Aggregate Effects of Special Economic Zones: A Firm Dynamics Model with Endogenous Entry, Exit and Location Choices*

Boyao Zhang
Universitat Autonoma de Barcelona
IAE-CSIC and Barcelona School of Economics

November 20, 2023

(Check the latest version.)

Abstract

I have created a novel geo-coded firm-level panel dataset that, for the first time, allows the analysis of the impact of Special Economic Zones (SEZs) on firm distribution. SEZs, locations offering reduced corporate taxes and enhanced access to credit contingent upon maintaining a minimum scale, are empirically shown in this study to be associated with better firm selection, higher capital, and better resource allocation within firms over time. Utilizing this new panel data, the paper found that firms in SEZs outperform those outside in terms of productivity, investment, and resource allocation, regardless of whether they originated in SEZs or relocated there. Firms that originate in SEZs tend to be permanently more productive (better selection), invest more (higher capital), and show a better correlation between Total Factor Productivity (TFP) and capital over time (better resource allocation) compared to firms established outside SEZs. Firms that relocate to SEZs from outside also gain these advantages after their move, unlike firms that remain outside. However, even if SEZs were randomly allocated across China, which is not the case, these differences are not solely due to SEZs, considering firms' endogenous operational choices. To understand the wider effects of SEZs on aggregate productivity and their optimal size, I have constructed a firm dynamics model, incorporating firms' entry, exit, and location choices, where empirical analysis for these two aspects is not feasible. Counterfactual experiments indicate that SEZs increase aggregate TFP by 25.7%. This increase is driven by both better selection, with an average TFP increase of 25.1% among firms, and better resource allocation, indicated by an 88% increase in the correlation between capital and

*I am grateful for my supervisor Raul Santaeulalia-Llopis, who guided me and provided me with kind support during my research period. I also thank Luis Rojas, Andre Groeger, Albert Marcet, and Barthélémy Bonadio for their helpful comments. I also thank the seminar participants at NYUAD, macro club at UAB, BSE Jamboree and ENTER Jamboree.

TFP. Isolating the role of financial frictions, I find that reduced financial frictions in SEZs contribute about half of the increase, inducing better selection and resource allocation for the endogenous distribution of firms.

Keywords: Firm dynamics, Economic Growth, Special Economic Zones, Selection, Resource Allocation, Investment, Productivity
JEL Classification:E22, E23, E44, E62, H32, I38, O47

1 Introduction

Special Economic Zones (SEZs), geographically defined areas administered by a single authority, typically offer firms within them incentives such as reduced corporate taxes and enhanced credit access, conditional on maintaining a minimum scale. Adopted in over 130 countries across both advanced and developing economies, SEZs are implemented under the prospects to increase economic growth. This widespread policy has gained global momentum with the number of SEZs increasing thirty-fold from 1975 to 2002. China, hosting more than a quarter of these SEZs, is a major contributor to this global trend.[1] Simultaneously, the allocation of capital across firms has proven relevant for the behavior of aggregate total factor productivity (TFP) and output growth in China (Song et al., 2011).

Here, I provide, for the first time, an study of how the SEZs shape the distribution of firms in order to assess the aggregate effects of SEZs in China. First, I construct a novel panel data set that allows me to document a set of stylized facts regarding the distribution of firms differs across zones. Second, I build a firm dynamics model with entry and exit extended to endogenous location (zone) choices that replicates the stylized facts and that I use to assess the aggregate effects of SEZs. Two main mechanisms are at work. First, selection through birth (across zones) and movers versus stayers (across zones). Second, dynamic resource allocation that combines investment choices with changes in within-firm resource allocation across time defined as the evolution of the within-firm covariance between TFP and capital.

First, I create a novel firm-level panel dataset in order to keep track of firm-level zone location across time. This dataset combines detailed information on China's economic zones, including their precise locations sourced from Baidu Map's API [2] and web scraping techniques, with standardized, firm-level data for Chinese manufacturing firms. That is, I add a SEZ status variable (i.e. whether firms are in a SEZ or not) to the standard Annual Survey of Industrial Enterprise in China (ASIE) that is used, among others, by Hsieh and Klenow (2009) in their analysis of factor misallocation in China. My new panel enables the first-ever comprehensive evaluation of the implications of SEZs on firm behavior and distribution. My dataset tracks the firm-specific SEZ status of 586,599 unique firms across 2,574 districts from 1998 to 2013. Using this data, I carefully document a new set of stylized facts that highlight the advantages associated with Special Economic Zones (SEZs). These advantages include improved firm selection, increased

[1] Source: ILO database on export processing zones.

[2] API stands for "Application Programming Interface," a set of tools that enables different software applications to communicate with one another. Baidu offers APIs that grant users access to their data and functions. By utilizing this API, I can obtain location data for the Special Economic Zones (SEZs) that are the focus of this paper.

capital investment, and enhanced efficiency in resource allocation within firms over time. Utilizing the new panel data, I find that firms in SEZs, whether born in SEZs or relocated there, perform better in several dimensions. Firstly, firms that originate in SEZs tend to be more productive, invest more, and show a better correlation between Total Factor Productivity (TFP) and capital over time compared to firms established outside SEZs. Secondly, firms that relocate to SEZs from outside also gain these advantages after their move, in contrast to firms that remain outside. However, it is important to note that even if these zones were randomly allocated across China, which show it is not the case[3], the improved performance of firms in SEZs cannot be entirely attributed to the SEZs themselves. The reason is that the firm's decision to operate in an SEZ is an endogenous choice, hence, not random. This is particularly relevant because out of the total number of firms in SEZs I find that 55% of firms in SEZs are (choose to be) born there or 45% relocate (choose to move) from areas outside SEZs, which further limits the empirical analysis.

Second, to address the empirical limitations on the assessment of SEZs (endogenous SEZ rollout policy and endogenous selection into SEZ "treatment"), I propose a firm dynamics with entry and exit model that further endogeneizes firms' location (zone) choices. My model enables me to study the aggregate effects of SEZ. I use this model to assess how each of the characteristics of Special Economic Zones (SEZs), such as reduced corporate taxes, more relaxed financial constraints, and higher minimal profit scale requirements, impacts on aggregate Total Factor Productivity (TFP) and output. Ultimately, I use the model to assess the optimal size of SEZs.

Specifically, I assume that firms in both SEZs and Non-SEZs operate under the same technology. However, they differ in terms of corporate tax rates, financial constraints, and minimal profit requirements across zones. Specifically, in alignment with SEZs features, firms in SEZs benefit from lower corporate tax rate, face less stringent financial constraints, and are obliged to satisfy a minimal profit scale requirement. Firms are heterogeneous in both productivity level and cash-on-hand positions. Hence, the firms in my model are categorized into the three types: those not operating in the economy, those operating in Non-SEZ areas, and those operating within SEZs. Their location choices are determined by their specific conditions to meet their operational needs.

My model yields several main predictions, which qualitatively align with the empirical findings. The first prediction suggests that firms are positively selected into SEZs compared to Non-SEZs firms in terms of productivity, asset and resource allocation. Given reduced corporate taxes and improved access to credit markets offered by SEZs, every firm aspires to establish itself or relocate

[3]The establishment of SEZs in the early 1980s, targeting cities like Shenzhen, Zhuhai, Shantou, and Xiamen, as experimental grounds for new market-oriented policies. With subsequent expansion in 1984 to 14 other coastal cities, then in 1991 extended this strategy to inland cities.

itself within SEZs to benefit from these favorable policies. However, due to the stringent minimal requirements of SEZs, only firms surpassing these criteria can operate within the zones, while those with fewer assets operate outside. There are two thresholds in equilibrium, each consist of a combination of values for productivity and assets (cash-on-hand). Firms below the lower-asset threshold exit the economy, while those above the higher-asset threshold become SEZ firms. Firms falling between these two thresholds operate in Non-SEZ areas. Moreover, SEZs' preferential policies attract more firms to establish themselves there or relocate from Non-SEZs to there, intensifying competition within SEZs. Through a "survival-of-the-fittest" market mechanism, more efficient firms survive, and less productive ones are pushed out of the market. This selection mechanism results in a more efficient allocation of capital among firms in SEZs. As productivity and financial development increase within SEZs, high-productivity, low-asset firms from Non-SEZ areas relocate into SEZs, reducing the marginal asset threshold levels due to higher returns on productivity. Reduced financial frictions in SEZs enable high-productivity firms to expand their businesses, increasing their capital investment. Both factors attract more productive firms from Non-SEZ areas into SEZs, intensifying competition within SEZs and displacing less productive firms with higher assets from SEZs, turning them into Non-SEZ firms.

However, reduced tax frictions in SEZs attract more firms to enter and obtain loans. Low tax rates may enable inefficient yet wealthier firms (due to minimal profit scale requirements) to easily secure loans, potentially leading to an imbalance in resource allocation. Less efficient firms might exploit lower tax rates to obtain more loans without effectively using these resources, offsetting the positive selection driven by reduced financial frictions. These opposing forces create an ambiguous situation where the existence of higher average productivity (better selection) and improved resource allocation depend on their interplay. The trade-off between lower taxes and better financial access highlights the need for further investigation into optimal taxation. In this context, it is expected that the dominance of positive over negative selection will lead to optimal aggregate productivity growth and efficient resource allocation.

This model was carefully calibrated to align with empirical findings. Through counterfactual experiments assuming a scenario without SEZs, this paper finds that SEZs increase aggregate TFP by 25.7%. Examining the mechanisms, I find that the rise in aggregate TFP stems from two main channels: an improvement in firm selection, with an average TFP increase of 25.1% among firms, and better resource allocation, indicated by an 88% increase in the correlation between capital and TFP. My decomposition of the SEZs characteristics that drives reveals that the reduction of financial frictions in SEZs accounts for about half of the increase in aggregate TFP. This reduction leads to better firm selection, particularly for firms newly established in or relocated to SEZs, and also results in better resource allocation for the endogenous distribution of firms. Note that ex-ante it is not obvious, through my model, that SEZs will increase aggregate TFP and output because the reduction of corporate taxes in SEZs enhances the entry of less productive firms in the economy and in the SEZs. It turns out that this effect is quantitatively dominated by the gains from the relaxation of financial constraints.

Related Literature

This paper relates to a broad literature investigating, typically empirically, the impact of Special Economic Zones (SEZs) policies at the county (or municipality) level (Schminke and Biesebroeck, 2011; Farole and Akinci, 2011; Wang, 2013; Alder et al., 2016; Lu et al., 2019).[4,5]. These studies have consistently found positive effects on local economies in terms of GDP growth, productivity, and investment. While these studies focus on regional city-level analyses, my research explores the effects of SEZs using novel firm-level data, enabling a detailed assessment of the interaction between the distribution of firms and SEZs through selection, either by birth, exit or moving across zones, and dynamic factor misallocation. This variation in the firm-level performance not only allows for the empirical exploration of micro-variations both between SEZs and non-SEZs within cities, and within SEZs, but I also use it to discipline a dynamic firms model that I use to assess the aggregate effects of SEZs.

My paper delves into the concept of agglomeration within new economic geography (NEG), focusing on specialization (Marshall, 1890) and diversity (Jacobs, 1969) for knowledge creation and diffusion. The Marshall-Arrow-Romer (MAR) model (Marshall, 1890; Arrow, 1962; Romer, 1986) suggests that regional industry concentration enhances knowledge spillovers and innovation (Glaeser et al., 1992), favoring local monopoly over competition. Conversely, the Jacobs (1969) diversity model emphasizes the role of industrial diversity in cities in promoting cross-sectoral idea sharing and innovation. Although agglomeration effects are widely studied, this paper focuses on the selection effect, which some studies Combes et al. (2012) suggest is weak. Various scholars have examined these effects across different industries and regions, yielding mixed findings (Arimoto et al., 2014; Behrens et al., 2014). This study assesses the SEZs' impact through both agglomeration and selection channels, uncovering a distinct narrative in Chinese SEZs where agglomeration is not the main driver of aggregate productivity growth. One reason why the agglomeration effects are potentially less powerful is that SEZs might increase price cooperation and reduce competition as put forward in Brooks et al. (2021).

[4] The remarkable 58% annual growth rate of Shenzhen since its first SEZ inception. In 2007, SEZs accounted for 46% of China's Foreign Direct Investment (FDI) (Wong, 1987), and over half of the country's high-tech firms are in these zones, making a substantial contribution to China's high-tech industrial output (Zeng, 2010).

[5] Between 2006 and 2010, in several SEZs, the industrial added value represents over 30% of the city's total, significantly contributing to the economic development of their respective regions.

My work assesses the influence of SEZs' preferential policies on productivity via resource allocation. The role of resource allocation across firms in China has been highlighted in Hsieh and Klenow (2009). Here, I build on their same dataset on manufacturing firms to add a geolocation variable that determines whether firms are in Special Economic Zones or not. I use this geolocation information to specifically quantify the effects of SEZs on dynamic resource allocation. More recently, König et al. (2022) also show the effects of R&D misallocation on TFP in China. I abstract from endogenous R&D and focus on firm selection across zones and their dynamic resource allocation in terms of TFP and capital.

Further contributing to the literature on firm dynamics, entry barriers, and market selection, Hopenhayn (1992) discusses the interplay between entry costs and selection, and various papers have explored different selection processes Khan and Thomas (2011); Gottlieb and Grobovsek (2016); Restuccia and Rogerson (2008); Lagakos and Waugh (2013). Unlike previous works, this paper introduces discrete SEZ location choices, which do not fit neatly into either the agriculture versus non-agriculture division Restuccia and Rogerson (2008) or occupational choices Lagakos and Waugh (2013). While Adamopoulos et al. (2017) establishes links between selection and misallocation, I add investment dynamics and agglomeration effects to the analysis of selection and misallocation.

Moreover, this paper also contributes to quantitative studies on financial frictions and economic development (Buera et al., 2011; Buera and Shin, 2013; Midrigan and Xu, 2014). It examines the misallocation caused by financial frictions, as highlighted by Midrigan and Xu (2014), and explores the growth impact of financial frictions (Buera et al., 2011) and the transition dynamics following reduced financial frictions (Buera and Shin, 2013). This study links misallocation to specific policies like SEZs in China, quantifying their effects on resource allocation, particularly in terms of improved credit market access and reduced tax frictions, and their impact on firms' performance and aggregate TFP growth.

The rest of the paper is structured as follows. The next section, 2, introduces the background of Special Economic Zones (SEZs) in China, outlining their evolution over time and across different locations. In Section 3, I describe the construction process of the novel geo-coded firm-level data from China, along with the variables used in the analysis, and present the main stylized facts. Section 4 details the framework of the firm dynamics model with endogenous entry, exit, and location choices. The calibration process and the model's performance are discussed in Section 5. Section 6 describes the counterfactual exercise and reports the main quantitative results. Section 7 provides further discussion of this study. The paper concludes in Section 8. Additional details and results are provided in the Appendix A.

2 Background and Context: SEZs and their Evolution in China

I begin this section by providing a definition of a Special Economic Zone (SEZ) in subsection 2.1. This is followed by an introduction to the institutional background of China's Special Economics Zones and a set of new stylized facts associated to the evolution of the distribution of firms in SEZs across time and space presented in subsection 2.2.

2.1 What Is a Special Economic Zone (SEZ)?

A SEZ is a geographically defined area, often securely enclosed, with a unified management system. Typically, SEZs typically operate under more liberal economic laws compared to the national standards. The term encompasses various forms of zones like free trade zones, export-processing zones, and industrial parks. In China, however, SEZs are distinctively multifunctional, covering larger areas than in other countries and a wider range of economic activities than typical zones. Here, I use the term SEZ to broadly include not only the seven comprehensive SEZs of China—Shenzhen, Zhuhai, Shantou, Xiamen, Hainan, Shanghai Pudong New Area, and Tianjin Binhai New Area—but also other forms such as Economic and Technological Development Zones (ETDZs), Free Trade Zones (FTZs), Export-Processing Zones (EPZs), and High-Tech Industrial Development Zones (HIDZs).

The SEZs employ diverse preferential policies to attract qualified firms and also impose obligations for the incumbent firms. In these terms, I define SEZs as a zone with these collection of these incentives and duties:

1. Lower corporate taxes: A clear incentive for being in SEZs is the reduced corporate income tax rate, which varies from 15% to 24% based on a firm's technological contributions, compared to the standard 33% outside SEZs.
2. Higher access to credit: Encourage national policy banks and commercial banks to increase credit issuance. Support qualified enterprises within SEZs for the issuance of corporate bonds, medium-term notes, short-term financing bonds to expand direct financing through capital markets [6].
3. Minimum scale requirement: Third, there is a cost for being in SEZs. Specifically, there's a rigorous system to monitor firm performance within these zones. Expert reviews or consult-

[6] According to Article 26 of State Council [2010] No.28: Support qualified enterprises in SEZs issuing corporate bonds, medium-term notes, short-term financing bonds, collective enterprise bonds, and public financing through listings.

ing firms evaluate firm's economic benefits. Firms compete to meet annual benchmarks, including minimum profit requirements, as part of their operational mandates.

2.2 Special Economic Zones in China

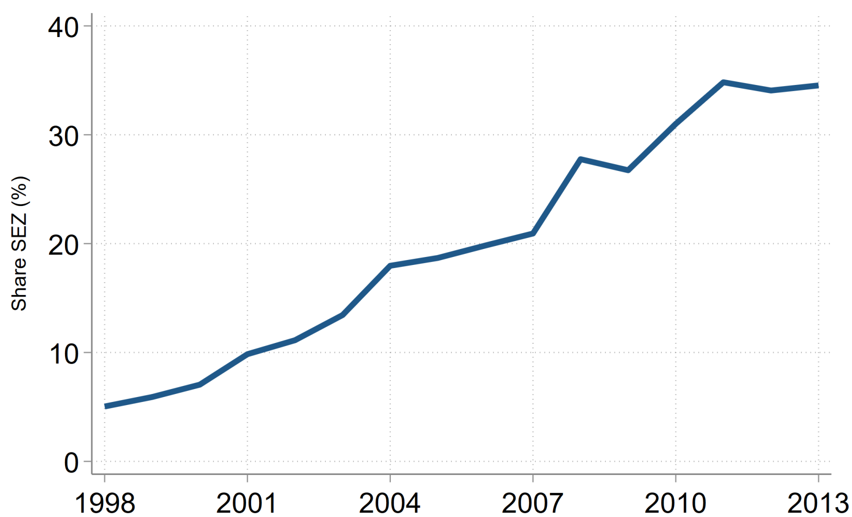
China's implementation of Special Economic Zones (SEZs) has been a dynamic process. The establishment of SEZs in the early 1980s, embodying Deng Xiaoping's pragmatic approach of *"crossing the river by touching the stones"*, targeting cities like Shenzhen, Zhuhai, Shantou, and Xiamen, as experimental grounds for new market-oriented policies and institutional models. With subsequent expansion in 1984 to 14 other coastal cities, each establishing its own economic and technological development zone. The establishment of these coastal SEZs coincided with a period of extraordinary growth in those regions, with Shenzhen's GDP surging at an annual rate of 58 percent in the early years, outpacing the national average annual GDP growth by approximately 10 percent.[7]

The State Council extended this strategy to inland cities starting in 1991, leading to a surge in SEZ establishments that exceeded 2,000 by 2018. I show the extent of this expansion in Figure 1, which illustrates the growing proportion of firms within SEZs over time, employing the cross-sectional dimension of the new constructed firm-level panel data that I construct in Section 3.1. The proportion of firms in SEZs climbs from a merely 5% in 1998 to approximately 35 percent in 2013. A spatial and temporal examination, as presented in Figure 2, shows that initially in panel (a) a modest number of firms within SEZs, with lighter shades indicating a prevalence ranging from 0-20% to 20-40%. These firms were predominantly situated in coastal cities, accounting for 20-40% of manufacturing firms. By 2013, Panel (b) reveals a significant evolution, with much darker shades dominating the map, signaling a denser concentration of SEZ firms. Their dominance had not only intensified in coastal areas, where 60-80% of manufacturing firms were located within SEZs, but there was also a remarkable expansion into the interior regions. The growth in northwestern cities is particularly notable, with SEZ firm representation soaring from under 20% to over 40%.[8] Further, the maps illustrate not just an increase in firm density within SEZs but also a delivered—non-random—expansion across the country over the 15-year span, especially into inland regions, highlighting a developmental strategy to incorporate these areas into the SEZ economic structure.

[7] By 1986, Shenzhen had begun to form basic markets for capital, labor, and technology. Other SEZs, such as Zhuhai (32 percent), Xiamen (13 percent), and Shantou (9 percent), also achieved remarkable growth rates by 1986 (Yeung et al., 2009).

[8] By 2007, the collective GDP of the main state-level SEZs constituted about 21.8 percent of the national GDP. That year also saw these SEZs drawing approximately 46 percent of China's total Foreign Direct Investment (FDI). Research and Development (R&D) expenditures within these zones tripled to RMB 105.4 billion, while high-tech industries within SEZs accounted for nearly 40 percent of the industrial output (Zeng, 2010).

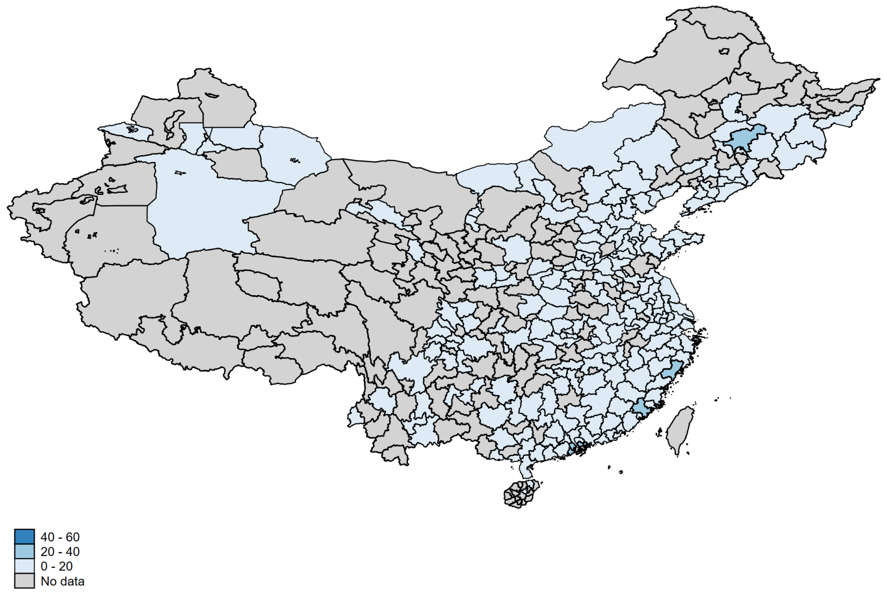
Figure 1: Share of SEZs Firm Across Time



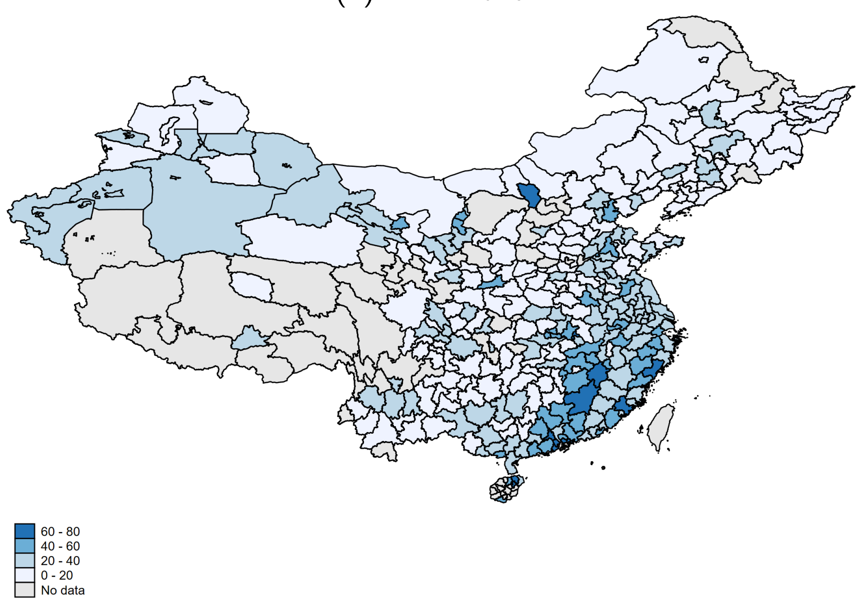
Notes: This figure plot the proportion of firms in Special Economic Zones (SEZs) over a period from 1998 to 2013 using the cross-sectional dimension of the new panel data set that I construct in Section 3.1. The y-axis, represents Share of SEZ firms in percentage, indicates the percentage of firms within SEZs relative to the total number in the country. The x-axis represents the years during which the data was collected.

Figure 2: Share of SEZs Firm Across Time and Space

(a) Year: 1998



(b) Year: 2013



Notes: The maps provide a visual representation of the spatial distribution and proportion of firms in Special Economic Zones (SEZs) across China for the years 1998 and 2013 using the cross-sectional dimension of the new panel data set that I construct in Section 3.1. In panel (a), there is a sparse distribution of SEZ firms, indicated by lighter shades that mostly represent a 0-20% and 20-40% SEZ firm presence, with limited regions showing a 40-60% range, predominantly in coastal cities. Panel (b) depicts a notable transformation with a significant increase in the prevalence of SEZs, as evidenced by the darker shades indicating a 40-60% and even 60-80% presence in certain areas.

3 New Firm-Level Panel Data and Stylized Facts

In this section, I introduce two primary data sources and describe the procedure I use to create a new firm-level panel dataset designed to tracks firms across zones in Section 3.1. I describe the measurement of firm-level Total Factor Productivity (TFP) and other relevant variables in Section 3.2. I then provide a set of new stylized facts using the new firm-level panel data to tracking firm dynamics between SEZs and Non-SEZs, including births and movers across zones in subsection 3.3.

3.1 Construction of a New Firm-level Panel Data Set

Section describes the process of merging data from the Annual Report of Industrial Enterprise Statistics (ASIE) with the China Development Zone Review Announcement List (2018), along with numerous other official sources, to construct this unique dataset. This new dataset is the first to keep track of firm-level data across zones and is crucial for examining differences in firms' productivity and performance before and after the implementation of Special Economic Zones (SEZs). The final dataset contains 586,599 distinct firms across 2,574 district-level regions from 1998 to 2013, and approximately one-fifth of the total number of firms are in SEZs.[9]

Annual Survey of Industrial Enterprise in China (ASIE)

The primary firm-level data for this study comes from the Annual Survey of Industrial Enterprise in China (ASIE), collected by the National Bureau of Statistics (NBS) of China, including firm-level variables from 1998 to 2013.[10] This dataset includes all state-owned and nonstate-owned industrial 'firms with annual sales above 5 million RMB (about 780,000 USD), and this threshold was raised to 20 million RMB in 2011. The databese contains 4-digit industry classifications that cover a variety of industries, such as mining, manufacturing, and the production and supply of electricity, gas, and water, where manufacturing firms account for over 90% of the dataset. As the database is predominantly composed of manufacturing firms, which aligns with industrial classifications in other countries, and variables such as output, capital, employment, and export delivery value are more easily measurable, I focus on the manufacturing firms in this study.[11] In addition, the dataset offering two main information types: basic company details and financial variables. These include crucial data like postal codes, addresses for locating firms in SEZs.

[9] For a detailed data cleaning procedure, please refer to Appendix **??**.

[10] Hsieh and Klenow (2009) use the same data source for their analysis of factor misallocation in China.

[11] Manufacturing coverage includes 30 major categories (two-digit industry sectors) ranging from processing of food from crop and animal husbandry products, food manufacturing, to arts and crafts and other manufacturing, and recycling of waste resources and materials, corresponding to the codes 13 to 43 (excluding 38) in the National Economic Industry Classification and Codes (GB/T4754-2002).

Moreover, financial variables like total industrial output and profit, industrial intermediate input, the total value-added, investments, fixed assets, accumulated depreciation, liabilities and so on, are essential for analyzing firm performance in this study.

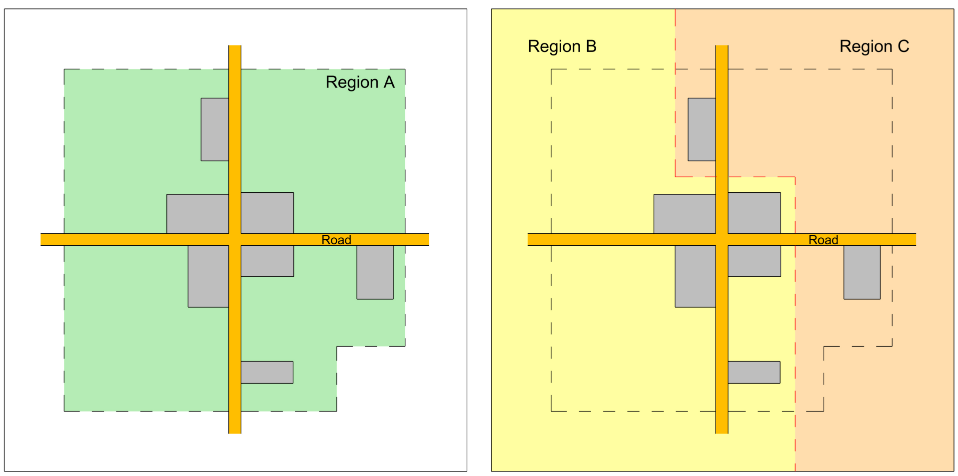
The use of this dataset presents a complex issue due to the dynamic nature of China's district-level administrative divisions, which can have significant effects on firms' reported locations. These administrative changes can lead to measurement errors in the identification of firms entering Special Economic Zones (SEZs), resulting in incorrect associations of firms with districts. As illustrated in the figure, 3 initially, all firms (shown in gray) within the boundaries of Region A have the same district-level code. However, after the administrative division, the firm in the center and the one below it are reassigned to Region B, while the remaining two are now part of Region C. To mitigate these mismatches, and address these issues, it is essential to standardize the district codes over time by aligning them with the single code from 2013 and then use this standardized code to merge the SEZs information into the firm-level data. This standardization process is underpinned by a thorough review of official documents that trace the historical changes in district boundaries. In addition, detailed local information such as street and community names, as reported by the firms themselves, is utilized to ensure accurate district location.

This meticulous approach enables the preservation of data integrity, ensuring that firms are accurately represented within the same district over time despite the administrative boundary modifications.Consequently, this enhances the credibility of the analysis related to Special Economic Zones.

Development Zone Review Announcement List

The Development Zone Review Announcement List version 2018 is an official documentation that record all registry development zones in China. This resource includes the zones' names, their sizes, the dates when they were approved, and their predominant industries. However, it lacks precise geographical information. To overcome this gap, especially when aligning SEZ data with firm-level data, I utilize Baidu Map's API, employing web scraping techniques to retrieve location information. By querying the API with the names of the SEZs, I can extract Points of Interest (POI) data, which provides the location details that are necessary to determine each SEZ's district location. After collecting this data, I assign accurate district codes to each SEZ, facilitating the integration of geographic data with the firm-level dataset.

Figure 3: County-level Administrative Area: Consolidation, Division



Notes: This figure represents the administrative changes over time within a given area, impacting firm location data. Initially, all firms (gray blocks) are within Region A (left panel). Subsequent redistricting divides the area into Region B (yellow) and Region C (orange).The central road remains the same through the reclassification. The division results in firms being categorized under new regional codes, despite no physical relocation.

Identification of SEZ firm

To accurately identify firms situated in China's economic development zones, I used their addresses from the ASIF database, which has detailed location info like town, streets and doorplate numbers. I conducted a text analysis on address variables for keywords indicative of SEZ locations using 17 key terms like "kaifa" for development zones or "gaoxin" for high-tech zones.[12] Additionally, I collected postal information from Development Zone Review to minimize identification errors. Matching the ASIF data with postal ZIP code confirmed whether firms identified by keyword searches were indeed within officially recognized SEZs. This method ensures that firms are correctly matched with the SEZs they are associated with, based on their physical locations.

Merged Firm-level Data

While firm-level data have large samples, numerous variables, and spans an extended period, it has gaps and errors in terms of missing values. I follow the protocol in Brandt et al. (2012) and exclude firms lacking essential financial details or firms with fewer than nine workers. In addition, I only focused on panel firms—excluding those with less than two years of observations, in order to further ensure reliability on estimates. The final dataset that merges SEZ information at the firm-level contains 586,599 unique firms within 2,574 district-level regions, spanning from 1998 to 2013. Then, by applying district codes, I accurately determine firms' location, in particular, whether they are situated within SEZs, have migrated from NSEZs to SEZs, or are newly born firms in either type of zone. This novel firm-level panel data is essential for describing the evolution of the distribution of firms in relation to SEZ policies. After data cleansing, the firm-level dataset contains 117,000 firms inside SEZs, representing 20% of the total firms (pooling all firms across time and space).

3.2 Measurement of TFP and Other Variables

In this Section I provide a brief description of the estimation of firm's total factor productivity. The Data Appendix A.1.1 provides more details on the variables used to calculate value-added to estimate firm-level TFP. For a more detailed information regarding these measurements of TFP and agglomeration are found in the respective Appendix A.1.2 and A.1.4.

TFP Measurement I use the Olley and Pakes (1996) method as my benchmark for the measurment of firm-level TFP. For robustness, I also employ alternative methods for measuring firm-

[12] 17 keywords that would indicate the location within a SEZ, like kaifa, gaoxin, jingkai, jingji, yuanqu, baoshui, bianjing, kejiyuan, chuangyeyuan, huojuyuan, huojuqu, gongyeyuan, chanyeyuan, gongyequ, gongyexiaoqu, and chukoujiagong, any of those keywords indicated the presence of any kind of economic development zone.

level TFP from specifying a Cobb-Douglas production function (OLS), a fixed effects (FE) model and the Levinsohn and Petrin (2003) approach. Detailed methodologies for these productivity estimations are provided in the Appendix A.1.2.

3.3 Stylized Facts

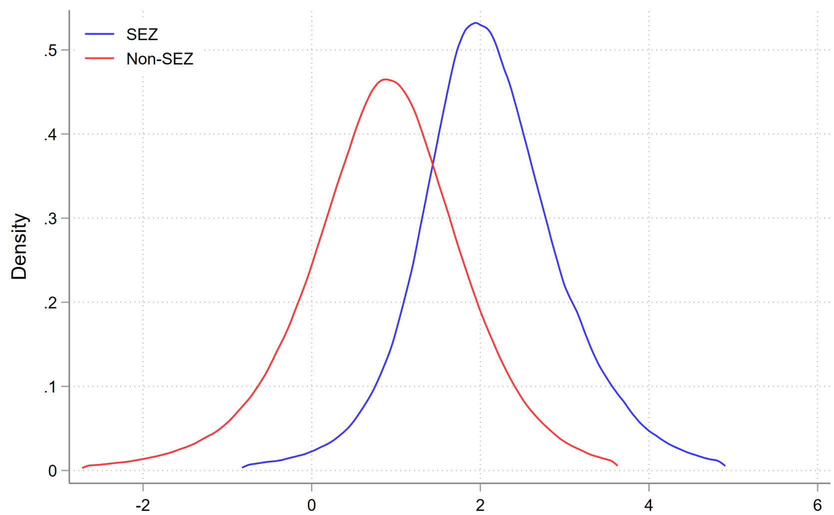
This subsection, utilizing the new panel dataset, presents a set of new stylized facts concerning firms in SEZs and non-SEZs, with particular focus on their performance in terms of productivity, capital investment, and the capital-productivity correlation within firms over time. It also examines the same comparative outcomes between firms born in SEZs and those originating in non-SEZs, as well as the performance of firms that have relocated from non-SEZs to SEZs relative to those that have remained outside the zones.

Stylized Fact I: Cross-Sectional Zone Differences In this part, I present the first stylized fact from a cross-sectional analysis of firms in Special Economic Zones (SEZs) and Non-SEZ areas, examining the firm-level productivity variations between the two. Accompanying data is detailed in a table that explores potential factors contributing to these differences, such as capital investment, and the covariance between productivity and capital.

Figure 4 illustrates the kernel density of firm-level Total Factor of Productivity (TFP) for SEZs versus Non-SEZs. The distribution for SEZs is skewed rightward, suggesting an average higher productivity among these firms. Notably, the distribution is broader and flatter for firms in Non-SEZs, implying a greater variance in productivity within these zones compared to SEZ regions. This observation raises questions about the mechanisms that may enhance productivity within SEZs. The preferential policies of SEZs aim to spur technological progress, leading to this productivity advantage in SEZ, open the productivity gap between two zones. Additionally, the financial incentives and lower corporate tax rates within SEZs are likely to foster a more streamlined environment for resource distribution, potentially facilitating greater capital investment. As a result, we expect a higher relationship between capital and productivity within SEZs. To gain insight into the contribution of these factors to the observed productivity advantage, we look into capital investment and capital-productivity correlation within firms across time between SEZs and Non-SEZ, and detailed statistics are reported in Table 1.

The Table 1 compares firm performance in Special Economic Zones (SEZs) with that in non-SEZ areas, using logarithmic measures for variables. The table shows that the average productivity of firms in SEZs, at 2.21, is approximately 136% higher than that of their non-SEZ counterparts, which is 0.85. This significant disparity highlights a better selection of firms within SEZs. In terms of capital, the average for SEZ firms is 9.48, which is 66% higher than the 8.82 average for

Figure 4: Firm-Level Productivity Across Zones ($log(z_i)$)



Notes: This figure plot the kernel density firm-level Total Factor of Productivity (TFP) in Special Economic Zones (SEZs) compared to Non-SEZ areas. The x-axis represents the TFP level $log(z_i)$ of the firms, and the y-axis shows the estimated density of firms at different TFP levels. Blue line represents for SEZs and red for Non-SEZs.

non-SEZ firms, suggesting that firms in SEZs may have better access to capital. Supporting this notion, the covariance between productivity and capital for SEZ firms is less negative (-0.0239) than that for non-SEZ firms, implying a marginally more effective use of resources—by about 2%—which could be attributed to more efficient resource allocation in SEZs.

Table 1: Firm-Level Productivity and Capital Across Zones

Indicator	SEZ	Non-SEZ
Avg Productivity (z_i)	2.21	.85
Avg Capital (k_i)	9.48	8.82
cov (z_i, k_i)	-.00005	-.0239

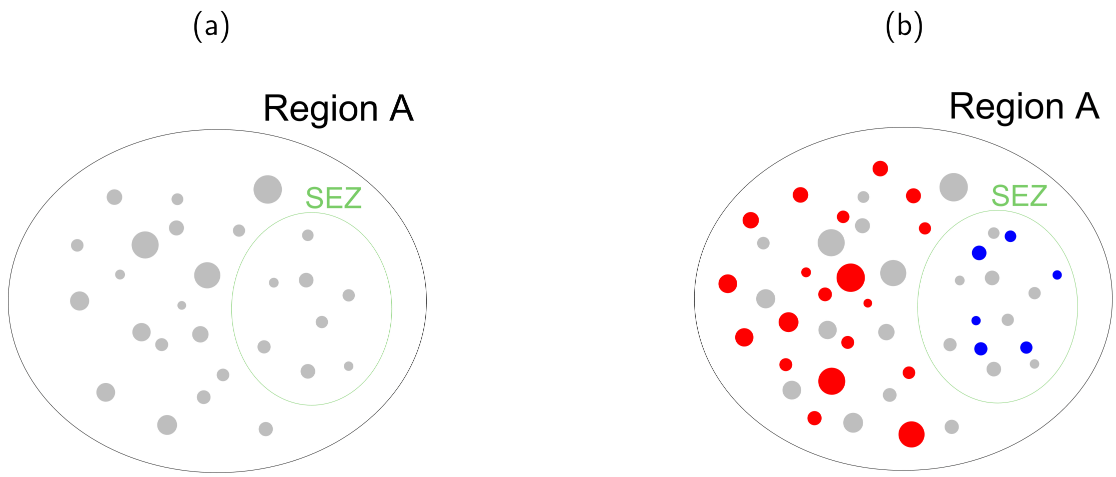
After documenting cross-sectional variances where SEZs exhibit a better selection, more efficient resource allocation—evidenced by higher average productivity, increased investment, and higher correlation of productivity and capital compared to Non-SEZs, we are compelled to investigate further the root causes of these differences. I aim to discern if the observed productivity advantage of firms in SEZs is a product of their inherent characteristics, implying that SEZs naturally attract 'diamond' firms that are intrinsically more productive and they are originating born there, or if it results from the relocation of already high-performing firms from Non-SEZs through this selection channel. These qualified Non-SEZ firms might migrate to SEZs seeking to benefit from a more supportive environment, especially if they were previously hindered by restrictive financial conditions. To unravel the reasons behind the heightened productivity of SEZs, we need to understand these dynamics more clearly. It includes examining the performance differences between firms originating born in SEZs and those originating born in Non-SEZs, as well as the firms performance that migrate from Non-SEZs to SEZs in the following sections.

After looking at the cross-sectional differences between Special Economic Zones (SEZs) and Non-SEZs, where SEZs tend to perform better with higher productivity, increased investment, and a higher correlation between productivity and capital, this paper want to dig deeper into what's causing these disparities. My goal is to figure out if the higher productivity of firms in SEZs is because they naturally attract more productive "diamond" firms or if it's because successful firms from Non-SEZs are moving to SEZs. It's possible that well-performing firms from Non-SEZs are relocating to SEZs to take advantage of a more supportive business environment, like lower tax burden and better access to the credit market, especially if they were previously struggling with financial constraints. To understand why SEZs are more productive, we need to explore these dynamics further. This involves comparing the performance of firms that originally started their business in SEZs with those from Non-SEZs and also looking at how firms from Non-SEZs perform after moving to SEZs in the next sections.

Stylized Fact II: Birth Differences by Zone This part presents the second stylized fact, focusing on the performance of originally established firms in both SEZs and non-SEZs. The Figure 5 visually compare firm dynamics across these zones. Panel (a) shows that in Region A, for example, the SEZ is marked with a green circle, and existing firms within it are shown as gray points. In panel (b), firms originating in the SEZ are represented by blue dots, while those from non-SEZ areas are red dots. My analysis mainly examines these originally established firms, comparing the red dots in SEZs with the blue dots outside. This comparison aims to demonstrate the effectiveness of SEZs in selecting qualified firms based on their performance for location within SEZs.

The Table 2 compares the performance of firms originating in Special Economic Zones (SEZs) and non-SEZ areas. It shows that SEZ firms are consistently more productive, with an average productivity rate of 2.21, compared to 1.03 in non-SEZ firms. SEZ firms also invest more, with an average capital that is 66% higher at 9.36, versus 8.70 for non-SEZ firms. This is accompanied by a stronger correlation between productivity and capital over time, suggesting better resource allocation within these firms. This likely results from improved access to financial resources in SEZs, leading to a 2.8% increase in resource allocation efficiency and a 118% increase in investment, which significantly enhances firm growth and productivity.

Figure 5: Firm Dynamics Across Zones: Born in SEZ vs. Born in Non-SEZ



Notes: The figure provides a clear visualization of firms dynamics originating in Special Economic Zones (SEZs) and non-SEZ areas. Panel (a) shows that within Region A, an established SEZ (green cycle), a subset of firms already exists, indicated by gray points. In panel (b), we observe that firms originating in the SEZ are symbolized by blue dots, while those originating in the Non-SEZ areas are marked with red dots.

These results highlight the key factors driving higher productivity in SEZs. They also guide us towards a deeper examination of how SEZ policies yield such effective results, offering solid evidence of the potential mechanisms at play.

Table 2: Firm Dynamics Across Zones: Born in SEZ vs. Born in Non-SEZ

Indicator	SEZ	Non-SEZ
Avg Productivity (z_i)	2.21	1.03
Avg Capital (k_i)	9.36	8.70
cov (z_i,k_i)	-.002	-.03

Stylized Fact III: Movers vs. Stayers As mentioned earlier, another factor contributing to the productivity advantage in Special Economic Zones (SEZs) may be the migration of high-performing firms from non-SEZ areas to SEZs. This migration is influenced by the selection process and resource allocation. Therefore, our next analysis will focus on firms moving from non-SEZs to SEZs compared to those that stay in non-SEZs. Similar to the previous section, I will first present a visual representation of firm dynamics within Region A. This will concentrate on the movement of non-SEZ firms into SEZs versus those remaining in non-SEZs. Understanding these migration patterns and identifying the types of firms involved is crucial for our comparative study.

In Figure 6, panel (a) sets the scene by showing an established Special Economic Zone (SEZ) in green within Region A, along with existing firms represented by gray dots. Panel (b) then illustrates firms that potentially moved into the SEZ, marked as blue dots, and compares them with firms that remained in non-SEZ areas, marked as red dots. This panel highlights the initial difference between movers and stayers. Panel (c) traces the paths of the moving firms, linking their origins in non-SEZ areas (red dots) to their new locations in the SEZ (blue dots). This shows the transition of certain firms from non-SEZ to SEZ areas. Finally, panel (d) depicts the region after these movements, with blue dots now within the SEZ and red dots indicating firms that stayed in non-SEZ areas, showing a later stage compared to panel (b).

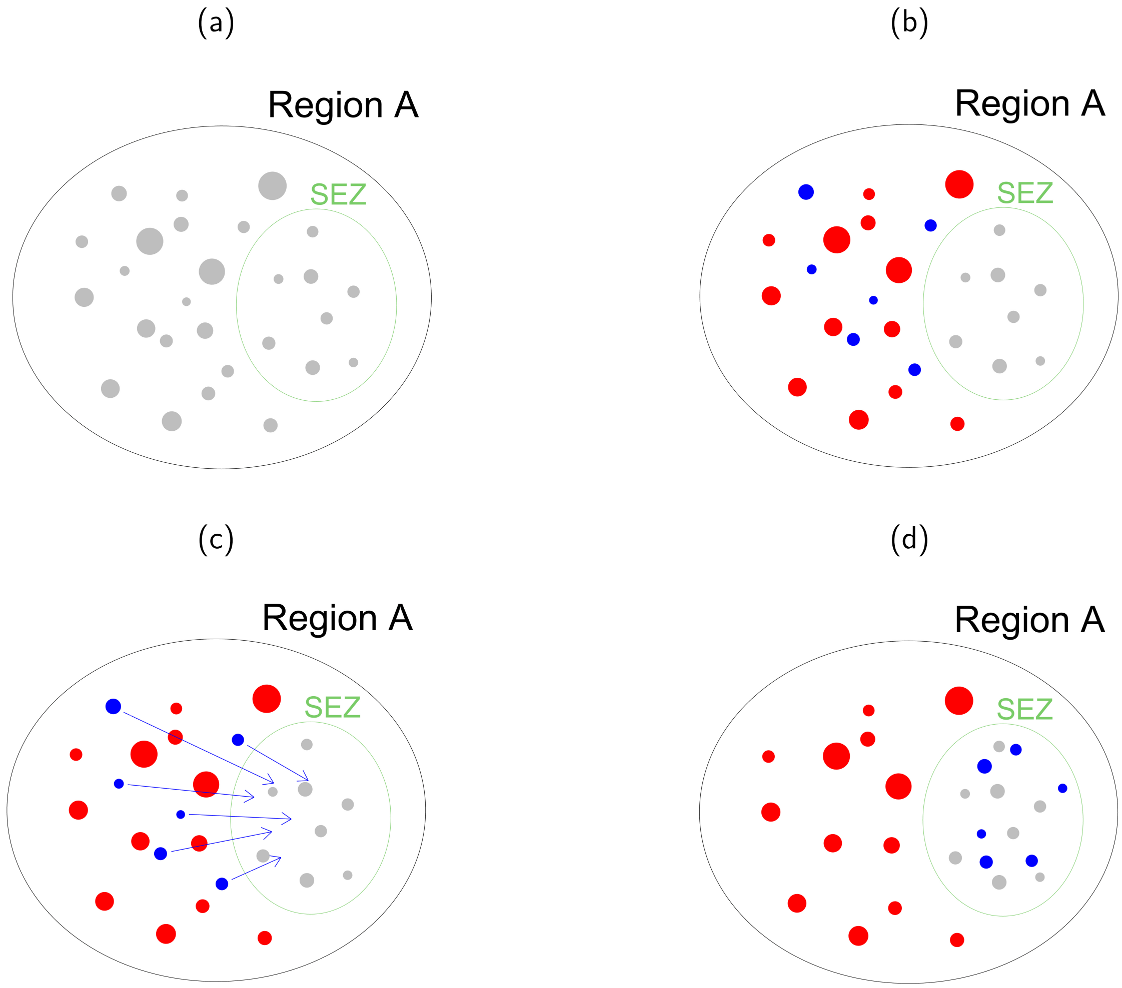
Table 3 presents a before-and-after analysis of firms that moving to Special Economic Zones (SEZs) compared to those remaining in Non-SEZ areas. It report various description statistics of productivity, capital, and the covariance between productivity and capital $Cov(lnz_i, lnk_i)$. Initially, firms that potentially move to SEZs already exhibited better performance with an average productivity 20% greater (0.84 compared to 0.64) than their Non-SEZ counterparts. After relocation, the difference in output and productivity between movers and stayers widened further, with SEZ movers experiencing a 78% increase in productivity. And the productivity of movers after moving into SEZ increases at 58%, only 26% driven by selection, suggesting that firms relocating to SEZs tend to outperform Non-SEZ stayers in terms of productivity.

Capital investment followed a similar trend, with mover firms in SEZs showing an increase in average capital from 9.02 to 9.73 post-move, while Non-SEZ firms only increased from 8.73 to 9.16. This growth in capital for SEZ movers, 28% of which is attributed to relocating in the SEZ, and 51% driven by selection, underscores the capacity of SEZs to attract or cultivate firms with higher capital investment. Notice that prior to the move, potential movers had a lower correlation between productivity and capital, indicating capital misallocation. After moving to SEZs, this correlation improved significantly, moving from -0.04 to 0.07, reflecting a more efficient capital allocation. In contrast, Non-SEZ firms saw a negligible change. The better resource allocation due to being located in SEZ was 10%, with 18% of the post-move improvement due to selection into SEZs.

Table 3: Firm Dynamics Across Zones: Movers (into SEZ) vs. Stayers (in Non-SEZ)

	Before Move		After Move		Difference			Proportion of the After SEZ-NSEZ
	SEZ	NSEZ	SEZ	NSEZ	Before Move	After Move	Effect of SEZ	driven by Selection
Avg. Productivity (z_i)	0.84	0.64	1.73	0.95	0.2	0.78	0.58	0.26
Avg. Capital (k_i)	9.02	8.73	9.73	9.16	0.29	0.57	0.28	0.51
cov (z_i, k_i)	-0.04	-0.02	0.07	-0.01	-0.01	0.09	0.10	0.18

Figure 6: Firm Dynamics Across Zones: Movers (into SEZ) vs. Stayers (in Non-SEZ)



Notes: The figure visually distinguishes between firms moving from non-Special Economic Zones (Non-SEZ) into SEZs and those staying in Non-SEZ areas. In panel (a), we see an established SEZ highlighted in green within Region A, along with a pre-existing population of firms represented by gray dots. Panel (b) shows firms that potentially moved into the SEZ (blue dots) and contrasts them with firms that stayed in Non-SEZ areas (red dots). Panel (c) illustrates the paths of the moving firms, linking their original locations in Non-SEZ areas (red dots) to their new positions in the SEZ (blue dots). Finally, panel (d) displays the region after these movements, with blue dots representing the movers into the SEZ and red dots showing the firms that remained in Non-SEZs, depicting a later stage than panel (b).

Summary of Stylized Facts The summary of the empirical evidence shows that firms in Special Economic Zones (SEZs) exhibit better performance in terms of productivity, capital, and the correlation between productivity and capital over time, compared to those in non-SEZ areas. This better selection of the firms in SEZs is observed not only in firms originally established within SEZs but also in those relocating from outside. Both type of firms demonstrate enhanced performance relative to their counterparts outside SEZs. Additionally, firms that relocate to SEZs from outside zone display similar advantages post-relocation, including improved firm selection evidenced by higher average productivity, increased capital, and more effective resource allocation, in comparison to firms that remain outside the SEZs.

3.4 Agglomeration

A large of literature explore the concept of agglomeration, suggest that regional industry concentration enhances knowledge spillovers,idea sharing and innovation[13]. This section empirically examines how agglomeration influences the productivity differences between SEZs and non-SEZ areas. I describe the measurement of agglomeration in Appendix A.1.4, here I present results on productivity variations at different levels of agglomeration across zones. To further analyze how SEZs influence on productivity difference in the context of agglomeration, a mediation effect econometric model is employed to investigate these potential agglomeration effects. For detailed statistic description of firm-level productivity between SEZs and Non-SEZs can be found in Appendix A.1.3.

TFP and Agglomeration The Figure 7 shows the kernel density of firm's TFP distribution between non-SEZs and SEZs by agglomeration. Comparing two gray lines, We can see that SEZs have a higher mean log TFP, and the total distribution shifts to the right, suggesting that SEZs are associated with higher levels of firm productivity compared to NSEZs. Additionally, within both zones, the figure shows that higher agglomeration correlates with higher productivity, firms in high-agglomeration areas (solid-colored lines) tend to have higher productivity levels than those in low-agglomeration areas (dashed-colored lines). This indicates that agglomeration contributes to productivity since a higher density of firms (which may lead to better knowledge spillovers, more specialized suppliers, and a larger labor pool) is associated with higher productivity.

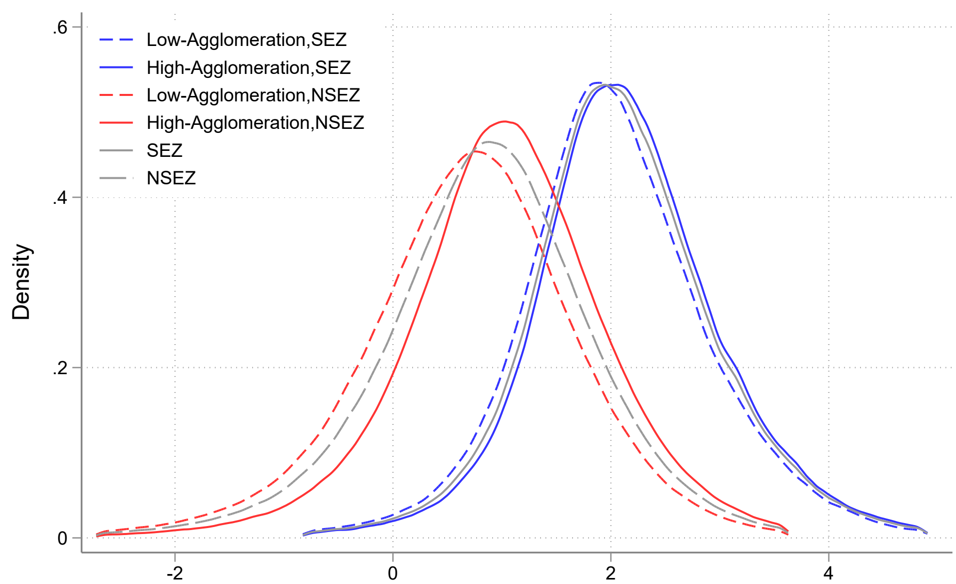
Moreover, comparing across zones, the productivity peaks for high-agglomeration areas in SEZs (solid blue line) are to the right of those in non-SEZs (solid red line), suggesting that firms in SEZs are more productive than those in non-SEZs at similar levels of agglomeration. This

[13] Agglomeration effects are widely studie (Marshall, 1890; Arrow, 1962; Romer, 1986), suggests that regional industry concentration enhances knowledge spillovers and innovation; (Jacobs, 1969) emphasizes the role of industrial diversity in cities in promoting cross-sectoral idea sharing and innovation

implies that while agglomeration is one of the factor in explaining productivity difference between zones, but it is not the only one, as SEZs have additional advantages. Additionally, the small gap between the dashed and solid lines within both SEZs and NSEZs suggests that there is an upper limit to the benefits of agglomeration. If agglomeration were the main mechanism, we might expect the productivity differences between low and high agglomeration to be more significant.

In summarized, figure 7 suggests agglomeration certainly impacts productivity, as evidenced by the differences within the zones, the fact that SEZs outperform NSEZs even at similar levels of agglomeration indicates that other factors are at play in enhancing productivity in SEZs. The exact impact of agglomeration relative to these other factors would require further analysis beyond the visual distribution presented in this figure. Thus, in the following section, I will evaluate the mediating role of agglomeration in explaining the higher productivity of firms in SEZs by employing a mediation effect econometric model.

Figure 7: Firm-Level Productivity by agglomeration Across Zones



Note: The figure plots the kernel density distribution of logarithmic firm-level Total Factor Productivity (TFP) in both SEZs and non-SEZs. The distributions across zones are showed using different line styles: a solid gray line for SEZs and a long-dashed gray line for non-SEZs. Additionally, the distribution of logarithmic TFP by agglomeration level is presented in blue for SEZs and red for non-SEZs. These blue and red lines vary in pattern to indicate the agglomeration level within each zone. A dashed line represents a low agglomeration level, indicating it is below the median, while a solid line represents a high agglomeration level, indicating it is above the median.

Event study and Mediation Effects Here, I use the staggered implementation of SEZs to conduct an event study. Of course, as documented in my previous Section 2.2 the evolution of the implementation of SEZs is not random, plus the firm entry into SEZs (either through birth or relocation from Non-SEZs) is also endogenous. Therefore, I cannot interpret the results of the event study as casual. With all these important caveats in mind, here I assess the presence of agglomeration effects using standard panel regression techniques. To examine the influence of SEZ policy on productivity through the lens of agglomeration, I pose a mediation effect model. Employing the econometric specification in equation (1), the model measures the agglomeration levels as a direct consequence of being in SEZs. Furthermore, equation (2) is used to determine the joint effect of the policy and the agglomeration mediator on firm productivity, thereby shedding light on the potential ways SEZ policies could be catalyzing productivity enhancements.

$$EG_{jrt} = \alpha_{rt} + \theta_j + \gamma D_{it} + \epsilon_{jrt} \tag{1}$$

$$TFP_{it} = \theta_i + \alpha_{rt} + \beta_1 D_{it} + \beta_2 EG_{jrt} + \delta X_{it} + \epsilon_{it} \tag{2}$$

Critical to this examination is the coefficient γ in the econometric model (1). A significantly positive γ would indicate that the SEZs has indeed intensified the degree of agglomeration. Subsequently, in equation (2), the focus shifts to β_1 and β_2, which evaluate the direct impact of the policy and the mediating effect of agglomeration on productivity, respectively. The interaction term, $\beta_2 \times \gamma$, is interpreted as the indirect effect of the SEZ policy on productivity via agglomeration, while β_1 represents the policy's direct effect on productivity. When both γ and β_2 are found to be significantly positive, it suggests that the observed increase in TFP is partly due to the enhanced agglomeration, indicative of a partial mediation effect. Conversely, a non-significant β_1 alongside a significant β_2 implies that the productivity gains are fully attributed to the heightened level of agglomeration, denoting a complete mediation effect. In this scenario, the agglomeration effect would be the only channel through which the SEZ policy exerts influence on productivity.

Table 4 displays results, where the columns indicate models that vary in their use of fixed effects: no fixed effects, fixed effects for region and year separately, and combined fixed effects for region and year. The top panel of the table details our baseline model outcomes, with the SEZ coefficient indicating the overall impact of the policy on productivity. The middle panel of the table reflects results from equation (1), where the mediator variable EG_{irt} is related to SEZ. The SEZ coefficient here indicates how the policy affects agglomeration levels. It appears that without fixed effects, the coefficient is positive, suggesting that the establishment of the economic zone slightly increases agglomeration by 0.002. However, this positive effect becomes statistically insignificant once fixed effects are included in the model. The bottom panel presents findings from Equation 2, showing both policy and agglomeration coefficients as positive and statistically significant. This implies that the policy effectively raises productivity and that denser, more agglomerated areas further enhance it. Nevertheless, due to the lack of significance in the policy's effect on agglomeration in the middle panel, the calculated indirect effect of economic zone establishment on productivity through agglomeration is not significant, indicating no mediation through this channel.

Table 4: Mediation Effect through Agglomeration on TFP

	(OLS)	(SepFE)	(corssFE)
Model with TFP regressed on SEZ (path c)			
SEZ	1.222 ***	0.909***	0.922***
	(775.16)	(316.53)	(321.20)
constant	.857***	0.912***	0.909***
	(1241.31)	(1367.41)	(1374.28)
Observations	2310570	2319020	2318971
R-sq	0.206	0.766	0.777
Model with mediator EG_irt regressed on SEZ (path a)			
SEZ	.002 ***	0.00000453	-0.0000282*
	(187.78)	(0.30)	(-1.86)
constant	.007***	0.00762***	0.00763***
	(1265.86)	(2175.36)	(2172.91)
Observations	2310570	2331564	2331508
R-sq	0.0150	0.881	0.884
Model with TFP regressed on mediator EG_irt and SEZ (paths b and c')			
Agglomeration	19.24***	6.502***	5.686***
	(242.25)	(45.73)	(40.33)
SEZ	1.176***	0.907***	0.920***
	(749.32)	(315.19)	(319.98)
constant	0.720***	0.868***	0.872***
	(810.88)	(681.96)	(690.01)
Observations	2310570	2294206	2294152
R-sq	0.226	0.766	0.777

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The figure shows how EG_{irt}, as a mediator of SEZs policy, affects on productivity. The column (1) represents the estimates when we control for separate fixed-effect of region and time, while column (2) allows each region control for their own trend, i.e. instead of control for separate fixed-effect here control for cross interact term between region and time.

The empirical evidence highlight not only a more effective selection process, favoring the most promising firms into the SEZs, but also a more efficient allocation of resources that enhances input utilization, fosters output growth, and increase capital accumulation and productivity gains. However, these observed differences cannot be solely attributed to the SEZ environment due to the endogeneity of firm entry. The decision for a new firm to be established in an SEZ or for an existing firm to migrate there is influenced by factors that are inherently linked to the firm's potential for success, thus complicating the causal attribution of performance improvements to the SEZ policy itself.

To address this complexity, a dynamic firm model accounting for endogenous entry, exit, and location choice is built. As I find in empirical part that in the context of chinese SEZs, the agglomeration effect doesn't play the main role in explaining the outperform of firms in SEZs compared to those in Non-SEZs. I abstain from explicitly modeling agglomeration effects. Such a model will enable such analysis of the aggregate effects of SEZs, understanding the mechanisms at play and quantifying the extent to which each contributes to the overall aggregate change.

Furthermore, the model will be instrumental in determining the optimal size of SEZs, essentially the ideal proportion of firms within SEZs. This involves examining whether it would be economically beneficial to include all firms within the SEZ framework or if a more selective approach is warranted. The goal is to identify an optimal SEZ firm share that maximizes economic benefits while minimizing potential drawbacks such as overcrowding or excessive competition.

4 Model

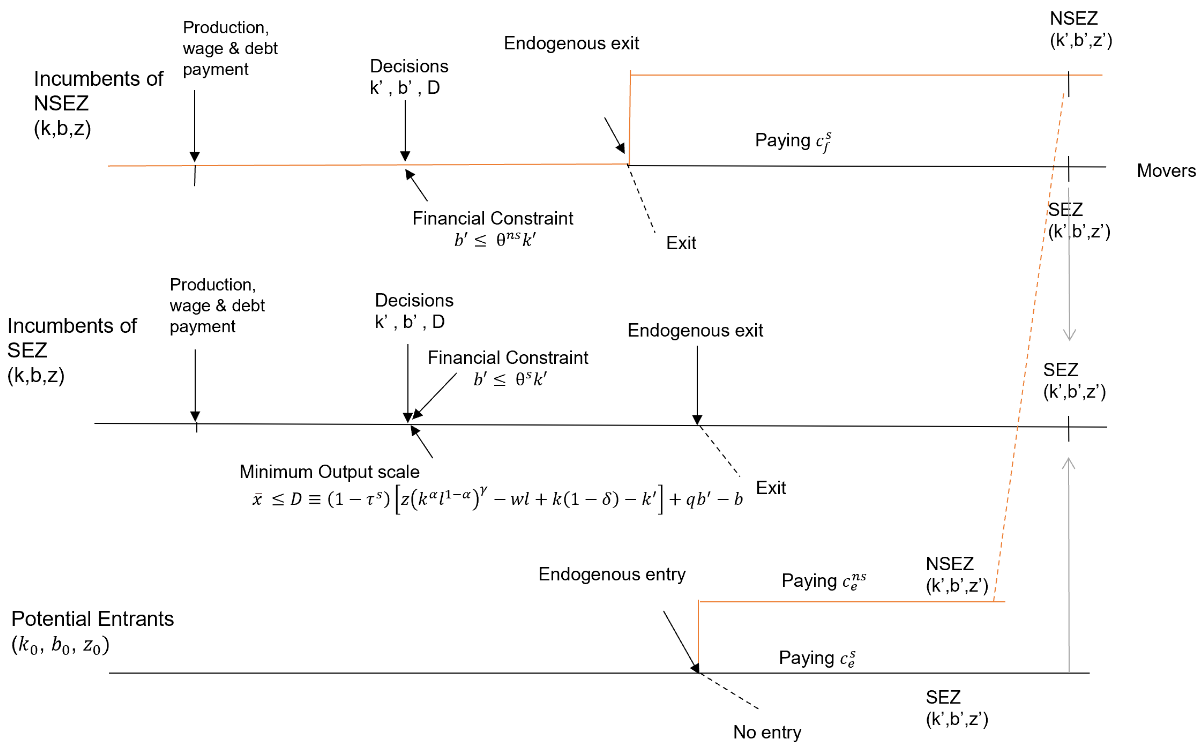
I build a firm dynamics model with endogenous entry and exit (Hopenhayn, 1992; Hopenhayn and Rogerson, 1993) enhanced with a discrete location choice. Specifically, in my model, firms are heterogeneous in terms of productivity and cash-at-hand (assets and bonds) and, at the same time that they face endogenous entry and exit decisions, they make location decisions between special economic zone (SEZs) and non-SEZs.[14]

To capture the distinctive characteristics of SEZs, I introduce a collateral constraint, with its tightness varying across different zones. Moreover, SEZ-based firms benefit from a reduced corporate tax rate and a minimal profit scale requirement, while firms in Non-SEZ have the option to endogenously move to SEZ. These three key features, combined with the endogenous entry and exit dynamics, enable us to employ this model for assessing the aggregate impact of SEZs on Total Factor Productivity (TFP), gaining insights into how SEZ contribute to overall TFP growth, and evaluating the optimality of SEZ policies.

In the following sections, we present our model economy and outline the assumptions underlying our model setup (Section 4.1). We then specify the timing of firms' decision-making processes (Section 4.1.1), followed by an examination of the optimization problem of firms (Section **??**). Next, we explore the household problem (Section 4.4) and, finally, provide the definition of a recursive competitive equilibrium (Section 4.5).

4.1 Model Setup

Time is discrete in infinite horizon. There are two zones in the economy, special economic zone (SEZ) and Non-special economic zone (Non-SEZ). The economy consists of a continuum of firms that are heterogeneous in productivity and producing a homogeneous good in two zones. Firms in both zones subject to persistent chocks to individual productivity. Each firm owns its predetermined capital, k, and bonds, b, from previous period, and hires labor, n. They pay different corporate tax for firms in SEZ and Non-SEZ, face different collateral constraints for firms in SEZ and Non-SEZ. Moreover, for firms in SEZ they need to keep up with the minimal profit scale in each period, while there is no such requirement for non-zone firms.



4.1.1 Timing

The life-cycle of firms is identical across zones, and in the following, we use $r \in \{S, NS\}$ to index firms in different zones. In each period, there is an exogenous measure M_e of potential entrants draw their initial state, (k_0, b_0, z_0), and decide whether to enter and which zone to enter in every period by paying the correspondence fixed entry cost of each zone. Upon entering the market, firm starts operating in the next period, and become an incumbent firm in the zone they decided to enter. At the end of each period, incumbent firms in SEZ choose whether to exit the market, but for incumbent firms in Non-SEZ they face two choices, they choose either to exit the market or move to SEZ by paying the moving cost to the SEZ.

All incumbent firms in both SEZ and non-SEZ solve a three-stage optimization problem after making the entry decision. In the first stage, conditional on their current period idiosyncratic productivity level z, predetermined capital,k, and the amount of debt,b, carried from the previous period, they choose optimal labor inputs to maximize their profits. In the second stage, conditional on their available resources, including after-tax profits, the firms in each zone decide whether to continue by paying fixed operation cost, c_f^r. For firms who decided to stay, they make

[14] Khan and Thomas (2011) pioneered the study discrete—non-convex adjustment cost—choices in the study of firm dynamics, further enhanced by entry and exit in Clementi and Hopenhayn (2006)

intertemporal decisions on investment, i, and borrowing, b, facing the collateral constraints and determining their current dividend, D, facing the minimal profit scale restriction for SEZ's firms. The capital accumulation of each firm is standard, $i = k' - (1-\delta)k$, with $\delta \in (0,1)$.

Markets are perfectly competitive, so firms take the wage rate, w, and the discount debt price, q, as given.

4.1.2 Productivity

Productivity distribution is firm-specific, denoted by $\mu_i \equiv \mu(z_i)$ for firm i. A firm's productivity $z_{it} = A_i v_{it}$ consist of a time-varing idiosyncratic component and a transitory component which capture the firms caracteristics that doesn't depend on time. The transitory component, A_i, follows Pareto distribution. The idiosyncratic component denoted by v_{it} and follows AR(1) process, given by

$$log(v_{it}) = \rho log(v_{it-1}) + \epsilon_{it}, \ \ \epsilon \sim N\left(0, \sigma^2\right) \tag{3}$$

4.2 Incumbent Firms: Location Choice

There is a continuum of incumbent firms, every of them make their location choice decision between SEZ and Non-SEZ and maximize their profit. Each firm i at time t use the same technology and exhibits decreasing returns to scale showing as follows:

$$y_{it} = z_{it}\left(k_{it}^{\alpha} l_{it}^{1-\alpha}\right)^{\gamma}$$

where $\gamma, \alpha \in (0,1)$. A higher v_{it} represents firm i is more productive than others, which is considered as relative productivity.

Assume that each period, firm draws its initial relative productivity from an exogenous distribution $\mu(z_i)$. Given the productivity z, predetermined capital k and bonds b, firms make production decisions, pay corporate tax τ^r and maximize their profit as follows:

$$\pi_{it}^{r}(k,b,z) = (1-\tau^r)[z_{it}(k_{it}^{\alpha} l^{l_{it}-\alpha})^{\gamma} - wl_{it} + k_{it}(1-\delta) - k'_{it}] + qb'_{it} - b_{it}$$

where corporate tax rate τ^r , $r \in \{s, ns\}$ are different between SEZ and Non-SEZ, according to the SEZ's policy where a reduced corporate tax rate is offered for SEZ's firms, then we have $\tau^s < \tau^{ns}$.

Given $z_i^r \in \mu(z)$, let $V_i^r(k,b,z)$ be the value of a firm i at the current period in region r, after firm decides the location to continue operating next period after paying the operating cost c_o^r. Firm choose the location $r \in \{s, ns\}$ in order to maximize their current value, if firm chooses to be in SEZ the continuation value of being in the SEZ should be higher than the continuing value in Non-SEZ. This location decision is described by the following discrete-choice problem:

$$v_i^j(k,b,z) = \max_{j\in\{s,ns\}} \{v_i^s(k,b,z), v_i^{ns}(k,b,z)\} \tag{4}$$

Where the first term in the curly bracket is the value of being in SEZ and the second term is the value of being in the Non-SEZ. Upon the current location, firm optimally chooses its labor, l, future capital k', and the optimal level of debt, b', to maximize the sum of the firm's current profit after paying the region-specific corporate tax rate, D, and the discounted expected values of next period, $V(k,b,z')$. Meanwhile financial frictions are introduced in a collateral constraint and firms facing this collateral constraint when they borrow in the above maximization problem. Additionally, for firms located in SEZs they also face a minimum profit scale requirement to keep up. Firm decides whether continue operating next period or exit from the market. They chooses to exit if the continuation value is lower than the exit liquidation value. If they exit next period they liquid the depreciated capital and pay bonds. Finally, conditional on the current location the value of firm's optimization problem can be defined by $V_i^r(k,b,z)$ as below.

$$V_i^r(k,b,z) \quad = \max_{l,b',k'} \pi_i^r(k,b,z) + \beta \mathbb{E}_{z'} \max\left\{V_i^X(k',b'), V_i^r(k',b',z') - \xi^r\right\} \tag{5}$$

s.t.

$$b' \le \theta^r k'$$

$$\bar{x}^s \le D \equiv (1-\tau^s)\left[z(k^\alpha l^{1-\alpha})^\gamma - wl + k(1-\delta) - k'\right] + qb' - b$$

where $\theta^s > \theta^{ns}$, $\bar{x}^{ns} = 0$, $\bar{x}^{ns} > 0$

$$V_i^x(k',b') = k'(1-\delta) - b'$$

Remember each period there is a group of potential entrants refer to the new firms born in SEZ and Non-SEZ, next we are going to define the value of those potential entrants indexed by (k_0, b_0, z_0).

4.3 New firms: Entry (Birth) and Location Choice

Potential Entrants (Birth)

The potential entrant makes a discrete location choice about in which zone to be born, whether to enter the economy and become an incumbent after paying the region-specific entry cost c_e^r , $r \in \{s, sn\}$. Once it enters a specific zone, the firm starts operation in the next period given its initial state, and the value of entry is denoted by the following entry condition 6.

$$V_e(k,b,z) \ = \max \left\{ \underbrace{0}_{\text{No Birth}}, \underbrace{V^{ns}(k,b,z) - c_e^{ns}}_{\text{Birth in NSEZ}}, \underbrace{V^{s}(k,b,z) - c_e^{s}}_{\text{Birth in SEZ}} \right\} \tag{6}$$

where V^s and V^{ns} is the value function for incumbent firms in SEZ and Non-SEZ. A potential entrant chooses to enter if the value of operation in one of the zone conditional on their initial state $V^r(b_0, k_0, z_0)$ is greater than the no birth value, 0.

Exit

An incumbent firm has the option to endogeneously exit the market. Each period, the firm upon its new productivity and decides whether to continue operating or exit. Upon exit, the firm's liquidation value is determined as the depreciated value of its capital k, minus any bond obligations. This discrete exit decision can be expressed as follows in the equation: 7:

$$\max \left\{ V_i^X(k', b'), V_i^r(k', b', z') - \xi^r \right\} \tag{7}$$

where $V_i^r(k', b', z')$ is the value of continuing production for next period, and $V_i^X(k', b')$ is liquidation value. A firm choose to continue operating if and only if $V_i^r(k', b', z') - \xi^r \geq V_i^X(k', b')$.

4.4 Households

"Households in this economic model are identical and live in infinite time horizon. In each time period, households make decisions regarding their consumption and labor, aiming to maximize their current utility function $U(C, 1-N)$, and the expected discounted utility from future periods. These decisions are made subject to a budget constraint that incorporates labor income, government transfers (T), and the returns on non-contingent discount bonds (b) from the previous period. Household value, denoted as $V^h(b)$, can be expressed using the following equation 8

$$V^h(\phi) = \max_{C^h, N^h, \phi'} U(C^h, 1 - N^h) + \beta V^h(\phi') \quad (8)$$

s.t.

$$C^h + q\phi' \leq wN^h + \phi + T$$

where

$$T = \int_{\{(k,b,z)|j(k,b,z)=s,ns\}} \tau^j(y - wl - k' + (1-\delta)k) d\mu^p(k,b,z)$$

and q denotes the bond price.

4.5 Industry Equilibrium

In the following we define a stationary competitive industry equilibrium of the model. We denote $\mu^p(k,b,z)$, $\mu^e(k,b,z)$ $\mu^{ex}(k,b,z)$ as distribution of producing firms, new birth firms and firms exit the market respectively. Given the time invariant distribution of capital, bonds and productivity $\mu(k,b,z)$, a stationary competitive equilibrium consists of prices (w,q), value functions $V^r(k,b,z), V_e^r(k,b,z), V^x(k,b,z), V^h(k,b,z)$; agents' policy functions (C^h, N^h, Φ^h); firms' policy functions $l(k,b,z)$, $k(k,b,z)$, $b(k,b,z)$, $j(k,b,z)$ such that:

1. V^{se}, V^{ns} solve incumbent firms' problem 4 - 5, and $l(k,b,z)$, $k(k,b,z)$, $b(k,b,z)$, $j(k,b,z)$ are the associated policy functions for firms;
2. $V_e^r(k,b,z)$ solve new firms' problem 6
3. V^h and (C^h, N^h, Φ^h) the associated policy functions solve household problem 8;
4. The labor market clears

$$N^h = \int_{\{(k,b,z)|j(k,b,z)=s,ns\}} l(k,b,z) d\mu^p(k,b,z)$$

5. Asset market clears

$$\phi^h = \int_{\{(k,b,z)|j(k,b,z)=s,ns\}} b(k,b,z)d\mu^p(k,b,z) - \int_{\{(k,b,z)|j(k,b,z)=s,ns\}} b(k,b,z)d\mu^{ex}(k,b,z)$$

6. The goods market clears.

$$C^h = \int_{\{(k,b,z)|j(k,b,z)=s,ns\}} \left[z(l^\alpha k^{1-\alpha})^\gamma - (k' - (1-\delta)k) - \xi^j\right] d\mu^p(k,b,z) + \int_{\{(k_0,b_0,z_0)|j(k_0,b_0,z_0)=s,ns\}} (k_0 - c_e^j)d\mu^e(k_0,b_0,z_0) - \int_{\{(k,b,z)|j(k,b,z)=s,ns\}} (1-\delta)kd\mu^{ex}(k,b,z)$$

7. Resource Constraint

$$T = \int_{\{(k,b,z)|j(k,b,z)=s,ns\}} \tau^j(y - wl - k' + (1-\delta)k)d\mu^p(k,b,z)$$

8. Distribution follow the law of motion:

$$\mu(k',b',z') = \int_{\{(k,b,z)|j(k,b,z)=s,ns\}} d\mu^p(k,b,z) + \int_{\{(k_0,b_0,z_0)|j(k_0,b_0,z_0)=s,ns\}} d\mu^e(k_0,b_0,z_0) - \int_{\{(k,b,z)|j(k,b,z)=s,ns\}} d\mu^{ex}(k,b,z)$$

4.6 Decisions Rules for Heterogeneous Firms

In this section, we characterize the decision rules for firms of different types by distinguishing them based on whether they are unconstrained firms, solely financially constrained firms, or firms constrained both financially and by minimal requirements. An unconstrained firm is one that never faces binding financial constraints and possesses sufficient wealth, thereby ensuring that minimal profit scale requirements are never limiting.

This approach facilitates the derivation of intertemporal decisions for capital k, bonds b, and labor l for each type of firm. It's worth emphasizing that a firm's classification may change over its lifecycle, contingent on its state variables. For a more detailed solution regarding decision rules for unconstrained firms are in Appendix **??**.

Solving the model with three state variables can be quite challenging. Since capital k and bonds b jointly determine the choices of k' and b', it allow us to collapse these two state variables into a new variable referred to as *"cash-on-hand"*, denoted as $m(k, b, z)$, and defined as follows:

$$m(k, b, z) \equiv (1 - \tau)\left[z(k^{\alpha}\hat{L}^{(1-\alpha)})^{\gamma} - w\hat{L} + (1 - \delta)k\right] - b$$

where $\hat{L}$ is the optimal static labor choice solved from the unconstrained firms problem in appendix **??**. A firm with (k, z) chooses optimal labor demand as

$$\hat{L}(k, z) = \left[\frac{(1 - \tau) * (zk^{\alpha\gamma}(1 - \alpha)\gamma)}{w}\right]^{\frac{1}{1-(1-\alpha)\gamma}}$$

Subsequently, upon firms' location we can reformulate their problem with this new collapsed state variable $m(k, b, z)$, as follows:

$$V^r(m, z) = \max_{k',b',D,m'_j}\left[D + \max\Big\{V_x(m), \beta\int_{z'} V(m', z')dG(z'|z)\Big\}\right] \tag{9}$$

$$\begin{aligned} s.t. \qquad & \bar{X}^r \leq D \equiv m - k'(1 - \tau^r) + qb' \\ & b' \leq \theta^r k' \\ & m' \equiv m(k', b', z') \\ & = (1 - \tau^s)\left[z'(k'^{\alpha}\hat{L}^{(1-\alpha)}(k', z'))^{\gamma} - w\hat{L}(k', z') + (1 - \delta)k'\right] - b' \end{aligned}$$

where $r \in \{s, ns\}$, $\bar{X}^{ns} = 0$ since there is no minimal profit scale requirement for firms.

For an unconstrained firm that follows unconstrained capital policy $\hat{K}'$ and bond policy $\hat{B}'$, solved from their maximization problem, maintains a current profit level of $\hat{D} = m - \hat{K}'(1-\tau^r) + q\hat{B}' \geq 0$. However, if a firm's available cash-on-hand m, falls below a specific threshold value, denoted as $\bar{m}$ where $\bar{m} \equiv \hat{K}'(1 - \tau^r) + q\hat{B}'$, then that firm faces a binding minimal profit scale requirement. Moreover, some of these firms face binding financial constraints, limiting their ability to invest up to the extent allowed by their collateral value k'. The upper limit on their capital choice, which serves as collateral, is established through the minimal profit scale requirement condition. It is defined as $\bar{K} \equiv \frac{m-\bar{X}}{(1-\tau^r)-q\theta^r}$. As the financial parameter θ^r approaches q^{-1}, this upper limit extends to infinity. Under these circumstances, we relax the financial constraints imposed on firms, effectively allowing them to operate as unconstrained firms. When these previously constrained firms adopt an optimal capital policy, with $k' = \bar{K}$, their optimal bond policy can be expressed as $b' = \frac{1}{q}(\bar{K}(1-\tau^r) + \bar{X} - m)$. This bond policy is derived from the binding profit scale condition. In summary, upon the location, depending on the firm's cash-on-hand situation, there are three types of firms that employ distinct decision rules for capital k' and bonds b:

- Unconstrained Firms with $m > \bar{m}$: These firms do not face a binding minimal profit scale requirement or financial constraints. They employ unconstrained policy functions for capital and bonds, denoted as $\hat{K}'$ and $\hat{B}'$, respectively.
- Only minimal profit scale binding $m < \bar{m}$: In this case, the minimal profit scale requirement is binding, but there are no financial constraints. These firms adopt an unconstrained capital policy, where $k' = \hat{K}$, and an optimal bond policy defined as $b' = \frac{1}{q}(\hat{K}(1-\tau^r) + \bar{X} - m)$.
- Both minimal profit scale and financial constraint binding $m < \bar{m}$: For these firms, they adopt a specific capital policy, where $k' = \bar{K} \equiv \frac{m-\bar{X}}{(1-\tau^r)-q\theta^r}$, and an optimal bond policy with upper bound capital employed $b' = \frac{1}{q}(\bar{K}(1-\tau^r) + \bar{X} - m)$.

5 Calibration and Estimation

In this section, I provide an overview of the model's parameterization, the calibration strategy, and the evaluation of the model's empirical performance. The key objectives in quantifying the model's parameters are threefold. First, the model aims to accurately replicate the distribution of firms' value-added, particularly focusing on the skewed nature of the distribution, where the top 5% of firms account for approximately 34% of the total value-added. Second, it is crucial that the model aligns with empirical evidence by ensuring consistency in several aspects. This includes the share of firms located in SEZs, the relative average productivity of SEZ firms compared to Non-SEZ firms, and the relative average productivity of firms born in SEZs versus those born in NSEZs. The third goal involves matching the relative average leverage levels, expressed as the ratio of debt to capital, for SEZ firms compared to NSEZ firms. This helps in characterizing different financial tightness levels across zones, and also the exit rate from SEZs, which captures the minimal profit scale requirement.

The model's parameters fall into three categories. The first group comprises parameters with values from existing literature and align with China's Special Economic Zones policy. The second group of parameters is determined independently to replicate the characteristics of firms in the Chinese economy. The third group includes internally-calibrated parameters, obtained through

Simulated Methods of Moments (SMM), as discussed in Section **??**. These parameters are matched to steady-state moments from the model with those observed in firm-level panel data spanning from 1998 to 2013. Overall, seven moments are selected to characterize the distinct behaviors of newly-born and incumbent firms across zones.

The remainder of this section will be the following. First, Section 5.1 presents parameters that are externally calibrated, with values directly sourced from data. Second, Section 5.2 discuss the internally calibrated parameters, which capture the characteristics of firms within and outside Special Economic Zones in the Chinese economy. These parameters are fine-tuned to ensure the model aligns with empirical data. Once the model is calibrated, Section **??** demonstrates how a firm's individual state influences its decisions regarding investment, borrowing, entry, and exit, shedding light on key model components. Additionally, I provide insights into the model's performance in depicting firm dynamics across different zones.

5.1 Externally Calibrated Parameters

We conduct model calibration on an annual basis. In Table 5 presents the externally calibrated parameters. The corporate income tax rate in Non-SEZ $\tau^{ns} = 0.33$, aligns with China's corporate income tax policy for firms in Non-SEZ, and the corporate income tax rate in SEZ $\tau^{s} = 0.195$, falls within the range of income tax rates for firms in SEZs in China, varying from 15% to 24%. The capital share parameter $\alpha = 0.37$ and the span of control $\gamma = 0.862$ are estimates obtained from the Olley & Parks method used to estimate firm's Total Factor Productivity (TFP) based on firm-level data. The time discount rate $\beta = 0.961$, is chosen so that the long-run equilibrium interest rate approximates the standard rate of about 4% per annum. The capital depreciation rate $\delta = 0.068$ is a commonly used value in the literature.

Productivity process

To establish the productivity distribution of firms, I used the Annual Survey of Industrial Enterprises from 1998 to 2013. The distribution of firms' productivity can be well captured by a Pareto distribution, The choice of the shape parameter, μ, is chosen to match the 95th, 75th, 50th and 25th percentiles of the productivity distribution among firms. The transition matrix for firm productivity is selected to capture certain dynamic characteristics observed in firm-level output data. Specifically, we aim to match the standard deviation of the change in log output across firms and the one-year autocorrelation of individual output. These moments serve as key reference points for calibration and have been employed in previous studies, such as Midrigan and Xu (2014), to calibrate the transition dynamics of firm productivity.

Table 5: Externally Calibrated Parameters

Parameter		Value
Corporate income tax rate NSEZ	τ^{ns}	0.33
Corporate income tax rate SEZ	τ^{s}	0.195
Discount factor	β	0.961
Capital Share	α	0.37
Depreciation rate	δ	0.068
Span of control	γ	0.862
Shock standard deviation	σ	0.0077
Shock persistence	ρ	0.7968
Pareto shape parameter	μ	8.6955

5.2 Internally Calibrated Parameters

In this section, I outline the seven remaining parameters listed in Table 6, which are internally estimated within the model. The selection of these parameters involves minimizing the distance between model statistics and their empirical counterparts from the data. While all of these parameters are jointly calibrated by using the SMM method, I will provide an explanation of why certain targeted moments are mostly influenced by specific parameters in the following.

I choose the parameter θ^r in order to match the average debt-to-capital ratio for firms in SEZs and Non-SEZs in the empirical data. A more tighter collateral constraint in Non-SEZs, characterized by a smaller θ^{ns}, reduces the marginal capital that can be employed to acquire bonds, consequently lowering the average bond-to-capital ratio for firms in Non-SEZs. It's important to note that this moment is also influenced by other factors, such as fixed operating costs and entry costs.

The parameter ξ^r, which represents fixed operating costs, predominantly affects the number of loss-making incumbent firms. Therefore, I calibrate this parameter to target the share of profit-making firms in each zone. Under my calibration, firms in SEZs incur lower fixed costs, suggesting that surviving firms in SEZs can be smaller than their counterparts in Non-SEZs, thanks to the favorable policies in SEZs.

Calibrating the minimal profit scale for firms in SEZs, denoted as $\bar{x}$, and the entering cost, c_e^r, is estimated by targeting various moments. These moments include the exit rate from SEZs, the relative initial assets for newly-born firms compared to incumbents, the relative average productivity for incumbent firms in SEZs versus firms in Non-SEZs, and the relative average productivity for newly-born firms in SEZs compared to those born in Non-SEZs. A higher minimal profit scale in SEZs indicates that firms in SEZs are expected to be more profitable. A smaller entry cost allows potential entrants with lower initial assets to survive in SEZs, and a more relaxed collateral constraint makes productive firms less constrained in SEZs. Consequently, both newly-born and incumbent firms in SEZs are more productive relative to firms in Non-SEZs.

Table 6: Internally Calibrated Parameters

Parameter		Value	
		SEZs	NSEZs
Collateral Constraint	θ^r	0.88	0.62
Fixed Operating cost	ξ^r	0.01	0.034
Minimal profit scale	$\bar{x}^r$	0.003	0
Entering cost	c_e^r	0.0081	0.0083

5.3 Model Performance

In this section, I show that my model predictions match reasonably well with the empirically observed important moments that I directly target in the calibration process, as shown in Table 7.Importantly, one of the key moment in this study to be targeted is that, the average productivity for firms in SEZ should be 2.5 times greater than that of NSEZ in the empirical data. The model successfully replicates this feature of the moment, albeit with a slightly smaller magnitude. Additionally, the average relative productivity for new-born firms in SEZ is 2.5 times that of new-born firms in NSEZ. My targeted moment captures the characteristic of higher productivity for new-born firms in SEZ, with a magnitude that exceeds the empirical finding. The average debt-to-capital ratio is lower in SEZ and my model appropriately captures this feature. While there is room for further improvement in these moment targets, they currently closely approximate the empirical data and capture the most critical features of the data.

5.4 Steady State: Firm Decision Rules

In this section, I present the distribution of firms from the model. Firstly, I showcase the distribution of productivity and output at the steady state, showing a highly skewed distribution with over 85 percent of firms exhibiting a particular output level. In the following, I show how heterogeneous firms adopt distinct decision rules that are endogenously determined by a combination of persistent productivity shocks, collateral constraints, and endogenous extensive margin adjusted through firm entry and exit.

Firms facing financial constraints are tend to be smaller firms with limited assets, characterized by lower cash-on-hand m. This limitation is imposed by their productivity and borrowing limits. To

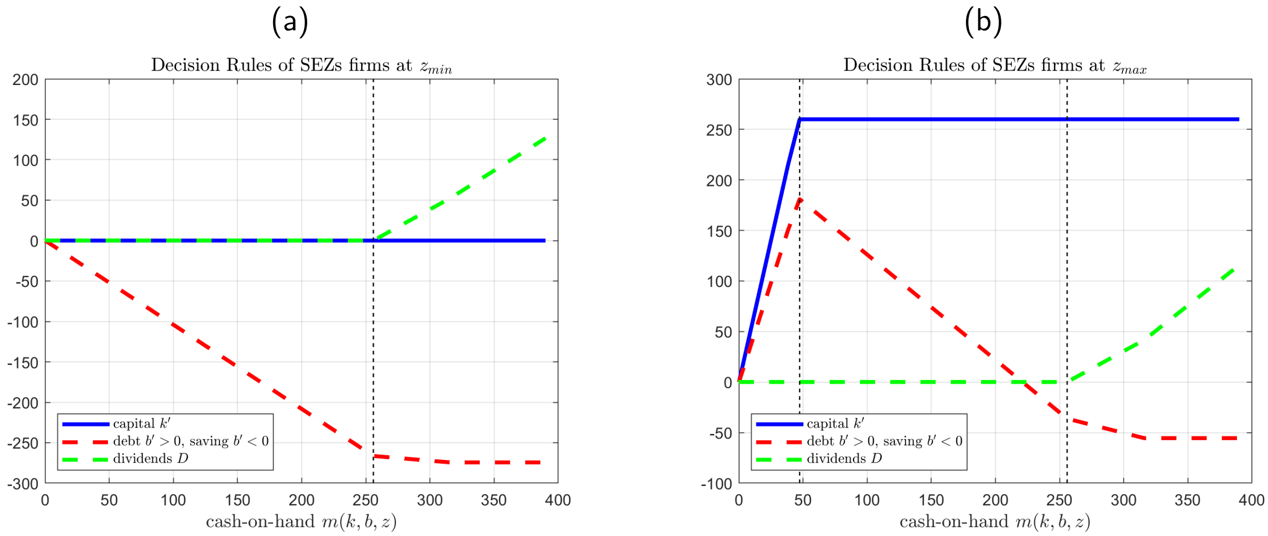
Table 7: Moments Used in Calibration

Moments				
Target Moments	Data		Model	
I/Y	.0847		3.2709	
wL/Y	.7012		.4585	
$AvgTFP^s/\ AvgTFP^{ns}$	2.4715		1.188	
$AvgTFP_0^s/\ AvgTFP_0^{ns}$	2.5305		8.98	
Exit rate from SEZ	.10		.2827	
New business (%)				
Relative B_0 to Incumb	.1827		.3034	
	SEZ		NSEZ	
	Data	Model	Data	Model
Average leverage (debt/capital)	.9590	0.7343	.9622	1.046

better understand this concept, Figure 8 illustrates the heterogeneous decision rules of incumbent firms on capital k', debt b', and dividends D as functions of cash-on-hand $m(k, b, z)$ for a given productivity value z. Panel (a) of the Figure 8 displays the decision rules for the least productive firm, while panel (b) shows the decision rules for the most productive firm. Two vertical dashed lines in the figure represent two thresholds for cash-on-hand, which help categorize firms into different types based on $m(k, b, z)$. Firms with cash-on-hand exceeding the upper threshold $\bar{m}$ around $m = 250$, are classified as unconstrained firms. Firms falling between the lower threshold $\underline{m}$ situated near $m = 50$, and the higher threshold $\bar{m}$, face solely profit scale constraints. Those firms with cash-on-hand lower than the lower threshold $\underline{m}$ contend with both profit scale and financial constraints.

In Panel (a), for surviving firms with the lowest productivity z_{min}, being considered unconstrained and their wealth is larger than the upper threshold m, such that $m > \bar{m}$. These unconstrained firms adopt a unconstrained optimal choice of capital $k' = \hat{K}$ and bonds $b' = \hat{B}$, while maintaining positive dividends $D > 0$. Firms with cash-on-hand less than $\bar{m}$ still employ to the optimal capital rule $\hat{K}$, but they follow a zero-profit decision. They gradually reduce debt and accumulate internal financial savings $b' < 0$, ultimately transitioning into unconstrained firms. It's important to note that firms with low cash-on-hand and low productivity may not survive in the economy if they are positively leveraged. Consequently, positive borrowing rules may not be observed in Panel (a). In contrast, firms with high productivity z_{max} in Panel (b) are able to invest up to their collateral value when their cash-on-hand is below the lower threshold $\underline{m}$, and their capital choice is constrained with positive borrowing. Analyzing the distribution of cash-on-hand

Figure 8: Decision Rules: Capital, Bond and Dividends



Notes: The figure plots decision rules of capital k', represented by solid blue line, bond $b' > 0$ and financial savings $b' < 0$, represented by red dash line and dividends D dashed green line. Panel (a) of the Figure 8 displays the decision rules for the least productive firm z_{min}, while panel (b) shows the decision rules for the most productive firm z_{max}. Two vertical dashed lines in the figure represent two thresholds for cash-on-hand to distinguish type of firms.

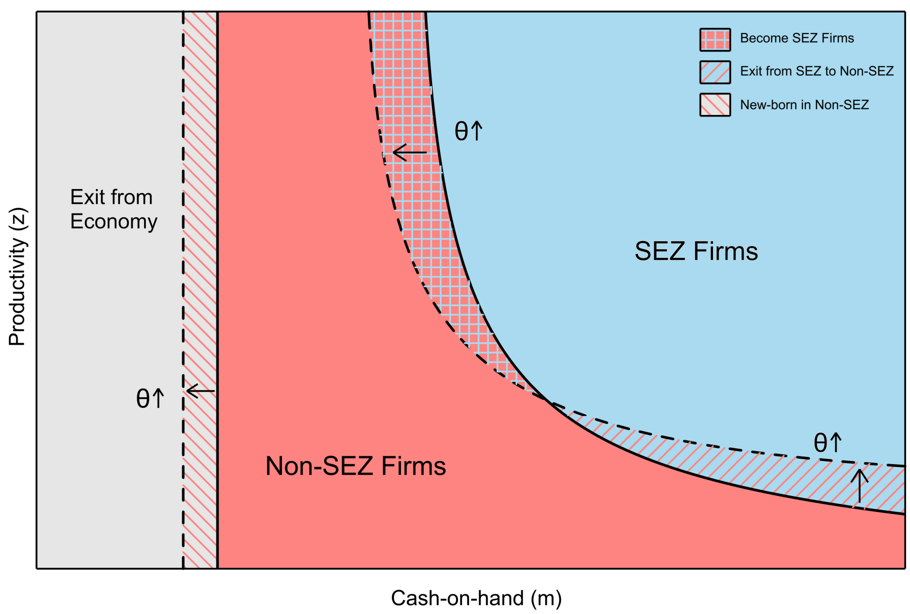
among firms, we observe that small firms with lower cash-on-hand are concentrated at the lower tail of the distribution while maintaining positive leverages.

Location choice

The choice of firms' location is jointly determined by their wealth and productivity. Figure 9 provides a visualization of how cash-on-hand and productivity influence a firm's location choice. It also demonstrates the impact of reducing financial frictions (reflected in an increase in θ) on firms' location choices.

By reducing financial frictions and relaxing financial constraints, highly productive firms in Non-SEZs, which have been constrained by borrowing limits, are empowered to invest more capital and expand their businesses in SEZs. This surge in investment contributes to an increase in the equilibrium interest rate. Simultaneously, less productive but wealthier SEZ firms opt to exit the SEZ, transitioning from borrowing to saving capital. In contrast, highly productive but low-wealth firms choose to operate within the SEZ, initiating capital borrowing and business operations. As a result of these extensive margin adjustments, we observe a higher average productivity for firms located in SEZs. This selection process plays an important role in reducing the misallocation of high-ability firms in the economy. In this new equilibrium, a set of firms with

Figure 9: Location Choice: θ Increases



Notes: X-axis presents different level of cash-on-hand m, and y-axis presents different level of productivity z. Firms not in the economy are represented by the gray area on the left, Non-SEZ firms are represented by the red area in the middle, and SEZ firms are represented by the blue area in the top-right corner. Solid line represent the distribution division from the benchmark, dash line represents the division after increase financial parameter θ.

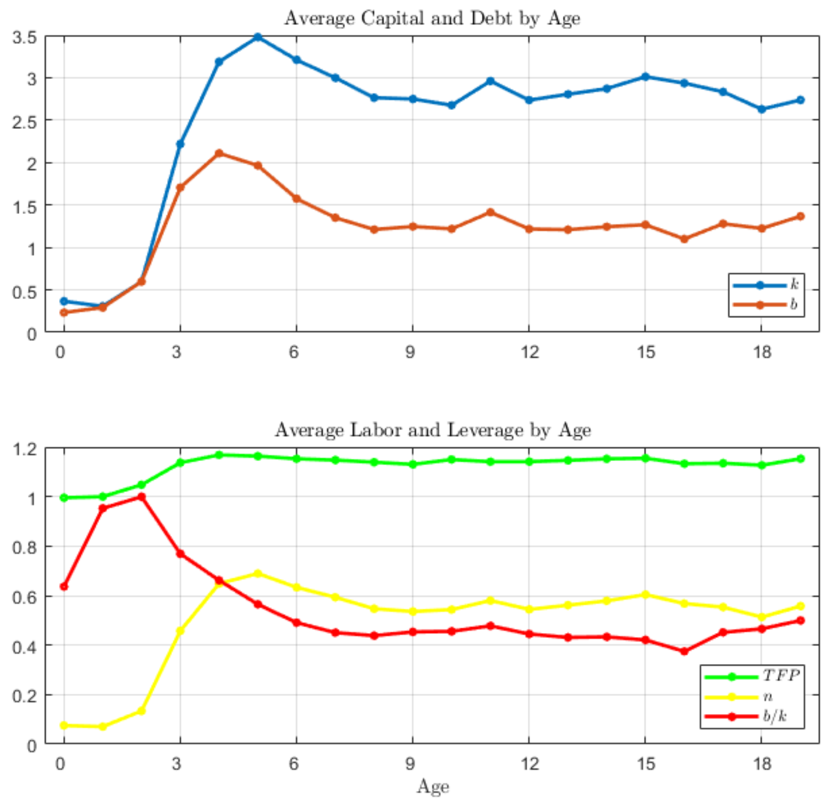
higher average productivity and lower average wealth emerges, which facilitates the entry of more firms with lower wealth into the economy as Non-SEZ firms, as indicated by the shaded area. On the aggregate level, following an increase in θ, output in the SEZ increases, more productive firms expand, and there is growth in aggregate output, productivity, and capital.

5.5 Firm Lifecycle Dynamics

Following the discussion on firms' location choice based on their cash-on-hand and productivity, this section shows the life-cycle dynamics of firms in the model. It provides insight into the heterogeneous dynamic paths observed across zones, showcasing how firms' output, productivity, capital, and leverage levels change over time.

As plotted in Figure 11, it is evident that all young firms start relatively small at the time of their birth and gradually accumulate capital as time goes by. At age 0, these firms experience financial constraints due to their limited capital for external financing. However, they continue to borrow funds until around age 4, at which point they begin deleveraging once capital is accumu-

Figure 10: Firm Lifecycle Dynamics: Average capital, Bond, Productivity and Labor

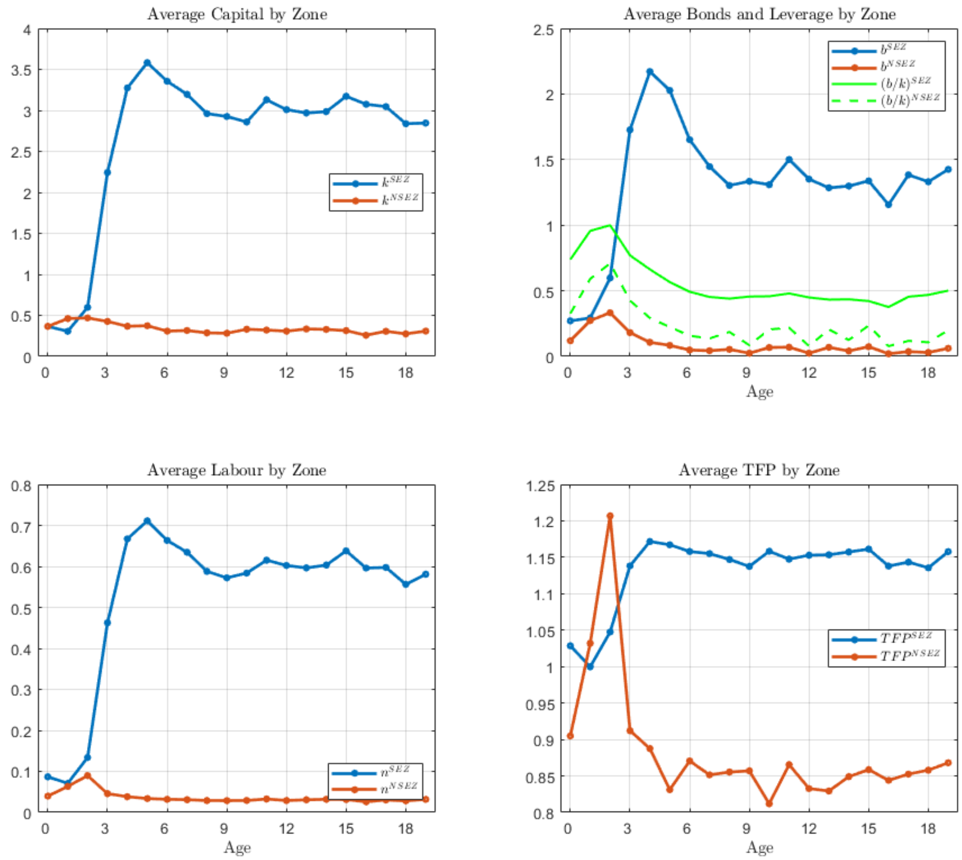


Notes: The upper panel illustrates the changes of average capital, represented by the blue line, and the average bond, indicated by the red line, throughout the firm's lifecycle. The lower panel plots the average productivity, indicated by the green line, labor, shown in yellow, and the leverage level, denoted by the bond-to-capital ratio, represented by the red line, changes over firm's age.

lated sufficiently. By around age 7, firms can adapt to the optimal unconstrained investment level $\hat{K}$, this featured is also showed in the lower panel of the figure by the hump-shaped red leverage curve. Due to the presence of SEZs, which effectively reduce financial frictions in the economy, enables firms to borrow more as they desire, as long as they possess enough capital. Consequently, as firms accumulate capital over time, they continue to borrow externally for investment rather than relying on internal savings. This results in the absence of unconstrained firms with positive profits. Additionally, new firms enter the market with higher productivity levels and initiate labor hiring for production. Productivity increases gradually as firms accumulate capital, maintaining a constant level once capital reaches its optimal level. The dynamics of firm lifecycles exhibit notable differences between firms located in SEZ and Non-SEZ. In both cases, young firms commence from a same low capital level. However, firms in SEZ begin accumulating capital at a faster rate, reaching a substantially higher level around age 3, followed by a significant surge thereafter. In contrast, firms in Non-SEZ maintain relatively small capital levels, as indicated in the top left panel of Figure 10. The increased capital accumulation in SEZ firms and more relaxed financial constraint enables them to borrow more and invest at a higher capacity, resulting in their becoming larger firms compared to those in Non-SEZ. Due to the higher financial frictions for firms in Non-SEZ, they are unable to secure external financing for investments at their desired levels. This discrepancy is illustrated in the top right panel of Figure 10, where SEZ firms exhibit higher leverage levels than their Non-SEZ counterparts.

Furthermore, new-born firms in SEZ regions already demonstrate higher productivity levels from their age 0, surpassing those in Non-SEZ. Their productivity gradually increases in SEZs, peaking around age 4 due to the substantial capital accumulation. As the SEZ firms get mature, their productivity stabilizes as they begin deleveraging and maintain an optimal unconstrained capital level. Due to the favorable conditions offered in SEZ for capital accumulation and productivity growth, enabling firms to expand rapidly. A relaxed financial constraint allows them to access external financing to invest at their desired levels, leading to increased leverage and ultimately higher productivity. In contrast, Non-SEZ face constraints that hinder capital accumulation and thus the ability to invest and grow at the same pace. Consequently, we observe a higher average capital, leverage levels, labor, and productivity for firms in SEZ. These differences in dynamics between SEZ and Non-SEZ highlight the impact of SEZ policy on firm growth, contributing to aggregate growth.

Figure 11: Firm Lifecycle Dynamics by zone: Average capital, Bond, Productivity and Labor



Notes: The upper panel illustrates the changes of average capital, represented by the blue line, and the average bond, indicated by the red line, throughout the firm's lifecycle. The lower panel plots the average productivity, indicated by the green line, labor, shown in yellow, and the leverage level, denoted by the bond-to-capital ratio, represented by the red line, changes over firm's age.

6 Quantitative Experiments

In this section, we present the main quantitative results of the paper. After calibrating the model to align with critical features like firm productivity, the distribution of bonds, and the capital for both incumbent and new-born firms in Special Economic Zones (SEZ) and Non-Special Economic Zones (Non-SEZ), using the model to connect with the micro evidence. I can then address questions related to aggregate-level analysis and quantify the aggregate impact of Special Economic Zones policies on output, capital, and productivity. Furthermore, we undertake a decomposition exercise of each of the features of the policy, provides me a comprehensive understanding of how each of the individual components of SEZ policy affect the economy. This analysis also contributes to discussions on optimal policy considerations, specifically concerning the maximum capacity and optimal share of firms located in SEZ. In the following a counterfactual exercise of absence of SEZ is detailed in Section 6.1. Then in Section 6.2, I dissect the impact of shutting down the credit subsidy channel for SEZ firms by tightening the financial constraint of SEZ to be the same as in NSEZ. Lastly, in Section 6.3, upon increased financial frictions in the previous section, I keep shutting down the tax subsidy policy for SEZ firms, increasing corporate tax rate to be the same as in Non-SEZ. This quantitative experiments quantify the effectiveness of each policy in rectifying resource misallocation across firms and in promoting the aggregate productivity and the total welfare.

6.1 No SEZs Counterfactual

First, I conduct a counterfactual exercise by examining the hypothetical scenario where Special Economic Zones (SEZs) do not exist in the economy. In this scenario, I align SEZ features to be the same as of NSEZ, which involves increasing financial frictions by setting down the financial constraint parameter θ^s equal to that of NSEZ θ^{ns}, raising the corporate tax rate τ^s, and eliminating the minimal profit scale requirement for firms in SEZ. This analysis provides insights into the significant impact of SEZs on several key economic variables. Particularly, I focus on Aggregate Total Factor Productivity (TFP), a engine that drives economic growth, as well as the allocation of capital, changes in output and input, and their aggregate implications. Furthermore, we investigate the fiscal implications of SEZs, through effective tax measurement and the effects of SEZs on total welfare.

Table 8 presents the outcomes of a counterfactual analysis, comparing the benchmark economy with Special Economic Zones (SEZs) to a scenario without SEZs. In the benchmark economy, aggregate total factor productivity (TFP) exhibits a remarkable increase of 25.7% relative to the counterfactual. The measurement of aggregate TFP, is defined as Solow residual of $\frac{Y}{(k^{\alpha}L^{(1-\alpha)})^{\gamma}}$.

It is calculated using aggregate values of output Y, capital K, and labor L, with fixed values of the parameters α and γ from the benchmark model.

There are three channels that contribute to the observed aggregate TFP gain. The first channel involves improved selection dynamics, focusing on changes in average productivity per firm and extensive margin adjustments in the benchmark economy. The rationale behind examining average productivity is to account for the differences in the number of firms between the benchmark and counterfactual scenarios. In an economic environment characterized by decreasing returns to scale technology in production, the absence of SEZs leads to an increased number of firms. However, this surge in the number of firms results in each firm operating at a sub-optimal scale, reducing production efficiency and ultimately affecting the TFP. Therefore, it is essential to understand whether the increased aggregate TFP is primarily due to the larger number of firms in the economy or whether it is a result of enhanced efficiency stemming from the selection channel. By comparing average productivity, which is computed by re-scaling aggregate variables by the number of firms, we can discern the true driver behind the increase in aggregate TFP. Regarding the extensive margin adjustment, the economy with SEZs exhibits a 2.82% reduction in the birth rate, allowing only more productive firms to enter the market. This leads to a 12.4% increase in the average TFP of newly born firms, accompanied by a higher death rate in the benchmark. This mechanism favors the inclusion of more productive firms born in the SEZs while expelling less productive firms from the market. As a result, there is an substantial increase in the average TFP by 25.10%.

The second channel contributing to the aggregate TFP gain results from a more efficient allocation of factor inputs, fostered by SEZs through a combination of credit subsidy policies that alleviate financial frictions and lower corporate taxes that reduce tax frictions. Focusing on the covariance of productivity and capital within firms, denoted as $cov(\ln k_i, \ln z_i)$, we observe a positive correlation among firms in SEZs. This correlation arises because SEZs facilitate the reallocation of capital, enabling more productive firms to secure loans and accumulate capital. Consequently, this correlation increases by 88% compared to the counterfactual scenario without SEZs. Furthermore, by examining the bond-to-capital ratio, which measures the average bond per unit of capital, we find that in the economy with SEZs, this ratio is 35.56% higher than that in NSEZs. The presence of SEZs indicates a better credit market, which reduces the share of firms facing financial constraints by 95%. This quantifies the effectiveness of SEZ policies in terms of addressing resource misallocation among firms.

Table 8: Aggregate and Distributional Effects of SEZs

	Benchmark			No-SEZ Scenario	Effects of SEZs
	NSEZs	SEZs	Overall	Overall	(%)
Aggregate TFP (Z)	.3563	.5271	.5305	.4221	25.70
TFP Distribution:					
Firm-Level TFP (Avg.)	.5262	.6252	.5284	.4221	25.10
Birth Rate	.9736	.8577	.9717	1.00	-2.83
Firm-Level TFP at Birth (Avg.)	.0712	.6397	.4849	.4314	12.40
Death Rate	.0147	.2827	.2974	.2028	46.64
Financial Constraint:					
$\rho_i(\ln z_{it}, \ln k_{it})$ (Avg.)	-.0214	.0346	.0281	-.0249	88.00
Bond-capital ratio (b_i/k_i) (Avg.)	1.0455	.7343	.7456	.5500	35.56
Financial const. firm (%)	.0019	.9997	.0366	.8210	-95.55
Corporate Taxation:					
Effective τ	.0049	.0562	0.18	.0001	1800.31

Notes: The No-SEZ scenario is computed by setting the SEZ values for $(\tau, \theta, \bar{X})$ to their NSEZ counterpart.

6.2 Financial Frictions

This sub-section presents the results of a counterfactual analysis in which I explore the effects of eliminating the credit subsidy channel for SEZ firms by imposing tighter financial constraints in SEZs, making them equivalent to those in NSEZ. This analysis examines the aggregate consequences of enhanced access to credit markets on the allocation of factor inputs, and, subsequently, its impact on aggregate productivity and output. The results of this quantitative experiment are presented in Table 9.

In the benchmark economy, we observe a 7.29% improvement in aggregate TFP. This improvement is primarily attributed to the reduced resource misallocation across firms. In the model, collateral constraints prevent small but productive firms from investing at their desired levels. The resulting lack of capital in these firms reduces aggregate TFP. Consequently, the presence of SEZs, which offer credit subsidies to these firms, addresses capital misallocation, leading to an increase in aggregate productivity. This effect is evidenced by a 92.9% reduction in the share of financially constrained firms, accompanied by a 79.36% increase in the correlation between capital and productivity, thus raising aggregate TFP. Moreover, the improved access to credit for productive firms also impacts the selection channel. As more productive newly-born firms enter SEZs, there is a 12.4 percent increase in their birth productivity, along with a 9 percent reduction in the death rate. These factors contribute to a 15.14 percent increase in average TFP in the benchmark.

Table 9: Change SEZ (θ) to Non-SEZ

	Benchmark			Counterfactual	Difference
	NSEZs	SEZs	Overall	Overall	%
Aggregate TFP (Z)	.3563	.5271	.5305	.4945	7.29
TFP Distribution:					
Firm-Level TFP (Avg.)	.5262	.6252	.5284	.4589	15.14
Birth Rate	.9736	.8577	.9717	.7195	35.06
Firm-Level TFP at Birth (Avg.)	.0712	.6397	.4849	.4314	12.40
Death Rate	.0147	.2827	.2974	.3271	-9.09
Financial Constraint:					
$\rho(\ln z_i, \ln k_i)$ (Avg.)	-.0214	.0346	.0281	-.0223	79.36
Bond-capital ratio (b_i/k_i) (Avg.)	1.0455	.7343	.7456	.6616	12.69
Financial const. firm (%)	.0019	.9997	.0366	.5156	-92.91
Corporate Taxation:					
Effective τ	.0049	.0562	.18	.1500	18.75

6.3 Corporate Taxes

In this sub-section, I perform a counterfactual analysis by eliminating the credit subsidy channel for SEZ firms and, simultaneously, increasing tax frictions. This involves raising the corporate tax rate for SEZ firms to be equivalent to that in NSEZ. The purpose of this analysis is to explore the aggregate effects resulting from the interplay of improved access to credit markets and reduced tax frictions on the allocation of factor inputs and their influence on aggregate productivity and output. The quantitative results of this experiment are presented in Table 10.

The combined effects of the credit subsidy and reduced corporate tax policies lead to a substantial increase in aggregate TFP by 65.15 percent in the benchmark economy. This increase is mainly attributed to a significant surge in the number of firms entering the market, particularly firms originating in SEZs. The lower tax rate lightens the fiscal burden on businesses and enables less efficient firms to take advantage of the reduced tax rates, access more loans and continue their operations. This is evident in the 9 percent reduction in the death rate in the benchmark. However, this scenario results in a less efficient selection process, leading to resource inefficiency and wastage. As a consequence, the correlation between productivity and capital diminishes compared to the scenario with only reduced financial frictions. The increased correlation between productivity and capital falls from 79 percent (as observed in Table 9) to 72 percent (as reported in Table 10). This decline in correlation results from the survival of low-productivity firms due to the favorable tax environment, leading to a 37 percent decrease in average productivity. Additionally, the tax's impact on firms is reduced by 41 percent in the benchmark.

Table 10: Change SEZ $(\theta + \tau)$ to Non-SEZ

	Benchmark			Counterfactual	Difference
	NSEZs	SEZs	Overall	Overall	%
Aggregate TFP (Z)	.3563	.5271	.5305	.3212	65.15
TFP Distribution:					
Firm-Level TFP (Avg.)	.5262	.6252	.5284	.8389	-37.01
Birth Rate	.9736	.8577	.9717	0.5717	69.66
Firm-Level TFP at Birth (Avg.)	.0712	.6397	.4849	0.0004	108698.04
Death Rate	.0147	.2827	.2974	0.0002	297
Financial Constraint:					
$\rho(\ln z_i, \ln k_i)$ (Avg.)	-.0214	.0346	.0281	-.0205	72.95
Bond-capital ratio (b_i/k_i) (Avg.)	1.0455	0.7343	.7456	.6616	24.06
Financial const. firm (%)	.0019	.9997	.0366	.5740	-93.63
Corporate Taxation:					
Effective τ	0.0049	.0562	0.18	.33	-41.01

6.4 Summarizing the effects of the SEZs

In this subsection, I summarize the effects of SEZs on aggregate productivity through the selection and resource allocation channels. Table 11 shows that SEZ promote aggregate TFP by 25.7%, and this rise is due to both better selection with an average firm TFP increased by 25.1%, and better resource allocation with an increase in correlation of capital and productivity by 88%. Isolating the role of financial frictions, I find that approximately half of the increase in the aggregate TFP is generated by the reduction of financial frictions in SEZs, which induces better selection and better resource allocation for the endogenous distribution of firms. Quantifying the roles played by each of these channels provides a clearer perspective on their respective contributions to promoting productivity growth. And this understanding also help me to further discussions about making the optimal share of SEZ firm decisions.

Table 11: Summarize effects of the SEZ

Aggregate TFP (Z)	benchmark	counterfactual	Difference (%)
Collateral constraint θ	100 (0.5305)	93.2058 (0.4945)	7.29%
+ Corporate income tax τ	100 (0.5305)	60.54 (0.3212)	65.15%
+ Minimal profit scale $\bar{X}$	100 (0.5305)	79.57 (0.4221)	25.7%

Average TFP (z_i)	benchmark	counterfactual	Difference (%)
Collateral constraint θ	100 (0.5284)	93.2058 (0.4589)	15.14%
+ Corporate income tax τ	100 (0.5284)	158.13 (0.8389)	-37.01%
+ Minimal profit scale $\bar{X}$	100 (0.5284)	79.57 (0.4221)	25.1%

$\rho(lnz_i, lnk_i)$	benchmark	counterfactual	Difference (%)
Collateral constraint θ	100 (0.0281)	-79.36 (-.0223)	79.36%
+ Corporate income tax τ	100 (0.0281)	-72.95 (-.0205)	72.95%
+ Minimal profit scale $\bar{X}$	100 (0.0281)	-88 (-.0249)	88%

7 Further Discussion

Reduced financial frictions within SEZs induces better selection and more efficient resource allocation. As productivity and financial development increase within SEZs, high-productivity firms from Non-SEZ areas relocate into SEZs, reducing the marginal asset threshold levels due to higher returns on productivity. Reduced financial frictions in SEZs enable high-productivity firms to expand their businesses, increasing their capital investment. Both factors attract more productive firms from Non-SEZ areas into SEZs, intensifying competition within SEZs and displacing less productive firms with higher assets from SEZs, turning them into Non-SEZ firms.

However, reduced tax frictions in SEZs attract more firms to enter and obtain loans. And this low tax rates allow inefficient and wealthier firms (due to minimal profit scale requirements) to easily secure loans, potentially leading to an imbalance in resource allocation, lower correlation between productivity and capital. This decline in correlation due less efficient firms survive with lower tax rates and obtain more loans without effectively utilizing these resources. This offsets the positive selection driven by reduced financial frictions, evidenced by a 37% decrease in average productivity. The opposing forces between reduced tax friction and reduced financial frictions create an ambiguous situation where the presence of higher average productivity (better selection) and improved resource allocation depends on the interplay between them. The trade-off between lower tax and better financial access highlights the need for further investigation into optimal taxation that lead to the optimal size of firms in SEZs. In this context, it is expected that the dominance of positive selection over negative selection will lead to optimal aggregate productivity growth and efficient resource allocation.

8 Conclusion

This paper construct an innovative geo-coded firm-level panel data to investigate the effects of China's Special Economic Zones (SEZs) on economic performance, specifically focusing on the mechanisms through which SEZs impact outcomes. The empirical analysis demonstrates that SEZs exhibit better performance in terms of output, productivity, capital, and resource allocation, with a notable contribution from new-born firms through selection channel. The study further incorporates a quantitative approach through a dynamic firm model incorporates endogenous entry, exit, and location choices. This model successfully replicates the empirical findings and quantifies an increase in aggregate Total Factor Productivity (TFP) by 25.7%. The increase is driven by improved selection, leading to a 25.1% rise in average firm TFP, and improved resource allocation, evidenced by a 35.6% rise in the bond-capital ratio and an 88% increase in the productivity-capital correlation, a significant 95% decrease in the proportion of firms facing financial constraints and a 12.8% increase in capital accumulation within SEZs. Remarkably, about half of the TFP growth is due to improved selection, resource allocation, and investment, are significantly influenced by the reduction of financial frictions.

This research contributes to the economic literature by elucidating the dynamics behind the success of SEZs, particularly emphasizing the crucial role of selection and financial policies in driving economic growth within these zones. The findings highlight the importance of firm selection processes and resource allocation efficiencies in shaping the economic of SEZs, offering valuable insights for policy-makers in making decisions to promote the productivity and economic growth.

References

Adamopoulos, T., Brandt, L., Leight, J., and Restuccia, D. (2017). Misallocation, Selection and Productivity: A Quantitative Analysis with Panel Data from China. NBER Working Papers 23039, National Bureau of Economic Research, Inc.

Alder, S., Shao, L., and Zilibotti, F. (2016). Economic reforms and industrial policy in a panel of Chinese cities. *Journal of Economic Growth*, 21(4):305–349.

Arimoto, Y., Nakajima, K., and Okazaki, T. (2014). Sources of productivity improvement in industrial clusters: The case of the prewar japanese silk-reeling industry. *Regional Science and Urban Economics*, 46:27–41.

Arrow, K. (1962). Economic welfare and the allocation of resources for invention. In *The Rate and Direction of Inventive Activity: Economic and Social Factors*, pages 609–626. National Bureau of Economic Research, Inc.

Behrens, K., Duranton, G., and Robert-Nicoud, F. (2014). Productive cities: Sorting, selection, and agglomeration. *Journal of Political Economy*, 122(3):507–553.

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

Brandt, L., Van Biesebroeck, J., and Zhang, Y. (2014). Challenges of working with the chinese nbs firm-level data. *China Economic Review*, 30:339–352.

Brooks, W. J., Kaboski, J. P., and Li, Y. A. (2021). Agglomeration, misallocation, and (the lack of) competition. *American Economic Journal: Macroeconomics*, 13(4):483–519.

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

Buera, F. J. and Shin, Y. (2013). Financial frictions and the persistence of history: A quantitative exploration. *Journal of Political Economy*, 121(2):221–272.

Clementi, G. L. and Hopenhayn, H. (2006). A theory of financing constraints and firm dynamics. *The Quarterly Journal of Economics*, 121:229–265.

Combes, P.-P., Duranton, G., Gobillon, L., Puga, D., and Roux, S. (2012). The productivity advantages of large cities: Distinguishing agglomeration from firm selection. *Econometrica*, 80(6):2543–2594.

Ellison, G. and Glaeser, E. (1997). Geographic concentration in u.s. manufacturing industries: A dartboard approach. *Journal of Political Economy*, 105(5):889–927.

Farole, T. and Akinci, G. (2011). The World Bank Group.

Glaeser, E., Kallal, H. D., Scheinkman, J., and Shleifer, A. (1992). Growth in cities. *Journal of Political Economy*, 100(6):1126–52.

Gottlieb, C. and Grobovsek, J. (2016). Communal Land and Agricultural Productivity. Edinburgh School of Economics Discussion Paper Series 270, Edinburgh School of Economics, University of Edinburgh.

Hopenhayn, H. and Rogerson, R. (1993). Job turnover and policy evaluation: A general equilibrium analysis. *Journal of Political Economy*, 101(5):915–938.

Hopenhayn, H. A. (1992). Entry, exit, and firm dynamics in long run equilibrium. *Econometrica*, 60(5):1127–1150.

Hsieh, C.-T. and Klenow, P. J. (2009). Misallocation and Manufacturing TFP in China and India. *The Quarterly Journal of Economics*, 124(4):1403–1448.

Jacobs, J. (1969). *The Economy of Cities*. Random House.

Jacobson, L. S., LaLonde, R. J., and Sullivan, D. G. (1993). Earnings losses of displaced workers. *The American Economic Review*, 83(4):685–709.

Khan, A. and Thomas, J. (2011). Credit shocks and aggregate fluctuations in an economy with production heterogeneity. *Journal of Political Economy*, 121.

König, M., Storesletten, K., Song, Z., and Zilibotti, F. (2022). From imitation to innovation: Where is all that chinese r&d going? *Econometrica*, 90(4):1615–1654.

Lagakos, D. and Waugh, M. E. (2013). Selection, agriculture, and cross-country productivity differences. *American Economic Review*, 103(2):948–80.

Levinsohn, J. and Petrin, A. (2003). Estimating Production Functions Using Inputs to Control for Unobservables. *The Review of Economic Studies*, 70(2):317–341.

Lu, Y., Wang, J., and Zhu, L. (2019). Place-based policies, creation, and agglomeration economies: Evidence from china's economic zone program. *American Economic Journal: Economic Policy*, 11(3):325–60.

Marshall, A. (1890). *The Principles of Economics*. McMaster University Archive for the History of Economic Thought.

Midrigan, V. and Xu, D. Y. (2014). Finance and misallocation: Evidence from plant-level data. *American Economic Review*, 104(2):422–58.

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

Restuccia, D. and Rogerson, R. (2008). Policy distortions and aggregate productivity with heterogeneous plants. *Review of Economic Dynamics*, 11(4):707–720.

Romer, P. (1986). Increasing returns and long-run growth. *Journal of Political Economy*, 94(5):1002–37.

Schminke, A. and Biesebroeck, J. (2011). Using export market performance to evaluate regional preferential policies in china. *Review of World Economics*, 149.

Song, Z., Storesletten, K., and Zilibotti, F. (2011). Growing like china. *American Economic Review*, 101(1):196–233.

Wang, J. (2013). The economic impact of special economic zones: Evidence from chinese municipalities. *Journal of Development Economics*, 101:133–147.

Wong, K.-Y. (1987). China's special economic zone experiment: An appraisal. *Geografiska Annaler. Series B, Human Geography*, 69(1):27–40.

Yeung, Y., Lee, J., and Kee, G. (2009). China's special economic zones. *Eurasian Geography and Economics - EURASIAN GEOGR ECON*, 50:222–240.

Zeng, D. (2010). *Building Engines for Growth and Competitiveness in China: Experience with Special Economic Zones and Industrial Clusters*. Directions in Development. World Bank.

A Online Appendix

A.1 Measurement of Variables

The analysis and discussion in this Section focuses on the estimation of the firm's total factor productivity and the measurement of agglomeration. The Data Appendix A.1.1 provides more details on the variables used to calculate value-added to estimate firm-level TFP. A preliminary results about the estimated TFP distribution between zones and by agglomeration level are provided in Appendix 3.4. A more detailed information regarding these measurements of TFP and agglomeration are found in the respective Appendix A.1.2 and A.1.4.

A.1.1 Value Added Measurement

Main variables used in this study are the following:
-1. Value added is used to estimate the TFP. It is not directly reported by firms and is calculated by NBS by using the expenditure approach, which defines the value added as the sum of the output net of goods purchased for resale and indirect taxes subtracts material inputs:

$$VA = Output - \text{intermediate Input} + \text{payable value added tax}$$

due to missing values in 2004 and 2008, despite this shortcoming, we construct an alternative value added measure by using income approach at firm level. Then the value added will be calculated by adding four components, labor compensation, profit, net indirect taxes (indirect taxes minus government subsidies) and depreciation. Among them labor compensation consists of wage, unemployment insurance, welfare expenditures, pension contributions (after 2003) and housing subsidy (after 2004). Indirect taxes consist of three accounting items in our data, which are sales tax, value added tax and "other taxes under management expenses".

$$VA = \text{labor compensation} + \text{profit} + \text{indirect taxes} + \text{depreciation}$$

where labor compensation = salary+unemployment insurance+welfare expenditures+pension contributions+ housing subsidy, Indirect taxes = sales tax+value added tax+other taxes under management expenses.

Investment estimated by using capital stock

A.1.2 TFP Measurement

Regarding firm productivity, a popular approach to its measurement is to use TFP. In the literature, there are always discussions about the measurement of TFP, and there are several ways to estimate it, like OLS, FE, Olley and Pakes (1996), Levinsohn and Petrin (2003), GMM. Generally considered, using OLS and FE to estimate production functions raises the selection bias and endogenous problem between input choices and productivity shocks. This paper followed Olley and Pakes (1996). This method considers the influence of TFP on firm investment decisions as well as the influence of firms' investment decisions and the TFP on their survival probability. Thus, this method resolves the two-way causality and sample selection problems that parametric and non-parametric methods are faced with and invert the investment equation nonparametrically to proxy for unobserved productivity to avoid endogeneity issue. Relevant to our estimation of the TFP, two specific points need to be clarified. First, the output we employed in the estimation of TFP is value-added and calculated by using the input-output method. Our estimation process improves the TFP estimation used by Brandt et al. (2012). For instance, we used officially reported price deflators, while Brandt et al. (2012) constructed deflators by using the nominal and real output reported by the firms. For the price deflators of inputs, we used input-output tables from 1997, 2002, and 2007, while Brandt et al. (2012) only used a table representing one year, and thus ignored any changes that occurred over time. We also carefully constructed firm-level capital stock (but we choose not to report the lengthy procedure here in order to save space– an appendix section is available at the end this study). Second, we estimated the output elasticity of capital, labor, and intermediate inputs for each 2-digit industry separately, thus allowing for variation of output elasticity of inputs among industries. Importantly, this method did not affect our empirical results because all of the regressions provided below control for industry-fixed effects. To reduce the influence from outliers, trimming 0.5% of observations at both extremes is applied to the firm-level TFP in both SEZ and Non-sez.

I am going to describe some of the variables that are essential for the TFP's calculation, but they are not accurately reported in the data, so in what follows I will detail the variables' processing. One of the primary variable for estimating the TFP is capital stock. The procedure to calculate this capital stock used the reported total fixed assets value in the data. According to accounting principles, this indicator refers to the tangible assets held by an enterprise to produce goods, providing labour services, renting or operating management with the use of life for more than one fiscal year. Moreover, it is "the original fixed assets" [15] after deduction and impairment provision with construction materials and ongoing construction materials; thus, it can provide us with a relatively accurate description of a firm's capital status. Further, the weakness of this

[15] The sum of all past investments at purchase price

data is the missing fixed investment, using the capital stock as an estimate of fixed investment, applying the macro accounting method as follows, $I_t = K_t - K_{t-1} + D_t$, where K_t is capital stock at time t, D_t is depreciation of fixed assets.

The industrial value-added variable is utilized as the dependent variable output in the process of estimating TFP; since the reported value-added indicator in the data contains many missing values, I construct a value-added estimate through the income approach as in (Brandt et al., 2014) by adding four components: labour income, profit, net indirect taxes and depreciation [16].

Furthermore, to make nominal variables comparable over time, all those variables in nominal term are obtained by deflating with their respective price deflators with the base year of 1998 to express their values in constant price. The real value-added variable obtained from nominal value-added after deflating by using output deflators by 4-digit industrial classification across years; a set of input deflators varies across 4-digit industries by year are calculated from the National Input-Output(IO) table is applying to convert raw materials and intermediate inputs to the real term. Table A-1 shows the descriptive statistics of main variables to estimate TFP.

Table A-1: Descriptive Statistics of main variables to estimate TFP.

	mean	sd	p25	p50	p75	max
Output	10.53	1.29	9.56	10.36	11.29	19.82
Value-Added	8.56	1.45	7.57	8.39	9.39	18.44
Input	9.87	1.25	8.97	9.69	10.57	19.01
Capital	8.85	1.77	7.70	8.75	9.93	20.38
Investment	7.26	2.12	5.93	7.27	8.62	19.81
Employment	5.05	1.10	4.28	5.01	5.73	12.20
age	14.98	11.55	7.00	12.00	19.00	64.00
Observations	2,257,829					

Note: All variable values are expressed in logarithmic.

The following table A-2 and table A-3 comparing the elasticity of labor and capital between four different ways to estimate TFP. We can see that in case of OLS and FE labor is more intensive in the production technology, and in SEZs the elasticity of capital is higher than it in Non-SEZs. However, these results are not robust after controlling the unobserved shocks to productivity, as the results showed in OP method and LP method, in Non-SEZs capital is more intensive than in SEZs.

[16] Labor compensation consists of wage, unemployment insurance, welfare expenditure, pension contributions and housing subsidy. Indirect tax includes sales tax, value-added tax and other taxes under management expenses. Net indirect taxes is the indirect tax after subtracting subsidy

Table A-2: Descriptive Statistics TFP.

	OLS			FE		
	Overall	SEZ	nSEZ	Overall	SEZ	nSEZ
lnK	0.377***	0.416***	0.364***	0.332***	0.357***	0.322***
	(808.47)	(408.25)	(691.91)	(405.74)	(184.53)	(345.45)
lnL	0.531***	0.522***	0.539***	0.445***	0.414***	0.451***
	(745.46)	(323.40)	(678.12)	(431.34)	(176.50)	(383.02)
Constant	0.968	2.055	2.434	4.410***	3.329***	4.089***
	(0.00)	(0.00)	(0.01)	(8.99)	(14.36)	(7.48)
Observations	2356689	435161	1921528	2356689	435161	1921528
R2	0.683	0.699	0.673	0.393	0.414	0.365
R2-adjusted	0.683	0.699	0.673	0.224	0.210	0.177

t statistics in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A-3: Descriptive Statistics.

	OP			LP		
	Overall	SEZ	nSEZ	Overall	SEZ	nSEZ
lnK	0.515***	0.408***	0.523***	0.339***	0.303***	0.343***
	(88.29)	(17.25)	(61.99)	(141.77)	(56.49)	(156.14)
lnL	0.582***	0.598***	0.585***	0.356***	0.380***	0.355***
	(190.11)	(95.87)	(117.65)	(444.81)	(178.21)	(317.95)
age	-0.00487***	-0.00533***	-0.00466***			
	(-13.15)	(-4.00)	(-12.00)			
Observations	255814	26963	228851	1645061	268771	1376290
R2						
R2-adjusted						

t statistics in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

A.1.3 Firm-level TFP Across Zones

In the table A-4, shows the statistic description of logarithm of TFP. It indicates that the productivity distribution of firms in SEZs is right-skewed, whereas the distribution of firms' productivity in non-SEZs is left-skewed, meaning that there are fewer inefficient firms in SEZs.This could be due to a selection effect that leads to a left truncation of firm productivity distributions in those SEZs, since competition in SEZs is tougher and fewer of the weaker firms survive there.

Table A-4: The descriptive statistics of log(TFP) between zones

Non-SEZ :								
	skewness	mean	sd	p25	p50	p75	min	max
TFP	-0.27	0.85	0.96	0.28	0.88	1.47	-2.72	3.63
Observations	1,902,313							
SEZ :								
TFP	0.13	2.11	0.85	1.58	2.07	2.62	-0.83	4.91
Observations	430,811							

A.1.4 Industrial Agglomeration Measurement

As meaure of the degree of agglomeration I use "EG" index proposed in (Ellison and Glaeser, 1997). This index incorporates both the Gini index and the Hirschman-Herfindahl index, taking into account the industrial size of firms as well as the size of their regional distribution. The strength of the EG index lies in its ability to capture the nuances of both urban and industrial diversity. The calculation of the index is defined as follows

$$\hat{\gamma}_i^{EG} = \frac{\sum_{i=1}^{J}\left(s_{ij}^c - s_{*j}\right)^2 - \left(1-\sum_{j=1}^{J} s_{*j}^2\right)\sum_{k=1}^{K}(z_{k\in i})^2}{\left(1-\sum_{j=1}^{J} s_{*j}^2\right)\left(1-\sum_{k=1}^{K}(z_{k\in i})^2\right)} = \frac{G_j - \left(1-\sum_{j=1}^{J} s_{*j}^2\right)H_j}{\left(1-\sum_{j=1}^{J} s_{*j}^2\right)(1-H_j)}$$

where $s_{ij}^c = \frac{x_{ij}}{\sum_{j=1}^{J} x_{ij}} = \frac{x_{ij}}{x_{i*}}$ represents the concentration of industry i in region j, relative to all regions, $s_{*j} = \frac{\sum_{i=1}^{I} x_{ij}}{\sum_{i=1}^{I}\sum_{j=1}^{J} x_{ij}} = \frac{x_{*j}}{x_{**}}$ denotes the share of employment in region j, $H_j = \sum_{k=1}^{K}(z_{k\in i})^2$ is Hirschman-Herfindahl index, based on the number of plants, K which reflects the plant size distribution in industry i. The employment share of plant k in industry i is given by $z_{k\in i}$. The advantage of the EG Index is that it provides an unbiased estimate of agglomeration forces and, it is also straightforward to interpret and compute using only spatial-unit level information about industry plant distribution. Further, the index can be consistently applied across industries with different firm size distributions while controlling for general agglomeration trends.

A.2 Staggered DID Estimates: Within Firms

Direct comparisons of key variables such as output, productivity, and capital, along with their covariance for firms relocating to Special Economic Zones (SEZs) versus those remaining in Non-SEZ areas, may be confounded by simultaneous policy interventions. These external influences could distort the true impact of SEZ policies on firm performance. To mitigate this and to address the possible endogeneity in firms' relocation decisions to SEZs, the Fixed Effect Staggered Difference-in-Differences (FE-DID) methodology is utilized. This method strengthens the robustness of the analysis by controlling for unobserved heterogeneity, thereby enabling a more precise attribution of observed performance shifts to SEZ policies. The Difference-in-Differences estimation specifically calculates the average change in a firm's Total Factor Productivity (TFP) before and after relocating to an SEZ, making within-firm comparisons to more accurately delineate the policy's impact.

As outlined in Section 2.2, the timeline for each firm's move to an SEZ varies. A conventional DID approach is inadequate for capturing the staggered nature of these relocations. Consequently, the focus is on a staggered adoption framework for DID, setting up the analysis to account for when a firm moves into the SEZ, in contrast with those that have yet to move but will do so later, as well as those that remain outside the SEZ. This approach acknowledges the varied timing of policy exposure, enabling a temporal comparison among movers, and stayers, to isolate and examine the incremental impact of being located within an SEZ.

$$Y_{it} = \theta_i + \alpha_{rt} + \beta D_{it} + \delta X_{it} + \epsilon_{it} \tag{10}$$

where y_{it} is a dependent variable that we are interested in, including output $ln(y_{it})$, productivity $ln(z_{it})$, capital $ln(k_{it})$ and correlation between them $cov(ln(k_{it}), ln(z_{it}))$; θ_i is an individual firm fixed effect to control for time-invariant, unobserved firm characteristics that shape productivity distribution across firms. $\alpha_{rt} = \alpha_r \times \alpha_t$ is an interact fixed effect term between region r and time t to control for region-wide shocks and trends that shape the firm's TFP distribution over time, such as business cycles, regional changes in regulations and laws γ_i is an industrial fixed effect; The variable of interest is $D_{it} = treat_i \times post_{i,t}$, called a SEZ indicator that is equal to one in the years after firm i moved into the SEZ and zero otherwise. The coefficient, β, therefore represents the impact of the SEZ on the firms output, productivity, capital and their correlation; If β is positive and significant suggests that firms move into the SEZ are tend to be more productive, have higher output and capital and use capital more efficiently than those firm outside of the zone; X_{it} is the set of time-varying, firm-levels variables that capture the firms characteristics[17]; ϵ_{it} is the error term. The estimates indicating that firms move into the SEZs are more likely to be around 97% more productive comparing those stayers in Non-SEZs, which is larger than the effect of SEZ on productivity across firms 58% (in Table 3). Detailed estimates results can be found in the following.

Table A-5, presents the estimates results we can see that SEZs policy significantly promoting TFP across all different models, regardless of whether control variables are included into the baseline specification to capture firms' time-varying characteristics, or region-wide fixed effects are applied in the baseline estimation or separate time and regional fixed effects are applied.

Table A-5: The effects of the SEZs on TFP

	(1) m1	(2) m2	(3) m3	(4) m4	(5) m5	(6) m6	(7) m7	(8) m8	(9) m9	(10) m10	(11) m11	(12) m12
SEZ	0.909***	0.898***	0.973***	0.972***	0.978***	0.957***	0.922***	0.905***	0.974***	0.974***	0.966***	0.959***
	(199.94)	(146.89)	(157.50)	(157.33)	(116.02)	(205.65)	(205.54)	(149.22)	(159.00)	(158.85)	(116.45)	(207.01)
size		0.127***	0.109***	0.109***				0.117***	0.0995***	0.0992***		
		(53.62)	(39.96)	(39.62)				(48.92)	(35.93)	(35.69)		
lnage		-0.0569***	-0.0310***	-0.0307***	-0.0509***	-0.0277***		-0.0508***	-0.0209***	-0.0208***	-0.0486***	-0.0207***
		(-25.32)	(-12.02)	(-11.83)	(-12.26)	(-15.98)		(-23.07)	(-8.30)	(-8.19)	(-11.97)	(-12.19)
lnROA		0.0695***	0.0144***	0.0145***				0.0645***	0.0129***	0.0129***		
		(125.15)	(26.37)	(26.34)				(118.92)	(23.74)	(23.74)		
ln(Debt ratio)		-0.0454***	-0.0342***	-0.0341***	-0.0217***	-0.0215***		-0.0374***	-0.0331***	-0.0329***	-0.0213***	-0.0204***
		(-22.59)	(-14.66)	(-14.60)	(-10.61)	(-14.94)		(-19.12)	(-14.40)	(-14.30)	(-10.51)	(-14.41)
Export		-0.00881***						0.000799				
		(-3.07)						(0.27)				
State-owned		-0.0700***	-0.0331***	-0.0331***	-0.00470	-0.0281***		-0.0586***	-0.0193**	-0.0198**	-0.00206	-0.0153***
		(-9.69)	(-3.96)	(-3.95)	(-0.40)	(-4.93)		(-8.19)	(-2.29)	(-2.34)	(-0.17)	(-2.72)
lnky			-0.721***	-0.720***	-0.623***	-0.621***			-0.718***	-0.718***	-0.625***	-0.623***
			(-341.86)	(-340.89)	(-255.78)	(-379.93)			(-338.14)	(-337.20)	(-256.90)	(-379.20)
ln(Export density)			-0.00644***	-0.00658***	-0.00674***	-0.00803***			-0.00621***	-0.00628***	-0.00678***	-0.00766***
			(-4.90)	(-4.98)	(-5.27)	(-8.86)			(-4.78)	(-4.81)	(-5.40)	(-8.56)
lnEG				0.00819**	0.00406	0.00700***				0.00775**	0.00611	0.00687***
				(2.41)	(1.08)	(3.14)				(2.29)	(1.62)	(3.10)
lnsales					0.0936***						0.0855***	
					(31.02)						(27.83)	
lnprofit_net					0.0618***	0.0648***					0.0604***	0.0618***
					(61.35)	(95.48)					(60.71)	(92.47)
lnY						0.0437***						0.0353***
						(22.77)						(18.48)
_cons	0.912***	-0.103***	0.0305	0.0744**	-0.174***	0.275***	0.909***	-0.0114	0.111***	0.151***	-0.0640*	0.372***
	(1050.59)	(-4.26)	(1.05)	(2.14)	(-4.56)	(11.59)	(1062.27)	(-0.47)	(3.76)	(4.35)	(-1.65)	(15.75)
Observations	2319020	777655	205890	205053	251997	413465	2318971	777497	205374	204540	251730	412935
R-sq	0.766	0.820	0.949	0.949	0.944	0.945	0.777	0.830	0.952	0.952	0.947	0.948

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

A.3 Dynamic SEZ Impact: Event Study

In this section, I perform an event study akin to the approach of (Jacobson et al., 1993), aiming to study the dynamic effect of SEZ on firms performance. To facilitate this analysis, I add a series

[17]Control variables used in this study are: (1) The capital-output ratio;(2) $state_{it}$ is a dummy variable indicate whether a firm is state-owned;(3) $EX_{it} = 1$ if a enterprise engaged in export activities; (4) the age of the firm; (5)Operating characteristics of the firm, like the return on asset($ROA = \frac{\text{profit\_net}}{\text{average asset}}$), debt asset ratio($lev = \frac{debt}{\text{total asset}}$), and the enterprise scale ($size = ln(\text{total asset}/ipi)$).

of period when the SEZs firms are not yet moved into the SEZs and also the period after they have already been in the SEZs. There adding those series of dummy variables that characterize the year allows me evaluate the dynamic effets of SEZs based on the following Equation.

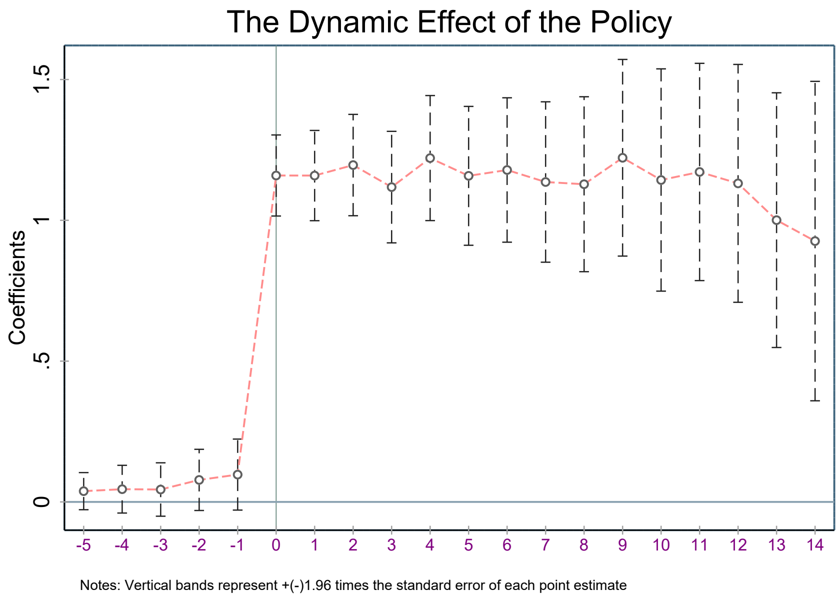
$$TFP_{it} = \theta_i + \alpha_{rt} + \beta_0 D_{it} + \sum_{s=1}^{S} D_{i,t+s}\beta_s + \sum_{m}^{M} D_{i,t-m}\beta_{-m} + \epsilon_{it} \tag{11}$$

Here $D_{i,t+s} = 1$, if a firm is in the SEZ after s years and β_s identifies the causal effect of the SEZ program s years following its occurrence; $D_{i,t-m} = 1$, if a treated firm is m years prior to the start of the policy and β_{-m} represents the impact of the establishment of SEZ on TFP comes up to m years before the program, moreover, if the parallel trend test holds, we expect that setting up an SEZ does not affect TFP before its occurrence time, in other words, those β_{-m} should be insignificant; Using clustered avoid the potential problems of serial correlation, ϵ_{it}. Cross fixed effect between region and time is included as control, since in this study the regions are different and they experience the different region-wide shocks, regional changes in regulations and laws, with a cross fixed effect to control for those region-wide shocks. The dummy for m = 5 is omitted so that the post-treatment effects are all relative to the five period before the start of the program. More detailed event study estimates can be found in Appendix A.3.

In the following visualize the dynamic effect, by showing the coefficient of the SEZs effect from Equation 11 along with the 95% confidence bands. In Figure A-1, we can observe that prior to the policy implementation, the estimated coefficients are close to zero and the confidence intervals cross the zero line, suggesting no significant effect on TFP. This implies that the common trend assumption holds, and there is no evidence of systematic differences in TFP trends between the treatment and control groups before the policy. After firms move into the SEZs, the coefficients show a substantial increase, indicating a positive effect of the SEZ policy on TFP. This effect appears to grow over the first few years following the policy's enactment, as seen by the upward trajectory of the coefficients. Over time, however, the effect sizes along with their confidence intervals suggest some variability, with some years showing a stronger policy effect than others.

Overall, the figure suggests that the SEZ policy has had a positive and dynamic effect on firms' productivity, with the magnitude of this effect changing over time. The initial positive impact post-policy implementation suggests that the SEZs may have provided an environment conducive to productivity enhancements among firms, although the long-term sustainability of this effect would require further analysis beyond what is shown in this figure. The event study depicted above demonstrates that the SEZ program influences not only the levels but also the trends of Total Factor Productivity (TFP). A detailed analysis of these effects on TFP trends is provided in Appendix A.4. The impact of SEZs on firm productivity is found to vary across

Figure A-1: Dynamic Effect of SEZ on Productivity



Notes: The figure plots the estimated impact of the Special Economic Zone (SEZ) policy on firms' Total Factor Productivity (TFP) over time. The horizontal axis represents the years relative to the implementation of the policy, with year 0 indicating the year the policy was enacted. The years preceding the policy are marked with negative numbers, and the years following the policy are marked with positive numbers. The vertical axis measures the coefficients, which represent the estimated effect of the SEZ policy on firms' TFP. Each point on the graph corresponds to a coefficient estimate for a particular year relative to the policy implementation. The vertical lines through each point represent confidence intervals, specifically 1.96 standard errors, which provide a 95% confidence range for the coefficient estimates.

different productivity levels of firms. To explore these varied impacts, Appendix A.5 study the heterogeneous effects of SEZs on firm productivity, providing a deeper understanding of how SEZs influence firms with different productivity profiles.

The following specification based on 11 adding time-varying firm-level control variables X_{it} to capture firm's characteristics.

$$TFP_{it} = \theta_i + \alpha_t + \gamma_r + \beta_0 D_{it} + \sum_{s=1}^{S} D_{i,t+s}\beta_s + \sum_{m}^{M} D_{i,t-m}\beta_{-m} + \delta X_{it} + \epsilon_{it} \tag{12}$$

$$TFP_{it} = \theta_i + \alpha_{rt} + \beta_0 D_{it} + \sum_{s=1}^{S} D_{i,t+s}\beta_s + \sum_{m}^{M} D_{i,t-m}\beta_{-m} + \delta X_{it} + \epsilon_{it} \tag{13}$$

Table A-6 below illustrates the estimates from Eq.(3)- Eq.(6). Column (1) reports the results of Eq.(3) condition on time and region fixed effect separately, while column (3) provide regression results controlling for region-wide shocks from Eq.(4). Column (2) provides the estimates obtained from Eq.(5) As shown in the table, the SEZ program increases total productivity.which increased by an average 96% in the year when the SEZ program is implemented relative to the five year before its start time. like double of it. Column (3) estimates the Eq.(2)

Table A-6: An event study: the effects of the SEZs on TFP

	(1) m1	(2) m2	(3) m3	(4) m4	(5) m5	(6) m6	(7) m7	(8) m8	(9) m9	(10) m10	(11) m11	(12) m12
pre4	0.0453***	0.0286***	-0.00606	-0.00706	-0.0423***	-0.0133**	0.0493***	0.0340***	-0.00798	-0.00924	-0.0410***	-0.00838
	(6.36)	(3.07)	(-0.70)	(-0.82)	(-4.35)	(-2.21)	(7.02)	(3.71)	(-0.94)	(-1.08)	(-4.28)	(-1.42)
pre3	0.0633***	0.0389***	0.00461	0.00349	0.0142	0.00990	0.0628***	0.0423***	0.00444	0.00315	-0.00678	0.0102
	(8.07)	(3.67)	(0.49)	(0.37)	(0.92)	(1.52)	(8.19)	(4.09)	(0.49)	(0.35)	(-0.45)	(1.60)
pre2	0.0882***	0.0424***	0.0120	0.0119	0.0290*	0.0167**	0.0901***	0.0453***	0.0105	0.0101	0.00329	0.0193***
	(10.63)	(3.77)	(1.23)	(1.20)	(1.67)	(2.39)	(11.10)	(4.13)	(1.12)	(1.07)	(0.20)	(2.82)
pre1	0.0805***	0.0298***	0.0180*	0.0173*	0.0195	0.0171**	0.0959***	0.0454***	0.0211**	0.0202**	-0.00575	0.0237***
	(9.65)	(2.62)	(1.78)	(1.71)	(1.13)	(2.35)	(11.76)	(4.10)	(2.17)	(2.08)	(-0.34)	(3.33)
current	0.963***	0.974***	1.016***	1.015***	1.008***	0.987***	0.975***	0.991***	1.017***	1.016***	0.979***	0.990***
	(113.93)	(82.41)	(85.40)	(85.20)	(58.59)	(121.61)	(117.96)	(85.57)	(88.72)	(88.46)	(58.75)	(124.61)
post1	0.991***	0.934***	0.988***	0.987***	1.020***	0.975***	1.011***	0.953***	0.994***	0.993***	0.992***	0.982***
	(115.64)	(78.19)	(84.78)	(84.50)	(58.97)	(118.17)	(120.26)	(81.35)	(88.87)	(88.51)	(59.13)	(121.36)
post2	0.991***	0.927***	0.985***	0.984***	1.028***	0.973***	1.011***	0.946***	0.992***	0.991***	0.997***	0.981***
	(113.41)	(75.98)	(83.62)	(83.36)	(59.25)	(116.83)	(117.57)	(78.65)	(87.02)	(86.70)	(59.02)	(119.24)
post3	0.956***	0.902***	0.975***	0.974***	1.023***	0.963***	1.000***	0.938***	0.987***	0.986***	0.993***	0.975***
	(106.63)	(72.46)	(80.74)	(80.52)	(58.32)	(111.38)	(112.64)	(76.10)	(83.88)	(83.60)	(58.07)	(113.55)
post4	0.958***	0.907***	0.975***	0.974***	1.000***	0.945***	1.008***	0.949***	0.991***	0.989***	0.972***	0.962***
	(100.95)	(69.42)	(77.15)	(76.93)	(55.75)	(105.39)	(106.86)	(73.07)	(80.33)	(80.05)	(55.47)	(107.70)
post5	0.972***	0.926***	0.971***	0.970***	1.012***	0.952***	1.032***	0.979***	0.990***	0.989***	0.983***	0.971***
	(98.89)	(67.91)	(73.34)	(73.13)	(55.22)	(101.67)	(105.36)	(71.76)	(76.15)	(75.89)	(54.71)	(103.74)
post6	0.977***	0.945***	0.989***	0.989***	1.030***	0.962***	1.054***	1.007***	1.013***	1.012***	1.001***	0.983***
	(95.15)	(66.63)	(70.31)	(70.11)	(54.53)	(96.62)	(102.40)	(70.68)	(73.48)	(73.24)	(53.83)	(98.95)
post7	0.984***	0.957***	1.000***	1.000***	1.069***	0.977***	1.074***	1.027***	1.029***	1.028***	1.038***	1.004***
	(91.74)	(64.86)	(70.64)	(70.47)	(55.33)	(96.31)	(99.46)	(68.84)	(72.33)	(72.12)	(54.31)	(97.61)
post8	0.979***	0.971***	1.013***	1.014***	1.089***	0.987***	1.082***	1.048***	1.051***	1.050***	1.061***	1.018***
	(86.67)	(62.18)	(68.27)	(68.14)	(55.07)	(93.29)	(94.93)	(66.14)	(70.33)	(70.18)	(54.06)	(94.65)
post9	0.981***	1.031***	1.033***	1.033***	1.131***	1.010***	1.139***	1.147***	1.082***	1.082***	1.106***	1.051***
	(82.45)	(60.74)	(63.82)	(63.75)	(55.89)	(90.87)	(93.73)	(65.95)	(66.21)	(66.10)	(54.71)	(92.83)
Observations	2319020	777655	205890	205053	251997	413465	2318971	777497	205374	204540	251730	412935
R2	0.767	0.821	0.949	0.949	0.945	0.945	0.778	0.830	0.952	0.952	0.947	0.948

t statistics in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The m1-m6 is the result from equation (5) and m7-m12 presents the results from equation (6)

A.4 SEZ effect on trends in TFP

The event study above ilustrated that the SEZ program has not only the impact on levels but also the trends in TFP. Therefore, in the following specification adding the post-SEZ trend to be $F_{it+s} = 1$ when a firm entered in SEZ after s years and zero otherwise.

$$TFP_{it} = \theta_j + \alpha_t + \gamma_r + \beta_0 D_{it} + \sum_s \beta_s F_{it+s} + \delta X_{it} + \epsilon_{it}$$

where $D_{it} = treat_i \times post_{i,t}$ is a treatment indicator that is equal to one if a firm i is entered in the SEZ experiment at time t and zero otherwise, the SEZ experiment on TFP is identified by β_0; $F_{it+s} = treat_i \times post_{i,t+s}$ is the post-SEZ trend, the effect on the trend of the TFP is identified by β_s, X_{it} are control variables as above.

The alternative specification including a cross fixed-effect

$$TFP_{it} = \theta_j + \alpha_{rt} + \beta_0 D_{it} + \sum_s \beta_s F_{it+s} + \delta X_{it} + \epsilon_{it}$$

The following table presents the results from estimations above,

A.5 Heterogeneous SEZ effects

The SEZs can affect the firm's productivity differently across different level of firm's productivity. Therefore, at each period before the policy implementation the sample will be divided into 5 quantiles and each quantile represents 20% of the firms at that pre period. Further, I run the baseline specification by using different sub-samples. The following table 8 shows the estimates across different samples in terms of heterogeneous productivity. It can be seen that there is a positive effect of policy on TFP in all specifications, but that the magnitudes are different. When compared to firms that are efficient, this policy has a greater impact on promoting productivity for low-productivity firms. In addition, the difference of the policy impact between the top 20% and bottom 20% also varies with time. At the time far from the policy implementation, this impact difference between productive and inefficient firms is getting smaller because the impact on productive firm is smaller and the inefficient firms are more effected.

Additionally, even when we include the time-varying control variables in the model, the facts that we have seen in table 8 seem to be fairly consistent. The inefficient firms are more likely to benefit from the SEZs policy for their productivity growth. The only difference exists between these results and those we observed before, which is that, the policy impact gap between high and low productive firms is more stable over time, and it varies less over the period.

Table 8: Heterogeneous effects of the SEZs on TFP

	(1) quantile20%	(2) 40%	(3) 60%	(4) 80%	(5) 100%
Pre1					
SEZ	1.415***	1.035***	0.905***	0.761***	0.487***
	(108.03)	(95.69)	(80.06)	(62.24)	(34.15)
constant	0.127***	0.809***	1.133***	1.468***	2.153***
	(12.03)	(92.62)	(125.78)	(152.22)	(193.97)
Observations	29178	28886	27235	25973	24360
R-sq	0.742	0.688	0.655	0.614	0.596
Pre2					
SEZ	1.275***	1.012***	0.910***	0.779***	0.583***
	(77.03)	(70.04)	(63.41)	(52.58)	(33.71)
constant	0.257***	0.795***	1.093***	1.430***	1.995***
	(23.13)	(83.01)	(116.61)	(151.27)	(185.39)
Observations	24079	24026	23496	22498	21489
R-sq	0.751	0.722	0.680	0.639	0.590
Pre3					
SEZ	1.176***	0.998***	0.918***	0.786***	0.692***
	(61.88)	(55.15)	(51.82)	(39.76)	(30.99)
constant	0.344***	0.806***	1.102***	1.420***	1.920***
	(31.22)	(78.18)	(111.24)	(130.46)	(158.05)
Observations	22272	21648	20876	19305	17981
R-sq	0.752	0.731	0.700	0.641	0.611
Pre4					
SEZ	1.068***	0.942***	0.929***	0.832***	0.817***
	(44.49)	(42.05)	(38.20)	(31.54)	(28.76)
constant	0.433***	0.875***	1.124***	1.426***	1.857***
	(37.84)	(82.23)	(98.92)	(116.60)	(138.61)
Observations	18391	18069	17085	16176	15333
R-sq	0.751	0.736	0.712	0.671	0.623
Pre5					
SEZ	0.893***	0.880***	0.839***	0.819***	0.779***
	(47.81)	(42.64)	(38.35)	(31.63)	(23.69)
constant	0.489***	0.858***	1.105***	1.361***	1.854***
	(70.64)	(119.07)	(153.38)	(169.72)	(198.73)
Observations	26298	19090	15910	13880	12184
R-sq	0.772	0.766	0.730	0.674	0.571

t statistics in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The m1-m6 is the result from equation (5) and m7-m12 presents the results from equation (6)

A.6 Agglomeration Effect: Sobel Test

In addition, a sobel test also permits us to determine whether the mediation agglomeration effect plays a role in how SEZs influence productivity. And the proportion of total effect that is mediated only accounts for 3.8%.

Table 10: Sobel Test: indirect effect

Sobel-Goodman Mediation Tests

	Coef	StdErr	Z	P>Z
Sobel	0.047	0.0003	148.4	0
Goodman-1	0.047	0.0003	148.4	0
Goodman-2	0.047	0.0003	148.4	0
	Coef	StdErr	Z	P>Z
η	0.002	0.000013	187.78	0
β_2	19.24	0.079	242.25	0
Indirect effect	0.047	0.000314	148.412	0
Direct effect	1.176	0.0016	749.32	0
Total effect	1.222	0.0016	775.159	0
Proportion of total effect that is mediated: 0.0381				
Ratio of indirect to direct effect: 0.0396				
Ratio of total to direct effect: 1.039				

In the appendix 11, we provide further study to check whether the average effect from table 15 at different periods before the policy implementation would be heterogeneous across sub-samples ranging from bottom 20% to top 20%. Moreover, a further study to assess the moderating effect of regional agglomeration on how SEZ policy impacts firm productivity is prodived in Appendix A.8 to better understand how agglomeration serves as a moderator, potentially intensifying the influence of SEZ policy on firms' productivity.

A.7 Heterogeneous Agglomeration Mediation Effect

For a robustness check here we provide the results from the specification with cross fixed-effect between region and year.

Based on the Sobel-Goodman Mediation tests, it shows that the indirect effect is 0.047 significant different from zero, which is computed by the product of the directly policy impact on agglomeration from Eq.(7) (0.002) and the partial policy impact on agglomeration from Eq(8) (19.24). Moreover the test illustrate that agglomeration effect can only explain 3.8% of the impact of SEZ policy on TFP and it is significant. Thus, from this test we can see that agglomeration effect therefore plays no role in explaining SEZs' positive effect on productivity growth.

Table 11: Mediation Effect through Agglomeration on TFP at Pre1

	(1) Bottom 20%	(2) 40%	(3) 60%	(4) 80%	(5) 100%
Model with TFP regressed on SEZ (path c)					
SEZ	1.415***	1.035***	0.905***	0.761***	0.487***
	(124.06)	(104.47)	(90.75)	(72.78)	(40.48)
constant	0.127***	0.809***	1.133***	1.468***	2.153***
	(12.94)	(94.84)	(133.39)	(166.38)	(214.02)
Observations	29178	28886	27235	25973	24360
R-sq	0.742	0.688	0.655	0.614	0.596
Model with mediator EG_irt regressed on SEZ (path a)					
SEZ	0.0000544	0.000354***	0.000173**	0.0000707	-0.000131
	(0.78)	(4.87)	(2.29)	(0.90)	(-1.55)
constant	0.00861***	0.00832***	0.00877***	0.00908***	0.0103***
	(142.59)	(132.70)	(136.66)	(137.43)	(145.07)
Observations	29408	28956	27321	26098	24643
R-sq	0.860	0.852	0.864	0.872	0.889
Model with TFP regressed on mediator EG_irt and SEZ (paths b and c')					
Agglomeration	5.161***	4.957***	4.867***	8.064***	7.098***
	(4.87)	(5.55)	(5.41)	(8.62)	(6.89)
SEZ	1.414***	1.034***	0.903***	0.758***	0.488***
	(123.86)	(104.27)	(90.53)	(72.58)	(40.59)
constant	0.0841***	0.768***	1.092***	1.396***	2.081***
	(6.25)	(67.88)	(94.13)	(114.12)	(142.60)
Observations	29074	28796	27161	25907	24292
R-sq	0.742	0.689	0.656	0.615	0.597

t statistics in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The figure shows how EG_{irt}, as a mediator of SEZs policy, affects on productivity at one period before the policy implementation across sub-samples ranging from bottom 20% to top 20%. The colunm (1) represents the estimates for sub-samples at bottom 20%, column (2) for sub-samples from 20% to 40% column (3) for sub-samples from 40% to 60%, cloumn (4) for sub-samples from 60% to 80% and column (5) for top 20% of the sample.

Table 12: Mediation Effect through Agglomeration on TFP at Pre2

	(1) Bottom 20%	(2) 40%	(3) 60%	(4) 80%	(5) 100%
Model with TFP regressed on SEZ (path c)					
SEZ	1.275***	1.012***	0.910***	0.779***	0.583***
	(97.25)	(93.37)	(83.38)	(67.85)	(43.73)
constant	0.257***	0.795***	1.093***	1.430***	1.995***
	(26.67)	(100.76)	(139.26)	(177.08)	(217.10)
Observations	24079	24026	23496	22498	21489
R-sq	0.751	0.722	0.680	0.639	0.590
Model with mediator EG_irt regressed on SEZ (path a)					
SEZ	0.000118	0.000157**	0.000326***	0.000157*	-0.000333***
	(1.45)	(2.02)	(4.02)	(1.87)	(-3.48)
constant	0.00847***	0.00818***	0.00827***	0.00867***	0.0102***
	(140.46)	(143.83)	(141.40)	(146.41)	(154.48)
Observations	24218	24039	23578	22551	21681
R-sq	0.837	0.840	0.841	0.859	0.875
Model with TFP regressed on mediator EG_irt and SEZ (paths b and c')					
Agglomeration	7.524***	8.277***	6.097***	4.587***	10.09***
	(6.64)	(8.38)	(6.30)	(4.53)	(9.57)
SEZ	1.276***	1.009***	0.908***	0.776***	0.586***
	(97.34)	(93.21)	(83.22)	(67.57)	(43.98)
constant	0.193***	0.728***	1.043***	1.392***	1.892***
	(14.18)	(64.52)	(93.01)	(116.62)	(133.70)
Observations	23962	23935	23444	22415	21421
R-sq	0.752	0.724	0.681	0.640	0.592

t statistics in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The figure shows how EG_{irt}, as a mediator of SEZs policy, affects on productivity at one period before the policy implementation across sub-samples ranging from bottom 20% to top 20%. The colunm (1) represents the estimates for sub-samples at bottom 20%, column (2) for sub-samples from 20% to 40% column (3) for sub-samples from 40% to 60%, cloumn (4) for sub-samples from 60% to 80% and column (5) for top 20% of the sample.

Table 13: Mediation Effect through Agglomeration on TFP at Pre3

	(1) Bottom 20%	(2) 40%	(3) 60%	(4) 80%	(5) 100%
Model with TFP regressed on SEZ (path c)					
SEZ	1.176***	0.998***	0.918***	0.786***	0.692***
	(79.26)	(78.89)	(71.07)	(55.33)	(41.48)
constant	0.344***	0.806***	1.102***	1.420***	1.920***
	(36.09)	(100.68)	(137.40)	(163.63)	(191.50)
Observations	22272	21648	20876	19305	17981
R-sq	0.752	0.731	0.700	0.641	0.611
Model with mediator EG_irt regressed on SEZ (path a)					
SEZ	0.000250***	0.000378***	0.000224**	0.000297***	-0.000152
	(2.75)	(4.08)	(2.28)	(2.97)	(-1.31)
constant	0.00803***	0.00811***	0.00823***	0.00846***	0.00952***
	(137.01)	(137.70)	(134.10)	(137.89)	(135.87)
Observations	22389	21703	20925	19367	18148
R-sq	0.823	0.827	0.825	0.858	0.865
Model with TFP regressed on mediator EG_irt and SEZ (paths b and c')					
Agglomeration	8.235***	7.413***	5.196***	7.166***	7.469***
	(6.93)	(7.33)	(5.24)	(6.38)	(6.33)
SEZ	1.173***	0.996***	0.917***	0.783***	0.693***
	(78.99)	(78.64)	(70.95)	(55.08)	(41.52)
constant	0.279***	0.746***	1.059***	1.360***	1.849***
	(20.69)	(65.15)	(92.59)	(105.72)	(122.95)
Observations	22174	21580	20812	19245	17925
R-sq	0.752	0.731	0.701	0.642	0.612

t statistics in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The figure shows how EG_{irt}, as a mediator of SEZs policy, affects on productivity at one period before the policy implementation across sub-samples ranging from bottom 20% to top 20%. The colunm (1) represents the estimates for sub-samples at bottom 20%, column (2) for sub-samples from 20% to 40% column (3) for sub-samples from 40% to 60%, cloumn (4) for sub-samples from 60% to 80% and column (5) for top 20% of the sample.

Table 14: Mediation Effect through Agglomeration on TFP at Pre4

	(1) Bottom 20%	(2) 40%	(3) 60%	(4) 80%	(5) 100%
Model with TFP regressed on SEZ (path c)					
SEZ	1.176***	0.998***	0.918***	0.786***	0.692***
	(79.26)	(78.89)	(71.07)	(55.33)	(41.48)
constant	0.344***	0.806***	1.102***	1.420***	1.920***
	(36.09)	(100.68)	(137.40)	(163.63)	(191.50)
Observations	22272	21648	20876	19305	17981
R-sq	0.752	0.731	0.700	0.641	0.611
Model with mediator EG_irt regressed on SEZ (path a)					
SEZ	0.000250***	0.000378***	0.000224**	0.000297***	-0.000152
	(2.75)	(4.08)	(2.28)	(2.97)	(-1.31)
constant	0.00803***	0.00811***	0.00823***	0.00846***	0.00952***
	(137.01)	(137.70)	(134.10)	(137.89)	(135.87)
Observations	22389	21703	20925	19367	18148
R-sq	0.823	0.827	0.825	0.858	0.865
Model with TFP regressed on mediator EG_irt and SEZ (paths b and c')					
Agglomeration	8.235***	7.413***	5.196***	7.166***	7.469***
	(6.93)	(7.33)	(5.24)	(6.38)	(6.33)
SEZ	1.173***	0.996***	0.917***	0.783***	0.693***
	(78.99)	(78.64)	(70.95)	(55.08)	(41.52)
constant	0.279***	0.746***	1.059***	1.360***	1.849***
	(20.69)	(65.15)	(92.59)	(105.72)	(122.95)
Observations	22174	21580	20812	19245	17925
R-sq	0.752	0.731	0.701	0.642	0.612

t statistics in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The figure shows how EG_{irt}, as a mediator of SEZs policy, affects on productivity at one period before the policy implementation across sub-samples ranging from bottom 20% to top 20%. The colunm (1) represents the estimates for sub-samples at bottom 20%, column (2) for sub-samples from 20% to 40% column (3) for sub-samples from 40% to 60%, cloumn (4) for sub-samples from 60% to 80% and column (5) for top 20% of the sample.

Table 15: Mediation Effect through Agglomeration on TFP at Pre5

	(1) Bottom 20%	(2) 40%	(3) 60%	(4) 80%	(5) 100%
Model with TFP regressed on SEZ (path c)					
SEZ	1.176***	0.998***	0.918***	0.786***	0.692***
	(79.26)	(78.89)	(71.07)	(55.33)	(41.48)
constant	0.344***	0.806***	1.102***	1.420***	1.920***
	(36.09)	(100.68)	(137.40)	(163.63)	(191.50)
Observations	22272	21648	20876	19305	17981
R-sq	0.752	0.731	0.700	0.641	0.611
Model with mediator EG_irt regressed on SEZ (path a)					
SEZ	0.000250***	0.000378***	0.000224**	0.000297***	-0.000152
	(2.75)	(4.08)	(2.28)	(2.97)	(-1.31)
constant	0.00803***	0.00811***	0.00823***	0.00846***	0.00952***
	(137.01)	(137.70)	(134.10)	(137.89)	(135.87)
Observations	22389	21703	20925	19367	18148
R-sq	0.823	0.827	0.825	0.858	0.865
Model with TFP regressed on mediator EG_irt and SEZ (paths b and c')					
Agglomeration	8.235***	7.413***	5.196***	7.166***	7.469***
	(6.93)	(7.33)	(5.24)	(6.38)	(6.33)
SEZ	1.173***	0.996***	0.917***	0.783***	0.693***
	(78.99)	(78.64)	(70.95)	(55.08)	(41.52)
constant	0.279***	0.746***	1.059***	1.360***	1.849***
	(20.69)	(65.15)	(92.59)	(105.72)	(122.95)
Observations	22174	21580	20812	19245	17925
R-sq	0.752	0.731	0.701	0.642	0.612

t statistics in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The figure shows how EG_{irt}, as a mediator of SEZs policy, affects on productivity at one period before the policy implementation across sub-samples ranging from bottom 20% to top 20%. The colunm (1) represents the estimates for sub-samples at bottom 20%, column (2) for sub-samples from 20% to 40%, column (3) for sub-samples from 40% to 60%, column (4) for sub-samples from 60% to 80% and column (5) for top 20% of the sample.

Table 16: Mediation Effect through Agglomeration on TFP at Pre5(CrossFE)

	(1) Bottom 20%	(2) 40%	(3) 60%	(4) 80%	(5) 100%
Model with TFP regressed on SEZ (path c)					
SEZ	0.867***	0.872***	0.807***	0.810***	0.772***
	(50.55)	(51.47)	(43.73)	(38.30)	(27.80)
constant	0.509***	0.871***	1.120***	1.364***	1.860***
	(68.34)	(124.04)	(154.82)	(174.03)	(199.31)
Observations	25309	18068	14970	13022	11355
R-sq	0.807	0.803	0.779	0.732	0.649
Model with mediator EG_irt regressed on SEZ (path a)					
SEZ	0.000370***	-0.0000468	0.000145	-0.0000371	-0.000884***
	(3.42)	(-0.36)	(1.02)	(-0.21)	(-4.19)
constant	0.00764***	0.00774***	0.00773***	0.00794***	0.00948***
	(161.18)	(140.90)	(138.19)	(123.07)	(131.43)
Observations	25328	18037	14991	13051	11425
R-sq	0.818	0.823	0.845	0.846	0.885
Model with TFP regressed on mediator EG_irt and SEZ (paths b and c')					
Agglomeration	5.930***	6.142***	2.678**	5.463***	4.911***
	(5.20)	(5.54)	(2.15)	(4.29)	(3.37)
SEZ	0.862***	0.868***	0.805***	0.811***	0.777***
	(50.19)	(51.23)	(43.54)	(38.31)	(27.95)
constant	0.469***	0.825***	1.102***	1.321***	1.813***
	(40.82)	(74.47)	(91.36)	(103.21)	(109.06)
Observations	25083	17931	14896	12964	11255
R-sq	0.807	0.803	0.780	0.732	0.651

t statistics in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table shows the sobel test results from the specification with controlling cross fixed-effect between region and time across sub-samples ranging from bottom 20% to top 20% at five-period before the policy implementation. The column (1) represents the estimates for sub-samples at bottom 20%, column (2) for sub-samples from 20% to 40%, column (3) for sub-samples from 40% to 60%, column (4) for sub-samples from 60% to 80% and column (5) for top 20% of the sample.

Table 17: Sobel Test: Indirect Effect through Agglomeration

Test	**Coefficient**	**Std Error**	**Z**	**P**$>|Z|$
Sobel	0.4660172	0.000314	148.4	0
Goodman-1 (Aroian)	0.4660172	0.000314	148.4	0
Goodman-2	0.4660172	0.000314	148.4	0

Effect	**Coefficient**	**Std Error**	**Z**	**P**$>|Z|$
a coefficient	0.002422	0.000013	187.776	0
b coefficient	19.2416	0.079428	242.253	0
Indirect effect	0.46602	0.000314	148.412	0
Direct effect	1.17573	0.01569	749.319	0
Total effect	1.22233	0.01577	775.159	0

Ratio Description	**Value**
Proportion of total effect that is mediated	0.3812523
Ratio of indirect to direct effect	0.3963638
Ratio of total to direct effect	1.0396364

Table 18: SobelTest: Heterogeneous indirect effect across sub-samples at Pre5

Sobel-Goodman Mediation Tests

	Botton 20%	40%	60%	80%	100%
Sobel	.041 ***	.0174 ***	.0111 ***	.0098 ***	.0066 ***
	(16.95)	(10.88)	(6.97)	(7.29)	(5.024)
Goodman-1	.041 ***	.0174 ***	.0111 ***	.0098 ***	.0066 ***
	(16.94)	(10.87)	(6.96)	(7.28)	(5.00)
Goodman-2	.041 ***	.0174 ***	.0111 ***	.0098 ***	.0066 ***
	(16.96)	(10.89)	(6.98)	(7.31)	(5.04)
η	.0029 ***	0.0022 ***	.0025 ***	.00196 ***	0.0018 ***
	(29.04)	(19.13)	(20.19)	(13.46)	(9.59)
β_2	14.11 ***	7.88 ***	4.421 ***	5.014 ***	3.57 ***
	(20.87)	(13.22)	(7.43)	(8.68)	(5.90)
Indirect effect	0.041 ***	.0174 ***	.0111 ***	.0098 ***	.0066 ***
	(16.95)	(10.88)	(6.97)	(7.29)	(5.024)
Direct effect	1.762 ***	1.456 ***	1.274 ***	1.068 ***	.638 ***
	(159.94)	(149.323)	(129.89)	(102.35)	(47.43)
Total effect	1.803 ***	1.473 ***	1.286 ***	1.078 ***	.645 ***
	(164.887)	(151.83)	(132.37)	(103.65)	(48.02)
Proportion of total effect that is mediated:	.0226	.0117	.0086	.0091	.0101
Ratio of indirect to direct effect:	.0231	.0119	.0087	.0092	.0103
Ratio of total to direct effect:	1.023	1.012	1.008	1.009	1.010

Note: The table shows the sobel test results from the specification without controlling fixed-effect of region and time across sub-samples ranging from bottom 20% to top 20% at five-period before the policy implementation. The column (1) represents the estimates for sub-samples at bottom 20%, column (2) for sub-samples from 20% to 40%, column (3) for sub-samples from 40% to 60%, column (4) for sub-samples from 60% to 80% and column (5) for top 20% of the sample.

The table 18 illustrates how indirect effects in terms of agglomeration differ across sub-samples. For more productive firms, the agglomeration effect contributes only 1% to the total impact of the policy on productivity, compared to 2% for inefficient firms.

For a robustness check we provide the sobel test results from specification controlling for FE of region and time. The following table 19 shows the results[18].

Table 19: sobeltest

Sobel-Goodman Mediation Tests

	Botton 20%	40%	60%	80%	100%
Sobel	.0036 ***	.0010	.0019 **	.0014	-.0030 **
	(4.10)	(1.19)	(2.46)	(1.46)	(-2.52)
Goodman-1	.0036 ***	.0010	.0019 **	.0014	-.0030 **
	(4.07)	(1.18)	2.42	(1.44)	(-2.47)
Goodman-2	.0036 ***	.0010	.0019 **	.0014	-.0030 **
	(4.13)	(1.20)	2.50	(1.49)	(-2.57)
η	.0005 ***	.0001	.0003 ***	.0002	-.0006 ***
	(5.18)	(1.21)	(3.03)	(1.53)	(-3.49)
β_2	7.45 ***	7.23 ***	4.93 ***	5.96 ***	5.04 ***
	(6.71)	(6.78)	(4.21)	(4.94)	(3.64)
Indirect effect	.004 ***	.0010	.0019 ***	.0014 ***	-.0030 ***
	(4.10)	(1.19)	(2.46)	(1.46)	(-2.52)
Direct effect	.888 ***	.876 ***	.835 ***	.8179 ***	.781 ***
	(57.99)	(57.82)	(49.95)	(43.19)	(32.94)
Total effect	.89 ***	.88 ***	.837 ***	.8192 ***	.778 ***
	(58.21)	(57.80)	(50.05)	(43.22)	(32.81)
Proportion of total effect that is mediated:	.00398	.0011	.00223	.0017	-.0039
Ratio of indirect to direct effect:	.0040	.0011	.00224	.0017	-.0039
Ratio of total to direct effect:	1.004	1.001	1.002	1.001	.996

Note: The table shows the sobel test results from the specification with controlling fixed-effect of region and time across sub-samples ranging from bottom 20% to top 20% at five-period before the policy implementation. The column (1) represents the estimates for sub-samples at bottom 20%, column (2) for sub-samples from 20% to 40%, column (3) for sub-samples from 40% to 60%, column (4) for sub-samples from 60% to 80% and column (5) for top 20% of the sample.

A.8 DDD estimation

In this section, the baseline model is expanded to a Triple Difference (DDD) estimation to assess the moderating effect of regional agglomeration on how SEZ policy impacts firm productivity. In this context, agglomeration serves as a moderator, potentially intensifying the influence of SEZ policy on a firm's Total Factor Productivity (TFP). The assumption is that a higher level of agglomeration within a city may amplify the SEZ's impact on firms' TFP. To clarify, while a mediator would help explain the process by which the SEZ policy and firm productivity are linked, a moderator like agglomeration level alters the intensity and possibly the direction of this relationship. This nuanced analysis will enable us to understand not just if the SEZ policy affects productivity, but also how the context of agglomeration influences this effect.

[18] notice that the coefficient of total effect and direct effect are not consistent with the estimates showed in table 14, thus the results showed here are not reliable. Need to find another better way to test the meditation effect in the case of FE.

The following Table 20 shows the results from the specification 14,

$$TFP_{it} = \theta_i + \alpha_r + \gamma_t + \beta_0 D_{it} + \beta_1 D_{it} \times EG^H_{irt} + \beta_3 treat_i \times EG^H_{irt} + \beta_4 post_{i,t} \times EG^H_{irt} + \epsilon_{it} \quad (14)$$

where EG^H_{irt} is an indicator variable, it is equal to 1 if agglomeration level in that city is above the median of the full sample, zero otherwise.

Table 20: How EG^H_{it} changes the heterogeneous effects of the SEZs on TFP

	(1) quantile20%	(2) 40%	(3) 60%	(4) 80%	(5) 100%
SEZ× High_EG	0.0557***	0.0644***	0.0655***	0.0192	0.00906
	(2.61)	(4.15)	(4.14)	(1.11)	(0.43)
SEZ	1.387***	1.002***	0.868***	0.748***	0.483***
	(84.79)	(74.91)	(61.65)	(47.22)	(25.68)
sez_ag	-0.0248	-0.0395**	-0.00881	0.00113	0.0429*
	(-1.07)	(-2.19)	(-0.49)	(0.06)	(1.81)
time_ag	0	0	0	0	0
	(.)	(.)	(.)	(.)	(.)
constant	0.138***	0.828***	1.137***	1.468***	2.127***
	(8.88)	(65.20)	(86.61)	(100.70)	(118.11)
Observations	29074	28796	27161	25907	24292
R-sq	0.742	0.688	0.656	0.614	0.596

t statistics in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The figure shows how the moderator variable EG^H_{it} changes the heterogeneous effect of SEZ on TFP. The columns from (1)-(5) represent the quantile of the sub-sample for those firms at 1 period before the policy implementation.

Then instead of using a dummy variable EG^H_{it} to indicate the cities' agglomeration level, I use $lnEG$ logarithm of EG index a continues variable to see how the different density of agglomeration varies the SEZ effect on firms' TFP. Specifically,

$$TFP_{it} = \theta_i + \alpha_r + \gamma_t + \beta_0 D_{it} + \beta_1 D_{it} \times lnEG_{irt} + \beta_3 treat_i \times lnEG_{irt} + \beta_4 post_{i,t} \times lnEG_{irt} + \epsilon_{it} \quad (15)$$

where $lnEG_{irt}$ is logarithm of EG index to measure the city agglomeration level.

The results from the table 21 show that the coefficient of the interact term between the SEZs policy and agglomeration is significant positive and if agglomeration level in the area increase by 1% the impact of SEZs on TFP will increase by 1.26%, that means a denser region will have a greater impact on TFP through the SEZ.

Table 21: How $lnEG_{irt}$ changes the effects of the SEZs on TFP

	(1) m1
SEZ	0.971***
	(35.55)
SEZ × lnEG	0.0126**
	(2.42)
$Treat_i$ × lnEG	0.0109*
	(1.86)
$Post_{i,t}$ × lnEG	0.0431***
	(16.83)
constant	1.155***
	(90.53)
Observations	2294206
R-sq	0.766

t statistics in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The figure shows how the moderator variable $lnEG_{irt}$ changes the effect of SEZ on TFP. The colunm m1 represent the estimates from Equation 15

A.9 Unconstrained firms

Labor and Capital Decision Rules Unconstrained firm, it never experiences binding borrowing constraints in any possible future state

- Optimal static labor choice. A firm with (k, z) chooses $\hat{L}(k, z) = \left[\frac{(1-\tau)*(zk^{\alpha\gamma}(1-\alpha)\gamma)}{w}\right]^{\frac{1}{1-(1-\alpha)\gamma}}$
- Current earnings with optimal labor hiring $\hat{l}$, then $\hat{\Pi} = (1-\tau)\left[z(k^{\alpha}\hat{L}^{(1-\alpha)})^{\gamma} - w\hat{L}\right]$
- Choice of future capital, k' by the unconstrained firms (collateral constraint is not binding), optimal level of $k' = \hat{K}(z)$, which is the solution of the following problem.

$$\max_{k'} \left[-(1-\tau)k' + \beta \sum_{j=1}^{N_z} \pi_{ij}^{z} \left(\hat{\Pi}(k', z_j) + (1-\delta)k'\right)\right]$$

Debt Decision Rules

- With policy functions $\hat{L}, \hat{K}$, the optimal debt policy $b' = \hat{B}(z)$ is defined by the following equations.
$$\begin{aligned}\tilde{B}(k, z_i) = (1-\tau)\Big[z_i \left(k^\alpha \hat{L}^{1-\alpha}\right)^\gamma - w\hat{L} + (1-\delta)k - \hat{K}(z_i)\Big] \\ + q \min\left\{\hat{B}(z_i), \theta\hat{K}(z_i)\right\}\end{aligned}$$

- Maximum level of debt of the unconstrained firm unaffected by the constraint over any future path of z.
$$\hat{B}(z_i) = \min\left(\tilde{B}\left(\hat{K}(z_i), z_j\right)\right)$$
where $\tilde{B}\left(\hat{K}(z_i), z_j\right)$ is the maximum level of debt that an unconstrained firm can hold in which $z' = z_j$ is realized.

Cash-on-hand and decision rules

- The incumbent firm's problem is a challenging object because of the occasionally binding constraints for b'
- Levels of k and b of firms do not separately determine the choices of k' and b'.
- Collapse two state variables into newly defined variable cash-on-hand, $m(k, b, z)$.
- $m(k, b, z)$ is defined as
$$m(k, b, z) \equiv (1-\tau)\left[z(k^\alpha \hat{L}^{(1-\alpha)})^\gamma - w\hat{L} + (1-\delta)k\right] - b$$
- $m' \equiv m(k', b', z')$
- Rewrite the incumbent firm's problem in SEZ.
$$\begin{aligned}V(m, z_i) = \max_{k', b', D, m'_j} &\left[D + \beta \max\left\{V_x(m), \sum_{j=1}^{N_z} \pi^z_{ij} V(m'_j, z_j)\right\}\right] \\ s.t.\ \ & 0 \le D \equiv m - k'(1-\tau) + qb' \\ & b' \le \theta k' \\ & m'_j \equiv m(k', b', z_j) \\ & = (1-\tau)\left[z_j(k'^\alpha \hat{L}^{(1-\alpha)}(k', z_j))^\gamma - w\hat{L}(k', z_j) + (1-\delta)k'\right] - b'\end{aligned}$$