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

Boyao Zhang
University of Konstanz

April, 2025

Abstract

This paper develops a novel geo-coded firm-level panel dataset to analyze how Special Economic Zones (SEZs) influence firm distribution and productivity. SEZs, offering reduced corporate taxes and enhanced access to credit upon maintaining minimum operational scale, are shown to foster better firm selection, higher capital investment, and improved resource allocation. Firms located within SEZs consistently outperform external firms regarding productivity and investment, regardless of whether they originate within SEZs or relocate there later. Firms initially founded within SEZs demonstrate permanently higher productivity, greater capital accumulation, and better alignment between Total Factor Productivity (TFP) and capital. Firms that relocate to SEZs also gain these productivity benefits after entry, unlike firms remaining outside. However, as SEZ placements aren't random, these advantages partly reflect firms' endogenous location choices. To quantify SEZs' broader aggregate effects and determine their optimal scale, the study builds a firm-dynamics model with endogenous entry, exit, and location decisions. Counterfactual simulations indicate that SEZs boost aggregate TFP by 25.7%, driven primarily by improved firm selection (average firm-level TFP increases by 25.1%) and better resource allocation (an 88% increase in capital–TFP correlation). Further analysis isolates financial frictions, revealing that reduced financial constraints within SEZs explain roughly half of these productivity gains by enabling more efficient firm selection and resource allocation.

Keywords: Firm dynamics, Economic Growth, Special Economic Zones, Selection, Resource Allocation, Investment, Productivity

JEL Classification: E22, E23, E44, E62, H32, I38, O47

*I am deeply grateful to my supervisor Raul Santaeulalia-Llopis for his invaluable guidance, insightful discussions, and generous support throughout my research. I express special appreciation to my committee members, Andre Groeger, Chaoran Chen, and Manuel García-Santana, for their constructive feedback and insightful suggestions. I also thank Luis Rojas, Albert Marcet, and Barthélémy Bonadio for their helpful comments. Furthermore, I am grateful to seminar and conference participants at New York University Abu Dhabi (NYUAD), the Macroeconomics Club at Universitat Autònoma de Barcelona (UAB), the Barcelona School of Economics (BSE) Jamboree, the ENTER Jamboree, the China Economics Summer Institute (CESI) at Beijing University, and the University of Konstanz for their valuable insights and discussions.

1. Introduction

Special Economic Zones (SEZs), geographically defined areas administered by a single authority, typically offer firms incentives such as reduced corporate taxes and enhanced access to credit, contingent upon maintaining a minimum operational scale. SEZs have been adopted in over 130 countries, spanning both advanced and developing economies, with the aim of promoting economic growth. This policy has gained significant global momentum, as the number of SEZs increased thirty-fold between 1975 and 2002. China, which hosts more than a quarter of the world's SEZs, has been a key contributor to this global trend.[1] Simultaneously, the allocation of capital across firms has proven to be a crucial factor influencing aggregate Total Factor Productivity (TFP) and output growth in China (Song et al., 2011).

In this study, I provide the first analysis of how SEZs influence firm distribution to assess their aggregate effects in China. First, I construct a novel panel dataset that enables the documentation of several stylized facts regarding how firm distribution differs across zones. Second, I develop a firm dynamics model with endogenous entry, exit, and location (zone) choices, which replicates the stylized facts and allows for the evaluation of the aggregate effects of SEZs. Two key mechanisms drive the results. The first is firm selection through birth (across zones) and the distinction between movers and stayers (across zones). The second is dynamic resource allocation, which integrates investment choices with changes in within-firm resource allocation over time, measured as the evolution of the within-firm covariance between TFP and capital.

First, the novel firm-level panel dataset I created tracks firm-level zone locations over 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. Specifically, I introduce a SEZ status variable (i.e., whether firms are in an SEZ or not) to the standard Annual Survey of Industrial Enterprises in China (ASIE), which has been used by Hsieh and Klenow (2009) and others in their analysis of factor misallocation in China. This new panel enables the first comprehensive evaluation of the implications of SEZs on firm behavior and distribution. My dataset tracks the SEZ status of 586,599 unique firms across 2,574 districts from 1998 to 2013.

Using this data, I document a new set of stylized facts that highlight the advantages associated with SEZs. These advantages include improved firm selection, increased capital investment, and enhanced resource allocation efficiency within firms over time. My findings show that firms in SEZs, whether established there or relocated, perform better across several dimensions. First, firms originating in SEZs tend to be more productive, invest more, and exhibit a stronger correlation between Total Factor Productivity (TFP) and capital over time compared to firms established outside SEZs. Second, firms relocating to SEZs from outside also acquire these advantages after their move, in contrast to firms that remain outside SEZs.

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

[2] API stands for "Application Programming Interface," a set of tools that enable 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 SEZs that are the focus of this paper.

However, it is important to note that even if SEZs were randomly allocated across China, which is demonstrably not the case,[3] the improved performance of firms in SEZs cannot be entirely attributed to SEZs alone. This is because a firm's decision to operate in an SEZ is an endogenous choice and thus not random. This is particularly relevant because I find that 55% of firms in SEZs are "born" there, while 45% relocate from areas outside SEZs, further complicating empirical analysis.

Secondly, to address the empirical limitations related to assessing SEZs (including the endogenous SEZ rollout policy and the endogenous selection into SEZ "treatment"), I propose a firm dynamics model that incorporates entry and exit, further endogenizing firms' location (zone) choices. This model enables an examination of the aggregate effects of SEZs. Using this framework, I aim to evaluate how specific characteristics of SEZs, such as reduced corporate taxes, relaxed financial constraints, and higher minimum profit scale requirements, impact aggregate Total Factor Productivity (TFP) and output. Ultimately, I employ the model to determine the optimal size of SEZs.

In particular, I assume that firms in both SEZs and non-SEZs operate under the same technology, but differ in terms of corporate tax rates, financial constraints, and minimum profit requirements. Specifically, in alignment with the features of SEZs, firms within SEZs benefit from lower corporate tax rates, face less stringent financial constraints, and are required to meet a minimum profit scale. Firms are heterogeneous in both productivity levels and cash-on-hand positions. Consequently, firms in my model are categorized into 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, as they seek to meet their operational requirements.

My model yields several key predictions, which qualitatively align with the empirical findings. The first prediction suggests that firms selected into SEZs outperform non-SEZ firms in terms of productivity, assets, and resource allocation. Given the reduced corporate taxes and improved access to credit markets offered by SEZs, every firm aspires to establish itself or relocate within SEZs to benefit from these favorable policies. However, due to the stringent minimum requirements in SEZs, only firms surpassing these thresholds can operate within the zones, while firms with fewer assets remain outside. In equilibrium, there are two thresholds, each defined by a combination of 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 between these two thresholds operate in non-SEZ areas.

Moreover, the preferential policies of SEZs attract more firms, either newly established or relocated from non-SEZ areas, intensifying competition within SEZs. Through a "survival-of-the-fittest" mechanism, more efficient firms thrive, while less productive ones are forced out of the market. This selection process leads to 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 to SEZs, lowering the marginal asset threshold due to higher returns on productivity. Reduced financial frictions in SEZs enable high-productivity firms to expand, increasing capital investment. These dynamics attract more productive firms from non-SEZ areas, intensifying competition and displacing less productive firms with higher assets from SEZs, effectively turning them into non-SEZ firms.

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

However, reduced tax frictions in SEZs also attract more firms to enter and secure loans. Low tax rates may allow inefficient but wealthier firms (due to minimum profit scale requirements) to easily obtain loans, potentially leading to misallocation of resources. Less efficient firms might exploit lower tax rates to acquire more loans without effectively utilizing them, which could counterbalance the positive selection effects of reduced financial frictions. These opposing forces create an ambiguous outcome, where the overall improvement in average productivity (through better selection) and enhanced resource allocation depends on the interaction of these factors. The trade-off between lower taxes and better financial access underscores the need for further investigation into optimal taxation. It is expected that the dominance of positive selection over negative effects will lead to optimal aggregate productivity growth and more efficient resource allocation.

This model was carefully calibrated to align with empirical findings. Through counterfactual experiments assuming a scenario without SEZs, the analysis reveals that SEZs increase aggregate Total Factor Productivity (TFP) by 25.7%. Examining the underlying mechanisms, I find that the rise in aggregate TFP results from two main channels: improved firm selection, contributing to a 25.1% average TFP increase among firms, and enhanced resource allocation, demonstrated by an 88% increase in the correlation between capital and TFP. A decomposition of the SEZ characteristics driving these effects reveals that the reduction of financial frictions within SEZs accounts for approximately 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 improved resource allocation for the endogenous distribution of firms.

It is important to note that ex-ante, it is not obvious from my model that SEZs would increase aggregate TFP and output, as the reduction of corporate taxes within SEZs encourages the entry of less productive firms into both the economy and SEZs. However, the results indicate that this negative effect is quantitatively outweighed 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 et al., 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

[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.

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).

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 geocoded 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.

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 this 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 consulting firms evaluate the 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 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

[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.

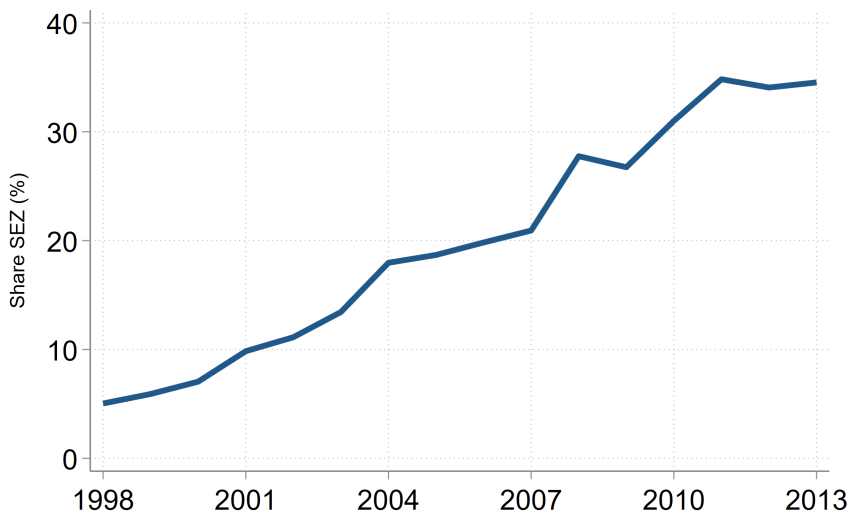
14 other coastal cities, each establishing its own economic and technological development zone. The establishment of these coastal SEZs coincided with a period of extraordinary growth in those regions, with Shenzhen's GDP surging at an annual rate of 58 percent in the early years, outpacing the national 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).

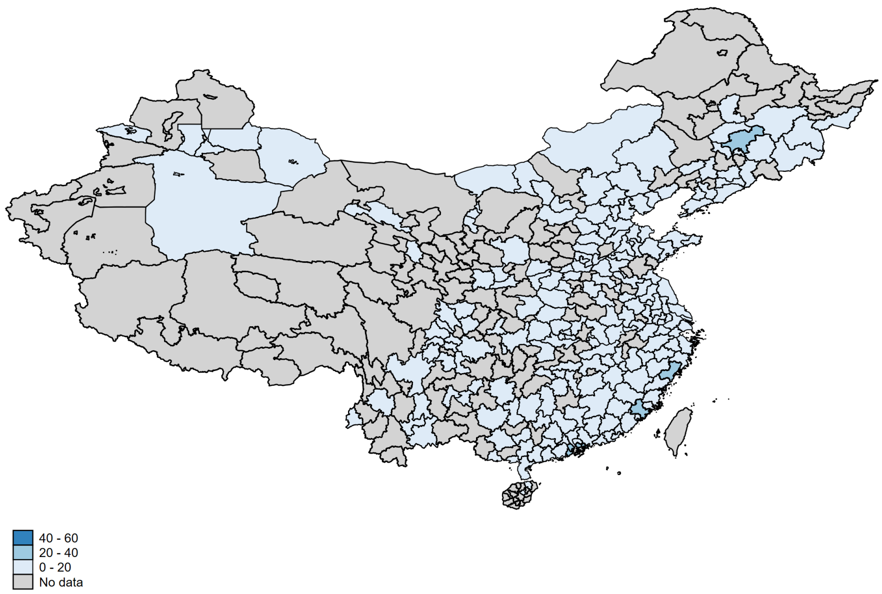
Figure 1: Share of SEZs Firm Across Time



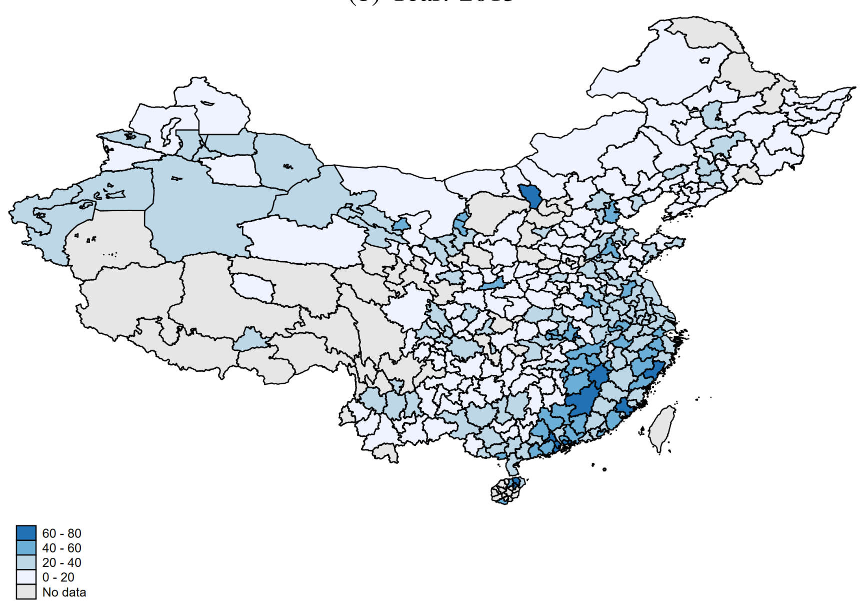
Notes: This figure plots the proportion of firms in SEZs over the period from 1998 to 2013, using the cross-sectional dimension of the new panel dataset constructed in Section 3.1. The y-axis, labeled ”Share of SEZ Firms (%),” represents the percentage of firms within SEZs relative to the total number of firms 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 SEZs across China for the years 1998 and 2013, using the cross-sectional dimension of the new panel dataset constructed 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 used to create a new firm-level panel dataset designed to track 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 track firm dynamics between SEZs and Non-SEZs, including births and movers across zones, in Section 3.3.

3.1. Construction of a New Firm-Level Panel Data Set

This section describes the process of merging data from the Annual Survey of Industrial Enterprises (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 track firm-level data across zones and is crucial for examining differences in firms' productivity and performance before and after the implementation of 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 Enterprises in China (ASIE), collected by the National Bureau of Statistics (NBS) of China, covering firm-level variables from 1998 to 2013.[10] This dataset includes all state-owned and non-state-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 database contains 4-digit industry classifications that span 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. Since 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 offers two main types of information: basic company details and financial variables. These include crucial data like postal codes and addresses for locating firms in SEZs. Moreover, financial variables like total industrial output, profit, industrial intermediate input, total value-added, investments, fixed assets, accumulated depreciation, liabilities, and more, are essential for analyzing firm performance in this study.

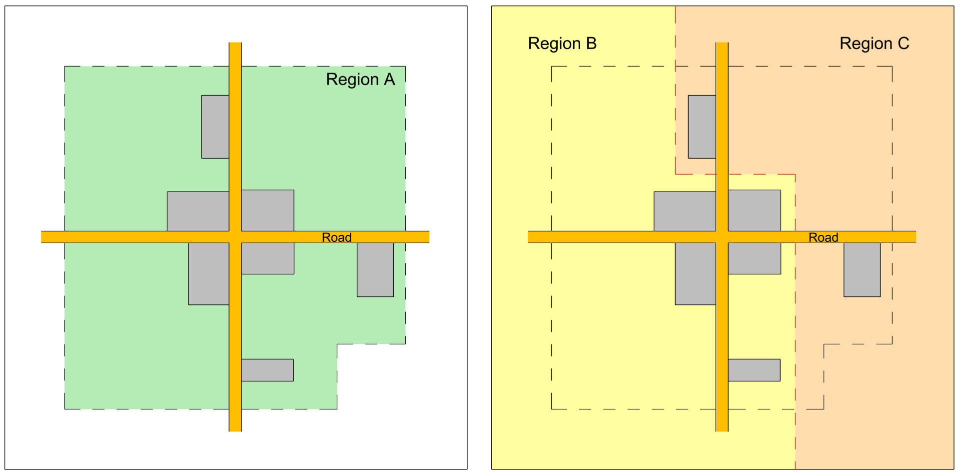
The use of this dataset presents a complex issue due to the dynamic nature of China's district-level administrative divisions, which can significantly affect firms' reported locations. These administrative changes can lead to measurement errors in identifying firms entering SEZs, resulting

[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).

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.

in incorrect associations of firms with districts. As illustrated in 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.

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 a firm's total factor productivity. The Data Appendix A1.1 provides more details on the variables used to calculate value-added to

[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.

estimate firm-level TFP. For a more detailed information regarding these measurements of TFP and agglomeration are found in the respective Appendix A1.2 and A1.4.

TFP Measurement I use the Olley and Pakes (1996) method as my benchmark for the measurement of firm-level TFP. For robustness, I also employ alternative methods for measuring firm-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 A1.2.

3.3. Stylized Facts

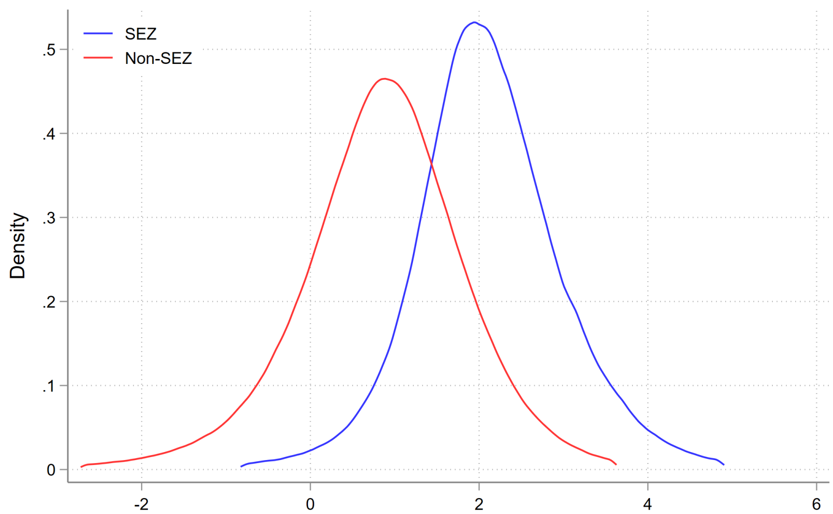
This subsection, utilizing the new panel dataset, presents a set of new stylized facts concerning firms in SEZs and Non-SEZs, with particular focus on their performance in terms of productivity, capital investment, and the capital-productivity correlation within firms over time. It also examines the 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 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 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, aimed at spurring technological progress, contribute to this productivity advantage in SEZs, widening the productivity gap between the 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 examine capital investment and the capital-productivity correlation within firms over time between SEZs and Non-SEZs, and detailed statistics are reported in Table 1.

Table 1 compares firm performance in 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 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 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

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



Notes: This figure plots the kernel density of firm-level Total Factor Productivity (TFP) in 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. The blue line represents SEZs, and the red line represents Non-SEZs.

a marginally more effective use of resources—by about 2%—which could be attributed to more efficient resource allocation in SEZs.

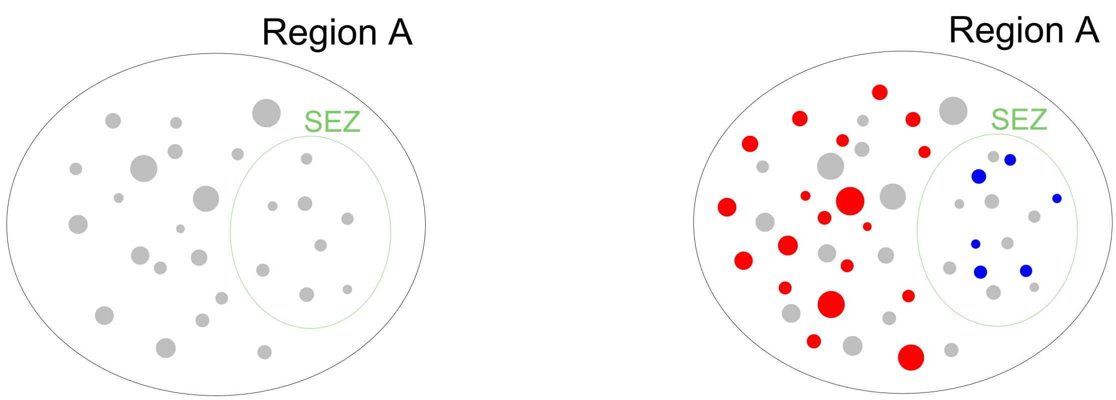
Table 1: Firm-Level Productivity and Capital Across Zones

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 better selection, more efficient resource allocation—evidenced by higher average productivity, increased investment, and a higher correlation between productivity and capital compared to Non-SEZs—we are compelled to investigate further the root causes of these differences. The aim is to discern whether 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 are originally born there, or if it results from the relocation of already high-performing firms from Non-SEZs through the selection channel. These qualified Non-SEZ firms may 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 in SEZs, it is necessary to understand these dynamics more clearly. This includes examining the performance differences between firms originally born in SEZs and those originally born in Non-SEZs, as well as the performance of firms that migrate from Non-SEZs to SEZs in the following sections.

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

(a) (b)



Notes: The figure provides a clear visualization of firm dynamics originating in SEZs and Non-SEZ areas. Panel (a) shows that within Region A, an established SEZ (green circle), 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 Non-SEZ areas are marked with red dots.

After reviewing the cross-sectional differences between SEZs and Non-SEZs, where SEZs tend to perform better in terms of higher productivity, increased investment, and a higher correlation between productivity and capital, this paper seeks to delve deeper into the causes of these disparities. The goal is to determine whether the higher productivity of firms in SEZs is due to their natural ability to attract more productive ”diamond” firms, or whether it stems from successful firms in Non-SEZs relocating to SEZs. It is possible that well-performing firms from Non-SEZs are moving to SEZs to take advantage of a more supportive business environment, such as a lower tax burden and better access to credit markets, especially if they were previously constrained financially. To understand why SEZs are more productive, we need to explore these dynamics further. This involves comparing the performance of firms that originated in SEZs with those from Non-SEZs and assessing how firms from Non-SEZs perform after relocating to SEZs in the subsequent sections.

Stylized Fact II: Birth Differences by Zone This section presents the second stylized fact, focusing on the performance of originally established firms in both SEZs and Non-SEZs. Figure 5 visually compares 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 depicted by 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.

Table 2 compares the performance of firms originating in 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 for Non-SEZ firms. SEZ firms also invest more, with average capital 66% higher at

9.36, compared to 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, significantly enhancing firm growth and productivity.

These results highlight the key factors driving higher productivity in SEZs. They also guide us toward 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

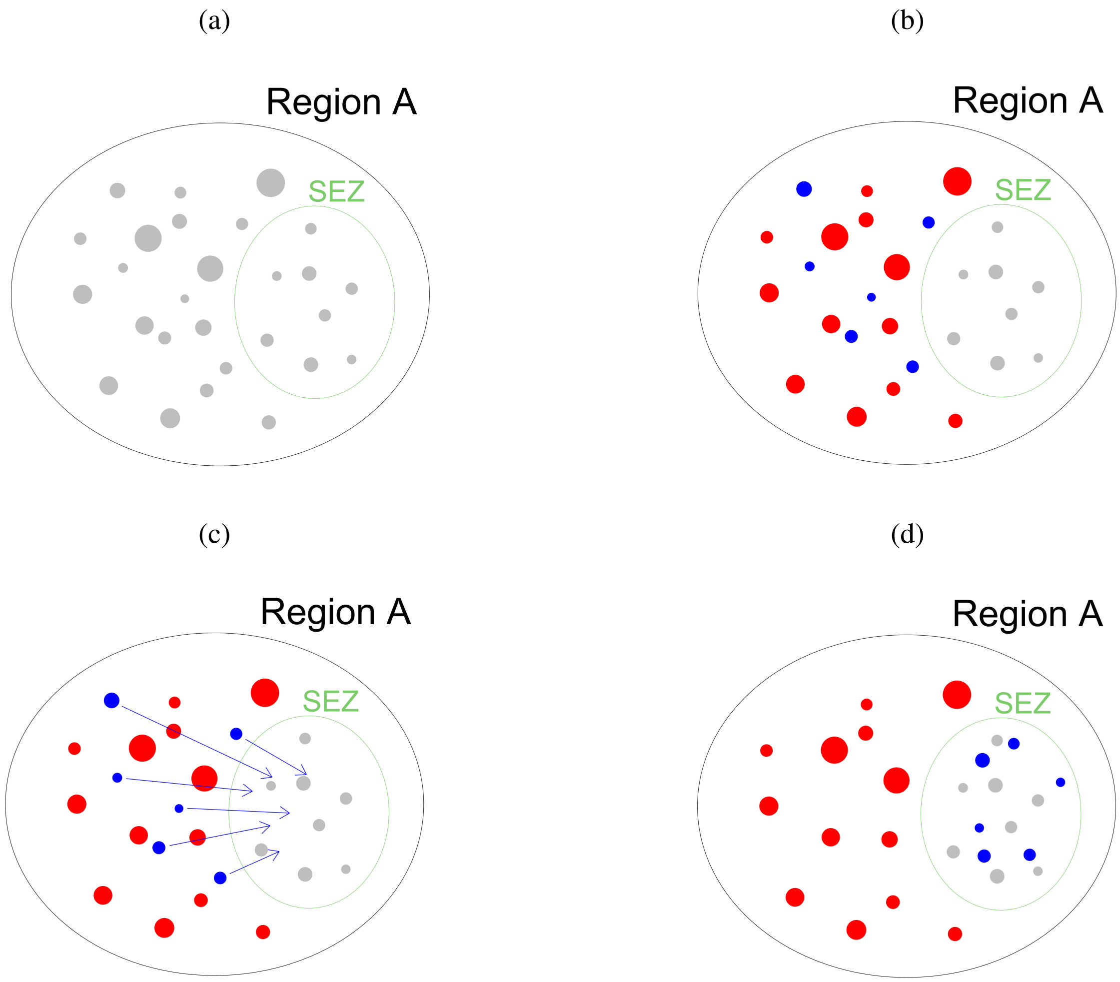
Stylized Fact III: Movers vs. Stayers As mentioned earlier, another factor contributing to the productivity advantage in SEZs may be the migration of high-performing firms from non-SEZ areas to SEZs. This migration is influenced by the selection process and resource allocation. Therefore, our next analysis 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 moved to SEZs compared to those remaining in Non-SEZ areas. It reports various descriptive statistics of productivity, capital, and the covariance between productivity and capital, $Cov(\ln z_i, \ln k_i)$. Initially, firms that potentially moved to SEZs already exhibited better performance, with 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. Of this increase, 58% was realized after moving into SEZs, and only 26% was 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

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 remaining in Non-SEZ areas. In panel (a), an established SEZ is 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).

to 9.16. This growth in capital for SEZ movers—28% of which is attributed to relocating to the SEZ, and 51% driven by selection—underscores the capacity of SEZs to attract or cultivate firms with higher capital investment. Notably, 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 more efficient capital allocation. In contrast, Non-SEZ firms saw a negligible change. The improved resource allocation due to being located in SEZs 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

Summary of Stylized Facts The empirical evidence shows that firms in SEZs exhibit superior 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 firms in SEZs is observed not only among firms originally established within SEZs but also among those relocating from outside. Both types of firms demonstrate enhanced performance relative to their counterparts outside SEZs. Additionally, firms that relocate to SEZs from outside zones display similar advantages post-relocation, including improved firm selection—evidenced by higher average productivity—increased capital, and more effective resource allocation, compared to firms that remain outside the SEZs.

3.4. Agglomeration

A large body of literature explores the concept of agglomeration, suggesting 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. The measurement of agglomeration is described in Appendix A1.4, and here I present results on productivity variations at different levels of agglomeration across zones.

To further analyze how SEZs influence productivity differences in the context of agglomeration, a mediation effect econometric model is employed to investigate these potential agglomeration effects. A detailed statistical description of firm-level productivity between SEZs and Non-SEZs can be found in Appendix A1.3.

TFP and Agglomeration The Figure 7 shows the kernel density of firms' TFP distribution between Non-SEZs and SEZs by agglomeration. Comparing the two gray lines, we observe that SEZs have a higher mean log TFP, and the total distribution shifts to the right, suggesting that SEZs

[13] Agglomeration effects are widely studied (Marshall, 1890; Arrow, 1962; Romer, 1986), suggesting 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.

are associated with higher levels of firm productivity compared to Non-SEZs. 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, as 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 implies that while agglomeration is a factor in explaining the productivity difference between zones, it is not the only one, as SEZs offer additional advantages. Furthermore, the small gap between the dashed and solid lines within both SEZs and Non-SEZs suggests that there is an upper limit to the benefits of agglomeration. If agglomeration were the main mechanism, we might expect more significant productivity differences between low and high agglomeration levels.

In summary, Figure 7 suggests that agglomeration certainly impacts productivity, as evidenced by the differences within the zones. However, the fact that SEZs outperform Non-SEZs even at similar levels of agglomeration indicates that other factors are 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.

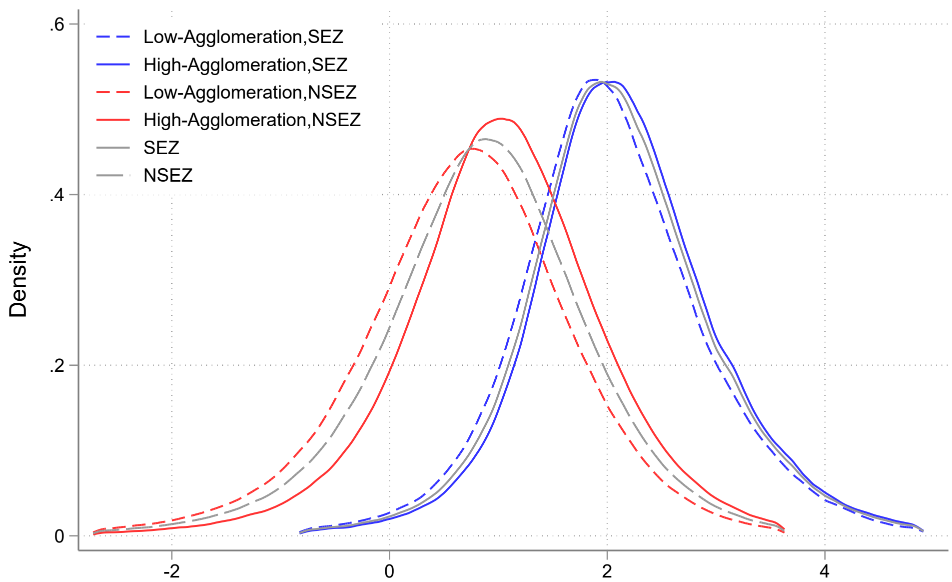
Event study and Mediation Effects Here, I use the staggered implementation of SEZs to conduct an event study. Of course, as documented in my previous Section 2.2, the evolution of the implementation of SEZs is not random, and firm entry into SEZs (either through birth or relocation from Non-SEZs) is also endogenous. Therefore, I cannot interpret the results of the event study as causal. With these important caveats in mind, 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 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 SEZs have 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.

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 shown 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 (below the median), while a solid line represents a high agglomeration level (above the median).

When both γ and β_2 are significantly positive, it suggests that the observed increase in TFP is partly due to 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 heightened agglomeration, denoting a complete mediation effect. In this scenario, the agglomeration effect would be the only channel through which the SEZ policy influences productivity.

Table 4 displays results, where the columns represent 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 the baseline model outcomes, with the SEZ coefficient indicating the overall impact of the policy on productivity. The middle panel presents results from Equation (1), where the mediator variable EG_{irt} is related to SEZ. The SEZ coefficient here reflects how the policy affects agglomeration levels. 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.

The empirical evidence highlights not only a more effective selection process, favoring the most promising firms into SEZs, but also a more efficient allocation of resources that enhances input utilization, fosters output growth, and increases 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 relocate there is influenced by factors inherently linked to the firm's potential for success, 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 the empirical part, in the context of Chinese SEZs, the agglomeration effect does not play the main role in explaining the outperformance of firms in SEZs compared to those in Non-SEZs. Therefore, I abstain from explicitly modeling agglomeration effects. Such a model enables the 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.

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 SEZ policy, affects productivity. Column (1) represents the estimates when we control for separate fixed effects of region and time, while column (2) allows each region to control for its own trend, i.e., instead of controlling for separate fixed effects, we control for the interaction term between region and time.

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-on-hand (assets and bonds). While facing endogenous entry and exit decisions, they also make location choices between Special Economic Zones (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 minimum profit scale requirement, while firms in Non-SEZs have the option to endogenously relocate to SEZs. These three key features, combined with the endogenous entry and exit dynamics, enable us to use this model to assess the aggregate impact of SEZs on Total Factor Productivity (TFP), gain insights into how SEZs contribute to overall TFP growth, and evaluate 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 4.2). Next, we explore the household problem (Section 4.4) and, finally, provide the definition of a recursive competitive equilibrium (Section 4.5).

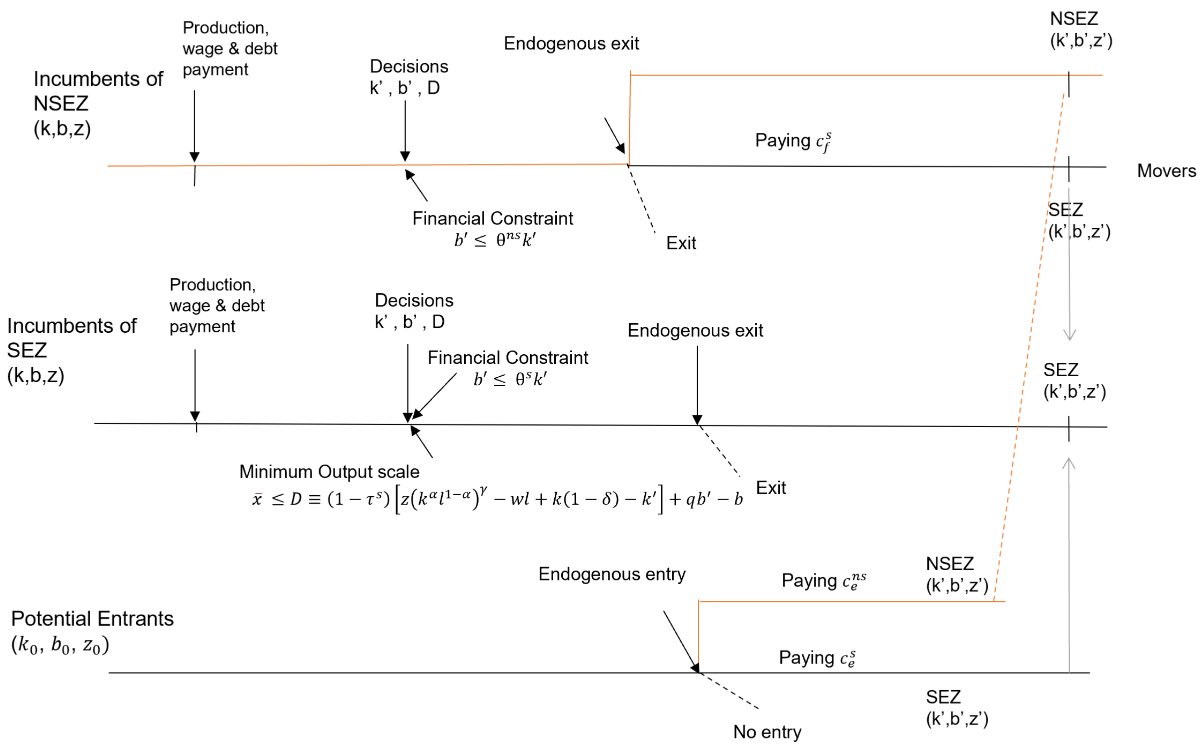
4.1. Model Setup

Time is discrete in an infinite horizon. There are two zones in the economy: the Special Economic Zone (SEZ) and the Non-Special Economic Zone (Non-SEZ). The economy consists of a continuum of firms that are heterogeneous in productivity and produce a homogeneous good in both zones. Firms in both zones are subject to persistent shocks to individual productivity. Each firm owns predetermined capital, k, and bonds, b, from the previous period, and hires labor, n. Firms pay different corporate tax rates depending on whether they are in the SEZ or Non-SEZ, and face different collateral constraints based on their location. Additionally, firms in the SEZ must meet a minimum profit scale requirement each period, whereas firms outside the zone are not subject to this requirement.

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 every period by paying the correspondence fixed entry cost of each zone. Upon entering the market, the firm starts operating in the next period and becomes 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 incumbent firms in Non-SEZ face two choices, they choose either to exit the market or move to SEZ by paying the moving cost to the SEZ.

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



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 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

The productivity distribution is firm-specific, denoted by $\mu_i \equiv \mu(z_i)$ for firm i. A firm's productivity $z_{it} = A_i v_{it}$ consists of a time-varying idiosyncratic component and a transitory component that captures firm-specific characteristics independent of time. The transitory component, A_i, follows a Pareto distribution. The idiosyncratic component, denoted by v_{it}, follows an 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, each of them makes their location choice decision between SEZ and Non-SEZ and maximizes 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 in each period, the 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 $V_i^r(k,b,z)$ represent the value of firm i in the current period in region r, after the firm decides its location to continue operating in the next period and pays the operating cost c_o^r. The firm chooses the location $r \in \{s, ns\}$ to maximize its current value. If the firm chooses to be in an SEZ, the continuation value of operating in the SEZ must be higher than the continuation value in a Non-SEZ. This location decision is described by the following discrete-choice problem:

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

where the first term in the curly brackets represents the value of being in an SEZ, and the second term represents the value of being in a Non-SEZ. At their current location, firms optimally choose 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 value of the next period, $V(k', b', z')$. Financial frictions are introduced through a collateral constraint, and firms face this constraint when borrowing in the above maximization problem. Additionally, firms located in SEZs must also meet a minimum profit scale requirement to continue operating. Firms decide whether to continue operating in the next period or exit the market. They choose to exit if the continuation value is lower than the exit liquidation value. Upon exiting, firms liquidate their depreciated capital and pay off any bond obligations. Finally, conditional on the current location, the value of the firm's optimization problem is defined by $V_i^r(k,b,z)$ as follows:

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

s.t.

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

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

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

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

Each period, there is a group of potential entrants, referred to as new firms born in SEZ and Non-SEZ. Next, we define the value of these 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 regarding which zone to enter, deciding whether to enter the economy and become an incumbent after paying the region-specific entry cost c_e^r, where $r \in \{s, ns\}$. Once the firm enters a specific zone, it begins operation in the next period, given its initial state. 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} represent the value functions for incumbent firms in SEZ and Non-SEZ, respectively. A potential entrant chooses to enter if the value of operating in one of the zones, conditional on their initial state $V^r(b_0, k_0, z_0)$, is greater than the no-entry value, which is 0.

Exit

An incumbent firm has the option to endogenously exit the market. Each period, the firm observes 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 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 the 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

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

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.

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

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:

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

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, 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 is worth emphasizing that a firm's classification may change over its lifecycle, contingent on its state variables. A more detailed solution regarding decision rules for unconstrained firms is provided in Appendix A9.

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', we can 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 firm's problem in Appendix A9. A firm with (k,z) chooses optimal labor demand as

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

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

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

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

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, including 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 Non-SEZs. The third goal involves matching the relative average leverage levels, expressed as the ratio of debt to capital, for SEZ firms compared to Non-SEZ firms. This helps characterize the differing levels of financial tightness across zones, and also captures the exit rate from SEZs, reflecting the minimal profit scale requirement.

The model's parameters fall into three categories. The first group comprises parameters with values from existing literature, aligning 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 Method of Moments (SMM), as discussed in Section 5.1. These parameters are matched to steady-state moments from the model and those observed in firm-level panel data from 1998 to 2013. In total, seven moments are selected to characterize the distinct behaviors of newly-born and incumbent firms across zones.

The remainder of this section is structured as follows. First, Section 5.1 presents parameters that are externally calibrated, with values directly sourced from data. Second, Section 5.2 discusses 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. Table 5 presents the externally calibrated parameters. The corporate income tax rate in Non-SEZs, $\tau^{ns} = 0.33$, aligns with China's corporate income tax policy for firms in Non-SEZs, while the corporate income tax rate in SEZs, $\tau^{s} = 0.195$, 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 & Pakes method used to estimate firms' 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 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 parameters are jointly calibrated using the SMM method, I will explain why certain targeted moments are primarily influenced by specific parameters.

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

The parameter ξ^r, representing 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, indicating that surviving firms in SEZs can be smaller than their counterparts in Non-SEZs due to the favorable policies in SEZs.

Calibrating the minimal profit scale for firms in SEZs, denoted as $\bar{x}$, and the entry cost c_e^r, is done by targeting various moments. These moments include the exit rate from SEZs, the relative initial assets of newly-born firms compared to incumbents, the relative average productivity of incumbent firms in SEZs versus firms in Non-SEZs, and the relative average productivity of newly-born firms in SEZs compared to those born in Non-SEZs. A higher minimal profit scale in SEZs implies that firms in SEZs are expected to be more profitable. A lower entry cost allows potential entrants with fewer initial assets to survive in SEZs, and a more relaxed collateral constraint enables productive firms to be less constrained in SEZs. As a result, 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 align reasonably well with the empirically observed moments that I directly target in the calibration process, as shown in Table 7. One of the key moments targeted in this study is that the average productivity for firms in SEZs should be 2.5 times greater than that of firms in Non-SEZs, according to the empirical data. The model successfully replicates this feature, albeit with a slightly smaller magnitude. Additionally, the average relative productivity for newly born firms in SEZs is 2.5 times that of newly born firms in Non-SEZs. My targeted moment captures the characteristic of higher productivity for new-born firms in SEZs, with a magnitude that exceeds the empirical findings. The model also accurately reflects the lower average debt-to-capital ratio in SEZs. While there is room for further refinement in these moment targets, they closely approximate the empirical data and capture the most critical features.

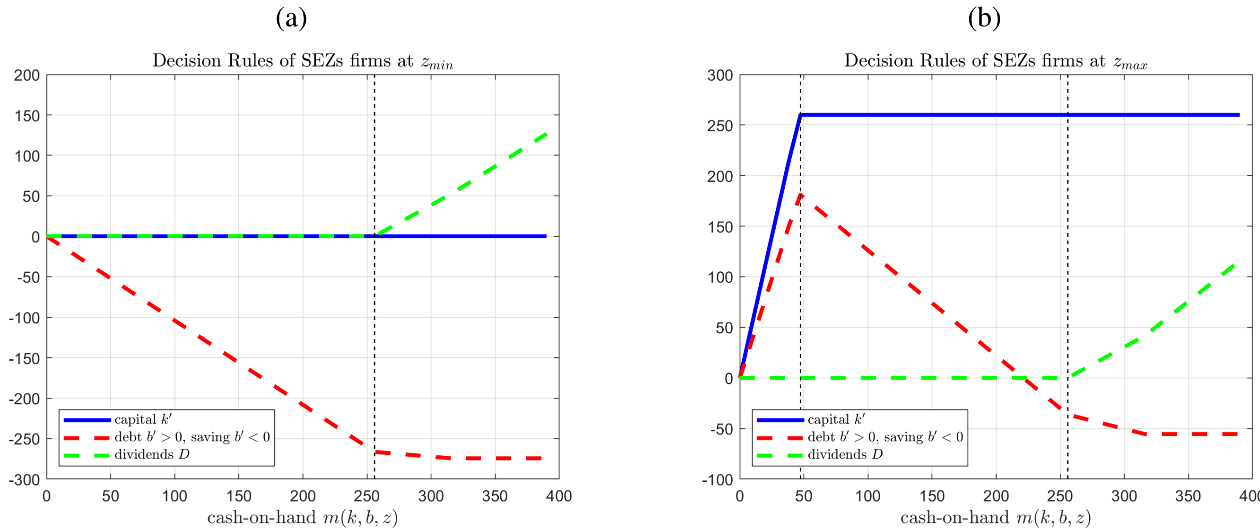
Table 7: Moments Used in Calibration

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

5.4. *Steady State: Firm Decision Rules*

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

Figure 8: Decision Rules: Capital, Bond and Dividends

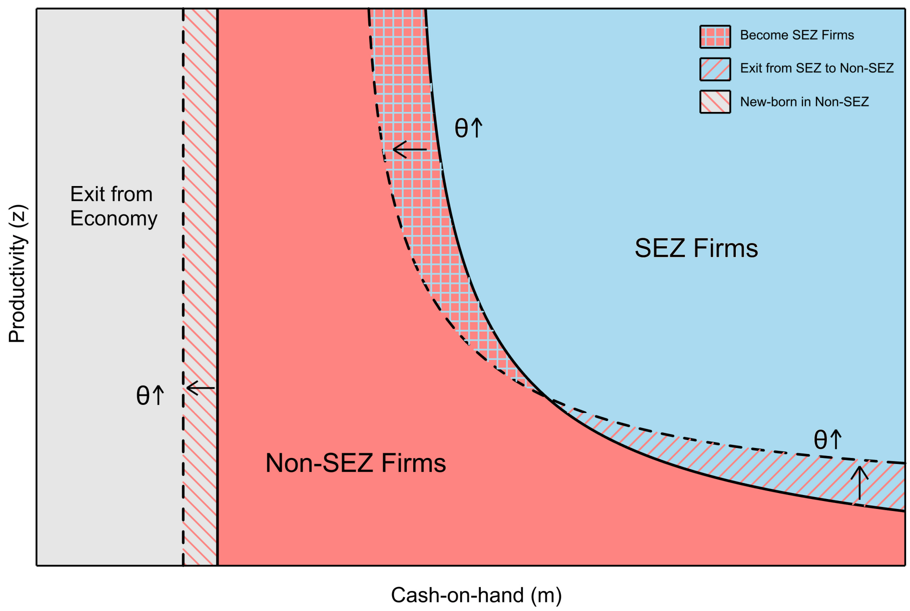


Notes: The figure plots the decision rules for capital k', represented by the solid blue line, bond $b' > 0$ and financial savings $b' < 0$, represented by the red dashed line, and dividends D, shown by the dashed green line. Panel (a) of Figure 8 displays the decision rules for the least productive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

Firms facing financial constraints 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 better understand this concept, Figure 8 illustrates the heterogeneous decision rules of incumbent firms regarding 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 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 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 only profit scale constraints. Firms with cash-on-hand below the lower threshold $\underline{m}$ are subject to both profit scale and financial constraints.

In Panel (a), surviving firms with the lowest productivity z_{min} are considered unconstrained if their wealth exceeds the upper threshold m, such that $m > \bar{m}$. These unconstrained firms adopt an 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 the optimal capital rule $\hat{K}$, but they follow a zero-profit decision. They gradually reduce debt and accumulate internal financial savings $b' < 0$, eventually transitioning into unconstrained firms. It is 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 by positive borrowing. Analyzing the distribution of cash-on-hand among firms, we observe that smaller firms with lower cash-on-hand are concentrated at the lower tail of the distribution while maintaining positive leverage.

Figure 9: Location Choice: θ Increases



Notes: The x-axis presents different levels of cash-on-hand m, and the y-axis presents different levels 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. The solid line represents the distribution division from the benchmark, while the dashed line represents the division after increasing the financial parameter θ.

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, previously 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 higher average productivity for firms located in SEZs. This selection process plays a crucial role in reducing the misallocation of high-ability firms in the economy.

In this new equilibrium, a set of firms with higher average productivity and lower average wealth emerges, facilitating the entry of more firms with lower wealth into the economy as Non-SEZ firms, as indicated by the shaded area. On an 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

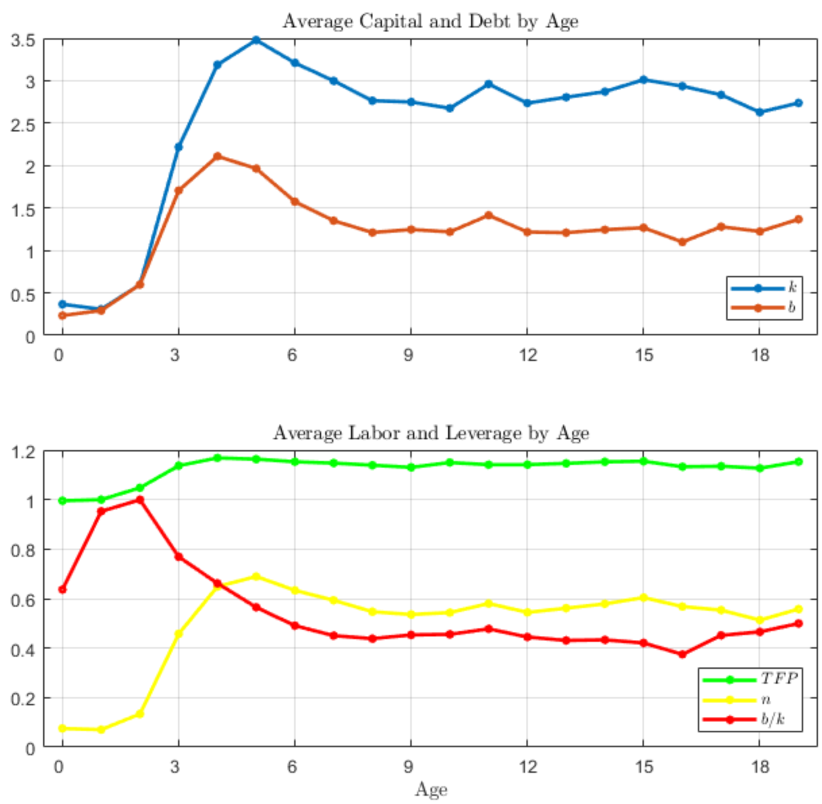
Following the discussion on firms' location choices based on their cash-on-hand and productivity, this section examines the life-cycle dynamics of firms in the model. It provides insights into the heterogeneous dynamic paths observed across zones, showcasing how firms' output, productivity, capital,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

As shown in Figure 11, all young firms start relatively small at the time of their birth and gradually accumulate capital over time. At age 0, these firms face financial constraints due to limited capital for external financing. However, they continue borrowing funds until around age 4, after which they begin deleveraging as sufficient capital is accumulated. By around age 7, firms reach the optimal unconstrained investment level $\hat{K}$, as illustrated by the hump-shaped red leverage curve in the lower panel of the figure. The presence of SEZs, which effectively reduce financial frictions in the economy, allows firms to borrow as needed, provided they have sufficient capital. As a result, firms accumulate capital over time and continue borrowing externally for investment rather than relying solely on internal savings. This leads to the absence of unconstrained firms with positive profits. Additionally, new firms enter the market with higher productivity levels and begin hiring labor for production. Productivity gradually increases as firms accumulate capital, reaching a constant level once capital hits its optimal value.

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.

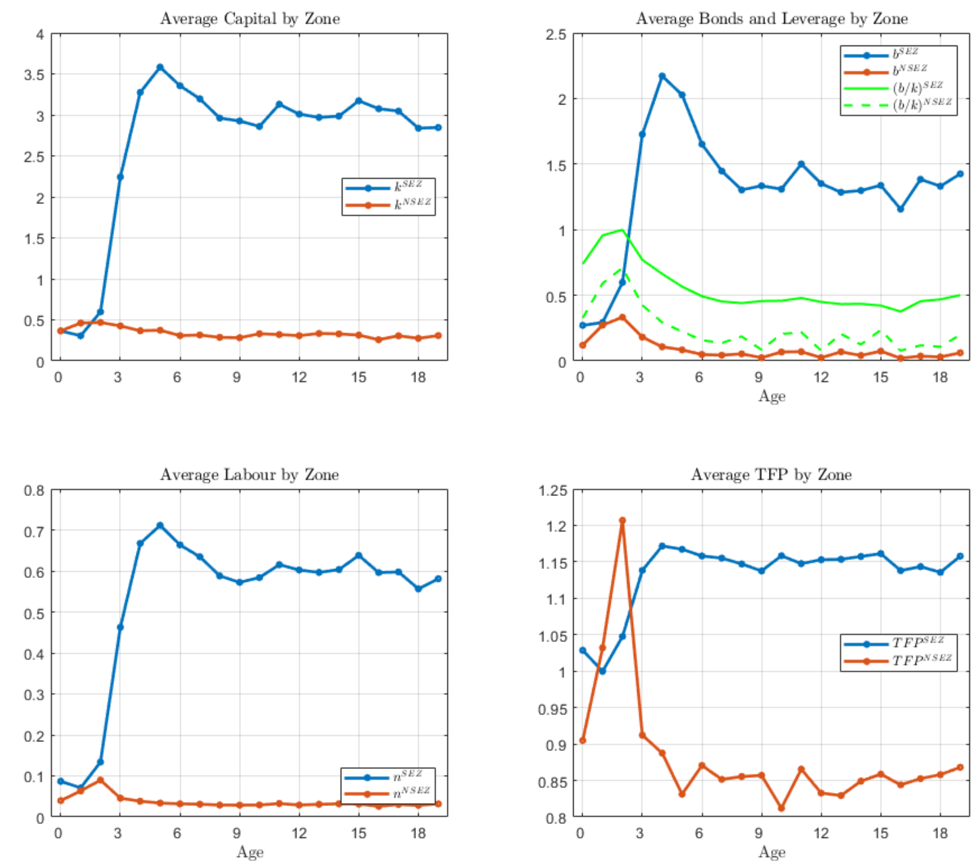
Furthermore, new-born firms in SEZ regions already demonstrate higher productivity levels from their age 0, surpassing those in Non-SEZ. Their productivity gradually increases in SEZs, peaking around age 4 due to the substantial capital accumulation. As the SEZ firms get mature, their productivity stabilizes 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 10: Firm Lifecycle Dynamics: Average capital, Bond, Productivity and Labor



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

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



Notes: The upper panel illustrates the changes in average capital, represented by the blue line, and average bond, indicated by the red line, throughout the firm's lifecycle. The lower panel plots the changes in 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, over the 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 such as firm productivity, the distribution of bonds, and capital for both incumbent and newly born firms in Special Economic Zones (SEZs) and Non-Special Economic Zones (Non-SEZs), the model is used to connect with the micro-level evidence. This allows me to address aggregate-level questions and quantify the impact of SEZ policies on output, capital, and productivity. Furthermore, we undertake a decomposition exercise for each policy feature, providing a comprehensive understanding of how each individual component of SEZ policy affects the economy. This analysis also contributes to discussions on optimal policy considerations, particularly regarding the maximum capacity and optimal share of firms located in SEZs.

In the following, a counterfactual exercise examining the absence of SEZs is detailed in Section 6.1. Then, in Section 6.2, I analyze the impact of shutting down the credit subsidy channel for SEZ firms by tightening the financial constraints in SEZs to match those in Non-SEZs. Lastly, in Section 6.3, after increasing financial frictions, I eliminate the tax subsidy policy for SEZ firms by raising the corporate tax rate to match that in Non-SEZs. These quantitative experiments quantify the effectiveness of each policy in addressing resource misallocation across firms and in promoting aggregate productivity and total welfare.

6.1. No-SEZs Counterfactual

First, I conduct a counterfactual exercise by examining the hypothetical scenario where SEZs do not exist in the economy. In this scenario, I align SEZ features with those of non-SEZs (NSEZ), which involves increasing financial frictions by setting 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 SEZs. This analysis provides insights into the significant impact of SEZs on several key economic variables. Specifically, I focus on Aggregate Total Factor Productivity (TFP), a key driver of economic growth, as well as the allocation of capital, changes in output and input, and their aggregate implications. Furthermore, I investigate the fiscal implications of SEZs by measuring their effective tax impact and evaluating the effects of SEZs on total welfare.

Table 8 presents the outcomes of this counterfactual analysis, comparing the benchmark economy with SEZs to the scenario without SEZs. In the benchmark economy, aggregate TFP exhibits a remarkable 25.7% increase relative to the counterfactual. Aggregate TFP is defined as the Solow residual, $\frac{Y}{(K^{\alpha}L^{(1-\alpha)})^{\gamma}}$, and 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.

Three channels 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. In an economic environment characterized by decreasing returns to scale technology, 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 TFP. To distinguish whether the increased aggregate TFP is primarily due to the larger number of firms or enhanced efficiency from the selection channel, I compare average productivity, computed by re-scaling aggregate variables by the number of firms. Regarding extensive margin adjustments, the economy with SEZs exhibits a 2.82% reduction in the birth rate, allowing only more productive firms to enter the market. This results in a 12.4% increase in the average TFP of newly born firms, along with a higher death rate in the benchmark. This mechanism favors the inclusion of more productive firms born in SEZs while expelling less productive firms from the market, leading to a substantial 25.10% increase in average TFP.

The second channel contributing to the aggregate TFP gain stems 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. Additionally, examining the bond-to-capital ratio, which measures the average bond per unit of capital, reveals that in the economy with SEZs, this ratio is 35.56% higher than in NSEZs. The presence of SEZs also reduces the share of financially constrained firms by 95%, demonstrating the effectiveness of SEZ policies in 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 subsection 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 non-SEZs. This analysis examines the aggregate consequences of enhanced access to credit markets on the allocation of factor inputs and its subsequent 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 reduction in resource misallocation across firms. In the model, collateral constraints prevent small but productive firms from investing at optimal levels. The resulting lack of capital in these firms reduces aggregate TFP. The presence of SEZs, which offer credit subsidies, addresses this 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, thereby raising aggregate TFP.

Furthermore, improved access to credit for productive firms also affects the selection channel. As more productive newly-born firms enter SEZs, there is a 12.4% increase in their birth productivity, along with a 9% reduction in the death rate. These factors contribute to a 15.14% 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 subsection, I perform a counterfactual analysis by eliminating the credit subsidy channel for SEZ firms and, simultaneously, increasing tax frictions by raising the corporate tax rate for SEZ firms to be equivalent to that in non-SEZs. The purpose of this analysis is to explore the aggregate effects resulting from the interplay between 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% in the benchmark economy. This increase is primarily driven by a significant rise in the number of firms entering the market, particularly those originating in SEZs. The lower tax rate reduces the fiscal burden on businesses, allowing less efficient firms to take advantage of the reduced tax rates, access more loans, and continue their operations. This is evident in the 9% reduction in the death rate in the benchmark. However, this scenario results in a less efficient selection process, leading to resource inefficiencies and wastage. Consequently, the correlation between productivity and capital decreases compared to the scenario with only reduced financial frictions. The increased correlation between productivity and capital drops from 79% (as observed in Table 9) to 72% (as reported in Table 10). This decline results from the survival of low-productivity firms due to the favorable tax environment, leading to a 37% decrease in average productivity. Additionally, the impact of taxes on firms is reduced by 41% 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 SEZs promote aggregate TFP by 25.7%, with this increase driven by both improved selection, resulting in a 25.1% rise in average firm TFP, and enhanced resource allocation, evidenced by an 88% increase in the correlation between capital and productivity. Isolating the role of financial frictions, I find that approximately half of the increase in aggregate TFP is attributable to the reduction of financial frictions in SEZs, which induces better selection and resource allocation for the endogenous distribution of firms. Quantifying the contributions of each channel provides a clearer perspective on their respective roles in promoting productivity growth. This understanding also informs further discussions on determining 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 induce better selection and more efficient resource allocation. As productivity and financial development increase within SEZs, high-productivity firms from non-SEZ areas relocate to SEZs, lowering marginal asset threshold levels due to higher returns on productivity. Reduced financial frictions in SEZs enable high-productivity firms to expand their operations and increase capital investment. Both factors attract more productive firms from non-SEZ areas, intensifying competition within SEZs and displacing less productive, higher-asset firms, which in turn relocate to non-SEZ areas.

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

8. Conclusion

This paper constructs an innovative geo-coded firm-level panel dataset to investigate the effects of China's Special Economic Zones (SEZs) on economic performance, focusing on the mechanisms through which SEZs impact outcomes. The empirical analysis demonstrates that SEZs perform better in terms of output, productivity, capital, and resource allocation, with a notable contribution from new-born firms through the selection channel. The study further incorporates a quantitative approach through a dynamic firm model that includes endogenous entry, exit, and location choices. This model successfully replicates the empirical findings and quantifies a 25.7% increase in aggregate Total Factor Productivity (TFP). The increase is driven by improved selection, leading to a 25.1% rise in average firm TFP, and enhanced resource allocation, evidenced by a 35.6% rise in the bond-capital ratio, 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. Notably, about half of the TFP growth is attributable to improved selection, resource allocation, and investment, which 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 economies of SEZs, offering valuable insights for policymakers to promote productivity and economic growth.

References

Adamopoulos, Tasso, Loren Brandt, Jessica Leight, and Diego Restuccia, "Misallocation, Selection and Productivity: A Quantitative Analysis with Panel Data from China," NBER Working Papers 23039, National Bureau of Economic Research, Inc January 2017.

Alder, Simon, Lin Shao, and Fabrizio Zilibotti, "Economic reforms and industrial policy in a panel of Chinese cities," *Journal of Economic Growth*, December 2016, *21* (4), 305–349.

Arimoto, Yutaka, Kentaro Nakajima, and Tetsuji Okazaki, "Sources of productivity improvement in industrial clusters: The case of the prewar Japanese silk-reeling industry," *Regional Science and Urban Economics*, 2014, *46*, 27–41.

Arrow, Kenneth, "Economic Welfare and the Allocation of Resources for Invention," in "The Rate and Direction of Inventive Activity: Economic and Social Factors," National Bureau of Economic Research, Inc, 1962, pp. 609–626.

Behrens, Kristian, Gilles Duranton, and Frédéric Robert-Nicoud, "Productive cities: Sorting, selection, and agglomeration," *Journal of Political Economy*, 2014, *122* (3), 507–553.

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

_ , _ , and _ , "Challenges of Working with the Chinese NBS Firm-Level Data," *China Economic Review*, 2014, *30*, 339–352.

Brooks, Wyatt J., Joseph P. Kaboski, and Yao Amber Li, "Agglomeration, Misallocation, and (the Lack of) Competition," *American Economic Journal: Macroeconomics*, October 2021, *13* (4), 483–519.

Buera, Francisco J. and Yongseok Shin, "Financial Frictions and the Persistence of History: A Quantitative Exploration," *Journal of Political Economy*, 2013, *121* (2), 221–272.

_ , Joseph P. Kaboski, and Yongseok Shin, "Finance and Development: A Tale of Two Sectors," *American Economic Review*, August 2011, *101* (5), 1964–2002.

Clementi, Gian Luca and Hugo Hopenhayn, "A Theory of Financing Constraints and Firm Dynamics," *The Quarterly Journal of Economics*, 02 2006, *121*, 229–265.

Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, Diego Puga, and Sébastien Roux, "THE PRODUCTIVITY ADVANTAGES OF LARGE CITIES: DISTINGUISHING AGGLOMERATION FROM FIRM SELECTION," *Econometrica*, 2012, *80* (6), 2543–2594.

Ellison, Glenn and Edward Glaeser, "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach," *Journal of Political Economy*, 1997, *105* (5), 889–927.

Farole, Thomas, Gokhan Akinci et al., *Special economic zones: progress, emerging challenges, and future directions*, World Bank Publications, 2011.

Glaeser, Edward, Hedi D. Kallal, Jose Scheinkman, and Andrei Shleifer, "Growth in Cities," *Journal of Political Economy*, 1992, *100* (6), 1126–52.

Gottlieb, Charles and Jan Grobovsek, "Communal Land and Agricultural Productivity," Edinburgh School of Economics Discussion Paper Series 270, Edinburgh School of Economics, University of Edinburgh February 2016.

Hopenhayn, Hugo A., "Entry, Exit, and firm Dynamics in Long Run Equilibrium," *Econometrica*, 1992, *60* (5), 1127–1150.

Hopenhayn, Hugo and Richard Rogerson, "Job Turnover and Policy Evaluation: A General Equilibrium Analysis," *Journal of Political Economy*, 1993, *101* (5), 915–938.

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

Jacobs, Jane, *The Economy of Cities*, Random House, 1969.

Jacobson, Louis S., Robert J. LaLonde, and Daniel G. Sullivan, "Earnings Losses of Displaced Workers," *The American Economic Review*, 1993, *83* (4), 685–709.

Khan, Aubhik and Julia Thomas, "Credit Shocks and Aggregate Fluctuations in an Economy with Production Heterogeneity," *Journal of Political Economy*, 08 2011, *121.*

König, Michael, Kjetil Storesletten, Zheng Song, and Fabrizio Zilibotti, "From Imitation to Innovation: Where Is All That Chinese R&D Going?," *Econometrica*, 2022, *90* (4), 1615–1654.

Lagakos, David and Michael E. Waugh, "Selection, Agriculture, and Cross-Country Productivity Differences," *American Economic Review*, April 2013, *103* (2), 948–80.

Levinsohn, James and Amil Petrin, "Estimating Production Functions Using Inputs to Control for Unobservables," *The Review of Economic Studies*, 04 2003, *70* (2), 317–341.

Lu, Yi, Jin Wang, and Lianming Zhu, "Place-Based Policies, Creation, and Agglomeration Economies: Evidence from China's Economic Zone Program," *American Economic Journal: Economic Policy*, August 2019, *11* (3), 325–60.

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

Midrigan, Virgiliu and Daniel Yi Xu, "Finance and Misallocation: Evidence from Plant-Level Data," *American Economic Review*, February 2014, *104* (2), 422–58.

Olley, G. Steven and Ariel Pakes, "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 1996, *64* (6), 1263–1297.

Restuccia, Diego and Richard Rogerson, "Policy Distortions and Aggregate Productivity with Heterogeneous Plants," *Review of Economic Dynamics*, 2008, *11* (4), 707–720.

Romer, Paul, "Increasing Returns and Long-run Growth," *Journal of Political Economy*, 1986, *94* (5), 1002–37.

Schminke, Annette and Johannes Biesebroeck, "Using Export Market Performance to Evaluate Regional Preferential Policies in China," *Review of World Economics*, 12 2011, *149.*

Song, Zheng, Kjetil Storesletten, and Fabrizio Zilibotti, "Growing Like China," *American Economic Review*, February 2011, *101* (1), 196–233.

Wang, Jin, "The economic impact of Special Economic Zones: Evidence from Chinese municipalities," *Journal of Development Economics*, 2013, *101*, 133–147.

Wong, Kwan-Yiu, "China's Special Economic Zone Experiment: An Appraisal," *Geografiska Annaler. Series B, Human Geography*, 1987, *69* (1), 27–40.

Yeung, Yue, Joanna Lee, and Gordon Kee, "China's Special Economic Zones," *Eurasian Geography and Economics - EURASIAN GEOGR ECON*, 03 2009, *50*, 222–240.

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

Online Appendix

A1. Measurement of Variables

The analysis and discussion in this section focus on the estimation of the firm's total factor productivity (TFP) and the measurement of agglomeration. The Data Appendix A1.1 provides more details on the variables used to calculate value-added to estimate firm-level TFP. Preliminary results regarding the estimated TFP distribution between zones and by agglomeration level are provided in Appendix 3.4. More detailed information regarding these measurements of TFP and agglomeration can be found in the respective Appendix A1.2 and A1.4.

A1.1. Value Added Measurement

Value added is used to estimate TFP. It is not directly reported by firms and is calculated by the NBS using the expenditure approach, which defines value added as the sum of output net of goods purchased for resale and indirect taxes, minus material inputs:

$$VA = \text{Output} - \text{Intermediate Input} + \text{Payable Value Added Tax}$$

Due to missing values in 2004 and 2008, we construct an alternative value-added measure using the income approach at the firm level. Value added is then calculated by summing four components: labor compensation, profit, net indirect taxes (indirect taxes minus government subsidies), and depreciation. Among them, labor compensation consists of wages, unemployment insurance, welfare expenditures, pension contributions (after 2003), and housing subsidies (after 2004). Indirect taxes consist of three accounting items in our data: sales tax, value-added tax, and "other taxes under management expenses":

$$VA = \text{Labor Compensation} + \text{Profit} + \text{Indirect Taxes} + \text{Depreciation}$$

where

$$\begin{aligned}\text{Labor Compensation} =& \text{Salary} + \text{Unemployment Insurance} + \text{Welfare Expenditures} \\ &+ \text{Pension Contributions} + \text{Housing Subsidy}\end{aligned}$$

and

$$\begin{aligned}\text{Indirect Taxes} =& \text{Sales Tax} + \text{Value Added Tax} \\ &+ \text{Other Taxes under Management Expenses.}\end{aligned}$$

Investment is estimated using capital stock.

A1.2. TFP Measurement

Regarding firm productivity, a widely used approach is to estimate Total Factor Productivity (TFP). In the literature, there is an ongoing debate about TFP measurement, with several methods available, such as Ordinary Least Squares (OLS), Fixed Effects (FE), the methods of Olley and Pakes

(1996) and Levinsohn and Petrin (2003), and Generalized Method of Moments (GMM). Commonly, estimating production functions using OLS and FE introduces issues such as selection bias and endogeneity between input choices and productivity shocks. This paper follows the method proposed by Olley and Pakes (1996), which accounts for the effect of TFP on firm investment decisions, and the reciprocal effect of investment decisions and TFP on firms' survival probability. This method effectively addresses the two-way causality and sample selection problems that both parametric and non-parametric approaches encounter. It does so by inverting the investment equation non-parametrically to proxy for unobserved productivity, thereby avoiding the endogeneity issue.

Two specific points regarding the estimation of TFP in this paper need clarification. First, the output used in the TFP estimation is value-added, which is calculated using the input-output method. Our estimation improves upon the TFP calculations used by Brandt et al. (2012). For example, we utilized officially reported price deflators, whereas Brandt et al. (2012) constructed deflators from nominal and real output reported by firms. For input price deflators, we used input-output tables from 1997, 2002, and 2007, unlike Brandt et al. (2012), who used only a single year's table, thus overlooking time-based changes. We also carefully constructed firm-level capital stock (details of which are provided in the appendix to save space).

Second, we estimated the output elasticity of capital, labor, and intermediate inputs separately for each 2-digit industry, allowing for variations in input elasticities across industries. Importantly, this method does not affect the empirical results, as all regressions control for industry-fixed effects. Additionally, to mitigate the influence of outliers, we applied a trimming procedure that removes the top and bottom 0.5% of firm-level TFP observations in both SEZ and non-SEZ regions.

I will describe several variables essential for calculating Total Factor Productivity (TFP), which are not accurately reported in the data. Below, I detail the processes used to address these issues.

One of the primary variables for estimating TFP is **capital stock**. The procedure to calculate capital stock relies on the reported total fixed assets value in the data. According to accounting principles, this indicator refers to tangible assets held by an enterprise to produce goods, provide labor services, or rent/operate for more than one fiscal year. Moreover, it represents the "original fixed assets" [15] after deducting impairments and provisions for construction materials and ongoing construction. Thus, this value offers a relatively accurate description of a firm's capital status.

However, the dataset lacks consistently reported values for fixed investment. To estimate fixed investment, I use the capital stock and apply the macro accounting method:

$$I_t = K_t - K_{t-1} + D_t$$

where K_t is the capital stock at time t, K_{t-1} is the capital stock from the previous period, and D_t represents the depreciation of fixed assets. This method provides a reliable estimate of investment values over time.

The **industrial value-added** variable is used as the dependent variable in the TFP estimation process. Since the reported value-added data contains many missing values, I construct an estimate following the income approach, as in Brandt et al. (2014). The value-added is calculated by summing four components: labor income, profit, net indirect taxes, and depreciation, as follows:

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

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

where labor compensation includes wages, unemployment insurance, welfare expenditures, pension contributions, and housing subsidies. Net indirect taxes include sales tax, value-added tax, and other management expenses, reduced by government subsidies.

To make nominal variables comparable over time, all nominal values are deflated using respective price deflators based on the year 1998 to express them in constant prices. The real value-added variable is obtained by deflating the nominal value-added using output deflators classified by 4-digit industrial codes. Additionally, input deflators, varying across 4-digit industries by year, are calculated from the National Input-Output (IO) tables to convert raw materials and intermediate inputs into real terms.

Table 12 shows the descriptive statistics for the main variables used to estimate TFP.

Table 12: 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.

Tables 13 and 14 compare the elasticity of labor and capital between four different ways of estimating TFP. We can see that in the case of OLS and FE, labor is more intensive in the production technology, and in SEZs the elasticity of capital is higher than in Non-SEZs. However, these results are not robust after controlling for unobserved shocks to productivity, as shown by the results in the OP and LP methods, where capital is more intensive in Non-SEZs than in SEZs.

A1.3. Firm-level TFP Across Zones

In the table 15, 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 13: 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 14: 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

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

Table 15: 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							

A1.4. Industrial Agglomeration Measurement

As measure 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.

A2. Staggered DID Estimate: Within Firms

Direct comparisons of key variables such as output, productivity, and capital, along with their covariance for firms relocating to 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 tend to be more productive, have higher output and capital and use capital more efficiently than those firm outside the zone;X_{it} is the set of time-varying, firm-levels variables that capture the firms characteristics[16]; 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 16, 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 16: 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
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

[16]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 ($size = ln(\text{total asset}/ipi)$).

A3. 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 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 effects of SEZs based on the following Equation.

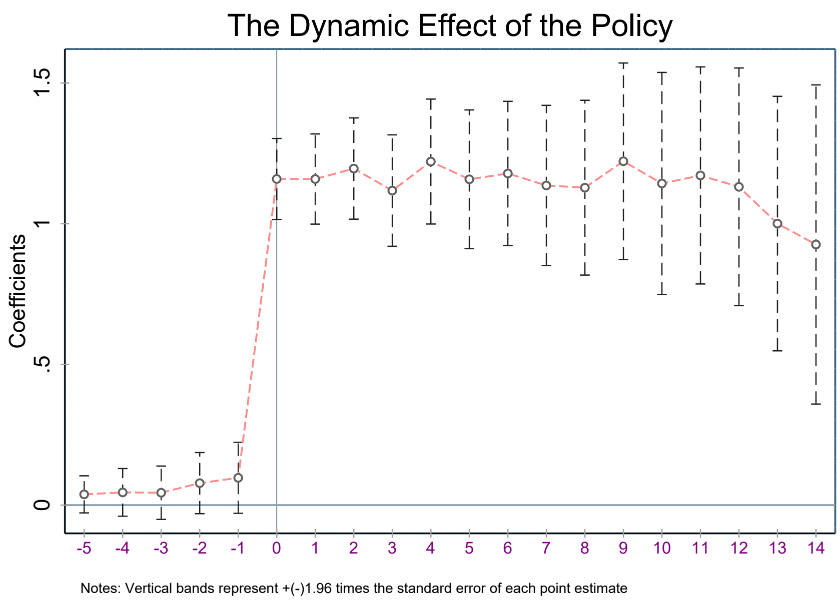
$$TFP_{it} = \theta_i + \alpha_{rt} + \beta_0 D_{it} + \sum_{s=1}^{S} D_{i,t+s}\beta_s + \sum_{m}^{M} D_{i,t-m}\beta_{-m} + \epsilon_{it} \quad (11)$$

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 A3.

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 12, 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 A4. The impact of SEZs on firm productivity is found to vary across different productivity levels of firms. To explore these varied impacts, Appendix A5 study the heterogeneous effects of SEZs on firm productivity, providing a deeper understanding of how SEZs influence firms with different productivity profiles.

Figure 12: Dynamic Effect of SEZ on Productivity



Notes: The figure plots the estimated impact of the 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.

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 17 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 17: 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)

A4. SEZ effect on trends in TFP

The event study above illustrated 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 18 presents the results from estimations above,

Table 18: The effects of the SEZs on TFP and its trend

	(1) m1	(2) m2	(3) m3	(4) m4	(5) m5	(6) m6	(7) m7	(8) m8	(9) m9	(10) m10	(11) m11	(12) m12
current	0.897***	0.936***	1.001***	1.001***	0.975***	0.972***	0.897***	0.938***	0.997***	0.997***	0.964***	0.969***
	(195.15)	(142.71)	(135.58)	(135.46)	(114.68)	(200.86)	(198.04)	(144.11)	(136.12)	(136.04)	(115.24)	(202.19)
PostSEZ	0.915***	0.883***	0.965***	0.965***	0.979***	0.952***	0.935***	0.893***	0.968***	0.968***	0.967***	0.956***
	(192.83)	(132.88)	(146.51)	(146.31)	(113.02)	(197.48)	(198.70)	(135.04)	(147.56)	(147.36)	(113.22)	(197.91)
size		0.127***	0.109***	0.109***				0.117***	0.0995***	0.0993***		
		(53.66)	(39.98)	(39.63)				(48.97)	(35.95)	(35.71)		
lnage		-0.0564***	-0.0306***	-0.0304***	-0.0512***	-0.0267***		-0.0505***	-0.0206***	-0.0205***	-0.0488***	-0.0201***
		(-25.10)	(-11.86)	(-11.67)	(-12.26)	(-15.36)		(-22.91)	(-8.20)	(-8.09)	(-11.95)	(-11.82)
lnROA		0.0695***	0.0145***	0.0145***				0.0645***	0.0129***	0.0129***		
		(125.20)	(26.43)	(26.39)				(118.98)	(23.79)	(23.79)		
lnlev		-0.0454***	-0.0342***	-0.0341***	-0.0217***	-0.0215***		-0.0374***	-0.0331***	-0.0329***	-0.0213***	-0.0205***
		(-22.56)	(-14.66)	(-14.59)	(-10.60)	(-14.94)		(-19.12)	(-14.40)	(-14.30)	(-10.50)	(-14.44)
Export		-0.00835***						0.000940				
		(-2.91)						(0.32)				
State-owned		-0.0698***	-0.0329***	-0.0329***	-0.00470	-0.0280***		-0.0585***	-0.0192**	-0.0197**	-0.00206	-0.0154***
		(-9.66)	(-3.93)	(-3.92)	(-0.40)	(-4.90)		(-8.17)	(-2.28)	(-2.33)	(-0.17)	(-2.72)
lnky			-0.721***	-0.720***	-0.623***	-0.621***			-0.718***	-0.718***	-0.625***	-0.623***
			(-341.80)	(-340.82)	(-255.92)	(-379.80)			(-338.09)	(-337.14)	(-257.00)	(-379.20)
lnexp			-0.00640***	-0.00654***	-0.00675***	-0.00798***			-0.00619***	-0.00625***	-0.00678***	-0.00763***
			(-4.87)	(-4.95)	(-5.28)	(-8.81)			(-4.76)	(-4.79)	(-5.41)	(-8.54)
lnEG				0.00823**	0.00406	0.00697***				0.00778**	0.00611	0.00689***
				(2.42)	(1.08)	(3.13)				(2.30)	(1.62)	(3.11)
lnsales					0.0935***						0.0855***	
					(31.00)						(27.80)	
lnprofit_net					0.0618***	0.0647***					0.0604***	0.0618***
					(61.36)	(95.45)					(60.71)	(92.47)
lnY						0.0440***						0.0356***
						(22.94)						(18.61)
_cons	0.911***	-0.103***	0.0307	0.0748**	-0.173***	0.270***	0.909***	-0.0115	0.111***	0.151***	-0.0633	0.369***
	(1040.88)	(-4.26)	(1.06)	(2.16)	(-4.53)	(11.36)	(1048.65)	(-0.48)	(3.77)	(4.36)	(-1.63)	(15.58)
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

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

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

A5. 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 19 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 19 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 19: 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)

Table 20: Heterogeneous effects of the SEZs on TFP: with control variables

	(1) quantile20%	(2) 40%	(3) 60%	(4) 80%	(5) 100%
Pre1					
SEZ	1.027***	1.005***	0.976***	0.900***	0.850***
	(29.11)	(42.19)	(47.26)	(45.53)	(37.59)
constant	0.260	-0.285	-0.248	-0.332	-0.518*
	(0.88)	(-0.93)	(-0.93)	(-1.48)	(-1.71)
Observations	5170	6341	6403	6297	5907
R-sq	0.902	0.889	0.890	0.888	0.895
Pre2					
SEZ	1.011***	0.969***	0.968***	0.913***	0.875***
	(27.01)	(39.01)	(38.04)	(39.55)	(35.72)
constant	-0.0226	-0.0594	-0.0438	-0.431	-0.338
	(-0.07)	(-0.19)	(-0.13)	(-1.62)	(-1.22)
Observations	4058	4972	5399	5651	5300
R-sq	0.911	0.918	0.898	0.906	0.903
Pre3					
SEZ	1.003***	0.958***	0.953***	0.938***	0.921***
	(23.49)	(31.44)	(33.57)	(33.38)	(29.83)
constant	0.474	0.299	-0.661**	-0.450	-0.447
	(1.16)	(0.85)	(-2.47)	(-1.54)	(-1.49)
Observations	3625	4525	4917	4669	4419
R-sq	0.914	0.921	0.914	0.917	0.903
Pre4					
SEZ	1.036***	0.925***	1.014***	0.999***	0.984***
	(26.24)	(24.84)	(30.14)	(30.25)	(25.28)
constant	0.379	-0.373	-0.361	-0.363	-0.567*
	(1.04)	(-1.02)	(-1.10)	(-1.21)	(-1.75)
Observations	3196	3711	3750	4015	3754
R-sq	0.922	0.916	0.926	0.923	0.917
Pre5					
SEZ	1.069***	1.003***	0.962***	1.030***	0.949***
	(33.16)	(26.05)	(27.87)	(23.98)	(18.12)
constant	-0.500	-0.0674	-0.844**	-0.606*	-0.306
	(-1.40)	(-0.16)	(-2.43)	(-1.79)	(-0.63)
Observations	4239	3269	2697	2294	1949
R-sq	0.929	0.933	0.932	0.934	0.915

t statistics in parentheses

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

Note: The m1–m6 are the results from equation (5), and m7–m12 present the results from equation (6).

A6. 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 21: 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 22, we provide further study to check whether the average effect from table 26 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 provided in Appendix A8 to better understand how agglomeration serves as a moderator, potentially intensifying the influence of SEZ policy on firms' productivity.

A7. 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 direct policy impact on agglomeration from Eq.(7) (0.002) and the partial policy impact on agglomeration from Eq(8) (19.24). Moreover, the test illustrates 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 22: 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 table shows how EG_{irt}, as a mediator of SEZs policy, affects productivity at one period before the policy implementation across sub-samples ranging from the bottom 20% to the top 20%. Column (1) represents the estimates for sub-samples at the 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 the top 20% of the sample.

Table 23: 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 table shows how EG_{irt}, as a mediator of SEZs policy, affects productivity at one period before the policy implementation across sub-samples ranging from the bottom 20% to the top 20%. Column (1) represents the estimates for sub-samples at the 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 the top 20% of the sample.

Table 24: 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 productivity at one period before the policy implementation across sub-samples ranging from the bottom 20% to the top 20%. Column (1) represents the estimates for sub-samples at the 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 the top 20% of the sample.

Table 25: 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 table shows how EG_{irt}, as a mediator of SEZs policy, affects productivity at one period before the policy implementation across sub-samples ranging from the bottom 20% to the top 20%. Column (1) represents the estimates for sub-samples at the 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 the top 20% of the sample.

Table 26: 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 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 27: 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 28: 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 29: 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 the top 20% of the sample.

The table 29 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 30 shows the results[17].

Table 30: Sobel Test Results

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 effects of region and time across sub-samples ranging from the bottom 20% to the top 20% at five periods before the policy implementation. Column (1) represents the estimates for sub-samples at the 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 the top 20% of the sample.

A8. 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.

The following Table 31 shows the results from the specification 14,

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

[17] 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.

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 31: 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 subsample 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 $lnEG$ the 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 the logarithm of EG index to measure the city agglomeration level.

The results from the table 32 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 32: 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 column m1 represent the estimates from Equation 15

A9. 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 $\hat{L}$ and $\hat{K}$, the optimal debt policy $b' = \hat{B}(z)$ is defined by the following equations:

$$\tilde{B}(k, z_i) = (1-\tau)\Big[z_i\left(k^{\alpha}\hat{L}^{1-\alpha}\right)^{\gamma} - w\hat{L} + (1-\delta)k - \hat{K}(z_i)\Big] + q\min\left\{\hat{B}(z_i), \theta\hat{K}(z_i)\right\}$$

- 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.

$$
\begin{aligned}
V(m, z_i) = & \max_{k', b', D, m'_j} \left[D + \beta \max\Big\{ V_x(m), \sum_{j=1}^{N_z} \pi^z_{ij} V(m'_j, z_j) \Big\} \right] \\
\text{s.t.} \quad & 0 \leq D \equiv m - k'(1-\tau) + qb' \\
& b' \leq \theta k' \\
& m'_j \equiv m(k', b', z_j) \\
& = (1-\tau) \left[z_j (k'^{\alpha} \hat{L}^{(1-\alpha)}(k', z_j))^{\gamma} - w\hat{L}(k', z_j) + (1-\delta)k' \right] - b'
\end{aligned} \tag{16}
$$