and significant (columns (1)-(3)). In the inland sample the estimate is positive and significant in column (4), and it is positive

[21] Arguably, inland capitals are *per se* a special group. Since the selection of treated cities was based on an administrative criterion (rather than on unknown, possibly heterogeneous criteria), we can better control for features making capital cities different from the control group. In Section 3.4.4 we allow cities to have year fixed effects that depend on such city characteristics, and we find that the results are similar.

[22] We also explored longer lags. The lags for five years prior to the reform are never significant in the full sample. In the inland sample some of the earlier lags become significantly negative but only in the specification in column (4) that does not control for changes in land area. Note that lags longer than three years are identified out of a significantly smaller set of reforming cities (since many cities were granted the SEZ status in the early 1990's, and our sample starts in 1988). For instance, in the full (inland) sample the first three lags are identified out of 75 (31) cities, while the fifth lag would only be identified out of 31 (18) cities.

but insignificant in columns (5)-(6). In summary, the results of Table 3.4 are reassuring and suggest that treated cities did not show higher economic performance already before the reform.[23]

Second, we consider a more flexible specification allowing treated cities to have different time trends from the non-reformers. This addresses the potential worry that in our baseline specification the positive effect of SEZ might arise spuriously due to the omission of pre-existing trends. The new specification allows the GDP of cities that are hosting a SEZ to have a linear time trend that differs from the control group's trend already before the reform. In some specifications, we even allow this trend to undergo a structural break at the time when the reform indicator switches on. More formally, we consider the following specification:

$$\begin{aligned} y_{ipt} &= \phi_i + \gamma_{tp} + \alpha_1 I\_Reform_{it} + \alpha_2 \left[(t - 1987) \times I\_Reformer_i\right] \\ &+ \alpha_3 \left[\max\left\{0, (t - ReformYear_i) \times I\_Reform_{it}\right\}\right] + X_{it}\beta + \varepsilon_{it}, \end{aligned} \tag{3.2}$$

where, as above, I_Reform_{it} is an indicator switching on in the first year after the reform. Moreover,

- $I\_Reformer_i$ is a dummy identifying cities that were reformed at any time. $t \geq 1988$ denotes the year of the observation. Therefore, α_2 captures the steepness of a linear trend specific to reformers, i.e., how many percentage points the growth rate differs between reformers and non-reformers.

- $ReformYear_i$ is the year in which the first SEZ was introduced in city i (if a city never became a SEZ, then we let $ReformYear_i = 0$). The interaction $[(t - ReformYear_i) \times I\_Reform_{it}]$

[23]Note also that the earliest zones (for example the CSEZ) introduced before 1989, likely the most selected group, are either excluded or exhibit no time-variation in the policy indicators in our sample period. Thus, they play no role in the identification of the treatment effect.

allows a differential trend (i.e., a trend break) starting as of the introduction of the first SEZ. The coefficient α_3 measures the steepness of such a trend break.

- α_1 captures a level shift as in the baseline specification of Equation (3.1).

The results for the full and restricted (inland) samples are shown in Table 3.5, columns (1)–(4) and (5)–(8), respectively. The results are robust to using GDP per capita as the dependent variable and controlling for population. Columns (1) and (5) of Table 3.5 reproduce columns (2) and (5) of Table 3.3 for comparison. In the regressions of columns (2) and (6) we add a linear trend specific to reformers. The estimated coefficient $\hat{\alpha}_2$ (time trend of reformers (state-level)) is statistically significant in both the full and the restricted sample. Interestingly, the coefficient $\hat{\alpha}_1$ continues to be highly significant in the full sample, although much of the effect is now absorbed by the trend. However, it becomes insignificant in the restricted sample. The trend in columns (2) and (6) does not distinguish pre- and post-reform periods. Thus, in columns (3) and (7) we allow a structural break in the trend of reformed cities, by including $\max\{0, (t - ReformYear_i) \times I\_Reform_{it}\}$ in the regression. Interestingly, the estimated coefficient $\hat{\alpha}_1$ remains almost unchanged in the full sample and increases in the restricted sample. Moreover, the estimated coefficient of the pre-reform trend, $\hat{\alpha}_2$, decreases and becomes insignificant in both samples. The post-reform trend, $\hat{\alpha}_3$, is positive but insignificant in the full sample and positive and significant in the inland sample. Altogether, the statistical specification studied so far suggests that the baseline model with a GDP level shift performs better than one allowing for a trend break implying a permanent GDP divergence between the treatment and control groups.

The specification of columns (2)–(3) and (6)–(7) – allowing for permanently diverging paths – may be too extreme. We consider, then, an alternative specification allowing SEZ to have a non-linear effect of the SEZ relative to the pre-reform trend. To avoid an over-parameterization, we omit the level shift, and we estimate the following alternative econometric specification:[24]

$$
\begin{aligned}
y_{ipt} &= \phi_i + \gamma_{tp} + \alpha_2 \left[(t-1987) \times I\_Reformer_i\right] \qquad (3.3) \\
&+ \alpha_3 \left[\max\left\{0, (t - ReformYear_i) \times I\_Reform_{it}\right\}\right] \\
&\quad +\alpha_4 \left[\max\left\{0, (t - ReformYear_i) \times I\_Reform_{it}\right\}\right]^2 + X_{it}\beta + \varepsilon_{it}.
\end{aligned}
$$

The regression results from this specification are provided in columns (4) and (8). In both cases, we find that $\hat{\alpha}_3 > 0$ and $\hat{\alpha}_4 < 0$, implying that the SEZ are associated with an acceleration of growth in the immediate post-reform years, but that the acceleration dies off in subsequent years. The coefficients are both individually and jointly statistically significant in the full sample, while the square term is negative but insignificant in the inland sample.[25] In summary, this specification suggests that the effect of SEZ is a significant gradual increase in the GDP level, rather than a permanent increase in growth (i.e., a *linear* trend break of the treated cities after the reforms).[26]

3.4.4 Heterogeneous City Characteristics

In the previous section, we allow different trends between treated cities and non-reformers. An alternative strategy is to control for differential trends associated with the initial characteristics of cities. This is an important check, since Table 3.2 shows that cities hosting a SEZ were more populated and had a higher initial development measured by GDP per capita than other cities. One might worry that the heterogeneity in these initial characteristics might be the actual driver of economic performance over time, and that our baseline specification

[24] It would be possible to also include the term $\alpha_1 I\_Reform_{it}$ to this specification. However, it is very difficult to identify separately all the effects in such a highly parameterized model. Therefore, we omit this term, and regard the current specification as a non-nested alternative to Equation (3.2).

[25] $\hat{\alpha}_3$ and $\hat{\alpha}_4$ are jointly significant at 5% in the full sample and at 10% in the inland sample.

[26] Clearly, the quadratic model is not a correct specification itself, since it would imply a negative long-run effect of SEZ. Given the short sample, the data only capture the increasing part of the quadratic relation. See Section 3.4.5 for a more general specification.

might spuriously attribute those effects to the establishment of SEZ.

To address this concern, we interact each year dummy with the log difference between certain city characteristics and their respective median values in the year they were first measured. We do this for the city characteristics GDP per capita, population, population density, and number of universities and include the interactions together in the regressions.[27] This allows the flexible growth path to depend on cities' initial characteristics and assumes this interaction to be log-linear.[28]

The results are shown in Panel A of Table 3.6. The coefficients of interest are similar to Table 3.3 in the full sample and larger in the inland sample. In both cases, they remain highly significant. In Panel B, we provide the results from an alternative specification where the year dummies are interacted with a set of indicators for whether a city has a GDP per capita, population, population density, number of universities, respectively, that is above the median in the year in which that variable first appears in the yearbooks for that city. The four sets of interaction effects are then included together in the regressions. The results of Panel A and Panel B are relatively similar.[29]

In summary, the effects of SEZ are robust to controlling in a flexible way for differential trends associated with heterogeneous initial conditions.

[27] We calculate for each year the median of the variables across all cities. When we restrict the regression sample to inland provinces, then we calculate the difference relative to the median in this restricted sample. Since our sample is an unbalanced panel, the year in which cities appear in our sample can vary. However, the results are robust to restricting the sample to a balanced panel of 172 cities. The sample size is reduced here because of missing data for the number of universities, but the results are also robust to excluding the interactions with the initial number of universities and thus using the larger sample.

[28] Consider for example a city i that enters our sample in 1988, and whose GDP p.c. is reported in the yearbook. The interaction effect between a year dummy (for example 1995) and the log difference between GDP p.c. in 1988 and the median in that year then is

$$D_{1995} \times [\log(GDP_{1988,i}) - \log(GDP_{1988,median})] \,.$$

The estimate on this interaction would capture how much higher GDP p.c. is in 1995 for city i when the log difference changes by some percentage. Therefore, cities with median initial characteristics have a time path as given by the main year dummies, and the interactions with initial characteristics allow differential relative paths for cities above or below the median.

[29] The difference between Panel A and B in the sample size is due to cities with zero universities in the first year, such that the log difference in Panel A is not defined.

3.4.5 Event Study

In this section, we perform a non-parametric analysis of the effects of the reform with the aid of a model that imposes no functional form restrictions on post- (and pre-) reform effects. All effects are captured by separate lag- or lead-specific dummies. More formally, we run the following regression:

$$y_{ipt} = \phi_i + \gamma_{t,p} + \sum_{n=-J_B}^{J_F} \alpha_n I_{it}^n\{(t - Reformyear_i) = n\} + X_{it}\beta + \varepsilon_{it}, \tag{3.4}$$

where positive values of $t - Reformyear_i$ measure how many years before year t city i became the host of a SEZ. Negative values measure how many years ahead of t city i will be reformed. Note that this specification allows us to identify some of the lagged effects out of reforms that took place before 1988. For instance, a city that hosted its first SEZ in 1984 will have variation for all leads ranging from 4 to 26 years. In our baseline specification, instead, such a city would display no within variation, and the reform indicator would be collinear with the city fixed effect. In our sample, the maximum number of post-reform leads, J_F, is 26, corresponding to cities which hosted their first SEZ in 1984. We also construct these indicators for the year of reform and the three years prior to the reform (i.e. $J_B = 3$), so we can test whether reforming cities already had a significantly different performance prior to the establishment of the first zone.[30] The omitted categories (for which all indicators are zero) are never-reforming cities and reformed cities more than three years before the reform. The controls include the logarithm of land area and the usual set of fixed effects.

The results for GDP are displayed in Figure 3.8. The results for GDP per capita are shown in Figure 3.8 and will be discussed in Section 3.5 where we decompose the effect. The graphs show the lead and lagged effects of the treatment n years after the reform (for

[30]For the same reasons described in the discussion of Table 3.4, we do not include more pre-reform indicators. When we include also indicators for four and five years prior to the reform, these indicators are marginally significant, but identified by only 27 observations.

instance, n=10 measures the effect ten years after the introduction of a SEZ). The upper graph in Figure 3.8 shows the effect on GDP in the full sample. This specification confirms the results of the previous section. In particular, there is a break in the GDP path a year after the reform, followed by a temporarily higher growth rate that levels off after about ten years. The size of the effect is comparable to that in the previous section. There is only some marginal, statistically insignificant evidence of a higher GDP growth in the three years before the reform, indicating a possibility for some minor positive selection. Note that the standard errors increase nineteen years after the establishment of the zone (corresponding to the vertical line added to each figure). This is due to a significant drop in the number of observations, since many cities were reformed in 1991 and 1992.[31]

We estimate the same regression for the restricted sample of inland provinces (excluding cities which had a reform but are not provincial capitals), see the lower graph in Figure 3.8. The qualitative pattern and the point estimates are similar, although the estimation is less precise.[32]

3.4.6 Different Types of SEZ

In this section, we attempt to disentangle the effects of the different types of state-level SEZ. To this aim, we create separate post-reform indicators for each of the three most important (and most common) SEZ: ETDZ, HIDZ, and EPZ. In addition, we create a single dummy for other types of state-level SEZ. Appendix Table A2 has the same structure as Table 3.3 but replaces the indicator for *any* state-level zone with the four separate indicators for each type of state-level SEZ. ETDZ and HIDZ individually appear to have a significant effect on

[31] When the cities reformed in 1991 and 1992 reach the year 2010, the subsequent number of cities that identify the individual coefficients drops from 54 to 9. The vertical dashed line in the figure marks this drop.

[32] The reforms in the inland provinces started almost a decade later than in the coastal provinces. The post-reform effects are therefore estimated for a shorter period and based on fewer observations. In separate regressions not shown here, we find that if residuals are clustered at the province×years of reform (instead of city) level, the effects after nine years are mostly statistically significant and positive in the inland sample. Two of the pre-reform indicators are also significant but negative.

the level of GDP. In the full sample, the effects of ETDZ and – to some degree – also HIDZ are relatively similar to those of the first zone in Table 3.3. The point estimates on ETDZ and HIDZ in the inland sample are relatively similar to the full sample, but less precisely estimated. The effects of ETDZ and HIDZ in the inland sample tend to be lower than for the first zone reported in Table 3.3. EPZ are insignificant in both samples, while OtherTypes are mostly significant and have particularly large estimates in the inland sample.[33] Overall, the disaggregation highlights the relative importance of the ETDZ and HIDZ, which are the two largest and most comprehensive types of zones in our sample, as well as those most explicitly emphasizing technology.

Since the effects of any zone has been shown to build up gradually during about ten years and then level off, we investigate whether the same pattern holds true for the individual types of zones. Since the pre- and post-reform effects of different types of zones often overlap (treated cities often had multiple zones of different kinds), the approach in Section 3.4.5 is demanding. Nevertheless, the resulting picture is reasonably clear. Figure B.4, which can be found in the appendix, plots the coefficients of the different types of zones (estimated in the same regression) over the years since the reform. The first panel shows that the pattern for ETDZ looks remarkably similar to that of Figure 3.8 (first zone reformed). The second panel shows that HIDZ also display a concave pattern, although the effect appears to decline after lag 13. EPZ and OtherTypes show a more mixed picture (the two lower panels in Figure B.4).[34] The standard errors are large and the effects are estimated imprecisely. In summary, most development effects appear to stem from ETDZ and HIDZ.[35]

[33]It should be noted that the estimates on OtherTypes are based on few observations. 14 cities have a zone type other than ETDZ, HIDZ, or EPZ, but in 11 of these the zone this is in conjunction with an ETDZ or HIDZ.

[34]The stark drop in OtherTypes is identified by only one observation. EPZ were established after 2000 and often inside an existing zone. Furthermore, the EPZ may have gained importance after the WTO accession in 2001, which could explain their upward trend (though insignificant).

[35]Recall that some zone types like ETDZ and HIDZ may target cities with certain characteristics such as having universities. This could raise concerns about selection and we address this in a similar fashion as in Section 3.4.4. When we include the interactions of year fixed effects with initial characteristics (GDP

3.5 Decomposing the Effects of SEZ

In this section, we investigate the channels through which the SEZ promote economic development by decomposing the effects on physical capital per capita, average human capital, population, and TFP. To construct TFP, we assume the aggregate technology to be described by a Cobb-Douglas production function in physical capital and efficiency units of labor (raw labor × average human capital). We use the local population size as a proxy for raw labor and the average years of schooling to measure human capital (see appendix for details). The aggregate production function is estimated using an OLS estimator from the panel of observations of output, capital, population, and average educational attainment of the population, including city fixed effects and province-time fixed effects.[36] We then use the estimated parameters to compute (the logarithm of) TFP as the residual component.[37]

In panel A of Table 3.7, we display the results of baseline difference-in-difference regressions analogous to those in Table 3.3, where, respectively, GDP per capita, capital-labor ratio, population, and human capital are used as the dependent variables. In both the full sample (column (1)) and the inland sample (column (5)), becoming the host of a state-level

p.c., population, density, and number of universities), then the estimates on these zone types are relatively similar. The two exceptions are that in column (5) ETDZ becomes significant while HIDZ loses significance and that in column (6) ETDZ becomes significant.

[36]The estimation of production functions can suffer from simultaneity bias, because profit-maximizing firms choose inputs after knowing the realization of productivity shocks, and selection bias, related to exit and survival of firms. In the firm-level literature, it is common to use the correction proposed by Olley and Pakes (1996). For example, Brandt et al. (2012) find that the TFP growth of Chinese firms is underestimated when the endogeneity bias is uncontrolled for. Martin et al. (2011) estimate a Cobb-Douglas production function using firm level data. They find that after controlling for simultaneity bias, TFP is still very close to the one obtained using a simple OLS estimation. Since we use aggregate data, we follow the traditional approach and use an OLS estimator. This is related to the growth accounting literature including Hall and Jones (1999) and Caselli (2005). See also Hsieh and Moretti (2015) for an application to city-level data.

[37]More formally, we let

$$\log TFP_{it} = logY_{it} - \hat{\alpha} \times logK_{it} - \hat{\beta} \times log\left(h_{it}L_{it}\right) - \hat{\gamma}_{pt} - \hat{\chi}_i, \tag{3.5}$$

where Y_{it} is GDP, K_{it} is physical capital stock, h_{it} is human capital, and L_{it} is population; $\hat{\alpha}$ and $\hat{\beta}$ are the estimated coefficients of the Cobb-Douglas production function; $\hat{\gamma}_{pt}$ is the estimated province-year dummy, and $\hat{\chi}_i$ is the estimated city fixed effect capturing, respectively, province-level common trends and city-level time-invariant components of productivity. TFP_{it} measures the city×time variation in TFP.

SEZ is associated with a significant and positive increase in the GDP per capita. This result is identical to that of columns (3) and (6) in Table 3.3. Columns (2) and (6) show that the establishment of SEZ is associated with an increase in the capital-labor ratio by 13.1% and 33.9% for the full and inland sample, respectively, both effects being highly significant. Columns (3)-(4) and (7)-(8) suggest that the SEZ have no significant effect on population and the human capital measure in the *China City Statistical Yearbook* data. However, both effects are positive and significant when one restricts the analysis to more precise population data from the decennial census, as is shown in panel B in columns (2)-(3) and (5)-(6). Population increases by 9.7% and 11.1% after the establishment of a SEZ in the full and the inland sample, respectively,[38] while the average years of schooling of the population above 6 increase by 0.18 years in the full sample and 0.36 years in the inland sample.[39] The increase in human capital can be explained by either selective immigration (i.e., cities with a SEZ attract more educated immigrants) or by stronger incentives for locals to accumulate human capital. Despite the higher population, GDP per capita increases after the introduction of a SEZ because GDP increases more than population. This is shown for the census sample in columns (1) and (3) in panel B.

The estimated effect of SEZ on TFP is shown in panel C of Table 3.7. In the specification of columns (1) and (4), TFP is estimated without imposing any restriction on the parameters of a Cobb-Douglas production function. The unconstrained estimation of the production function yields output elasticities of capital and labor of 0.3 and 0.6, respectively. In columns (2) and (5), we impose constant returns to scale, obtaining elasticities of 0.35 and 0.65. Since there is some evidence that the labor share has been declining in China (see en Bai and Qian,

[38]The difference is likely due to non-hukou population which is captured in the census data but not in the yearbook data. Since in panel A we compute population for the years in between the census based on the growth rate in the yearbooks, the annual variation does not fully reflect non-hukou migrants and is subject to measurement error. See also robustness section 3.7.4, where we discuss the use of census data.

[39]Ideally, we would prefer to use the educational attainment of the working population (age 25-64). However, this is not available in the census. In Appendix Table A4 we break down the result by different educational levels. The most salient effect is the increase in the share of college graduates.

2010), in columns (3) and (6) we estimate the production function separately for pre- and post-1995 subperiods.[40] In all specifications of the full sample, the SEZ have a positive and significant effect on TFP (columns (1)-(3)). As shown in columns (4)-(6), the estimated effects on TFP in the inland sample are positive but insignificant, except in column (6), where TFP is estimated separately for pre- and post-1995 sub-samples.

Figure 3.8 shows the effect of SEZ on GDP per capita, capital-labor ratio, human capital, and TFP, respectively, as an event study. The effects on GDP per capita and on the capital-labor ratio are concave over time. Both paths feature a break one year after the reform. In particular, the effects of SEZ on GDP per capita and on the capital-labor ratio become statistically significant around seven years after the reform. There appears to be some concavity in the effect on TFP as well, although less clearly and not statistically significant in the individual years. Human capital appears to be higher in cities with SEZ (in this case, some effects are already detected prior to the reform).

3.6 Spillovers

In this section we study whether the effects of SEZ spill over to other locations. SEZ could have negative spillovers on other cities if the policy attracts investments and workers away of other areas (beggar-thy-neighbor effect). Positive spillovers could accrue from the diffusion of knowledge and an increase in market access. Investigating the spillover effects of the SEZ is important to assess the overall effect of SEZ on economic development. The existing literature on spillover effects on non-treated locations is ambiguous (see Neumark and Simpson (2015) for an overview of the evidence).

In order to estimate the spillovers of SEZ on other cities, we make the identifying assumption that spillovers – either positive or negative – are decreasing in the distance from

[40] The result is similar when we split the sample into a pre- and post-2000 period.

SEZ.[41] This assumption is motivated by the evidence documented in the previous literature that distance plays a crucial role for spillovers. For example, Jaffe et al. (1993) and Keller (2002) find that spillover of knowledge significantly decreases with geographic distance.[42] We consider various alternative measures of geographical distance (as described below) in order to test the robustness of our results. To estimate the spillovers based on these distance measures, we first assume that the spillovers decay log linearly in distance from the closest SEZ. We then use a non-parametric approach based on various distance bands and more comprehensive measures of exposure to other cities' SEZ. As provincial borders may act as barriers, we also compare our results when restricting to spillovers within provincial borders. It is important to note that all of these variables are time-varying because they depend on the introduction of SEZ in other cities. Thus, identification hinges on this time variation. Note also that we always include the cities' own policy indicator for SEZ in the regression. This allows us to test whether the own reform effect changes when we allow for spillovers from other cities.

3.6.1 Measures of Distance and Transportation Costs Between Cities

Our first and simplest measure of distance between cities is the geodesic kilometer distance between all the city centers in our sample. This measure does not take into account geographical barriers between cities or transportation infrastructure. The second measure of distance is the driving time between cities derived from Google Maps.[43] The advantage of this mea-

[41] This is consistent, among others, with Rosenthal and Strange (2004). Geographic distance (or transportation costs) plays also a central role in the literature on trade and economic geography (Fujita et al., 1999). An alternative measure of distance is used by Bloom et al. (2013) who argue that cross-firm spillovers depend on the distance in technology and product markets. Neumark and Kolko (2010) also use the identification assumption that the effect of place-based policy on non-targeted areas differs in the distance to the treated areas.

[42] Jaffe et al. (1993) find that patent citations are highly spatially clustered, which implies that there is a distance decay in the knowledge diffusion. Keller (2002) finds that the benefit of technology spillover is halved with a distance of 1200 kilometers.

[43] We use the tool *traveltime3* in Stata that accesses the Google maps. Since only a limited number of queries can be submitted and there are more than 75'000 routes, we measured the distance of each bilateral

sure is that it captures how well cities are connected through road infrastructure, which is likely to be an important determinant of the interactions between them. The drawback is that it focuses on road transportation and that it is based on the current transportation network, which is potentially endogenous to the zone locations.

The third set of distance measures is based on the topography of the Chinese terrain. This has the important advantage of being entirely based on exogenous factors. We have detailed information on slope and land cover that allows us to construct a local measure of transportation costs on 10×10 kilometer cells throughout China. We then use a shortest-path algorithm in ArcGIS to find the shortest route between cities through this cost surface and we measure the total cost along this route.[44] Since we must make a number of assumptions for how to map slope into transportation costs, we investigate the robustness of the result to alternative ways to compute this measure. The first mapping of terrain slope to driving speeds is based on a scale that relates slope to driving speed in the US and has 10 different levels (AASHTO, 2001). The second mapping is based on a similar scale for China and has 7 different levels. In a further variation of this approach, we use a higher resolution for the transport cost cells (3 km instead of 10 km), and we exclude larger water bodies. All measures based on topography are normalized so that they have the same median as the driving times according to the Google maps. This is to facilitate comparisons between the different specifications.[45]

3.6.2 Results on Spillover Effects Across Cities

We use three complementary empirical strategies.

connection in only one direction and imposed symmetry.

[44]The tool in ArcGIS is *cost distance* and is an implementation of the Dijkstra algorithm. See for example Alder (2015) for a description of this method and the data.

[45]We assume that all distance measures have a linear relationship with effective transport costs. While this is only an approximation, it facilitates the comparison across the various distance measures.

Distance to closest SEZ Our first approach to estimate SEZ spillovers is based on the distance of each city from the closest city hosting a SEZ (excluding zones in the own city). This variable varies over time; the establishment of a new SEZ that is closer than the previous ones causes a reduction in this measure. Our regression equation is as follows:

$$y_{ipt} = \phi_i + \gamma_{pt} + \alpha I\_Reform_{it} + \lambda \ln(DistClosestZone_{it}) + X_{it}\beta + \varepsilon_{it},$$

where $DistClosestZone_{it}$ is the distance to the closest city that has a state-level SEZ in year t. The distance to the next SEZ is an *inverse* measure of the spillover intensity. Therefore, if spillovers were negative, we would expect $\lambda > 0$. On the contrary, the results in Panel A of Table 3.8 suggest that there are positive spillovers, since a longer distance to the next SEZ is associated with a lower GDP, controlling for land area and the usual fixed effects. The spillovers are especially large in the inland sample. Remarkably, the reform effect of the own SEZ, α, remains large and significant.

Indicator for SEZ within radius Our second approach is to include a binary indicator for having a zone within a given radius. We report the results for a 150 kilometer radius (or for the equivalent in driving time).[46] The regression equation is

$$y_{ipt} = \phi_i + \gamma_{pt} + \alpha I\_Reform_{it} + \mu I\_ReformRadius_{it} + X_{it}\beta + \varepsilon_{it},$$

where $I\_ReformRadius_{it}$ is the indicator for having another city with a SEZ within the specific radius. If spillovers were negative, we would expect $\mu < 0$. Instead, we typically find positive estimates of μ. Panel B of Table 3.8 reports the results for a radius in kilometers and minutes in columns (1) and (2), respectively. The remaining columns use the distance

[46]This is approximately the median distance to the next SEZ. The results are similar for a radius of 100. The coefficients vary more when we use a variety of different radii between 20 and 900 km, but we never find significant negative spillover effects. In robustness checks, we also computed the distance to the closest zone in the same province, and the results are qualitatively similar.

measures based on topography, which are normalized such that their median is equal to the median travel time in minutes. All estimates are positive, although the estimated coefficients are sometimes insignificant.

We also perform a similar analysis where, instead of one indicator for 150 kilometers, we simultaneously include multiple indicators for various rings (excluding the own zone): 0-50, 50-100, 100-200, and 200-400. These indicators take on the value 1 if there is at least one zone within the corresponding ring. The omitted group consists of cities for which the closest SEZ is more than 400 kilometers (or the corresponding alternative distance measures based on driving time or topography) away. The results are shown in Appendix Table A5. Most of the indicators have a positive coefficient, and in some cases they are statistically significant. We only observe negative coefficients for the geodesic distance beyond 50 kilometers, but the estimates are relatively small and insignificant. For all other distance measures, we find positive effects that tend to be larger and more significant for zones that are closer. This analysis suggests that the positive spillover effects of the zones on cities within a radius of up to 100 kilometers is not driven by reallocation from areas between 100 and 400 kilometers.

Exposure measure Our final spillover measure is inspired by Briant et al. (2015) and mimics the idea of a market access measure such as[47]

$$MA_{it} = \sum_{j \neq i} \frac{GDP_{j,t}}{dist(i,j)}. \tag{3.6}$$

[47] Such measures of market access or market potential appear in models of trade and economic geography, see for example Fujita et al. (1999).

We adjust this measure by summing only over cities that have a SEZ in that year.[48] Our measure of exposure to other cities with SEZ is therefore given by

$$B_{it} = \sum_{j \neq i} \frac{GDP_{j,t}}{dist(i,j)} \mathbb{I}\{I\_Reform_{j,t} = 1\},$$

where $\mathbb{I}\{I\_Reform_{j,t} = 1\}$ is an indicator function for cities that have a SEZ (are reformed) at time t. This measure allows us to capture the exposure by taking into account both the economic size of other cities with SEZ and the distance from them.

This exposure measure varies over time because of the introduction of SEZ in other cities, but also because of GDP growth in these cities. The latter channel implies the risk that this measure may confound the effect of other zones with growth in market access.[49] In order to control for growth in neighboring cities in general, we therefore also control for the logarithm of market access, which is measured across all cities in our sample as shown in Equation (3.6). The regression equation then becomes

$$y_{ipt} = \phi_i + \gamma_{pt} + \alpha I\_Reform_{it} + \xi \ln(B_{it}) + \eta \ln(MA_{it}) + X_{it}\beta + \varepsilon_{it}.$$

We would typically expect a positive coefficient on market access and, in the presence of negative spillover effects, a negative coefficient on the exposure to other SEZ, hence $\xi < 0$ and $\eta > 0$. The results are shown in Appendix Table A6. The coefficients on exposure are always positive but not significant. The measure of market access shows a negative estimate in the full sample and a positive estimate in the inland sample. However, it is generally insignificant except for column (1). The result is broadly consistent with the one from the

[48] Briant et al. (2015) weigh by population instead of GDP. The results are robust to using population.

[49] For example, if several cities in the close neighborhood experience GDP growth but only one of them has a SEZ, then this measure of exposure may partly capture the general increase in market access. Although we control for province-time interactions in all of our regressions and therefore absorb much of the regional growth trends, this measure gives higher weight to close neighbors and hence may capture spatial trends at the local level.

two previous specifications, and it indicates that there is no evidence of negative spillovers. The comparison to the market access measure in fact suggests that proximity to a reformed city is more beneficial than higher market access in general. Interestingly, the effect of the own zone remains large and significant in all specifications.

The analysis of potential spillover effects based on various distance measures and identification strategies suggests the existence of positive spillovers across cities. Although these effects are not always significant, the fact that we never find significant negative effects provides strong evidence against the presence of negative spillovers.

3.6.3 Spillovers Over Time

In Section 3.4.5, we observed that the effect of SEZ on the own city tends to flatten out over time (see Figure 3.8). One possible explanation for this pattern could be that the effect of the SEZ spills over to other cities as time goes by. In this section we investigate how the spillover effects evolve over time. The two upper graphs in Figure 3.8 show the estimates of a regression where the spans of the own SEZ are included in a regression together with the spans of the first zone that is established within 150 minutes driving time. The two lower graphs show the results from an analogous specification with a 150 kilometer radius. The point estimates on the spans for the neighboring zone generally become significant at the 5% level when a 150 minutes driving time radius is used, but not (or only marginally so) when a 150 kilometer radius is used. In both cases the patterns suggest that the spillover effects become stronger during the first ten years. The diffusion of positive spillovers could reduce the difference between treated and neighboring cities, which can potentially explain why the effect of the own zone flattens out over the years.

3.6.4 Decomposition of Spillover Effects

In this section, we decompose spillovers into investment, TFP, and population spillovers. Negative investment spillovers would indicate that the SEZ attract investments at the expense of neighboring cities. Positive investment spillovers would instead arise if firms choose to locate geographically close to their suppliers and customers. This would lead to higher investments in cities located near to growing SEZ. A similar argument applies to population and TFP. If innovative firms are attracted by the SEZ, this could yield a negative selection in nearby cities and lower TFP. Conversely, technological diffusion could induce positive spillovers. This could in turn trigger more investments in nearby cities.

Appendix Table A7 shows that the spillover effects on investment are insignificant in the full sample (columns (1)–(4)). The point estimates are small and – depending on the distance measure – either positive or negative. In the inland sample (columns (5)–(8)), the spillovers tend to be more significant and are always positive (since the effect of distance is negative). Appendix Table A8 shows that the spillovers in TFP are in all cases positive and mostly significant.[50] The coefficient on the indicator of a city's own SEZ remains stable. Appendix Table A9 shows that the effect of SEZ on population in nearby cities is positive, but usually not significant.[51]

One possible channel for productivity spillovers is foreign direct investment (see Gorodnichenko et al., 2014), an explicit target of SEZ. Appendix Table A10 shows that the onset of SEZ increases the FDI flows to the cities hosting SEZ, consistent with the results of Wang (2013). The spillover effect on neighboring cities is positive but insignificant. There is no evidence that SEZ have negative spillover effects on FDI in other cities.[52]

[50] Here TFP is constructed using the full-sample unrestricted production function estimation. The other two measures of TFP give similar results.

[51] When we restrict the sample to the years when we have better population data from the census, then the signs of the coefficients vary and they are never significant.

[52] Different from us, Wang (2013) finds some evidence of negative FDI spillovers in neighboring cities. A potential explanation for the difference in the results is that she does not distinguish between state-level and province-level zones and only considers the spillover effect of FDI on neighboring cities.

3.6.5 Spillover Effects in the Periphery of Cities

We have so far investigated cross-city spillover effects. Our data additionally allow us to explore the effect of SEZ on neighboring non-urban areas. Our baseline specification focuses on the entire area of cities, which include an urban core (where all state-level SEZ in our sample were established) and the periphery around the urban core. To investigate whether and how SEZ affect economic activity in the area surrounding the center, we re-run our baseline regressions of Section 3.4.2, using two distinct geographical definitions of GDP as the dependent variables. First, we use the logarithm of GDP of the urban core only as the dependent variable (see Appendix Table A11, Panel A).[53] Then, we use the logarithm of GDP of the periphery only, i.e. excluding the urban core (Appendix Table A11, Panel B). The effects for the urban core are comparable in magnitude to those obtained above for the combined area. Moreover, the results hold up when we consider only the periphery. In summary, there is no evidence that SEZ impoverished neighboring non-core city areas.[54]

3.7 Robustness

In this section we perform a number of robustness exercises.

3.7.1 Satellite Light as an Alternative Measure of GDP

Chinese price data are generally regarded as problematic, especially at the local level. Our empirical methodology has the advantage of not relying on any price deflator. Differences in

[53] The strategy of estimating the effects at different levels of aggregation in order to verify the presence of spillovers from the treated location to neighboring areas is also applied in Criscuolo et al. (2012) in their analysis of place-based policies in the UK.

[54] The positive effect may be due to firms active in SEZ setting up facilities in the periphery. To the extent to which firms do not benefit from special exemptions for the activities performed outside of the SEZ, we regard this as as a spillover. However, one might conjecture that firms located inside the SEZ can benefit from special treatment even if they locate some facilities in neighboring areas. We could not find any precise information in this regard.

price levels are filtered out by city fixed effects, whereas province $\times$ year fixed effects filter out cross-province inflation differentials. Yet, one might worry that within each province cities might experience different inflation rates. In particular, our estimated treatment effect would be biased upwards if the establishment of a SEZ causes systematically higher inflation. The existing price data do not suggest any such pattern. However, one might also worry that the local authorities over-report the *nominal* GDP in cities hosting SEZ, in order to meet the expectation of the central government regarding their success.

To address this issue, in this section we use light intensity measured by weather satellites as a proxy for GDP. A number of recent papers have argued that nighttime light intensity measured by weather satellites is a good proxy for GDP.[55] Most economic activities such as production, transport, and consumption produce light as a by-product. Therefore, light intensity is positively correlated with the intensity of local economic activity. We calculate the average light intensity within the geographical boundaries of the urban cores and use this as a proxy for economic activity. The light data has the advantage that it can be measured within the same administrative boundaries over time. We can use digital maps from 2010 to calculate the light statistics for all years. The change in administrative borders – which are relatively frequent for urban cores – are therefore not a concern.[56] A drawback of the light data is that it is only available from 1992 on.

In column (1) of Table 3.9 we re-run our baseline regression with the logarithm of the average light intensity as the dependent variable. The estimate suggests that SEZ have a

[55]Elvidge et al. (1997) are among the first to discuss the relationship between light and economic activity. See also Henderson et al. (2012) and Chen and Nordhaus (2011) and the literature cited there on the use of light to measure economic activity. Ma et al. (2012) and Halg (2012) discuss the use of light data for Chinese cities. See also the Online Appendix for further details on the data source.

[56]When there are no data constraints due to border changes, then the urban core is a reasonable unit of analysis, since the SEZ in our sample were located in the urban cores. The analysis using light data exploits this advantage, but we have also done the analysis for the larger definition of a city that includes the periphery, which is the unit that the baseline GDP results are based on. The effects of SEZ at that level are smaller and insignificant. We have no explanation for the difference in the result between urban core and the area that includes the periphery. It appears to be specific to the light data, since such large differences were not observed for other data.

positive and significant effect on economic activity as measured by light intensity. However, the point estimate of about 5% is lower than what we observed in the baseline regressions using GDP as the dependent variable. The point estimate for the inland sample is similar in magnitude, albeit statistically insignificant. The lower point estimate could be due to the sample period starting in 1992, because only one-third of the (first) SEZ were established after that year. Moreover, even for later reformers we lose annual observations that would be useful for a precise estimation of the within-city effect of the establishment of a SEZ.[57]

We also check the robustness of our results by using electricity consumption as a proxy of economic activity (see, Rawski, 2001). Data on electricity consumption by households and firms are reported in the same statistical yearbooks as GDP and are available only for the urban core. In column (3) of Table 3.9 we re-run our baseline regression using the logarithm of electricity consumption as the dependent variable. The result shows that the establishment of a SEZ is associated with a 15.7% increase in electricity consumption.[58]

3.7.2 Controlling for Local Government Spending and Road Infrastructure

One might conjecture that the establishment of a SEZ may be associated with additional transfers from the central or the provincial government. SEZ may also have triggered government investments in infrastructure. Although one might regard both transfers and investments in infrastructure as being part of the place-based policy, one may be interested in

[57] This loss of precision is confirmed by the observation that if we run the baseline regression of Section 3.4.2 with GDP as the dependent variable for the post-1992 period we obtain a positive (0.043) but statistically insignificant point estimate.

[58] However, we find no significant effect in the inland sample. We suspect that this is due to the poor quality of electricity data in this subsample, for which we have no explanation. We calculated the correlation between GDP and electricity separately in four sub-samples: inland reformers, inland non-reformers, coastal reformers and coastal non-reformers. The correlation is high and significant in all subsamples except for that of inland reformers, where the elasticity of GDP with respect to electricity is very low (0.02) and statistically insignificant. Interestingly, the correlation between GDP and satellite light intensity is instead consistent and significant across the four sub-samples, suggesting that the source of the problem is not the GDP statistics but rather the electricity data.

estimating the *net* effects after controlling for them.

While we have no information on transfers, we observe the area of finished roads in the urban cores in each year, which is an important component of infrastructure investments. Furthermore, we observe the overall *expenditures* of the local government for a subset of the years in our sample. Finally, we can also control for government *income* and hence the deficit of the local government. These measures can be used as a proxy for the contribution of public investments to GDP. The disadvantage of including these variables is twofold. First, we lose some observations. Second, causation could run in the opposite direction: government expenditure might have increased because the GDP expansion caused by the SEZ increased the tax revenue accruing to the local authorities. We have therefore also used one-year lags of government spending and income. The estimates reported below are for the contemporaneous years, but the results are relatively similar when using lags.

Table 3.10 shows that the reform effects are robust to the inclusion of controls for local road infrastructure (columns (1) and (4)) and government expenditure (columns (2) and (5)). The results are also robust to controlling for the log difference between government spending and government income – a proxy for the deficits of local governments (columns (3) and (6)). The effect of the reform remains positive and highly significant in both samples.

3.7.3 Earlier GDP Data

Our main analysis focuses on the period 1988-2010, for which the NBS provides a sample of cities that allows us to also track border changes over the years.[59] This approach entails the cost of losing variation in the reform variable, since some SEZ were established before 1988. We re-estimate our baseline specification for a subset of cities for which GDP is

[59]It is important to note here that the city size could vary over time, and there were changes in the administrative system. The yearbooks allow us to match the city names over the years and control for these border changes by including land area as an explanatory variable.

also available for earlier years.[60] In this case, we cannot control for changes in land area, government spending, and population as these data are missing for the earlier years. The reform effect estimated with this subsample is a 16.8% increase in the level of GDP in the full sample, and the estimated coefficient is highly significant. This estimate is similar to our baseline results reported in Table 3.3. In the inland sample, the estimate is 32.7%, which is higher than our baseline result.

3.7.4 Population from Census

In our analysis so far, we have combined population data from the census and from the City Statistical Yearbooks. Using the yearbook data allowed us to calculate the annual fluctuations for the years between the three censuses (1990, 2000, and 2010). To the best of our knowledge, the yearbook data cover only the registered population in the city, that is, people with "*hukou*." The existence of a large number of non-resident immigrant workers in the cities could potentially bias our estimation. To address this issue, we first check the City Statistical Yearbook data against the population census that in principle should record the entire resident population at the city level. We find that there is a gap between the two data sources. In particular, if the census is correct, then the population growth rate is overestimated by an annual 24 basis points in non-reforming cities, and underestimated by 35 basis points in reforming cities in the city statistics. The observation that the population is underestimated in the treatment group and overestimated in the control group is not surprising, as the treatment cities are likely to have attracted many *non-hukou* workers from the control group.

By relying on the census data in 1990, 2000, and 2010 and using the yearbook data only to infer the population growth rates for the years in between, we have already attempted to address this concern. To further test the robustness of our results, we repeat the baseline,

[60]Please see the Online Appendix B.2 for more detailed descriptions of the data source.

regressions of Table 3.3 and restrict our sample to the three census years, using only population census data. Table A3 simply replicates the results in Table 3.3 for the restricted sample. The estimates are somewhat larger compared to our baseline and they remain significant. This is the case for all specifications and in both samples. It is important to note that by restricting the sample to only three years, we lose some time variation in the treatment effect. However, the baseline results do appear to be robust to using the resident population data from the census.[61]

3.7.5 Population and Population Density

Our results suggest that SEZ have a positive effect on both GDP and population, but the effect on GDP is larger than the effect on population. This is consistent with the increase in GDP per capita shown in columns (3) and (6) of Tables 3.3 and A3. These specifications for GDP per capita also control for population in order to account for agglomeration effects, but this raises the concern that population is endogenous. For instance, an increase in productivity and wages can induce immigration. The typical instruments proposed in the literature are time-invariant, and it is difficult to find time-varying instruments that fit in our difference-in-difference framework.[62] To mitigate the concern, we adopt two strategies. First, we show that the results are robust to a specification where we use the lagged values instead of current values of population (and its density). The results in the Appendix Table A12 show that the reform effects on GDP per capita and TFP are robust when we include

[61]The same holds true for the capital-labor ratio and for TFP (result not shown).

[62]The literature finds a relatively small endogeneity bias in the coefficient for population density. For example, Combes and Gobillon (2015) document that the endogeneity bias on the elasticity of density is between 10% and 20%, sometimes the bias is close to zero and even negative. Combes et al. (2010) provide a detailed comparison of different identification strategies. In particular, they note how difficult it is to find valid time-varying instruments (most attempts in the existing literature have resulted in weak instrumentation). An example for time-invariant instruments is given in Ciccone and Hall (1996), who study the effect of density by using historical population as an instrument. Combes et al. (2008), Duranton and Puga (2004), and Glaeser and Gottlieb (2009) provide a more general discussion of spatial concentration and productivity. An example of an analysis of agglomeration forces in China is Combes et al. (2013), who use Chinese household survey data.

population or population density together with land area with one period lag. The results are also robust to using the lagged population (and its density) as an instrument for current population (and its density).[63]

Second, we explore other specifications where we do not control for population (so, the results are gross of agglomeration effects). Column (2) in the baseline regression of Table 3.3 already shows that the results are robust to a specification that includes changes in city areas but not in population. However, one might worry that a specification where the effect of border changes is log linear in land area is overly restrictive. Changes in land areas reflect changes in borders, and the effects are likely to be heterogeneous across cities. To address this concern, we propose a specification that controls for border changes in a more flexible way by allowing each city's fixed effect to undergo a structural break whenever the land area of a city changes in our data - indicating a change in city borders. In other words, we replace city fixed effects with city-land area fixed effects.[64] For instance, if a border change brings a poorer periphery into the city, this effect is absorbed by the new, and more flexible fixed effects. The results are shown in Appendix Table A13. Columns (1) – (2) for GDP show that the point estimates are similar to Table 3.3 both in the full sample and the inland sample. Columns (3) – (4) show that the estimates are also positive for TFP. In the inland sample the estimate is also positive but lower than in column (4) of Table 3.7 (Panel C) and marginally insignificant.

3.7.6 Heterogeneity in the Treatment Effect

The literature on place-based policies suggests that the effectiveness of such policies may vary with location characteristics such as city size, density, or market access (see Briant et al., 2015; Devereux et al., 2007). In this section, we include in our baseline specification

[63] See for example Martin et al. (2011) for a panel analysis where lagged variables are used as instruments.

[64] An average city then has roughly three different fixed effects over the years because of changes in the land area variable.

interaction terms of our reform indicators with indicators for whether initial population, population density, GDP p.c., and market access were above the median value of reformers.[65] Since our sample is unbalanced, we take as the initial year for each city the year in which the corresponding variable is reported the first time in our sample. The results are shown in Appendix Table A14. In the full sample there is evidence for interaction effects with population and population density, but the main effect remains positive and significant. In the inland sample the interaction effects are stronger and the main effect is reduced when including the interaction with population and population density. Interestingly, the interaction effect with initial GDP per capita is negative, suggesting that SEZ in relatively less developed capital cities were particularly effective in inland provinces.

3.7.7 Placebo Analysis

Our estimation exploits both the time and spatial variation in the establishment of SEZ. Since the establishment of the SEZ is staggered, but clustered in a few years, there could be a concern about the extent to which the exact timing of the reform matters for the identification of the reform effect. Furthermore, we would like to rule out that our reform indicators pick up shocks unrelated to SEZ that could be present also in other cities. In order to deal with these concerns, we run three placebo exercises based on the specification in column (3) of Table 3.3, but assign reform years randomly.

In a first exercise, we assign the actual number of new zone establishments in each year to a random selection of cities. The resulting placebo distribution is the same as the true distribution over time, but SEZ are assigned artificially to random cities. We repeat this exercise 1000 times. We find that in no case are the absolute t values and the R-squared of the placebo regressions larger than those of the true reform.[66] This suggests that the spatial

[65] We compare the characteristics to reformer cities because for some variables all reformers are above the median, such that the interaction effect would be collinear with the main effect.

[66] The mean estimate of the placebo reform is 0.0003, and is never significant and higher than the one of

distribution of SEZ indeed drives our result.

In a second more demanding placebo test, we assign the random reforms only to reformers, again holding the distribution of reforms across years constant. However, the timing of the treatment is scrambled across cities. This allows us to assess the extent to which the time dimension of the reform matters, because we are only randomizing the year of the reform but not the treated city. We find that the absolute t-values are higher when using the year of the true reform than in the placebo regressions in all but 5% of the cases.[67] This indicates that the actual year in which the SEZ were implemented is critical for our results, and supports our identification strategy based on within-city variation.

Finally, we use the random assignment of reforms from above and include the true reform year and the placebo reform year in the same regression.[68] While the estimate for the true reform is always significant at 5%, the placebo reforms are significant in only 33% of the cases. Overall, these placebo exercises strengthen our confidence in the empirical strategy used. Both the spatial and the temporal variation of the SEZ appear to be important for the results.

3.7.8 Effects by Year of Reform

In Section 3.4.5 we allowed the effect of SEZ to depend on the number of years since the reform. However, the reform effect may also depend on the year in which a city received the SEZ. Late SEZ could for instance imply a less intense treatment, since the Chinese economy was altogether more liberalized than in earlier periods. In this section we investigate whether there are significant differences between early and late SEZ. To this aim, we first construct separate policy indicators for early and late SEZ introductions. We use 1992 as the threshold

the true reform.

[67] The mean estimate of this placebo reform is 0.0911. In only 13% of the draws does the placebo specification yield significant coefficients that are higher than the actual coefficients.

[68] The assignment of random reform years among reformers implies that a placebo reform year is likely to coincide with the true reform year. This is the case in 36% of the draws.

year after which we label SEZ as "late" reforms. Note that 1992 is the median reform year in our sample. The policy indicators for early and late reforms are then used together in our baseline regression to replace the single indicator for reform. Table A15 columns (1) and (4) show the results with the two indicators. Early and late reformers both show positive point estimates, but the effect of the early reformers is larger and more precisely estimated. This is true for both the full sample and the inland sample, although the difference is smaller in the inland sample. This could suggest that earlier reforms had a larger impact, but it could also be driven by the fact that for earlier reformers the effect had more time to accumulate over the years since the reform. This seems particularly relevant in light of the patterns we observe for the flexible reform effects in Section 3.4.5, which suggest that the reform effect accumulates over about ten years. We then test for this pattern separately for early and late reformers in order to investigate how reform effects may differ between early and late reformers.

We parameterize the pattern we found in Section 3.4.5 by allowing a linear increase during ten years and then a constant effect during all following years.[69] Furthermore, we impose that the linear trend after ten years is equal to the constant effect after ten years. More precisely, we impose the restriction that $\beta_{11-30} = 10 * \beta_{trend}$, where β_{11-30} is the effect after ten years and β_{trend} is the coefficient on the linear time trend. We then run this regression separately for early and late reformers. The results are shown in Table A15 columns (2)–(3) and columns (5)–(6) for the full sample and the inland sample, respectively. We see that the pattern in the full sample is relatively similar for early reformers (column (2)) and late reformers (column (3)), but the coefficients are less precisely estimated for the late reformers. In the inland sample (columns (5)–(6)) we observe that the coefficients on the trend are similar but again less precisely estimated for late reformers. We also note that in

[69] The fully flexible specification with separate indicators for each year is very demanding, and yields imprecise estimates.

the inland sample we cannot identify the effect after ten years for the late reformers because there were no SEZ established in the inland sample within the period 1993–2010. The lower precision for late reformers is not surprising, since there are fewer observations to identify the effect until the end of the sample. However, these specifications suggest that the patterns and broad magnitudes are comparable for early and late reformers.

3.7.9 Alternative Clustering Strategies

In our main analysis we cluster standard errors at the city level, to allow for observations within a given city to be correlated as well as for heteroskedasticity. Our results are robust to alternative clustering strategies. First, we cluster the standard errors by province and year of reform (i.e., the first year in which a city hosts a SEZ). This strategy takes account of the fact that the introduction of SEZ is highly clustered in time. Many HIDZ were introduced in 1991–92, and many ETDZ were introduced in 2001–03, implying that different cities in these years cannot be treated as independent observations. The baseline results (not shown but available upon request) are essentially unchanged, and in some cases the statistical significance of the coefficients of interest is even strengthened.

We also run the regressions clustering standard errors at the province level (instead of province×year of first reform). This strategy is even more demanding, and runs into potential problems since we have only 28 provinces in the full sample and 18 provinces in the inland sample, and so the number of clusters is small. The results are robust to even this demanding approach. The coefficients of interest remain significant, although in the inland sample for GDP per capita only at 10%.

3.8 Conclusion

The place-based industrial policy is a building block of the development strategy pursued by the Chinese government. According to Naughton (2007): "Bold, fragmented, open to outside investment, but with a strong role for government: Special Economic Zones typify much of the Chinese transition process" (p. 410). This paper estimates the effect of SEZ on local economic performance. The results suggest that the establishment of SEZ has yielded large positive effects on GDP and GDP per capita for the cities in which these were located. Although our estimates are smaller than those found by the earlier literature based on cross-sectional growth regressions (typically on a smaller set of cities and years), the effects are sizeable and robust. We also find that the SEZ generated positive spillovers to neighboring areas.

What can we learn from the Chinese experience about the role of economic reform and industrial policy during the process of development? Existing theoretical and empirical work suggests that policies and institutions should be appropriate to the stage of development, and particularly to the stage of the process of technological convergence (Acemoglu et al.; Konig et al., 2016). The Chinese reform process was characterized by a mixture of elements of market liberalization and an active role of the government in promoting investment and technology adoption. Rodrik (2006) argues that the active role of the government was crucial for China's development because it supported a fast move towards more modern and productive sectors which have positive externalities on the whole economy. The results of our empirical analysis suggest that the industrial policy may have indeed been a catalyst of the development process. At the same time, the estimated effects are quantitatively not very large relative to the high growth rates experienced by China in this period.

Our analysis is subject to a number of limitations that we leave to future work to address. First, cities assigned to SEZ were not randomly allocated. To alleviate concerns for endogeneity, we focus on a subsample of inland cities where the allocation was driven by rigidly selected criteria, and we compare pre- vs. post-reform trends in treated cities. However, ideally one would like to have valid instruments for the spatial and time distribution of the policy intervention. This is very hard to find in the context of Chinese SEZ.

Second, we did not attempt a proper assessment of the welfare effect of SEZ. On the one hand, this would require a quantification of the budgetary costs of the policy. On the other hand, the local gains from spillovers through the agglomeration of labor may be partially offset by losses in other locations that experience an outflow of firms. As pointed out in Greenstone et al. (2010) and Kline and Moretti (2014a), whether the gains offset the losses depends on the shape of the agglomeration force.

Third, it would be interesting to disentangle which of the different components of the policy package had the largest effects. While the reduction in tax wedges must have been important, there are other channels through which SEZ may affect local and regional outcomes. As discussed in Kline and Moretti (2014a), place-based policies may also be used to reduce frictions such as excessive regulation. There may be political constraints that prevent the central government from implementing a reform nationally, such that a reduction within a subset of cities may be the best achievable alternative. In this case, firms would again relocate towards the SEZ and increase total factor productivity further through a larger labor pool. The reduction in frictions may also reduce prices, which could explain why prices are not increasing as much in our data as we may expect based on a simple spatial equilibrium framework. Another friction that can be relaxed by SEZ is the hukou system that restricts labor mobility, as SEZ may make it easier for workers to move there.

Finally, although the establishment of SEZ appears to have generated positive spillovers outside of the areas where they were introduced, we cannot rule out that the industrial policy drew resources away from locations that are remote and far away from the SEZ. In spite of these *caveats*, our results provide the basis for a realistic assessment of the effects of industrial policy in China, and some useful new evidence in the debate on place-based industrial policy in different countries.

Table 3.1: Descriptive statistics

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**N**
Real GDP (mil)	20938.83	30120.88	210	451172.34	5289
Growth of real GDP (%)	13.02	17.71	-52.19	594.78	4736
Real GDP per capita	5228.68	5066.3	220.26	51513.13	5252
Growth of real GDP per capita (%)	10.8	12.12	-75.77	391.7	4969
Land area (sq km)	14059.24	17185.24	137	253356	5335
Growth of land area (%)	16.58	276.78	-59.63	9423.84	5055
Population (mil)	3.85	2.92	0.1	48.51	5306
Growth of population (%)	2.26	16.28	-77.31	347.56	5030
Electricity consumption (GWh)	3.08	4.72	0.01	56.3	5210
Growth of electricity consumption (%)	17.67	198.6	-98.97	13486.34	4914
Mean light intensity	11.74	10.31	0.1	56.04	4730
Growth of mean light intensity per satellite (%)	4.38	14.48	-45.64	117.23	4274

Notes: The table shows the descriptive statistics of our main variables in our sample of 276 cities in 25 provinces from 1988 to 2010. Real GDP is derived from city-level nominal GDP and provincial deflators. Land area is the official size of the cities. Population includes registered residents only. Electricity consumption is by households and firms. Mean light intensity is the average brightness of pixels in the city.

Table 3.2: Descriptive statistics (at beginning of sample period) by region and reform year

Region:		Coast			Inland		
Reform year:	All	< 1988	1988-2010	Never	< 1988	1988-2010	Never
Population (1,000 people)	3113.0	5282.5	3913.0	2968.8	-	3381.2	2319.7
Land area (km2)	10266.6	9811.4	8059.6	10247.1	-	10414.4	11384.2
Population density (people/km2)	402.3	561.1	553.9	353.9	-	375.3	344.9
Real GDP (millions)	4870.1	11363.7	7936.9	4215.4	-	5290.5	2515.9
Real GDP p.c.	1777.7	2072.9	2290.0	1722.7	-	1867.1	1441.0
Nominal GDP p.c.	1862.0	2180.9	2408.7	1816.4	-	1945.3	1502.9
Government spending/GDP (%)	8.7	7.5	7.0	8.4	-	9.3	9.6
# of universities (per mil. people)	2.0	1.6	1.8	1.4	-	3.3	1.4

Notes: The table shows the mean values of selected city characteristics at the beginning of our sample (averaged over 1988 and 1989), separately for reformers before 1988, reformers between 1988 and 2010, and cities that never had a reform. The table also distinguishes inland and coastal cities. Note that no inland city was reformed before 1988. We restrict the sample to a balanced panel.

Table 3.3: Baseline regression

	(1)	(2)	(3)	(4)	(5)	(6)
State-level SEZ	0.156***	0.116***	0.0927***	0.213***	0.175***	0.130**
	(0.0330)	(0.0292)	(0.0283)	(0.0693)	(0.0560)	(0.0532)
Province-level SEZ	0.0217	-0.00166	-0.0113	0.0209	-0.0106	-0.00580
	(0.0226)	(0.0182)	(0.0165)	(0.0310)	(0.0252)	(0.0232)
Dependent variable (log)	GDP	GDP	GDPpc	GDP	GDP	GDPpc
Controlling for log land area	No	Yes	Yes	No	Yes	Yes
Controlling for log population	No	No	Yes	No	No	Yes
Sample	Full	Full	Full	Inland	Inland	Inland
N	5392	5321	5269	2864	2798	2768
Adj. Rsq.	0.960	0.975	0.974	0.949	0.972	0.971

Notes: The dependent variable is the logarithm of annual GDP or GDP per capita. State-level (respectively province-level) SEZ is a dummy switching on in the year after the introduction of any SEZ at that level. All specifications include city fixed effects and the interaction of province-year dummies. Standard errors are clustered at the city level: * $p \leq 0.10$, ** $p \leq 0.05$, *** $p \leq 0.01$.

Table 3.4: Pre-reform trends

	(1)	(2)	(3)	(4)	(5)	(6)
3 years before state-level SEZ	-0.00836	-0.00322	-0.0129	0.128	-0.104	-0.0109
	(0.0327)	(0.0284)	(0.0299)	(0.0972)	(0.0932)	(0.0801)
2 years before state-level SEZ	-0.0286	-0.0190	-0.0278	0.0765	-0.135	-0.0479
	(0.0334)	(0.0285)	(0.0315)	(0.0954)	(0.0944)	(0.0787)
1 year before any state-level SEZ	-0.0221	-0.0164	-0.0291	0.0508	-0.165*	-0.0836
	(0.0342)	(0.0287)	(0.0322)	(0.0984)	(0.0968)	(0.0821)
Year of state-level SEZ reform	-0.00807	-0.00737	-0.0154	0.0689	-0.167*	-0.0787
	(0.0346)	(0.0290)	(0.0342)	(0.0950)	(0.0954)	(0.0830)
State-level SEZ (all post-reform years)	0.143***	0.107***	0.0758*	0.280***	0.0551	0.0834
	(0.0445)	(0.0377)	(0.0417)	(0.0985)	(0.0989)	(0.0920)
Province-level SEZ (all post-reform years)	0.0212	-0.00198	-0.0119	0.0227	-0.0134	-0.00684
	(0.0228)	(0.0183)	(0.0167)	(0.0314)	(0.0256)	(0.0234)
Dependent variable (log)	GDP	GDP	GDPpc	GDP	GDP	GDPpc
Controlling for log land area	No	Yes	Yes	No	Yes	Yes
Controlling for log population	No	No	Yes	No	No	Yes
Sample	Full	Full	Full	Inland	Inland	Inland
N	5392	5321	5269	2864	2798	2768
Adj. Rsq.	0.960	0.975	0.974	0.949	0.972	0.971

Notes: The dependent variable is the logarithm of annual GDP or GDP per capita. State-level (respectively province-level) SEZ is a dummy switching on in the year after the introduction of any SEZ at that level. Lags are defined as described in the table. All specifications include controls for land area, city fixed effects, and the interaction of province-year dummies. Standard errors are clustered at the city level: * $p \leq 0.10$, ** $p \leq 0.05$, *** $p \leq 0.01$.

Table 3.5: Differential trends for reformers

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
State-level SEZ	0.116***	0.0588**	0.0511*		0.175***	0.0259	0.0738	
	(0.0292)	(0.0278)	(0.0286)		(0.0560)	(0.0627)	(0.0607)	
Province-level SEZ	-0.00166	-0.00112	-0.00200	-0.000637	-0.0106	-0.00913	-0.0135	-0.0143
	(0.0182)	(0.0182)	(0.0182)	(0.0184)	(0.0252)	(0.0254)	(0.0256)	(0.0257)
Time trend of reformers (state-level)		0.00711***	0.00410	0.00433		0.0130**	-0.0122	-0.0116
		(0.00260)	(0.00330)	(0.00328)		(0.00590)	(0.0131)	(0.0122)
Post-reform trend (state-level)			0.00470	0.0168**			0.0264*	0.0455**
			(0.00403)	(0.00658)			(0.0145)	(0.0202)
Sq. post-reform trend (state-level)				-0.000558*				-0.000986
				(0.000289)				(0.000813)
Sample	Full	Full	Full	Full	Inland	Inland	Inland	Inland
N	5321	5321	5321	5321	2798	2798	2798	2798
Adj. Rsq.	0.975	0.976	0.976	0.976	0.972	0.972	0.972	0.972

Notes: The dependent variable is the logarithm of annual GDP. State-level (respectively province-level) SEZ is a dummy switching on in the year after the introduction of any SEZ at that level. The trend variables are described in Equation (3.2). All specifications include controls for land area, city fixed effects, and the interaction of province-year dummies. Standard errors are clustered at the city level: * $p \leq 0.10$, ** $p \leq 0.05$, *** $p \leq 0.01$.

Table 3.6: Controlling for differential trends

Panel A : Year fixed effects interacted with initial characteristics (relative to median)

	(1)	(2)	(3)	(4)	(5)	(6)
State-level SEZ	0.119***	0.111***	0.0936***	0.257***	0.238***	0.257***
	(0.0389)	(0.0370)	(0.0341)	(0.0685)	(0.0908)	(0.0680)
Province-level SEZ	0.0273	0.0116	-0.0154	0.0148	0.00248	-0.0180
	(0.0215)	(0.0181)	(0.0162)	(0.0345)	(0.0281)	(0.0269)
Dependent variable (log)	GDP	GDP	GDPpc	GDP	GDP	GDPpc
Controlling for log land area	No	Yes	Yes	No	Yes	Yes
Controlling for log population	No	No	Yes	No	No	Yes
Sample	Full	Full	Full	Inland	Inland	Inland
N	4727	4663	4618	2375	2315	2292
Adj. Rsq.	0.968	0.977	0.978	0.959	0.975	0.974

Panel B: Year fixed effects interacted with initial characteristics (dummies for above median)

	(1)	(2)	(3)	(4)	(5)	(6)
State-level SEZ	0.140***	0.102***	0.0772**	0.241***	0.182**	0.187***
	(0.0363)	(0.0333)	(0.0314)	(0.0886)	(0.0737)	(0.0667)
Province-level SEZ	0.0246	-0.00204	-0.0148	0.0197	-0.00532	-0.00683
	(0.0217)	(0.0180)	(0.0164)	(0.0321)	(0.0270)	(0.0237)
Dependent variable (log)	GDP	GDP	GDPpc	GDP	GDP	GDPpc
Controlling for log land area	No	Yes	Yes	No	Yes	Yes
Controlling for log population	No	No	Yes	No	No	Yes
Sample	Full	Full	Full	Inland	Inland	Inland
N	5392	5321	5269	2864	2798	2768
Adj. Rsq.	0.963	0.976	0.976	0.953	0.973	0.971

Notes: The dependent variable is the logarithm of annual GDP or GDP per capita. State-level (respectively province-level) SEZ is a dummy switching on in the year after the introduction of any SEZ at that level. The specifications in Panel A control for year dummies interacted with the logarithm of population, GDP per capita, population density, and number of universities relative to the median in the year in which a city enters the sample. The specifications in Panel B control for year dummies interacted with indicators for population, GDP per capita, population density, and number of universities being above the median in the year in which a city enters the sample. The specifications also include city fixed effects and the interaction of province-year dummies. Standard errors are clustered at the city level: * $p \leq 0.10$, ** $p \leq 0.05$, *** $p \leq 0.01$.

Table 3.7: Decomposition of the reform effect

Panel A : GDP per capita, capital-labor ratio, population, and human capital

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
State-level SEZ	0.0927***	0.131***	0.0453	0.00295	0.130**	0.339***	0.0599	0.00278
	(0.0283)	(0.0404)	(0.0274)	(0.00367)	(0.0532)	(0.0614)	(0.0374)	(0.00816)
Province-level SEZ	-0.0113	-0.00402	0.0187	0.00214	-0.00580	-0.0385	-0.00177	0.00390
	(0.0165)	(0.0297)	(0.0130)	(0.00159)	(0.0232)	(0.0525)	(0.0123)	(0.00271)
Dependent Variable	log(Y/L)	log(K/L)	log(L)	log(h)	log(Y/L)	log(K/L)	log(L)	log(h)
Sample	Full	Full	Full	Full	Inland	Inland	Inland	Inland
N	5269	4495	5275	4561	2768	2219	2769	2261
Adj. Rsq.	0.974	0.965	0.822	0.961	0.971	0.960	0.883	0.950

Panel B: GDP per capita, population and human capital (census years only)

	(1)	(2)	(3)	(4)	(5)	(6)
State-level SEZ	0.141***	0.0965***	0.182***	0.175**	0.111***	0.360***
	(0.0374)	(0.0359)	(0.0485)	(0.0720)	(0.0368)	(0.0807)
Province-level SEZ	-0.0259	0.0527**	0.0747**	-0.0238	-0.00423	0.102**
	(0.0308)	(0.0246)	(0.0318)	(0.0408)	(0.0193)	(0.0476)
Dependent Variable	log(Y/L)	log(L)	avg. sch.	log(Y/L)	log(L)	avg. sch.
Sample	Full	Full	Full	Inland	Inland	Inland
N	694	695	582	366	366	303
Adj. Rsq.	0.984	0.784	0.979	0.984	0.892	0.979

Panel C: Total factor productivity

	(1)	(2)	(3)	(4)	(5)	(6)
State-level SEZ	0.0617**	0.0553**	0.137***	0.0631	0.0471	0.189***
	(0.0274)	(0.0274)	(0.0346)	(0.0455)	(0.0446)	(0.0527)
Province-level SEZ	-0.00370	-0.00395	-0.0144	0.0203	0.0210	0.0336
	(0.0185)	(0.0187)	(0.0212)	(0.0259)	(0.0268)	(0.0291)
Dependent Variable	log(TFP)	log(TFP)	log(TFP)	log(TFP)	log(TFP)	log(TFP)
Sample	Full	Full	Full	Inland	Inland	Inland
N	4019	4019	4019	1895	1895	1895
Adj. Rsq.	0.958	0.949	0.992	0.952	0.942	0.992

Notes: State-level (respectively province-level) SEZ is a dummy switching on in the year after the introduction of any SEZ at that level. In Panel A, the dependent variable is the logarithm of GDP per capita, capital labor ratio, population, or human capital based on the yearbook data. In Panel B, the dependent variable is the logarithm of GDP per capita, population, and average schooling years based on the census data. In Panel C, the dependent variable is the logarithm of TFP. In columns (1) and (2), TFP is estimated using the whole sample without restrictions on the return to scale. In columns (3) and (4), they are estimated in the whole sample while imposing the constant return to scale restriction. In columns (5) and (6), they are estimated in the pre- and post-1995 sample separately, without imposing the restriction on the return to scale. All specifications include city fixed effects, the interaction of province-year dummies, and controls for land area. All regressions except columns (3) and (7) in Panel A and columns (2) and (5) in Panel B also include controls for population. Standard errors are clustered at the city level: * $p \leq 0.10$, ** $p \leq 0.05$, *** $p \leq 0.01$.

Table 3.8: Spillovers across cities

Panel A : Distance to closest SEZ

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
State-level SEZ	0.124***	0.123***	0.123***	0.122***	0.213***	0.200***	0.197***	0.193***
	(0.0305)	(0.0299)	(0.0297)	(0.0296)	(0.0493)	(0.0526)	(0.0501)	(0.0507)
Province-level SEZ	-0.00390	-0.00217	-0.00174	-0.00151	-0.0120	-0.0152	-0.0145	-0.0134
	(0.0181)	(0.0182)	(0.0182)	(0.0182)	(0.0251)	(0.0249)	(0.0250)	(0.0250)
Log km distance next SEZ	-0.0528**				-0.111***			
	(0.0249)				(0.0342)			
Log driving time next SEZ		-0.0816***				-0.124***		
		(0.0269)				(0.0390)		
Log transport costs next SEZ (10 cat)			-0.0694***				-0.110***	
			(0.0210)				(0.0307)	
Log transport costs next SEZ (7 cat)				-0.0655***				-0.105***
				(0.0210)				(0.0309)
Sample	Full	Full	Full	Full	Inland	Inland	Inland	Inland
N	5254	5321	5321	5321	2775	2798	2798	2798
Adj. Rsq.	0.977	0.976	0.976	0.976	0.972	0.972	0.972	0.972

Panel B: Effect of other SEZ within a 150 distance radius

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
State-level SEZ	0.121***	0.116***	0.117***	0.117***	0.183***	0.173***	0.175***	0.177***
	(0.0295)	(0.0295)	(0.0297)	(0.0294)	(0.0538)	(0.0527)	(0.0538)	(0.0545)
Province-level SEZ	-0.000753	-0.00228	-0.00261	-0.00139	-0.00957	-0.0138	-0.0152	-0.0107
	(0.0183)	(0.0182)	(0.0181)	(0.0182)	(0.0253)	(0.0252)	(0.0246)	(0.0252)
Neighboring SEZ in 150 radius (distance)	0.0514*				0.0688			
	(0.0306)				(0.0452)			
Neighboring SEZ in 150 radius (driving time)		0.102***				0.102**		
		(0.0325)				(0.0438)		
Neighboring SEZ in 150 radius (transport costs, 10 cat)			0.0685*				0.128***	
			(0.0354)				(0.0475)	
Neighboring SEZ in 150 radius (transport costs, 7 cat)				0.0286				0.0560
				(0.0330)				(0.0414)
Sample	Full	Full	Full	Full	Inland	Inland	Inland	Inland
N	5321	5321	5321	5321	2798	2798	2798	2798
Adj. Rsq.	0.975	0.976	0.975	0.975	0.972	0.972	0.972	0.972

Notes: The dependent variable is the logarithm of annual GDP. State-level (respectively province-level) SEZ is a dummy switching on in the year after the introduction of any SEZ at that level. In Panel A, the additional independent variable is *distance to the next zone* and is an inverse measure of spillover intensity, and a negative coefficient therefore implies a positive spillover effect. In Panel B, the additional independent variable is *indicator for other zone in radius* and is expected to have a positive coefficient in the case of positive spillovers. All specifications include controls for land area, city fixed effects, and the interaction of province-year dummies. Standard errors are clustered at the city level: * $p \leq 0.10$, ** $p \leq 0.05$, *** $p \leq 0.01$.

Table 3.9: GDP proxied by light and electricity

	(1)	(2)	(3)	(4)
State-level SEZ	0.0501**	0.0460	0.157***	0.0389
	(0.0254)	(0.0537)	(0.0555)	(0.0722)
Province-level SEZ	-0.00754	-0.0322	0.0344	0.0339
	(0.0183)	(0.0281)	(0.0358)	(0.0425)
Dependent variable (log)	Light	Light	Electricity	Electricity
Controlling for log land area	No	No	Yes	Yes
Sample	Full	Inland	Full	Inland
N	4730	2570	5207	2718
Adj. Rsq.	0.836	0.817	0.792	0.755

Notes: The dependent variable is the logarithm of light intensity at the city level or the logarithm of electricity consumption in the urban core of the city. State-level (respectively province-level) SEZ is a dummy switching on in the year after the introduction of any SEZ at that level. All specifications include city fixed effects and the interaction of province-year dummies. Standard errors are clustered at the city level: * $p \leq 0.10$, ** $p \leq 0.05$, *** $p \leq 0.01$.

Table 3.10: Controlling for road infrastructure and government spending

	(1)	(2)	(3)	(4)	(5)	(6)
State-level SEZ	0.146***	0.111***	0.109***	0.203***	0.186***	0.159***
	(0.0299)	(0.0299)	(0.0284)	(0.0542)	(0.0533)	(0.0428)
Province-level SEZ	-0.00786	0.00163	-0.00130	-0.0138	0.00201	-0.0140
	(0.0177)	(0.0183)	(0.0178)	(0.0236)	(0.0261)	(0.0241)
Log road area	0.0761***	0.0433***	0.0329**	0.0628**	0.0270	0.0128
	(0.0215)	(0.0161)	(0.0150)	(0.0302)	(0.0220)	(0.0158)
Log government spending		0.480***	0.505***		0.432***	0.466***
		(0.0573)	(0.0602)		(0.0628)	(0.0483)
Log government spending/income			-0.281***			-0.319***
			(0.0458)			(0.0458)
Sample	Full	Full	Full	Inland	Inland	Inland
N	4423	2633	2632	2336	1427	1427
Adj. Rsq.	0.978	0.988	0.990	0.975	0.986	0.989

Notes: The dependent variable is the logarithm of annual GDP. State-level (respectively province-level) SEZ is a dummy switching on in the year after the introduction of any SEZ at that level. All specifications include city fixed effects, the interaction of province-year dummies, and controls for land area. Standard errors are clustered at the city level: * $p \leq 0.10$, ** $p \leq 0.05$, *** $p \leq 0.01$.

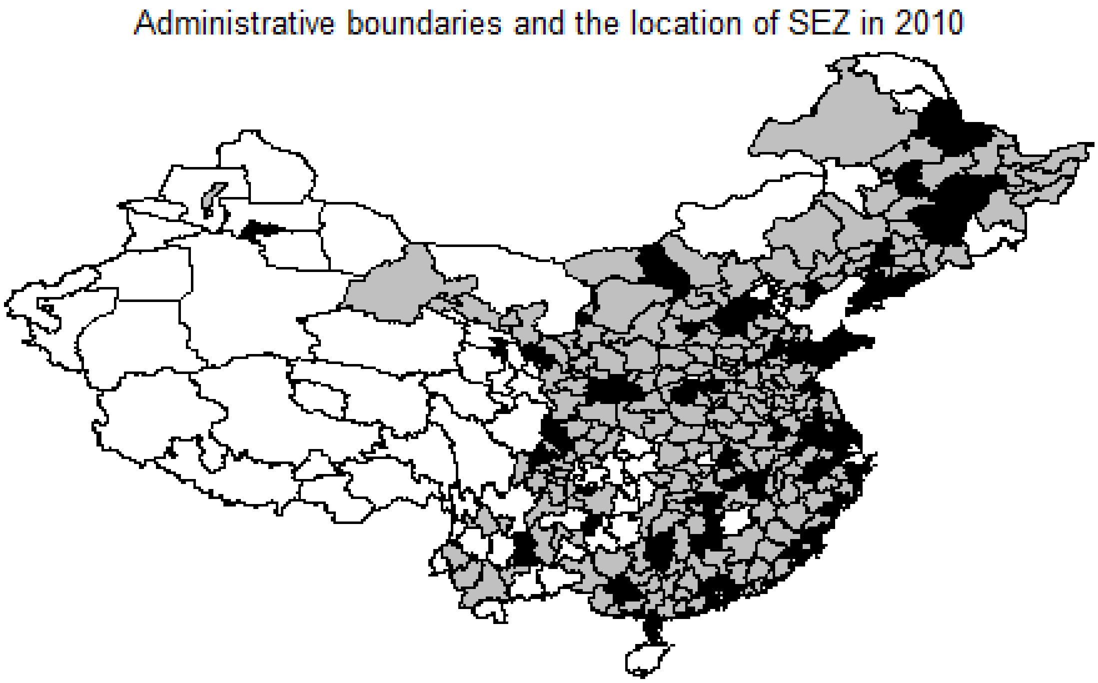


Figure 3.1: Location of treated cities in 2010

Notes: The cities in our sample with at least one state-level SEZ in 2010 are marked in black (a city may have more than one zone). The cities in our sample without a SEZ in 2010 are marked in grey.

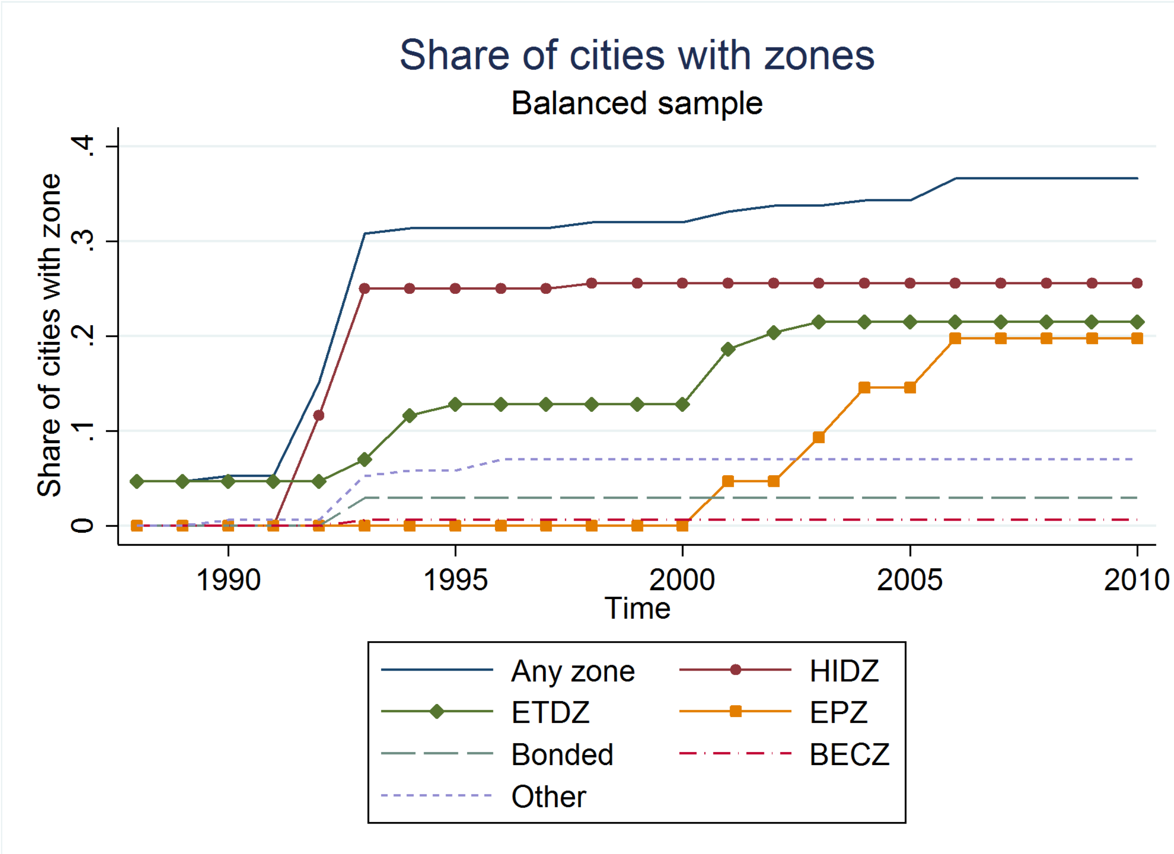


Figure 3.2: Share of cities with different types of zones

Notes: The figure shows the share of cities which have a state-level SEZ. The figure also shows the different types of SEZ: Hightech Industrial Development Zones, Economic and Technological Development Zones, Export Processing Zones, Bonded Zones, Border Economic Cooperation Zones, and other types. The sample is restricted to 172 cities that are observed in all years between 1988 and 2010.

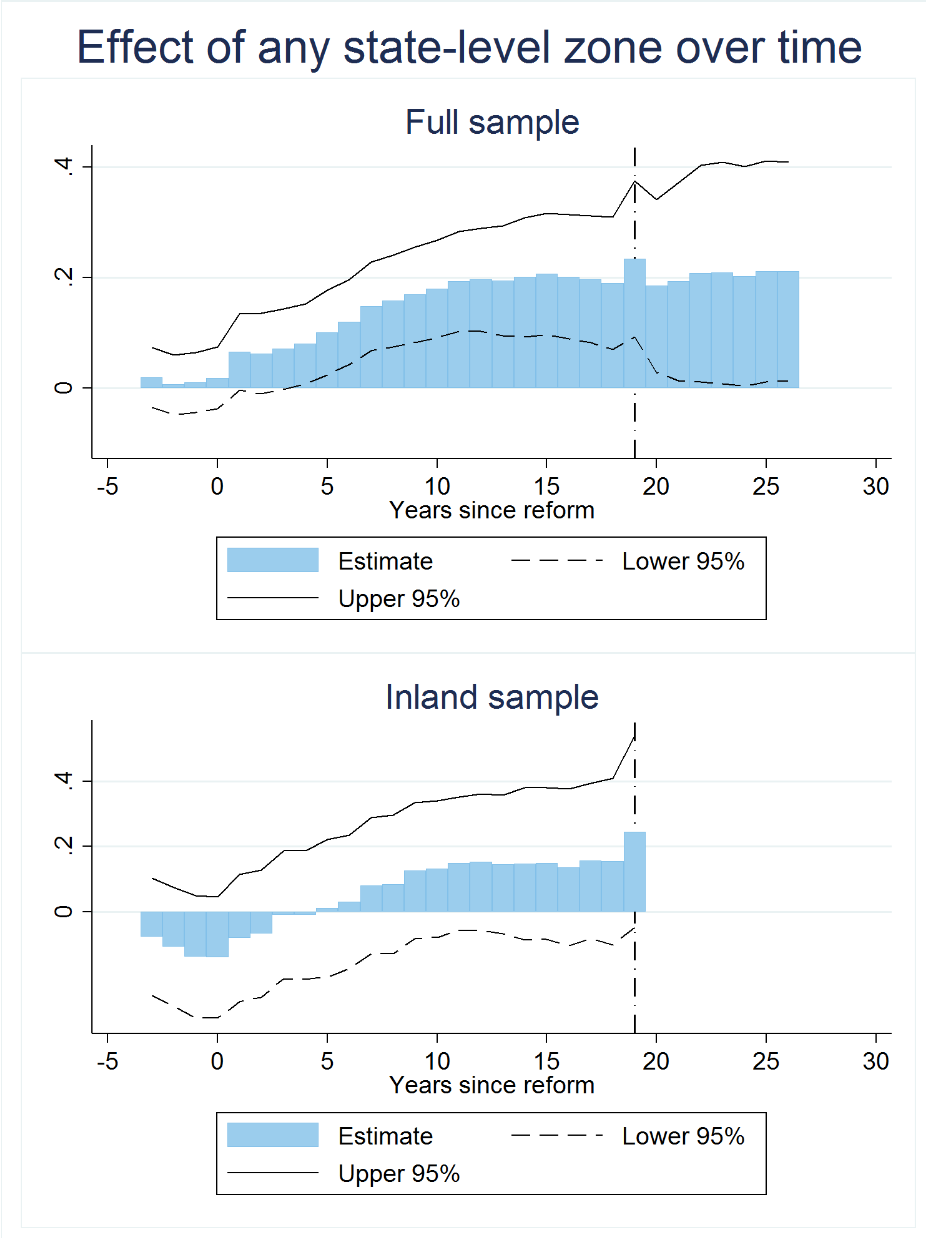


Figure 3.3: Reform effects over time

Notes: The bars show the coefficients of a regression of the logarithm of nominal GDP on indicators for years before and after the first zone. The solid and dashed lines show the confidence interval. The vertical dashed line at 19 shows when the reformers from 1991 reach 2010 and subsequently the number of observations to identify post-reform indicators drops to 9. The regressions also control for an indicator for province-level zones, land area, city fixed effects, and province-time fixed effects. Standard errors are clustered by city.

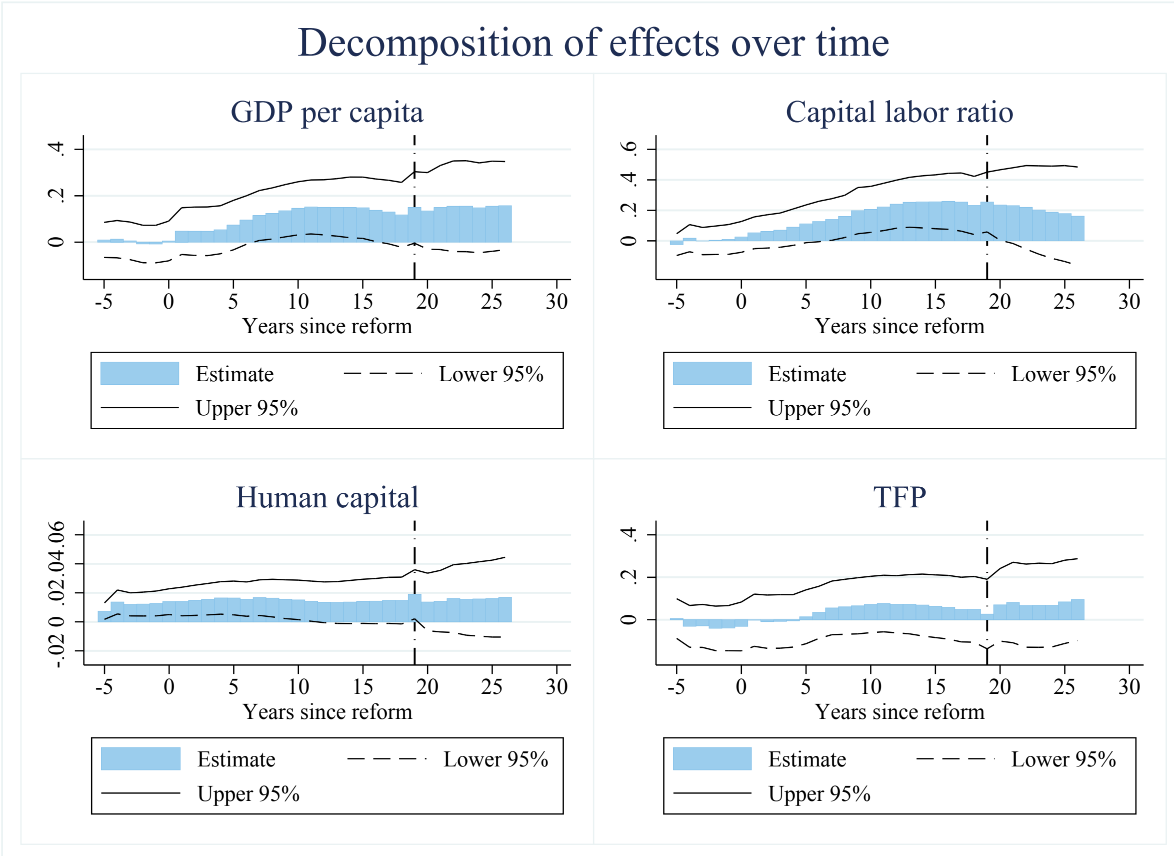


Figure 3.4: Decomposition of the reform effects over time

Notes: The bars show the coefficients of a regression of the four dependent variables on indicators for years before and after the first SEZ was established. TFP is computed using a full-sample unrestricted production function estimation. The solid and dashed lines show the confidence interval. The vertical dashed line at 19 shows when the reformers from 1991 reach 2010 and subsequently the number of observations to identify post-reform indicators drops to 9. The regressions control for an indicator for province-level zones, population, land area, city fixed effects, and province-time fixed effects. Standard errors are clustered by city.

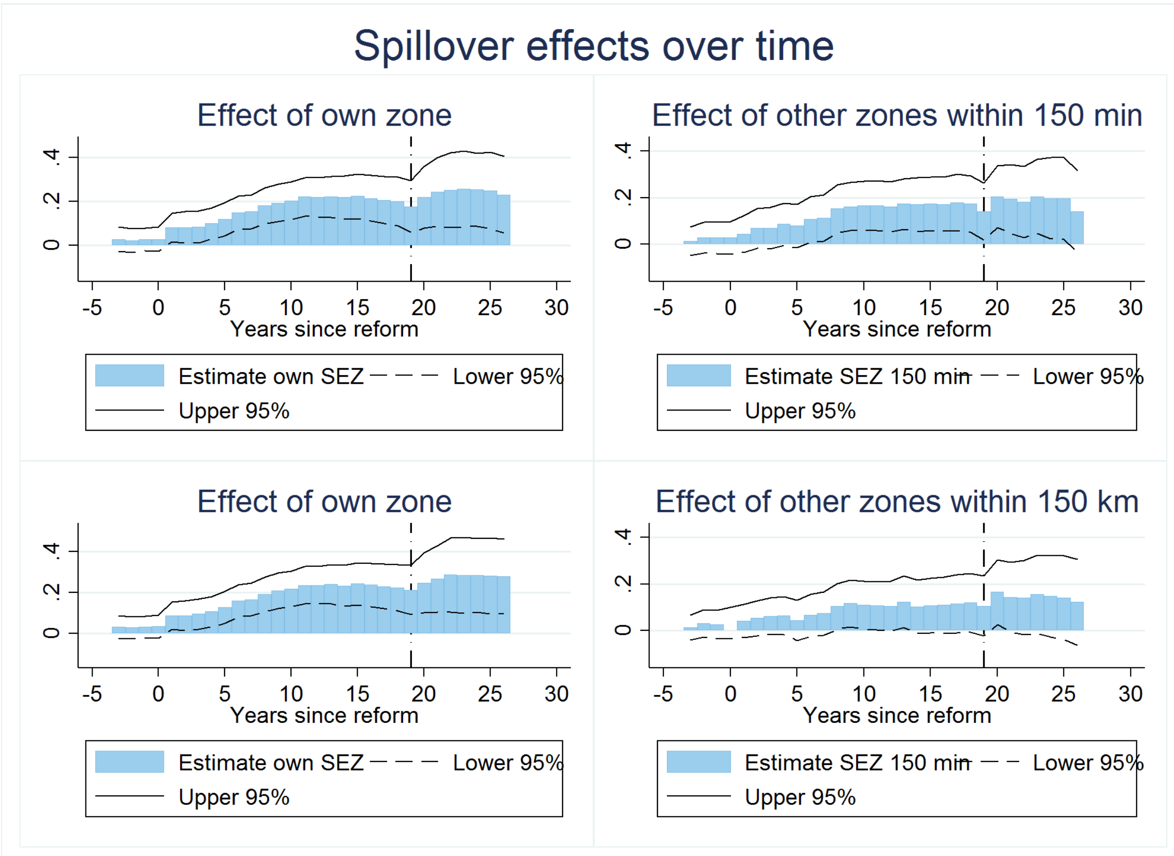


Figure 3.5: Spillover over time

Notes: The upper two graphs show the estimates from a regression of the logarithm of nominal GDP on indicators for years before and after the first zone in the own city and zones within 150 minutes driving time. The effect of the time lag since the own reform year (left graph) is included together with the time lags from the first zone within 150 minutes driving time (right graph). The lower two graphs show the same when a radius of 150 kilometers is used. The solid and dashed lines show the confidence interval. The vertical dashed line at 19 shows when the reformers from 1991 reach 2010 and subsequently the number of observations to identify post-reform indicators drops to 9. The regressions also control for an indicator for province-level zones, land area, city fixed effects, and province-time fixed effects. Standard errors are clustered by city.

References

AASHTO. Policy on geometric design of highways and streets. Technical report, American Association of State Highway and Transportation Officials, Washington, DC, 2001.

Daron Acemoglu, Philippe Aghion, and Fabrizio Zilibotti. Distance to frontier, selection, and economic growth. *Journal of the European Economic Association*, 4(1):37–74.

Philippe Aghion, Robin Burgess, Stephen J. Redding, and Fabrizio Zilibotti. The unequal effects of liberalization: Evidence from dismantling the license raj in India. *American Economic Review*, 98(4):1397–1412.

Gokhan Akinci and James Crittle. Special economic zones: Performance, lessons learned, and implications for zone development. *Foreign Investment Advisory Service (FIAS) occasional paper: Washington, DC: World Bank*, 2008.

Simon Alder. Chinese roads in India: The effect of transport infrastructure on economic development. *mimeo*, 2015.

Franklin Allen, Jun Qian, and Meijun Qian. Law, finance, and economic growth in China. *Journal of Financial Economics*, 77:57–116, 2005.

Heitor Almeida and Murillo Campello. Financial constraints, asset tangibility, and corporate investment. *Review of Financial Studies*, 20(5):1429–1460, 2007.

Cristina Arellano, Yan Bai, and Patrick J. Kehoe. Financial frictions and fluctuations in volatility. *mimeo*, 2012.

Meghana Ayyagari, Asli Demirguc-Kunt, and Vojislav Maksimovic. Formal versus informal finance: Evidence from China. *Review of Financial Studies*, 23(8):3048–3097, 2010.

Thorsten Beck and Ross Levine. Legal institutions and financial development. *NBER Working Paper*, (10126), 2003.

Ben Bernanke, Mark Gertler, and Simon Gilchrist. The financial accelerator in a quantitative business cycle framework. *mimeo*, 1989.

Bruno Biais and Christian Gollier. Trade credit and credit rationing. *Review of Financial Studies*, 10(4):903–937, 1997.

Saki Bigio. Endogenous liquidity and the business cycle. *American Economic Review*, forthcoming.

Saki Bigio and Jennifer La'O. Financial frictions in production networks. *mimeo*, 2014.

Nicholas Bloom, Mark Schankerman, and John Van Reenen. Identifying technology spillovers and product market rivalry. *Econometrica*, 81(4):1347–1393, 2013.

Frederic Boissay and Russell Cooper. Inside and outside collateral and financial market instability. *NBER Working Paper*, 2014.

Loren Brandt and Xiaodong Zhu. Accounting for China's growth. *mimeo*, 2010.

Loren Brandt, Johannes Van Biesebroeck, and Yifan Zhang. Creative accounting or creative destruction? firm-level productivity growth in Chinese manufacturing. *Journal of Development Economics*, 97(2):339–351, 2012.

Anthony Briant, Miren Lafourcade, and Benoit Schmutz. Can tax breaks beat geography? lessons from the french enterprise zone experience. *American Economic Journal: Economic Policy*, 7(2):88–124, 2015.

Wyatt J. Brooks, Joseph P. Kaboski, and Yao Amber Li. Agglomeration and (the lack of) competition. *mimeo*, 2015.

Francisco J. Buera, Joseph P. Kaboski, and Yongseok Shin. Finance and development: A tale of two sectors. *American Economic Review*, 101(5):1964–2002, 2011.

Francisco J. Buera, Roberto N. Fattal-Jaef, and Yongseok Shin. Anatomy of a credit crunch: From capital to labor markets. *Review of Economic Dynamics*, 18(1):101–117, 2015a.

Francisco J. Buera, Joseph P. Kaboski, and Yongseok Shin. Entrepreneurship and financial frictions: A macrodevelopment perspective. *Annual Review of Economics*, 7:409–436, 2015b.

Mike Burkart and Tore Ellingsen. In-kind finance: A theory of trade credit. *American Economic Review*, 94(3):569–590, 2004.

Matias Busso, Jesse Gregory, and Patrick Kline. Assessing the incidence and efficiency of a prominent place based policy. *American Economic Review*, 103(2):897–947, 2013.

Francesco Caselli. Accounting for cross-country income differences. In Philippe Aghion and Steven Durlauf, editors, *Handbook of Economic Growth*, volume 1, chapter 9, pages 679–741. Elsevier, New York, 2005.

Sudheer Chava and Michael Roberts. How does financing impact investment? the role of debt covenant. *Journal of Finance*, 63(5):2085–2121, 2008.

Xi Chen and William D. Nordhaus. Using luminosity data as a proxy for economic statistics. *Proceedings of the National Academy of Sciences*, 108(21):8589–8594, 2011.

Leonard K. Cheng and Yum K. Kwan. What are the determinants of the location of foreign direct investment? the Chinese experience. *Journal of International Economics*, 51(2): 379–400, 2000.

Anton Cheremukhin, Mikhail Golosov, Sergei Guriev, and Aleh Tsyvinski. The economy of People's Republic of China from 1953. *mimeo*, 2015.

Raj Chetty, Adamimeo Guren, Day Manoli, and Andrea Weber. Are micro and macro labor supply elasticities? a review of evidence on the intensive and extensive margins. *American Economic Review Papers and Proceedings*, 101(3):471–475, 2011.

Gabriel Chodorow-Reich. The employment effects of credit market disruptions. *Quarterly Journal of Economics*, 129(1):1–59, 2014.

Jae Ho Chung and Tao chiu Lam. China's "city system" in flux: Explaining post-Mao administrative changes. *The China Quarterly*, 180:945–964, 2004.

Antonio Ciccone and Robert E. Hall. Productivity and the density of economic activity. *American Economic Review*, 86(1):54–70, 1996.

Pierre-Philippe Combes and Laurent Gobillon. The empirics of agglomeration economies. In Gilles Duranton, J. Vernon Henderson, and William C. Strange, editors, *Handbook of Regional and Urban Economics*, volume 5, chapter 5, pages 247–348. Elsevier, Amsterdam, 2015.

Pierre-Philippe Combes, Thierry Mayer, and Jacques-Francois Thisse. *Economic geography: The integration of regions and nations*. Princeton University Press, 2008.

Pierre-Philippe Combes, Gilles Duranton, and Laurent Gobillon. The identification of agglomeration economies. *Journal of Economic Geography*, 11(2):253–266, 2010.

Pierre-Philippe Combes, Sylvie Démurger, and Shi Li. Urbanisation and migration externalities in China. *CEPR Discussion Paper*, (9352), 2013.

Anna M. Costello. Trade credit policy in long-term supply contract. *mimeo*, 2014.

Chiara Criscuolo, Ralf Martin, Henry Overman, and John Van Reenen. The causal effects of an industrial policy. *NBER Working Paper*, (17842), 2012.

Vicente Cunat. Trade credit: Suppliers as debt collectors and insurance providers. *Review of Financial Studies*, 20(2):491–527, 2007.

Vicente Cunat and Emilia Garcia-Appendini. *The Oxford Handbook of Entrepreneurial Finance*, chapter Trade Credit and its Role in Entrepreneurial Finance, pages 526–557. Oxford University Press, New York, 2012.

Rose Cunningham. Trade credit and credit rationing in Canadian firms. *Bank of Canada Working Paper*, 2004.

Hans Degryse, Liping Lu, and Steven Ongena. Informal or formal financing? or both? first evidence on the co-funding of Chinese firms. *mimeo*, 2013.

Sylvie Démurger, Jeffrey D Sachs, Wing Thye Woo, Shuming Bao, and Gene Chang. The relative contributions of location and preferential policies in China's regional development: Being in the right place and having the right incentives. *China Economic Review*, 13(4): 444–465, 2002.

Michael P. Devereux, Rachel Griffith, and Helen Simpson. Firm location decisions, regional grants and agglomeration externalities. *Journal of Public Economics*, 91(3-4):413–435, 2007.

Mathias Dewatripont and Gérard Roland. The design of reform packages under uncertainty. *American Economic Review*, 85(5):1207–1223, 1995.

Xiao-Yuan Dong and Louis Putterman. Soft budget constraints, social burdens, and labor redundancy in China's state industry. *Journal of Comparative Economics*, 31(1):110–133, 2003.

Gilles Duranton and Diego Puga. Micro-foundations of urban agglomeration economies. In J. Vernon Henderson and Jacques-Francois Thisse, editors, *Handbook of Regional and Urban Economics*, volume 4, pages 2063–2117. Elsevier, Amsterdam, 2004.

Christopher D. Elvidge, Kimberly E. Baugh, E. A. Kihn, H. W. Kroehl, and E. R. Davis. Relation between satellite observed visible-near infrared emissions, population, economic activity and electric power consumption. *International Journal of Remote Sensing*, 18(6): 1373–1379, 1997.

Chong en Bai and Zhenjie Qian. The factor income distribution in China: 1978-2007. *China Economic Review*, 21(4):650–670, 2010.

Joseph Fewsmith. *China Since Tiananmen: The Politics of Transition*. Cambridge University Press, Cambridge, 2001.

Masahisa Fujita, Paul Krugman, and Anthony J. Venables. *The Spatial Economy: Cities, Regions, and International Trade*. MIT Press, Cambridge, 1999.

Mark Gertler and Nobuhiro Kiyotaki. Financial intermediation and credit policy in business cycle analysis. *mimeo*, 2010.

Edward L. Glaeser and Joshua D. Gottlieb. The economics of place-making policies. *Brookings Papers on Economic Activity*, pages 155–239, 2008.

Edward L. Glaeser and Joshua D. Gottlieb. The wealth of cities: Agglomeration economies and spatial equilibrium in the United States. *Journal of Economic Literature*, 47(4):983–1028, 2009.

Yuriy Gorodnichenko, Jan Svejnar, and Katherine Terrell. When does FDI have positive spillovers? Evidence from 17 emerging market economies. *Journal of Comparative Economics*, 42(4):954–969, 2014.

Michael Greenstone, Richard Hornbeck, and Enrico Moretti. Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings. *Journal of Political Economy*, 118(3):536–598, 2010.

Jeremy Greenwood, Zvi Hercowitz, and Gregory W. Huffman. Investment, capacity utilization, and the real business cycle. *American Economic Review*, 78(3):402–417, 1988.

Veronica Guerrieri and Guido Lorenzoni. Credit crises, precautionary savings, and the liquidity trap. *mimeo*, 2012.

Florian Halg. Assessing economic growth in China using satellite data on lights at night. Master's thesis, University of Zurich, 2012.

Robert E. Hall and Charles I. Jones. Why do some countries produce so much more output per worker than others? *Quarterly Journal of Economics*, 114(1):83–116, 1999.

Keith Head and John Ries. Inter-city competition for foreign investment: Static and dynamic effects of China's incentive areas. *Journal of Urban Economics*, 40(1):38–60, 1996.

J. Vernon Henderson, Adam Storeygard, and David N. Weil. Measuring economic growth from outer space. *American Economic Review*, 102(2):994–1028, 2012.

Chang-Tai Hsieh and Peter J. Klenow. Misallocation and manufacturing TFP in China and India. *Quarterly Journal of Economics*, 124(4):1403–1448, 2009.

Chang-Tai Hsieh and Enrico Moretti. Why do cities matter? Local growth and aggregate growth. *NBER Working Paper*, (21154), 2015.

Tor Jacobson and Erik von Schedvin. Trade credit and the propagation of corporate failure: An empirical analysis. *Econometrica*, 83(4):1315–1371, 2015.

Adam B. Jaffe, Manuel Trajtenberg, and Rebecca Henderson. Geographic localization of knowledge spillovers as evidenced by patent citations. *Quarterly Journal of Economics*, 108(3):578–598, 1993.

Urban Jermann and Vincenzo Quadrini. Macroeconomic effects of financial shock. *American Economic Review*, 102(1):238–71, 2012.

Shang jin Wei. Open door policy and China's rapid growth: Evidence from city-level data. *NBER Working Paper*, (4602), 1993.

Charles I. Jones. Intermediate goods and weak links in the theory of economic development. *American Economic Journal: Macroeconomics*, 3(2):1–28, 2011.

Derek C. Jones, Cheng Li, and Ann L. Owen. Growth and regional inequality in China during the reform era. *China Economic Review*, 14(2):186–200, 2003.

Sebnem Kalemli-Ozcan, Se-Jik Kim, Hyun Song Shin, and Bent E. Sorensen. Financial shocks in production chains. *mimeo*, 2014.

Michael Keane and Richard Rogerson. Micro and macro labor supply elasticities: A reassessment of conventional wisdom. *Journal of Economic Literature*, 50(2):464–476, 2012.

Wolfgang Keller. Geographic localization of international technology diffusion. *American Economic Review*, 92(1):120–142, 2002.

Se-Jik Kim and Hyun Song Shin. Sustaining production chains through financial linkages. *American Economic Review Papers and Proceedings*, 102(3):402–06, 2012.

Nobu Kiyotaki and John Moore. Credit chains. *mimeo*, 1997a.

Nobu Kiyotaki and John Moore. Credit cycles. *Journal of Political Economy*, 105(2):211–248, 1997b.

Patrick Kline and Enrico Moretti. People, places, and public policy: Some simple welfare economics of local economic development policies. *Annual Review of Economics*, 6:629–662, 2014a.

Patrick Kline and Enrico Moretti. Local economic development, agglomeration economies, and the big push: 100 years of evidence from the tennessee valley authority. *Quarterly Journal of Economics*, 129(1):275–331, 2014b.

Michael Konig, Jan Lorenz, and Fabrizio Zilibotti. Innovation vs. imitation and the evolution of productivity distributions. *Theoretical Economics, forthcoming*, 2016.

Jeongmin Lee. Collateral circulation and repo spread. *mimeo*, 2015.

Ross Levine. Finance and growth: Theory and evidence. In Philippe Aghion and Steven N. Durlauf, editors, *Handbook of Economic Growth*, volume 1A, chapter 12, pages 866–910. Elsevier B.V., 2005.

Kui Wai Li, Lihong Yun, and Gilbert C. S. Lui. Economic performance of human capital in post-reform China. *Chinese Economy*, 42(1):40–61, 2009.

Yi Lu, Jin Wang, and Lianming Zhu. Do place-based policies work? Micro-level evidence from China's economic zones program. *mimeo*, 2015.

Ting Ma, Chenghu Zhou, Tao Pei, Susan Haynie, and Junfu Fan. Quantitative estimation of urbanization dynamics using time series of DMSP/OLS nighttime light data: A comparative case study from China's cities. *Remote Sensing of Environment*, 124:99–107, 2012.

Andreas Madestam. Informal finance: A theory of moneylenders. *Journal of Development Economics*, 107(157-174), 2014.

Yue man Yeung, Joanna Lee, and Gordon Kee. China's special economic zones at 30. *Eurasian Geography and Economics*, 50(2):222–240, 2009.

Philippe Martin, Thierry Mayer, and Florian Mayneris. Spatial concentration and plant-level productivity in France. *Journal of Urban Economics*, 69(2):182–195, 2011.

Philippe Martin, Thierry Mayer, and Florian Mayneris. Public support to clusters: A firm level study of French local productive systems. *Regional Science and Urban Economics*, 41(2):108–123, 2011a.

Allan H. Meltzer. Mercantile credit, monetary policy, and size of firms. *The Review of Economics and Statistics*, 42(4):429–437, 1960.

Atif Mian and Amir Sufi. House prices, home equity-based borrowing, and the U.S. household leverage crisis. *American Economic Review*, 101(5):2132–56, 2011.

Virgiliu Midrigan and Thomas Philippon. Household leverage and the recession. *mimeo*, 2011.

John J Murphy. The legal evolution of accounts receivable financing, 1850–1939. *The Secured Lender*, 48(4), 1992.

Barry Naughton. *The Chinese Economy: Transitions and Growth*. MIT Press, Cambridge, 2007.

David Neumark and Jed Kolko. Do enterprise zones create jobs? Evidence from California's enterprise zone program. *Journal of Urban Economics*, 68(1):1–19, 2010.

David Neumark and Helen Simpson. Place-based policies. In Gilles Duranton, J. Vernon Henderson, and William C. Strange, editors, *Handbook of Regional and Urban Economics*, volume 5, chapter 18, pages 1197–1287. Elsevier, Amsterdam, 2015.

Ministry of Commerce. Report on the development of state-level economic and technological development zones (in Chinese). 2006.

State Administration of Taxation. Notice of investigation of preferential tax policies in development zones. 2004.

G. Steven Olley and Ariel Pakes. The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, 64(6):1263–1297, 1996.

Ralph Ossa. A quantitative analysis of subsidy competition in the US. *NBER Working Paper*, (20975), 2015.

Dwight Heald Perkins. Reforming China's economic system. *Journal of Economic Literature*, 26(2):601–645, 1988.

Mitchell A. Petersen and Raghuram G. Rajan. Trade credit: Theories and evidence. *Review of Financial Studies*, 10(3):661–691, 1997.

Claudio Raddatz. Credit chains and sectoral comovement: Does the use of trade credit amplify sectoral shocks? *Review of Economics and Statistics*, 92(4):985–1003, 2010.

Thomas G. Rawski. What is happening to China's gdp statistics? *China Economic Review*, 12(4):347–354, 2001.

Thomas G. Rawski and Loren Brandt, editors. *China's Great Economic Transformation*. Cambridge University Press, Cambridge, 2008.

Stephen J. Redding. Goods trade, factor mobility and welfare. *NBER Working Paper*, (18008), 2012.

Dani Rodrik. Industrial policy for the twenty-first century. *CEPR Discussion Paper*, (4767), 2004.

Dani Rodrik. What's so special about China's exports? *China & World Economy*, 14(5): 1–19, 2006.

Stuart S. Rosenthal and William C. Strange. Evidence on the nature and sources of agglomeration economies. In J. Vernon Henderson and Jacques-Francois Thisse, editors, *Handbook of Regional and Urban Economics*, volume 5, chapter 49, pages 2119–2171. Elsevier, Amsterdam, 2004.

Annette Schminke and Johannes Van Biesebroeck. Using export market performance to evaluate regional preferential policies in China. *Review of World Economics*, 149(2):343–367, 2013.

Lin Shao. Trade credit in production chains. *mimeo*, 2016.

Michael Siemer. Firm entry and employment dynamics in the great recession. *mimeo*, 2014.

Zheng Song, Kjetil Storesletten, and Fabrizio Zilibotti. Growing like China. *American Economic Review*, 101(1):202–241, 2011.

Kjetil Storesletten and Fabrizio Zilibotti. China's great convergence and beyond. *Annual Review of Economics*, 6(1):333–362, 2014.

Amir Sufi. Information asymmetry and financing arrangements: Evidence from syndicated loan. *Journal of Finance*, 62(2):629–668, 2007.

Kellee Sing Tsai. Beyond banks: The local logic of informal finance and private sector development in China. *Conference on Financial Sector Reform in China*, 2001.

Ezra F. Vogel. *Deng Xiaoping and the Transformation of China*. Belknap Press of Harvard University Press, Cambridge, 2011.

Jin Wang. The economic impact of special economic zones: Evidence from Chinese municipalities. *Journal of Development Economics*, 101:133–147, 2013.

WEFore. Report: Analysis of the risks of development zones (in Chinese). *Company Report*, 2010.

Harry X. Wu. Accounting for China's growth in 1952-2008: China's growth performance debate revisited with a newly constructed dataset. *RIETI Discussion Paper Series*, 2011.

Chenggang Xu. The fundamental institutions of China's reforms and development. *Journal of Economic Literature*, 49(4):1076–1151, 2011.

Kei-Mu Yi. Can vertical specialization explain the growth of world trade? *Journal of Political Economy*, 111(1):52–102, 2003.

Alwyn Young. Gold into base metals: Productivity growth in the People's Republic of China during the reform period. *Journal of Political Economy*, 111(6):1220–1261, 2003.

Douglas Zhihua Zeng. How do special economic zones and industrial clusters drive China's rapid development? In Douglas Zhihua Zeng, editor, *Building Engines for Growth and Competitiveness in China: Experiences with Special Economic Zones and Industrial Clusters*. The World Bank, Washington, D.C., 2010.

Ariel Zetlin-Jones and Ali Shourideh. External financing and the role of financial frictions over the business cycle: Measurement and theory. *mimeo*, 2016.

Shengxing Zhang. Collateral risk, repo rollover and shadow banking. *mimeo*, 2014.

Siqi Zheng, Weizeng Sun, Jianfeng Wu, and Matthew E. Kahn. The birth of edge cities in China: Measuring the spillover effects of industrial parks. *NBER Working Paper*, (21378), 2015.

Appendix

A Appendix to Chapter 1

A.1 An alternative way of modeling trade credit

In section 1.3.3, we model trade credit as an explicit loan extended from intermediate goods entrepreneurs to final goods entrepreneurs. However, in reality, trade credit does not involve explicit loan that changes hands. Instead, extending trade credit means delay of cash flow. In this section, we show an alternative way of modeling trade credit that is more in accordance with trade credit in reality.

Consider a different timing setting of the model, in which the trade credit is repaid in the next period instead of at the end of the current period. Intermediate goods entrepreneurs enter each period t with wealth a_t and accounts receivable from period AR_{t-1}. Final goods entrepreneurs, on the other hand, enter each period t with wealth a_t and accounts payable AP_{t-1} from the previous period.

After the productivity is realized, the entrepreneurs choose optimally the input goods k_t, l_t, $x_{1,t}$, trade credit AR_t and AP_t, consumption c_t, and next period wealth a_{t+1}.

The need for intra-temporal loan are

$$\begin{aligned} intermediate: \quad & m_{1,t} = a_{t+1} + AR_t - a_t + c_t + r(k_t - a_t) + \delta k_t + wl_t, \\ final: \quad & m_{2,t} = a_{t+1} - a_t + c_t + r(k_t - a_t) + \delta k_t + wl_t + p_1 x_{1,t} - AP_t. \end{aligned}$$

Using budget constraint of the two entrepreneurs,

$$
\begin{aligned}
intermediate: \quad & c_t + a_{t+1} + AR_t = \\
& (1+r)a_t + p_1 A_1 z F_1(k_t, l_t) - (r+\delta)k_t - wl_t + (1+r^{tc})AR_{t-1}, \\
final: \quad & c_t + a_{t+1} - AP_t = \\
& (1+r)a_t + A_2 z F_2(k_t, l_t, x_{1,t}) - (r+\delta)k_t - wl_t - p_1 x_{1,t} - (1+r^{tc})AP_{t-1}.
\end{aligned}
$$

we derive the need for intra-temporal loan for intermediate goods entrepreneur is $m_{1,t} = p_1 A_1 z F_1(k_t, l_t) + (1+r^{tc})AR_{t-1}$ and for final goods entrepreneurs $m_{2,t} = A_2 z F_2(k_t, l_t, x_{1,t}) - (1+r^{tc})AP_{t-1}$.

A.2 Proofs

In order to prove the propositions, we first lay out the optimization problem of the entrepreneurs and derive the FOCs. For each proposition, we prove the first part of each proposition regarding the intermediate goods entrepreneur. The proof of the second part regarding the final goods entrepreneurs is very similar and hence is omitted. We assume that the constraint $a' \geq 0$ never binds when $a > 0$ and show at the end of the proof that it is indeed the case.

Intermediate goods entrepreneurs Consider the following problem.

$$
\begin{aligned}
V_1(a,z) \quad = \quad & \max_{c,k,l,AR,a'} log(c) + \beta \mathbb{E}_{z'} V_1(a', z') \\
s.t. \quad & c + a' = (1+r)a + p_1 A_1 z F_1(k,l) - (r+\delta)k - wl + r^{tc} AR, & (3.7) \\
& p_1 A_1 z F_1(k,l) + (1+r^{tc})AR \leq \gamma_1 a' + \gamma_2 AR, & (3.8) \\
& 0 \leq AR \leq p_1 A_1 z F_1(k,l), & (3.9) \\
& a' \geq 0.
\end{aligned}
$$

The Lagrangian of the problem can be written as,

$$
\begin{aligned}
\mathcal{L} \;=\; & log((1+r)a + p_1 A_1 z F_1(k,l) - (r+\delta)k - wl + r^{tc}AR - a') \\
& + \mu(\gamma_1 a' + \gamma_2 AR - p_1 A_1 z F_1(k,l) - (1+r^{tc})AR) \\
& + \chi_1(p_1 A_1 z F_1(k,l) - AR) + \chi_2 AR \\
& + \tau a'.
\end{aligned}
$$

The FOCs are:

$$
\begin{aligned}
k: \quad & p_1 A_1 z F_{1,k} = \frac{r+\delta}{1 - c\mu - c\chi_1}, && (3.10) \\
l: \quad & p_1 A_1 z F_{1,l} = \frac{w}{1 - c\mu - c\chi_1}, && (3.11) \\
AR: \quad & \frac{1}{c} r^{tc} = \mu(1 + r^{tc} - \gamma_2) + \chi_1 - \chi_2, && (3.12) \\
a': \quad & \frac{1}{c} = \beta \mathbb{E}_{z'} V_{1,a'} + \mu\gamma_1 + \tau. && (3.13)
\end{aligned}
$$

Together with the Envelop condition $V_{1,a} = \frac{1}{c}(1+r)$, we derive the the Euler Equation,

$$
\frac{1}{c} = \beta \mathbb{E}_{z'}[\frac{1}{c'}(1+r)] + \mu\gamma_1 + \tau. \tag{3.14}
$$

In addition, according to Kuhn-Tucker condition, the Lagrangian multipliers and the constraints have the following properties,

$$
\begin{aligned}
& \mu \geq 0, \gamma_1 a' - p_1 A_1 z F_1(k,l) - (1 + r^{tc} - \gamma_2)AR \geq 0, \\
& \chi_1 \geq 0, p_1 A_1 z F_1(k,l) - AR \geq 0 \\
& \chi_2 \geq 0, AR \geq 0 \\
& \tau \geq 0, a' \geq 0,
\end{aligned}
$$

with complementary slackness.

Before proceeding to the proofs of the propositions, we discuss some properties the the value function and policy function in the following Lemma.

Lemma 1 *Given any z, $c(a, z)$ is continuous and strictly increase in a. Given any z, $\mu(a, z)$ is a nonnegative, continuous, and decreasing function in a. In particular, $\mu(a, z)$ is strictly decreasing if $\mu(a, z) > 0$.*

Proof: Notice that the entrepreneur's problem is a standard stochastic dynamic problem. Theorem 9.8 of Stokey and Lucas (1989) implies that $c(a, z)$ is continuous in a and $V_1(a, z)$ is strictly concave in a. The envelope condition $V_{1,a}(a, z) = \frac{1}{c(a,z)}(1 + r)$ then implies that $c(a, z)$ strictly increases in a.

According to the first order condition 3.13 and the non-negativity condition of the Lagrangian multiplier, $\mu(a, z) = max\{\frac{1}{c(a,z)} - \beta\mathbb{E}_{z'}V_{1,a'}, 0\}$. Notice for any given z, $\beta\mathbb{E}_{z'}V_{1,a'}$ is constant in a, it follows that $\mu(a, z)$ is continuous since $c(a, z)$ is continuous. The fact that $c(a, z)$ is strictly increasing in a then implies that $\mu(a, z)$ is strictly increasing in a if $\mu(a, z) > 0$.

A.2.1 Proof of Proposition 1

We prove this proposition by construction. For any given z, define a set $\mathbf{U}^z = \{a : \mu(a, z) = 0\}$. We first show that $\mathbf{U}^z$ is not empty. To see this, consider the unconstraint optimization problem of the entrepreneur. Since the production function $F_1(k, l)$ is decreasing return to scale, the solution to the unconstrained problem exists and is characterized by setting $\mu = 0$ in the first order conditions 3.10, 3.11, 3.12, and 3.13. Let (k^*, l^*, AR^*) denote the unconstrained solution. It is easy to see from the first order conditions that k^*, l^*, AR^* all only depend on z but not on a. Therefore given z, $p_1A_1zF_1(k^*, l^*) + (1 + r^{tc} - \gamma_2)AR^*$ is constant. Notice that $\lim_{a\to\infty} a'(a, z) = \infty$ because otherwise the Envelope condition can not hold with $c = \infty$ and $a' < \infty$. Therefore, there exists a a^* such that $\gamma_1 a'(a^*, z) > p_1 A_1 z F_1(k^*, l^*) + (1 + r^{tc} - \gamma_2) AR^*$. Hence $a^* \in \mathbf{U}^z$.

It is easy to show the set $\mathbf{U}^z$ is an interval and unbounded from above. To see this, take any given $a < \hat{a}$, if $\mu(a, z) = 0$ then $\mu(\hat{a}, z) = 0$ since $\mu(a, z)$ is a decreasing function in a and is bounded below by 0.

Define $g_1(z) = inf \mathbf{U}^z$. By the definition of least lower bound, we know that if $a < g_1(z)$, $a \notin \mathbf{U}^z$, therefore $\mu(a, z) > 0$. And since $\mathbf{U}^z$ is an interval, it has to be true that for any $a > g_1(z)$, $a \in \mathbf{U}^z$, which means that $\mu(a, z) = 0$.

A.2.2 Proof of Proposition 2

The choice of trade credit for intermediate goods entrepreneur is characterized by the first order condition 3.12.

Together with the properties regarding the Lagrangian multipliers and the complementary slackness condition, we can derive the following equations,

$$
\begin{aligned}
c(a, z)\chi_1(a, z) &= \max\{0, r^{tc} - c(a, z)\mu(a, z)(1 - \gamma_2 + r^{tc})\}, \\
c(a, z)\chi_2(a, z) &= \max\{0, c(a, z)\mu(a, z)(1 - \gamma_2 + r^{tc}) - r^{tc}\}.
\end{aligned}
$$

Use the Euler equation, we can show that $c(a, z)\mu(a, z) = 1 - c(a, z)\beta \mathbb{E}_{z'} V_{1,a'}$. Since $c(a, z)$ is a strictly increasing and continuous function in a, it is clear that $c(a, z)\mu(a, z)$ is a decreasing and continuous function in a. Furthermore, $c(a, z)\mu(a, z)$ is a strictly decreasing function if $\mu(a, z) > 0$.

Consider a function of a: $q(a) = r^{tc} - c(a, z)\mu(a, z)(1 + r^{tc})$. It is easy to see that $q(a) = 0$ implies $c(a, z)\mu(a, z) = \frac{r^{tc}}{1+r^{tc}} > 0$. Therefore, there exists an unique a^* such that $c(a^*, z)\mu(a^*, z) = \frac{r^{tc}}{1+r^{tc}}$. Define $h_1(z) = a^*$.

For any $a < h_1(z)$, we know that $c(a, z)\chi_2(a, z) > 0$. Since $c(a, z) \geq 0$, this means that $\chi_2(a,z) > 0$. Using the complementary slackness condition, this means $AR(a,z) = 0$. Similarly, for any $a > h_1(z)$, $c(a,z)\chi_1(a,z) > 0$, therefore $\chi_1(a,z) > 0$. It then follows that $AR(a,z) = p_1 A_1 z F_1(k(a,z), l(a,z))$.

A.2.3 Proof of Proposition 3

For any given z such that $\mu(g_1(z),z) = 0$, we know that $c(g_1(z),z)\chi_1(g_1(z),z) = r^{tc} > 0$. Since $\gamma_2 \in [0,1]$ and $c(a,z)\mu(a,z)$ is strictly decreasing and continuous in a, there exists $\hat{a} < g(z)$ such that $r^{tc} - c(\hat{a},z)\mu(\hat{a},z)(1-\gamma_2+r^{tc}) > 0$, therefore $AR(\hat{a},z) = p_1 A_1 z F_1(k(\hat{a},z), l(\hat{a},z))$. Notice $r^{tc} - c(h_1(z),z)\mu(h_1(z),z)(1-\gamma_2+r^{tc}) > 0$. Since $c(a,z)\mu(a,z)$ decreases with a, we know that $h_1(z) \leq \hat{a}$. Then it follows that $h_1(z) < g_1(z)$.

A.3 Computation of the benchmark model

In this section, we describe the algorithms of computing the benchmark model. Section A.3.1 contains the algorithms to compute the stationary equilibrium. Section A.3.2 contains the algorithms to compute the transitional dynamics. The algorithms to compute the counterfactual model is very similar to the benchmark model, only with different set of FOCs, budget constraint, and working capital constraint. Hence they are omitted here.

A.3.1 Stationary equilibrium

- Guess equilibrium prices r, w, p_1, r^{tc}.
- Given the prices, solve the household problem.
- Given the prices, solve the entrepreneurs problem as follows,

  1. Discretize the state space.
  2. Guess policy function $c(a,z)$.

3. For each (a, z), assume that the entrepreneur is unconstrained, i.e. $\mu(a, z) = 0$. Solve for the system of equations that consists of FOCs and budget constraint.
4. Check whether the working capital constraint is satisfied with the solution to the above system of equations.
5. If the working capital constraint is not satisfied, it means that $\mu(a, z) > 0$ and working capital constraint holds with equality. Solve the systmem of equations that consists of FOCs, budget constraint, and working capital constraint (with equality).
6. Use the Euler equation to update the policy function $c(a, z)$ until it converges.

- Given any arbitrary distribution of (a, z), iterate using the policy functions derived above until a stationary distribution is reached.
- Generate the aggregate statistics of the four markets: capital, labor, intermediate goods and trade credit market.
- Update (r, w, p_1, r^{tc}) until the four markets clear simultaneously.

A.3.2 Transitional dynamics

To compute the transitional dynamics of the economy, we consider a transition path of $T = 100$ periods. The economy is at the initial stationary equilibrium level in period $t = 1$ and we assume that it converges back to the initial stationary equilibrium at period $t = T$.

- Guess a sequence of prices $\{r_t, w_t, p_{1,t}, r_t^{tc}\}_{t=2}^{T-1}$.
- Backward induction. For each $t = T - 1, T - 2, ..., 2$.
  - Discretize the state space.
  - Given prices, solve the household problem for period t.

- Given prices, solve the entrepreneurs policy functions for period t.

  1. Guess $c_t(a,z)\mu_t(a,z) = 0$, solve the system of equations that consists of FOCs of period t, budget constraint, and Euler equations (with the next period policy function $c_{t+1}(a,z)$ known).

  2. Check whether the working capital constraint is satisfied under the above solution.

  3. If the working capital is not satisfied, $c_t(a,z)\mu_t(a,z) > 0$ and the working capital constraint holds with equality. Solve the system of equations that consists of FOCs of period t, budget constraint, Euler equations (with the next period policy function $c_{t+1}(a,z)$ known), and working capital constraint with equality.

- Forward induction. The first period stationary distribution $\Phi_{1,1}(a,z)$ and $\Phi_{2,1}(a,z)$ is set to be the stationary equilibrium distribution. Using the policy functions for period $t = 2, ..., T-1$, compute the distribution along the transition path ($\Phi_{1,t}(a,z)$ and $\Phi_{2,t}(a,z)$)

- Generate aggregate statistics for the four markets in every period $t = 2, ..., T-1$ using the policy functions and the distributions.

- Update $\{r_t, w_t, p_{1,t}, r_t^{tc}\}_{t=2}^{T-1}$ until the four markets clear simultaneously in each period $t = 2, ..., T-1$.

B Appendix to Chapter 3

B.1 Data Sources

Official Statistics The main source for the official statistics of cities is the *China City Statistical Yearbooks (1988–2010)*.[70] In addition, we include three other city statistical collections. First, we take the GDP data for the years 1992 and 1993 from the *New China City in 50 Years Statistical Collection*, since these years are missing in the *China City Statistical Yearbooks*. Second, for a subset of cities, we collect GDP and investment data for the period of 1978-1988 from the *New China in 60 Years Provincial Statistical Collection*.[71] Third, we obtain population and educational attainment data from the *China Population Census* for the years 1990, 2000, and 2010.

The main source for province-level statistics is the *New China in 60 Years Statistical Collection*. We obtain the province-level price indexes, including the GDP and investment deflator, from this data set.

Light and Digital Maps Light intensity at night, an alternative measure for local economic activity, is provided by the National Geographical Data Center.[72] The data is available in cleaned form (taking into account clouds, forest fires, gas flaring, etc.) and on an annual basis from 1992 to 2013, but we focus on the period until 2010 for which we also have the official yearbook data. The light images show the intensity of light on each pixel (approximately one square kilometer) on an integer scale from 0 (no light) to 63 (maximum light). In order to map the light intensity of pixels to the administrative entities of cities, we use digital maps of Chinese cities from 2010.[73] This allows us to measure light within adminis-

[70] We obtained the data from China National Knowledge Infrastructure (CNKI).

[71] These are cities from the following provinces: Fujian, Guizhou, Hebei, Heilongjiang, Henan, Inner Mongolia, Jiangsu, Shaanxi, Shandong, and Shanxi.

[72] See http://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html

[73] The digital maps for several levels of administrative units of the People's Republic of China were obtained from the Asian Spatial Information and Analysis Network (ACASIAN), where they were produced by Dr.

trative boundaries that are held constant over time in order to rule out the possibility that changes in administrative boundaries mechanically affect the outcome measure. Changes in administrative boundaries could especially be a concern for the statistical yearbook data on urban cores, where such changes are more frequent than for the entire city that includes the periphery. Light measured within constant boundaries does not have this problem, and the analysis of the effect of SEZ on light exploits this advantage and focuses on the urban core of each city rather than the area that includes the periphery.

Light is measured by different satellites over time, and they show different light intensities because of differences in their calibration. These differences do not matter for our regression analysis since they affect the entire country within a given year and the calibration is therefore absorbed by year fixed effects. For the descriptive statistics, we calculate the growth rate for each satellite separately and then report the average ("within satellite") growth rate.[74]

Establishment of SEZ We collect data on the establishment of SEZ from the following sources: 1) the website of the Ministry of Commerce,[75] 2) the website of China's Development Zones,[76] and 3) the Report on SEZ (of Commerce, 2006). These sources provide information on the year in which SEZ were established, their types, and their location.

B.2 Sample Selection

In our main estimations, we focus on a sample of cities for the years 1988–2010. The sample is unbalanced because of the creation of new cities: in the year 1988 the sample has 171 cities, and this number increases to 276 in the year 2010.[77] Our sample covers all provinces

L. W. Crissman.

[74]For example, if a satellite delivers light images for four years, then we can calculate the growth rate for all but the first year. We do this for each satellite and then average over all satellites and years.

[75]See http://english.mofcom.gov.cn.

[76]See http://www.cdz.cn/.

[77]See Table 1 in Chung and chiu Lam (2004) for a more detailed assessment of the increase in the number of cities in China.

in China except for Tibet, Hainan, and the province-level cities Beijing, Shanghai, Tianjin, and Chongqing. We also exclude the cities of the first wave of CSEZ, Shenzhen, Zhuhai, Shantou and Xiamen.

We discuss below in detail our sample selection criteria.

Prefecture- v.s. County-level Cities The *China City Statistical Yearbooks* contain statistics for both prefecture-level and county-level cities. Our sample focuses on prefecture-level cities. Therefore our sample does not include cities that are county-level throughout the sample period. Some of the prefecture-level cities were promoted from county-level cities. For those cities, we only keep the years after the promotion in our sample.

Prefecture-Level Cities and their Urban Core The *China City Statistical Yearbooks* report data separately for the entire prefecture-level city (*shi*) and its urban core (*shixiaqu*). Our estimations are, unless stated otherwise, carried out for the entire city.

The urban core is more industrialized than the entire area of a city that includes the periphery. Most SEZ are located inside and at the edge of the urban core (Zeng, 2010). In fact, in our sample, all state-level SEZ reside inside of urban cores. However, focusing the analysis on the entire city has two advantages. First, the administrative boundary of cities change less frequently than that of urban cores. Although we can control for changes in land area as reported in the statistical yearbooks, border changes could remain a concern especially if they correlate with the establishment of SEZ. Second, it is difficult to account for workers who live in the periphery of cities but commute to their work in the urban core. This issue is less severe at the city-level since the cross-city commuting population is smaller.

Sample Period Our main estimation covers the period of 1988 to 2010. The GDP data in the *China City Statistical Yearbooks* only go back to the year 1988. For a subset of cities, we can obtain GDP data from the *New China in 60 Years Provincial Statistical Collection* for the period of 1978–1988. We do not use the early data in our main analysis due the lack of information on population and land area. Instead, we use them in a robustness check of the effects of SEZ and also in the decomposition in order to get a better estimate of initial capital stock of these cities in the year 1988.

Inland Sample When we restrict the sample to cities from inland provinces, we define the following provinces as inland: Anhui, Gansu, Guangxi, Guizhou, Heilongjiang, Henan, Hubei, Hunan, Jiangxi, Jilin, Ningxia, Qinghai, Shaanxi, Shanxi, Sichuan, Xinjiang, Yunnan, and Inner Mongolia. This classification was not purely based on access to the sea, but also considers whether the provinces were part of the reform wave targeted towards inland regions.

Population One concern with the population data from the *China City Statistical Yearbooks* is that it only covers the registered population, that is, people with *hukou*. This would lead to a bias in the estimation if reformed cities attracted more (or less) migrants than non-reformed cities.

To address this issue, we make use of the fact that population data in the *China Population Census* covers both *hukou* and *non-hukou* residents. For 1990, 2000, and 2010, we can take the population data from the *China Population Census.* We then interpolate the years in between the *Census* using the relative year-by-year population growth rate from the *China City Statistical Yearbooks*.[78] This imputed population data is used in our main estimations and decomposition exercise.

Measurement Errors Some cities in our sample change borders. Along with the border change there are changes in population size and land area which can be measured using

[78] We apply the Stata command *ipolate* to observations *between* the census years. This command interpolates the missing observations in the census data based on population data from the yearbooks, but whenever available it uses the data from the census. We construct the years before the first and after the last census year by directly applying the growth rate of the yearbook population data to the observation from the census.

the yearbook data. Changes in these variables should happen in the same year, but we found cases where there is clearly a delayed discrete jump in some of the variables relative to others. Therefore, we drop two observations where we observed a jump in land area but no corresponding jump in population.[79]

B.3 Decomposition

The following paragraphs provide information on the construction of physical capital per capita and efficiency unit of labor (raw labor × average human capital) used in the decomposition exercise.

Physical Capital Stock We apply the perpetual inventory approach to construct the physical capital stock in each city. The physical capital (K_{ipt}) is the sum of the physical capital stock after depreciation and new investment (I_{ipt}), such that

$$K_{ipt} = (1 - \delta_k)K_{ipt-1} + I_{ipt}/deflator_{pt}^{Inv}.$$

The deflator for new investment, $deflator_{pt}^{Inv}$, is province-specific. We set δ_k, the annual depreciation rate for physical capital, to be 0.08.[80]

In order to carry out the perpetual inventory approach, we need a reasonable estimate for the physical capital stock of the initial year, which is the year of 1988 given our sample period.

For a subset of cities whose investment data go back to 1978, we derive the capital stock

[79] The two observations are Chuzhou in the year 1993 and Dazhou in the year 1998.

[80] Given the large amount of creative destruction that took place in China, we pick the number to be higher than other cross-country growth accounting exercises.

for those cities in the year 1978 as follows

$$K_{ip1978} = \frac{I_{ip1978}}{g_{1978} + \delta_k},$$

where I_{ip1978} is the new investment in the year 1978 and g_{1978} is the average growth rate of real physical capital stock before 1978.[81] This is the steady state formula for the physical capital stock of a Solow-type growth model (Caselli, 2005).

For those cities whose investment data begins in 1988, we approximate the initial physical capital stock in 1988 using the same formula,

$$K_{ip1988} = \frac{I_{ip1988}}{g_{1988} + \delta_k},$$

where g_{1988} is the average growth rate of physical capital stock before 1988.

Size of Labor Force We use population as an approximation for employment in each city because the number of employed persons reported in the *China City Statistical Yearbooks* has some drawbacks. The most important drawback is that there is a substantial drop in the number of employed persons in the year 1998, the reason for which is unclear to us. Two reasons could potentially contribute to this large drop. The first reason is that the reform of state-owned enterprises laid off a large number of redundant workers around 1998.[82] The second reason is that perhaps the definition of employed persons changes in 1998. Specifically, before 1998, the employed persons include people who are registered as workers. After 1998, the number only includes people who are registered and are currently working in that city.[83]

[81] The growth rate of the real physical capital stock, g_{1978}, is calculated using the national physical capital stock. See the website of Kuai Wai Li (http://personal.cb.cityu.edu.hk/efkwli/ChinaData.html) and Li *et al.* (2009) for the detailed construction of the data.

[82] According to Dong and Putterman (2003), the labor redundancy rate of SOEs is 30% in 1992.

[83] Wu (2011) provides a detailed discussion of the issues with the employment data.

Human Capital Following Hall and Jones (1999), we use the average educational attainment (years of schooling) as an approximation for the level human capital of the cities, such that

$$h_{ipt} = e^{\phi_t(s_{ipt})},$$

where s_{ipt} is the average years of schooling and $\phi_t(.)$ is a piece-wise linear function whose slopes represent the return to schooling. To construct $\phi_t(.)$, we take the estimation for the return to schooling in China over the period 1988-2009 from Li et al. (2009).[84]

The only data source that reports educational attainment at the city level is the *China Population Census*, and it is only available for the years 1990, 2000 and 2010. We do a simple linear interpolation (extrapolation if needed) to obtain the approximation of human capital for the other years in our sample period.

B.4 Additional Tables and Figures

Table A1: State- and province-level zones in selected provinces

Province	#S	#P	Avg indus-output share of S	Avg indus-output share of P
Jiangsu	12	113	3.13%	0.55%
Guangdong	14	56	4.89%	0.56%
Zhejiang	8	57	4.09%	1.18%

Source: WEFore (2010). The table displays the number of state-level development zones (#S) and province-level development zones (#P) in three provinces: Jiangsu, Guangdong, and Zhejiang. In the last two columns, it also displays the average share of the state-level and province-level zones in the industrial output of each province. The data is for the year 2009.

[84]The estimation is not available for the year 2010. We simply assume that the return to schooling did not change between 2009 and 2010, i.e. $\phi_{2010}(.) = \phi_{2009}(.)$.

Table A2: Different types of zones

	(1)	(2)	(3)	(4)	(5)	(6)
ETDZ	0.155***	0.147***	0.108***	0.143*	0.104	0.0738
	(0.0454)	(0.0389)	(0.0386)	(0.0777)	(0.0695)	(0.0667)
HIDZ	0.137***	0.0760**	0.0743**	0.0718	0.104*	0.0688
	(0.0360)	(0.0305)	(0.0293)	(0.0663)	(0.0528)	(0.0471)
EPZ	-0.0130	0.0236	-0.0102	0.0275	0.0832	0.0386
	(0.0388)	(0.0326)	(0.0312)	(0.0963)	(0.0983)	(0.0874)
Other types	0.0840	0.113**	0.0797**	0.271***	0.205**	0.281***
	(0.0540)	(0.0447)	(0.0401)	(0.0770)	(0.0923)	(0.0654)
Province-level SEZ	0.0259	0.00183	-0.00770	0.0208	-0.0122	-0.00669
	(0.0223)	(0.0181)	(0.0164)	(0.0309)	(0.0251)	(0.0233)
Dependent variable (log)	GDP	GDP	GDPpc	GDP	GDP	GDPpc
Controlling for log land area	No	Yes	Yes	No	Yes	Yes
Controlling for log population	No	No	Yes	No	No	Yes
Sample	Full	Full	Full	Inland	Inland	Inland
N	5392	5321	5269	2864	2798	2768
Adj. Rsq.	0.961	0.976	0.975	0.949	0.972	0.971

Notes: The dependent variable is the logarithm of annual GDP. The reform indicators for each type of zone take on value 1 in the year after the introduction of the zone. All specifications include city fixed effects and the interaction of province-year dummies. Standard errors are clustered at the city level: * $p \leq 0.10$, ** $p \leq 0.05$, *** $p \leq 0.01$.

Table A3: Baseline specification (census years only)

	(1)	(2)	(3)	(4)	(5)	(6)
State-level SEZ	0.222***	0.177***	0.141***	0.280***	0.250***	0.175**
	(0.0427)	(0.0373)	(0.0374)	(0.0858)	(0.0715)	(0.0720)
Province-level SEZ	0.0633	0.00771	-0.0259	0.0681	-0.0103	-0.0238
	(0.0429)	(0.0328)	(0.0308)	(0.0601)	(0.0442)	(0.0408)
Dependent variable (log)	GDP	GDP	GDPpc	GDP	GDP	GDPpc
Controlling for log land area	No	Yes	Yes	No	Yes	Yes
Controlling for log population	No	No	Yes	No	No	Yes
Sample	Full	Full	Full	Inland	Inland	Inland
N	723	708	694	393	378	366
Adj. Rsq.	0.979	0.985	0.984	0.974	0.984	0.984

Notes: The dependent variable is the logarithm of annual GDP or GDP per capita. State-level (respectively province-level) SEZ is a dummy switching on in the year after the introduction of any SEZ at that level. All specifications include city fixed effects and the interaction of province-year dummies. Standard errors are clustered at the city level: * $p \leq 0.10$, ** $p \leq 0.05$, *** $p \leq 0.01$. The sample is restricted to the census years 1990, 2000, and 2010.

Table A4: Effect on human capital (census years only)

	(1)	(2)	(3)	(4)	(5)	(6)
State-level SEZ	-0.0102*	-0.0203***	0.0306***	-0.0147	-0.0444***	0.0591***
	(0.00606)	(0.00681)	(0.00538)	(0.0113)	(0.0103)	(0.00643)
Province-level SEZ	-0.00664	0.00778	-0.00114	-0.0174***	0.0160**	0.00143
	(0.00526)	(0.00626)	(0.00352)	(0.00606)	(0.00688)	(0.00516)
Dependent Variable	share low edu.	share mid. edu.	share high edu.	share low edu.	share mid. edu.	share high edu.
Sample	Full	Full	Full	Inland	Inland	Inland
N	582	582	582	303	303	303
Adj. Rsq.	0.973	0.884	0.933	0.976	0.908	0.947

Notes: The dependent variable is the share of the population over 6 with a low level education (primary school or lower), share of the population over 6 with an intermediate level education (junior and senior high school), or share of the population over 6 with a high level education (college or above). All specifications include city fixed effects, the interaction of province-year dummies, controls for land area, and controls for population. Standard errors are clustered at the city level: * $p \leq 0.10$, ** $p \leq 0.05$, *** $p \leq 0.01$.

Table A5: Spillovers within rings around the city

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
State-level SEZ	0.123***	0.120***	0.118***	0.118***	0.180***	0.185***	0.194***	0.179***
	(0.0301)	(0.0299)	(0.0296)	(0.0289)	(0.0571)	(0.0537)	(0.0561)	(0.0520)
Province-level SEZ	-0.00325	-0.00219	-0.00188	-0.00275	-0.0136	-0.0139	-0.0156	-0.0134
	(0.0182)	(0.0183)	(0.0183)	(0.0184)	(0.0255)	(0.0255)	(0.0258)	(0.0255)
SEZ in 0-50 ring (distance)	0.00961				0.187*			
	(0.0863)				(0.101)			
SEZ in 50-100 ring (distance)	-0.0128				0.135			
	(0.0792)				(0.103)			
SEZ in 100-200 ring (distance)	-0.0711				0.0594			
	(0.0742)				(0.100)			
SEZ in 200-400 ring (distance)	-0.0798				-0.00959			
	(0.0675)				(0.0831)			
SEZ in 0-50 ring (driving time)		0.149				0.250**		
		(0.0957)				(0.112)		
SEZ in 50-100 ring (driving time)		0.123				0.207**		
		(0.0771)				(0.103)		
SEZ in 100-200 ring (driving time)		0.0902				0.169		
		(0.0748)				(0.105)		
SEZ in 200-400 ring (driving time)		0.0644				0.117		
		(0.0660)				(0.0761)		
SEZ in 0-50 ring (transport costs, 10 cat)			0.191**				0.185	
			(0.0752)				(0.136)	
SEZ in 50-100 ring (transport costs, 10 cat)			0.135*				0.163	
			(0.0701)				(0.118)	
SEZ in 100-200 ring (transport costs, 10 cat)			0.0908				0.139	
			(0.0643)				(0.113)	
SEZ in 200-400 ring (transport costs, 10 cat)			0.0820				0.0399	
			(0.0543)				(0.110)	
SEZ in 0-50 ring (transport costs, 7 cat)				0.188**				0.176*
				(0.0752)				(0.0999)
SEZ in 50-100 ring (transport costs, 7 cat)				0.123				0.164
				(0.0745)				(0.110)
SEZ in 100-200 ring (transport costs, 7 cat)				0.0654				0.0827
				(0.0727)				(0.107)
SEZ in 200-400 ring (transport costs, 7 cat)				0.0546				0.0170
				(0.0635)				(0.0925)
Sample	Full	Full	Full	Full	Inland	Inland	Inland	Inland
N	5321	5321	5321	5321	2798	2798	2798	2798
Adj. Rsq.	0.975	0.975	0.975	0.976	0.972	0.972	0.972	0.972

Notes: The dependent variable is the logarithm of annual GDP. The independent variables are indicators for zones within the corresponding distances. State-level (respectively province-level) SEZ is a dummy switching on in the year after the introduction of any SEZ at that level. All specifications include city fixed effects, the interaction of province-year dummies, and controls for land area. Standard errors are clustered at the city level: * $p \leq 0.10$, ** $p \leq 0.05$, *** $p \leq 0.01$.

Table A6: Exposure to other zones and market access

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
State-level SEZ	0.124***	0.124***	0.120***	0.120***	0.121***	0.121***	0.217***	0.197***	0.192***	0.187***	0.186***	0.187***
	(0.0300)	(0.0295)	(0.0297)	(0.0296)	(0.0296)	(0.0295)	(0.0580)	(0.0619)	(0.0594)	(0.0596)	(0.0592)	(0.0592)
Province-level SEZ	-0.00253	-0.000214	-0.00127	-0.00120	-0.000867	-0.000838	-0.0124	-0.0118	-0.0114	-0.0103	-0.0104	-0.0104
	(0.0183)	(0.0183)	(0.0183)	(0.0183)	(0.0183)	(0.0183)	(0.0251)	(0.0249)	(0.0244)	(0.0244)	(0.0247)	(0.0247)
Exposure to other SEZ (distance)	0.0646						0.311					
	(0.0767)						(0.229)					
Market access (distance)	-1.073**						-0.669					
	(0.509)						(0.962)					
Exposure to other SEZ (driving time)		0.0623						0.309				
		(0.0804)						(0.283)				
Market access (driving time)		-0.811						-0.845				
		(0.521)						(1.174)				
Exposure to other SEZ (transport costs, 10 cat)			0.0513						0.340*			
			(0.0639)						(0.204)			
Market access (transport costs, 10 cat)			-0.331						0.344			
			(0.457)						(0.695)			
Exposure to other SEZ (transport costs, 7 cat)				0.0462						0.279		
				(0.0657)						(0.205)		
Market access (transport costs, 7 cat)				-0.349						0.339		
				(0.455)						(0.705)		
Exposure to other SEZ (transport costs, 7 cat, 3km)					0.0491						0.243	
					(0.0604)						(0.197)	
Market access (transport costs, 7 cat, 3km)					-0.405						0.220	
					(0.417)						(0.631)	
Exposure to other SEZ (transport costs, 7 cat, 3km)						0.0573						0.245
						(0.0596)						(0.197)
Market access (transport costs, 7 cat, 3km)						-0.385						0.218
						(0.415)						(0.632)
Sample	Full	Full	Full	Full	Full	Full	Inland	Inland	Inland	Inland	Inland	Inland
N	5254	5321	5321	5321	5321	5321	2775	2798	2798	2798	2798	2798
Adj. Rsq.	0.977	0.975	0.975	0.975	0.975	0.975	0.972	0.972	0.972	0.972	0.972	0.972

Notes: The dependent variable is the logarithm of annual GDP. The independent variable *exposure to other zones* is a measure of spillover intensity, and we expect a positive coefficient in case of positive spillovers. State-level (respectively province-level) SEZ is a dummy switching on in the year after the introduction of any SEZ at that level. All specifications include city fixed effects and the interaction of province-year dummies. All specifications also control for the logarithm of land area. Standard errors are clustered at the city level: * $p \leq 0.10$, ** $p \leq 0.05$, *** $p \leq 0.01$.

Table A7: Spillover effect on investment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
State-level SEZ	0.149***	0.146***	0.147***	0.147***	0.414***	0.399***	0.400***	0.396***
	(0.0497)	(0.0491)	(0.0491)	(0.0490)	(0.0692)	(0.0727)	(0.0696)	(0.0701)
Province-level SEZ	0.0310	0.0316	0.0315	0.0315	-0.0229	-0.0255	-0.0256	-0.0242
	(0.0317)	(0.0317)	(0.0317)	(0.0317)	(0.0455)	(0.0465)	(0.0469)	(0.0467)
Log km distance next SEZ	-0.0262				-0.149**			
	(0.0437)				(0.0674)			
Log driving time next SEZ		0.0132				-0.117		
		(0.0518)				(0.0859)		
Log transport costs next SEZ (10 cat)			0.00301				-0.126**	
			(0.0400)				(0.0605)	
Log transport costs next SEZ (7 cat)				0.00555				-0.115*
				(0.0412)				(0.0649)
Sample	Full	Full	Full	Full	Inland	Inland	Inland	Inland
N	4950	5016	5016	5016	2587	2609	2609	2609
Adj. Rsq.	0.957	0.957	0.957	0.957	0.954	0.954	0.954	0.954

Notes: The dependent variable is the logarithm of fixed asset investment. Investment is measured in current prices. State-level (respectively province-level) SEZ is a dummy switching on in the year after the introduction of any SEZ at that level. All specifications include city fixed effects and the interaction of province-year dummies. All specifications also control for the logarithm of population and land area. Standard errors are clustered at the city level: * $p \leq 0.10$, ** $p \leq 0.05$, *** $p \leq 0.01$.

Table A8: Spillover effect on TFP

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
State-level SEZ	0.0707**	0.0651**	0.0645**	0.0637**	0.0831*	0.0675	0.0654	0.0640
	(0.0276)	(0.0272)	(0.0270)	(0.0270)	(0.0425)	(0.0437)	(0.0417)	(0.0423)
Province-level SEZ	-0.00252	-0.00256	-0.00207	-0.00177	0.0192	0.0187	0.0190	0.0199
	(0.0185)	(0.0185)	(0.0184)	(0.0185)	(0.0258)	(0.0259)	(0.0259)	(0.0258)
Log km distance next SEZ	-0.0320				-0.0493			
	(0.0202)				(0.0308)			
Log driving time next SEZ		-0.0693***				-0.0591*		
		(0.0244)				(0.0332)		
Log transport costs next SEZ (10 cat)			-0.0627***				-0.0643**	
			(0.0187)				(0.0275)	
Log transport costs next SEZ (7 cat)				-0.0607***				-0.0601**
				(0.0186)				(0.0270)
Sample	Full	Full	Full	Full	Inland	Inland	Inland	Inland
N	3953	4019	4019	4019	1872	1895	1895	1895
Adj. Rsq.	0.959	0.958	0.958	0.958	0.953	0.953	0.953	0.953

Notes: The dependent variable is the logarithm of TFP. TFP is computed using full-sample unrestricted production function estimation. State-level (respectively province-level) SEZ is a dummy switching on in the year after the introduction of any SEZ at that level. All specifications include city fixed effects, the interaction of province-year dummies, controls for land area, and controls for population. Standard errors are clustered at the city level: * $p \leq 0.10$, ** $p \leq 0.05$, *** $p \leq 0.01$.

Table A9: Spillover effect on population

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
State-level SEZ	0.0451	0.0474*	0.0467*	0.0467*	0.0651	0.0652	0.0627	0.0632
	(0.0286)	(0.0281)	(0.0282)	(0.0281)	(0.0406)	(0.0409)	(0.0397)	(0.0395)
Province-level SEZ	0.0151	0.0186	0.0188	0.0188	-0.00202	-0.00270	-0.00225	-0.00224
	(0.0118)	(0.0130)	(0.0130)	(0.0130)	(0.0125)	(0.0125)	(0.0124)	(0.0124)
Log km distance next SEZ	-0.0433*				-0.0397*			
	(0.0225)				(0.0228)			
Log driving time next SEZ		-0.0231				-0.0254		
		(0.0250)				(0.0263)		
Log transport costs next SEZ (10 cat)			-0.0149				-0.0136	
			(0.0187)				(0.0202)	
Log transport costs next SEZ (7 cat)				-0.0162				-0.0182
				(0.0191)				(0.0200)
Sample	Full	Full	Full	Full	Inland	Inland	Inland	Inland
N	5207	5275	5275	5275	2746	2769	2769	2769
Adj. Rsq.	0.836	0.822	0.822	0.822	0.884	0.883	0.883	0.883

Notes: The dependent variable is the logarithm of population. State-level (respectively province-level) SEZ is a dummy switching on in the year after the introduction of any SEZ at that level. All specifications include city fixed effects, the interaction of province-year dummies, and controls for land area. Standard errors are clustered at the city level: * $p \leq 0.10$, ** $p \leq 0.05$, *** $p \leq 0.01$.

Table A10: Spillover effect on FDI

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
State-level SEZ	0.461***	0.459***	0.459***	0.458***	0.904***	0.918***	0.905***	0.903***
	(0.145)	(0.143)	(0.142)	(0.142)	(0.302)	(0.302)	(0.299)	(0.300)
Province-level SEZ	0.129*	0.127*	0.127*	0.128*	0.0767	0.0710	0.0739	0.0752
	(0.0756)	(0.0745)	(0.0745)	(0.0745)	(0.123)	(0.123)	(0.123)	(0.123)
Log km distance next SEZ	-0.102				-0.102			
	(0.140)				(0.261)			
Log driving time next SEZ		-0.0846				-0.218		
		(0.145)				(0.270)		
Log transport costs next SEZ (10 cat)			-0.0832				-0.151	
			(0.129)				(0.231)	
Log transport costs next SEZ (7 cat)				-0.0754				-0.137
				(0.130)				(0.240)
Sample	Full	Full	Full	Full	Inland	Inland	Inland	Inland
N	4491	4544	4544	4544	2305	2318	2318	2318
Adj. Rsq.	0.702	0.702	0.702	0.702	0.637	0.635	0.635	0.635

Notes: The dependent variable is the logarithm of effective use of FDI. FDI is measured in current prices. State-level (respectively province-level) SEZ is a dummy switching on in the year after the introduction of any SEZ at that level. All specifications include city fixed effects and the interaction of province-year dummies. All specifications also control for the logarithm of land area. Standard errors are clustered at the city level: * $p \leq 0.10$, ** $p \leq 0.05$, *** $p \leq 0.01$.

Table A11: Effect on core and periphery

Panel A: Urban core

	(1)	(2)	(3)	(4)	(5)	(6)
State-level SEZ	0.190***	0.147***	0.0952***	0.268***	0.211***	0.155***
	(0.0417)	(0.0343)	(0.0285)	(0.0550)	(0.0533)	(0.0554)
Province-level SEZ	-0.000288	-0.00683	-0.00108	-0.0153	-0.0301	-0.0171
	(0.0256)	(0.0228)	(0.0201)	(0.0289)	(0.0275)	(0.0268)
Dependent variable (log)	GDP	GDP	GDPpc	GDP	GDP	GDPpc
Controlling for log land area	No	Yes	Yes	No	Yes	Yes
Controlling for log population	No	No	Yes	No	No	Yes
Sample	Full	Full	Full	Inland	Inland	Inland
N	5152	5138	5109	2688	2684	2677
Adj. Rsq.	0.960	0.969	0.963	0.961	0.966	0.959

Panel B: Periphery only

	(1)	(2)	(3)	(4)	(5)	(6)
State-level SEZ	0.219**	0.107***	0.157***	0.310	0.141**	0.248***
	(0.0984)	(0.0328)	(0.0314)	(0.246)	(0.0580)	(0.0531)
Province-level SEZ	0.0797	-0.00517	-0.00518	0.0929	0.00796	0.000886
	(0.0501)	(0.0222)	(0.0176)	(0.0658)	(0.0324)	(0.0224)
Dependent variable (log)	GDP	GDP	GDPpc	GDP	GDP	GDPpc
Controlling for log land area	No	Yes	Yes	No	Yes	Yes
Controlling for log population	No	No	Yes	No	No	Yes
Sample	Full	Full	Full	Inland	Inland	Inland
N	4943	4913	4863	2560	2546	2519
Adj. Rsq.	0.865	0.966	0.977	0.865	0.962	0.976

Notes: The dependent variable is the logarithm of annual GDP or GDP per capita. State-level (respectively province-level) SEZ is a dummy switching on in the year after the introduction of any SEZ at that level. All specifications include city fixed effects and the interaction of province-year dummies. Standard errors are clustered at the city level: * $p \leq 0.10$, ** $p \leq 0.05$, *** $p \leq 0.01$. Panel A reports the results for the urban core, and Panel B reports the results for the periphery only (the city excluding the urban core).

Table A12: Lags of population, population density, and land area

	(1)	(2)	(3)	(4)	(5)	(6)
State-level SEZ	0.0702**	0.0508	0.0718**	0.0583**	0.0553**	0.0599**
	(0.0301)	(0.0331)	(0.0294)	(0.0275)	(0.0268)	(0.0264)
Province-level SEZ	-0.0150	-0.0247	-0.0150	-0.00428	-0.00696	-0.00305
	(0.0163)	(0.0171)	(0.0162)	(0.0187)	(0.0185)	(0.0185)
1-period lag of log population	-0.209***			-0.0564		
	(0.0561)			(0.0617)		
1-period lag of log population density		-0.0297	-0.314**		-0.0725	-0.144
		(0.108)	(0.129)		(0.0898)	(0.136)
1-period lag of log landarea			-0.224***			-0.0638
			(0.0600)			(0.0626)
Dependent variable	GDPpc	GDPpc	GDPpc	TFP	TFP	TFP
N	5019	4990	4990	3830	3827	3827
Adj. Rsq.	0.972	0.970	0.973	0.956	0.957	0.957

Notes: The dependent variables are the logarithm of annual GDP per capita and the logarithm of TFP. TFP is computed using a full-sample unrestricted production function estimation. State-level (respectively province-level) SEZ is a dummy switching on in the year after the introduction of any SEZ at that level. All specifications include city fixed effects, the interaction of province-year dummies, and controls for land area. Standard errors are clustered at the city level: * $p \leq 0.10$, ** $p \leq 0.05$, *** $p \leq 0.01$.

Table A13: Allowing different city fixed effects after border changes

	(1)	(2)	(3)	(4)
State-level SEZ	0.111***	0.162***	0.0543*	0.0455
	(0.0304)	(0.0569)	(0.0294)	(0.0504)
Province-level SEZ	-0.0156	-0.00627	-0.00861	0.0222
	(0.0186)	(0.0247)	(0.0198)	(0.0274)
Sample	Full	Inland	Full	Inland
Dependent variable (log)	GDP	GDP	TFP	TFP
Controlling for log population	No	No	Yes	Yes
N	5321	2798	4019	1895
Adj. Rsq.	0.991	0.991	0.979	0.974

Notes: The dependent variable is the logarithm of annual GDP per capita or TFP. TFP is computed using a full-sample unrestricted production function estimation. State-level (respectively province-level) SEZ is a dummy switching on in the year after the introduction of any SEZ at that level. All specifications include separate fixed effects for cities when they change borders. All regressions also control for the interaction of province-year dummies. Standard errors are clustered at the city level: * $p \leq 0.10$, ** $p \leq 0.05$, *** $p \leq 0.01$.

Table A14: Interactions of the reform indicator with initial characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
State-level SEZ	0.0742*	0.0709*	0.139***	0.0983**	-0.00393	0.0346	0.305***	0.122
	(0.0391)	(0.0406)	(0.0405)	(0.0447)	(0.0824)	(0.0933)	(0.0802)	(0.0896)
Post-reform indicator x Above-median population	0.0923*							
	(0.0518)							
Post-reform indicator x Above-median population density		0.0865*						
		(0.0509)						
Post-reform indicator x Above-median GDP p.c.			-0.0473					
			(0.0559)					
Post-reform indicator x Above-median market access				0.0369				
				(0.0548)				
Post-reform indicator x Above-median population (inland)					0.318***			
					(0.0996)			
Post-reform indicator x Above-median population density (inland)						0.205*		
						(0.112)		
Post-reform indicator x Above-median GDP p.c. (inland)							-0.196*	
							(0.106)	
Post-reform indicator x Above-median market access (inland)								0.119
								(0.104)
Sample	Full	Full	Full	Full	Inland	Inland	Inland	Inland
N	5321	5321	5321	5321	2798	2798	2798	2798
Adj. Rsq.	0.975	0.975	0.975	0.975	0.972	0.972	0.972	0.972

Notes: The dependent variable is the logarithm of annual GDP. State-level (respectively province-level) SEZ is a dummy switching on in the year after the introduction of any SEZ at that level. The treatment indicators are interacted with the logarithms of population, GDP per capita, population density, and market access relative to the median in the year in which a city enters the sample. The indicators for above-median values in the inland sample are calculated based on a comparison to the inland sample median only. All specifications include city fixed effects, the interaction of province-year dummies, and controls for land area. Standard errors are clustered at the city level: * $p \leq 0.10$, ** $p \leq 0.05$, *** $p \leq 0.01$.

Table A15: Early and late reformers

	(1)	(2)	(3)	(4)	(5)	(6)
Post-reform indicator for SEZ established prior to 1993	0.140***			0.182***		
	(0.0342)			(0.0602)		
Post-reform indicator for SEZ established 1993 - 2010	0.0473			0.137		
	(0.0436)			(0.137)		
Linear trend during 10 years after any state-level zone		0.0192***	0.0153		0.0255***	0.0236
		(0.00458)	(0.0116)		(0.00770)	(0.0195)
Indicator for 11 to 30 years after any state-level zone		0.192***	0.153		0.255***	
		(0.0458)	(0.116)		(0.0770)	
Regional Sample	Full	Full	Full	Inland	Inland	Inland
Reformer Sample	All	pre-1993	post-1992	All	pre-1993	post-1992
N	5321	4654	4086	2798	2729	2453
Adj. Rsq.	0.975			0.972		0.971

Notes: The dependent variable is the logarithm of annual GDP. Columns (2), (3), and (5) impose the constraint $\beta_{11-30} = 10 * \beta_{trend}$. Column (6) does not include an indicator for the years 11-30 years after the reform because in the inland sample after 1992 such observations do not exist. All specifications include city fixed effects, the interaction of province-year dummies, and controls for land area. Standard errors are clustered at the city level: * $p \leq 0.10$, ** $p \leq 0.05$, *** $p \leq 0.01$.

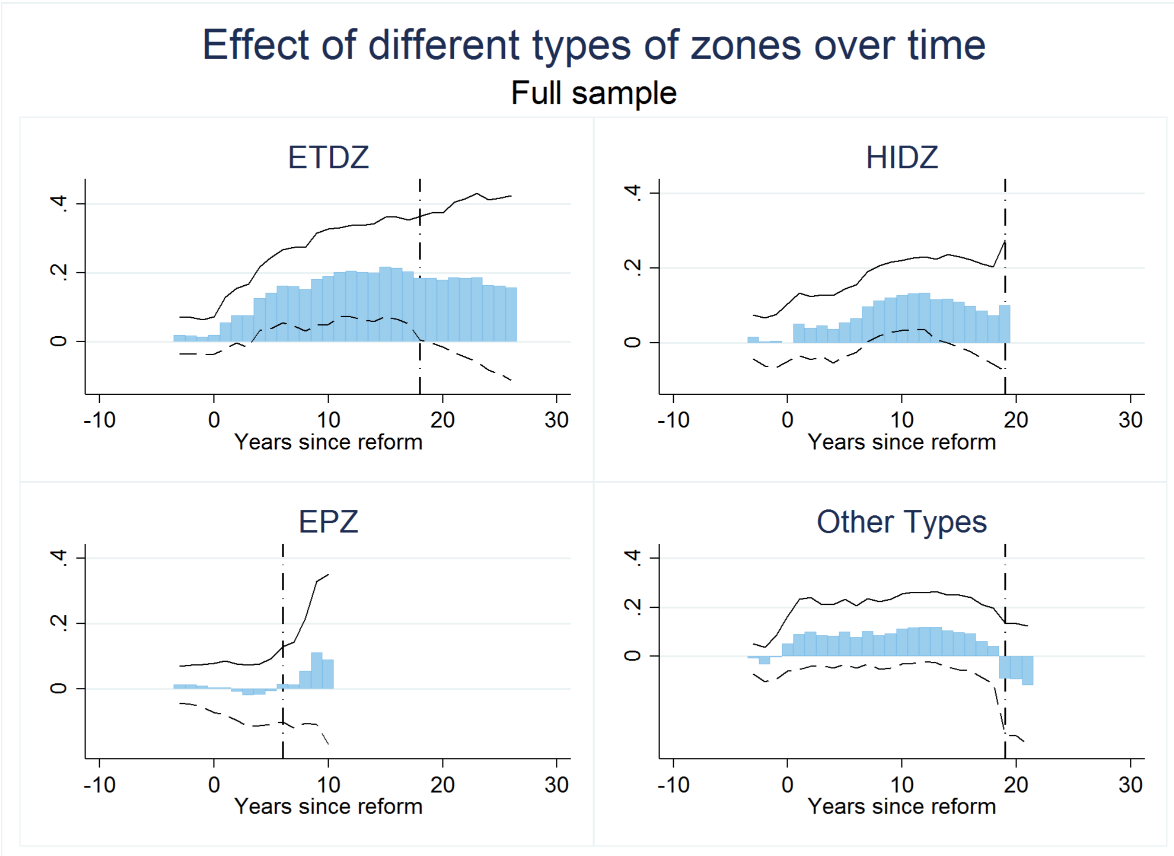


Figure A1: Effects of different types of zones over time

Notes: The four panels show the coefficients of different policy variables estimated in the same regression. The bars show the coefficients of a regression of the logarithm of nominal GDP on indicators for years before and after a type of zone was established. The solid and dashed lines show the confidence intervals. The vertical dashed line shows the lag at which the number of observations drops due to the first zones reaching the end of the sample period. The regression also controls for an indicator for province-level zones, land area, city fixed effects, and province-year fixed effects. Standard errors are clustered by city.