KIEL

WORKING PAPER

Place-based policies and agglomeration economies: Firm-level evidence from Special Economic Zones in India

No. 2209 revised version November 2022

Holger Görg and Alina Mulyukova

Kiel Institute for the World Economy
ISSN 1862–1155

KIEL **WORKING PAPER** NO. 2209 REVISED VERSION | NOVEMBER 2022

ABSTRACT

PLACE-BASED POLICIES AND AGGLOMERATION ECONOMIES: FIRM-LEVEL EVIDENCE FROM SPECIAL ECONOMIC ZONES IN INDIA*

Holger Görg and Alina Mulyukova

This paper exploits time and geographic variation in the adoption of Special Economic Zones in India to assess the direct and spillover effects of the program. We combine geocoded firm-level data and geocoded SEZs using a concentric ring approach. Our analysis yields that conditional on controlling for initial selection, SEZs induced negative effects on the productivity growth of within SEZ firms and no evidence for spillovers. In further analysis, we find that the significant negative effects on firms disappear once only looking at commercial SEZs, supporting the idea that government interference plays a role. We also show that the directors of firms located inside the zones experienced a significant increase in their total remuneration, suggesting excessive rent-seeking as a possible explanation for negative productivity effects. Additionally, we estimate the effect only for relatively large, i.e. above mean area, SEZs. Interestingly, we find a strong positive and sizable in magnitude productivity growth increase for inside SEZ firms. These results are in line with the idea that the inefficiency of the program may be due to one peculiarity of the Indian program design, where SEZs can be single-firm entities, which may make political interference and rent-seeking more likely than in a large SEZ with multiple firms.

Keywords: Special Economic Zones, India, TFP growth, firm performance, spillovers, time-varying treatment

JEL classification: O18, O25, P25, R10, R58, R23, F21, F60

Holger Görg
Kiel Institute for the World Economy and Kiel University
Kiellinie 66
D-24105 Kiel, Germany
Email: holger.goerg@ifw-kiel.de
www.ifw-kiel.de

Alina Mulyukova (corresponding author)
Kiel Institute for the World Economy
Kiellinie 66
D-24105 Kiel, Germany
Email: alina.mulyukova@ifw-kiel.de
www.ifw-kiel.de

*The authors are very grateful to Aradhna Aggarwal, Eckhardt Bode, Carsten Eckel, Hartmut Egger, Robert Gold, Tobias Heidland, Cecilia Hornok, Tobias Seidel, Farid Toubal and Gerald Willmann for very helpful comments, and the Riksbankens Jubileumsfond for financial support through the project "SEZs: A force for good to reduce inequality?".

1 Introduction

Place-based policies - a governmental tool used to enhance the economic growth of a particular area - have become increasingly popular among many policy-makers worldwide in the past few decades. Much of the research has focused on analyzing the effectiveness of these programs in developed countries, where the public resources target predominantly distressed regions (Busso et al., 2013; Kline and Moretti, 2014). There are only a few studies that evaluate such policies in developing countries, possibly due to lack of data. However, insights gained from programs in lagging regions in developed countries may not hold when examining programs in emerging economies since policies there generally target the most advantageous areas.

In this paper, we evaluate one of the most popular industrial policy tools used in the last two decades: Special Economic Zones (SEZs). SEZs constitute geographically delineated areas where fiscal incentives and regulatory frameworks are provided with the main goal to attract investments and generate additional economic activity in the region. A World Bank report states that within the zones, governments aim to create new firms and jobs and facilitate the skills and technology transfers. Outside the zones, SEZs are expected to generate synergies, networks and knowledge spillovers to stimulate the economic growth of the region (World Bank, 2017). However, such benefits may not materialize if SEZs lead to a misallocation of resources, in particular if this is due to political interference (Alkon, 2018).

This paper provides novel evidence on the effects of SEZs in one of the fastest growing emerging economies - India. Partly in response to the apparent success of China's SEZs, the government of India introduced the 2005 SEZs Act with the view to attract investments, generate a big push for infrastructure development and thus facilitate economic growth. Over fifteen years since the launch of the program, 354 SEZs have been notified hosting over 5.600 units that provide employment to 2.5 million people (Factsheet on SEZs, Department of Commerce).[1]

We use firm-level data to estimate the impact of SEZs on productivity growth of firms. We distinguish two types of firms - namely those located inside an SEZ, and those in the vicinity. We are thus able to differentiate between a direct effect on insider firms, and spillover effects on others located in the vicinity. While there is a small but growing literature evaluating the impact of SEZs in various countries, most studies use data at some aggregated administrative unit level (e.g., Wang (2013) and Alkon (2018)).[2] Only very few papers use firm level data (Brooks et al., 2021; Steenbergen and Javorcik, 2017; Nazarczuk, 2018) and, to the best of our knowledge, none does so for India. Zooming in on the firm level arguably allows a more precise estimation of the impact of SEZs, taking into account firm heterogeneity, and distinguishing firms within and around the SEZ.

[1] Available under: `http://sezindia.nic.in/upload/634908b5af04cImage_002.pdf` Accessed on 29.02.2020.

[2] With data aggregated at such a geographical level, a distinction between inside and outside SEZs is not possible, of course.

While governments clearly expect SEZs to yield positive effects on development, whether such effects in fact materialize is an open question. Conceptually, firms within an SEZ may be expected to be able to boost their growth performance. They are able to benefit from certain incentives (e.g., tax reductions) which allow them to generate a surplus vis-a-vis non-SEZ firms, which can be invested in innovation or other productivity-enhancing improvements. If such positive direct benefits exist, firms in the vicinty may also be affected. Either positively, if some of the new knowledge generated in the SEZ firms dissipates, or negatively if there is competition for scarce resources, e.g., on the labour market. However, positive direct benefits may not necessarily materialize, if the establishment of an SEZ leads to a misallocation of productive resources or excessive rent-seeking on the part of the firm's owners or managers. This may particularly be the case if there is government interference in the process of establishing SEZs, reflecting political motives in setting up SEZs, or corruption (Alkon, 2018). If firms inside the SEZ do not benefit, there is of course also no potential for spillovers on firms in the vicinity.

This indicates the need for careful economic analysis of the potential benefits of SEZs at the level of the firm. An important challenge confronting such research aiming to assess the SEZs impact on firm performance is the unavailability of data on firms operating inside the zones. For India, detailed information on the actual SEZ (e.g. location, size, establishment year, etc) is publicly available, though there is no information on which firms are located within the SEZs. We overcome this issue by firstly, geocoding the notified SEZs, and the firms in our data set. We then combine these two data sources based on the geocoding using a concentric ring approach. Thus, spatial rings around the centroid of SEZs are created using information on the size of the zone. This allows us to approximate firms inside an SEZ. Subsequently, the radius is increased by 5 kilometers to create the following distance bands: $inside$, $0-5km$, $5-10km$ and $10-15km$. This enables the estimation of the spatial extent of any potential spillovers stemming from SEZs. The use of these fine-grained spatial data thus allows us to identify any potential effect on firms inside SEZs and adjacent non-SEZ firms.

To the best of our knowledge, this study is the first to assemble a representative geocoded firm-level data set with an assigned SEZ status for India. The final dataset consists of an unbalanced panel of firms which includes information on firm characteristics, a firm's SEZ status for each respective distance band, the industry in which SEZ specializes and the date of notification of the SEZ. Firm-level data are obtained from Prowess - a database on the financial performance of Indian companies, collected by the Centre for Monitoring the Indian Economy (CMIE). The data cover periods before the implementation of the program, starting from 1988, and after the SEZs creation, up until 2020.

We exploit the longitudinal structure of the data and compare the performance of firms before and after the introduction of the program. To do so we proceed in two steps. Firstly, we estimate event studies to look at the development of the variables of interest in the years preceding and following the implementation of SEZs, distinguishing firms inside

SEZ and those in distance bands around it. In a second step, a difference-in-differences methodology combined with an inverse probability weighting technique is applied in an attempt to more formally identify the SEZ effect.

Since SEZs are established in more developed regions, a simple mean comparison of treated and untreated firms would lead to biased estimates due to a positive selection bias. Another estimation issue is that, because applications for developing an SEZ are reviewed on a rolling basis, firms are treated in different years over the period 2006-2020, resulting in a staggered treatment introduction. Thus, to correct for selection bias, we utilize a recently developed methodology for time-varying treatments, employing an inverse propensity score re-weighting approach, where weights are created at each point in time conditional on the development of the outcome variable as well as other time-varying variables. In this way, we create a pseudo-population where the treatment assignment at each point in time is orthogonal to the potential outcomes conditional on the pre-treatment covariates (Thoemmes and Ong, 2016; Girma and Görg, 2022). The control group is restricted to the sub-sample of firms located further than 40 kilometers away from the zones, to alleviate concerns that the outcomes of the untreated control group are affected by the treatment.

The analysis yields the following results. Conditional on controlling for initial selection, we do not find that the establishment of an SEZ yields positive growth effects on firms inside an SEZ. By contrast, we indeed find that the establishment of SEZs decreased the productivity growth of firms located inside the zones. In our baseline specification, the magnitude of this effect is around 15%, on average. We do not find any consistent evidence for spillovers, either positive or negative, on firms in the vicinity of the SEZ.

These results for India contrast with those found for China, where generally positive effects are found (e.g., Wang (2013), Lu et al. (2019)). In order to provide potential explanations for the estimated effects, we firstly distinguish SEZs into those developed by the state and those with private developers. We find that the significant negative effects on firms disappear once only looking at commercial SEZs, supporting the idea that government interference plays a role. In order to consider excessive rent-seeking on the part of the firm, we look at the impact on the remuneration of directors. Here we find that the directors of firms located inside the zones experienced a significant increase in their total remuneration growth, despite negative productivity effects. In particular, non-executive directors enjoyed significant gains in compensations paid. This indeed points towards rent-seeking on the part of owners and senior managers of the firm, inhibiting the growth potential.

Additionally, we estimate the effect only from relatively large, i.e. above mean area, SEZs. Interestingly, we find a strong positive and sizable in magnitude productivity growth increase for inside SEZ firms. These results are in line with the idea that the inefficiency of the program may be due to one peculiarity of the Indian program design, where SEZs can be single-firm entities, which may make political interference and rent-seeking more likely than in large SEZs with multiple firms. Once forming a cluster with greater industrial area and hosting multiple units, SEZs appear to have positive growth effects on firms inside the zone.

Our paper contributes to a broader literature examining the effects of place-based policies in the presence of agglomeration economies, which focus mostly on developed countries, (Ham et al., 2011; Busso et al., 2013; Kline and Moretti, 2014; Criscuolo et al., 2019). The focus of this paper is on the SEZs program, a popular policy tool in developing countries used to attract investments and stimulate economic activity in the region. A number of related studies have evaluated SEZs in China (Wang, 2013; Lu et al., 2019), showing that the SEZs establishment increased capital investment, employment, wages and productivity of firms by achieving agglomeration economies. These papers identify SEZs at a more aggregated level than is done in our paper.

In the context of India, there have been several studies evaluating the impact of place-based policies. A paper that most closely relates to our work is Alkon (2018) who also examines SEZs in India and finds no evidence for positive developmental spillovers.[3] The primary difference between our work and his paper is that we use detailed firm-level data and identify treatment at the firm level which allows to separate direct effects and spillovers, and also enables us to look at firm heterogeneity. Other related work is by Blakeslee et al. (2022) who examine the effects of the Industrial Areas program in one of the Indian states and find a significant increase in firm creation and employment in the affected villages. Hasan et al. (2021) focus on industrial backward districts and find a short-run effect of a tax-exemption program in the better-off backward districts. The effect disappears after the program ends. Shenoy (2018) also evaluates the developmental effect of investment subsidies to a newly created Indian state and finds improvements in nightlight activity and household welfare. While these studies bring important insights to understanding the effects of place-based policies in India, they focus primarily on one targeted state, whereas the SEZ program was open to all states. Moreover, these studies evaluate the developmental aspect of the program, whereas we are interested in firm-level outcomes.

The rest of the paper is structured as follows. Section 2 describes the background information on the SEZs program in India. In Section 3 the data are introduced. Section 4 presents the event study. Section 5 proceeds with describing the methodology of time-varying treatment, estimates the effect, and presents the results, heterogeneity analysis, and robustness checks. Section 6 concludes.

2 Background on Indian SEZs

India was one of the first countries in Asia to recognize the importance of Export Processing Zones (EPZs) for promoting exports, with Asia's first EPZ being established in the port city of Kandla, Gujarat state in 1965. The absence of modern infrastructure, an unstable fiscal regime as well as the complexities related to customs controls and clearance led to

[3] Several other papers provide descriptive evidence evaluating the efficiency of the SEZ program in India, e.g., (Aggarwal, 2007, 2012).

the reorganization of export-promoting policies. Motivated by the success of the SEZs in China, the Indian Government announced the launch of the "SEZs policy" in April 2000. The policy aims at enabling the establishment of SEZs in the private sector and making them an engine for economic growth by offering high-quality infrastructure, attractive fiscal incentives and minimum regulations. SEZs provide multiple new features as compared to the existing EPZs, which, among others, are no minimum export performance requirement and provision of social infrastructure in SEZs, whereas EPZs comprised only industrial activity (Aggarwal, 2012). While EPZs were predominantly viewed as export-promoting tools, SEZs' focus was shifted to the generation of additional economic activity and the advancement of infrastructure.

The "SEZs Act" was passed by Parliament in May 2005, receiving Presidential assent on the 23rd of June 2005. The Act came into effect on February 10th, 2006 with the main objectives of: (i) generating additional economic activity, (ii) promoting exports of goods and services, (iii) promoting investment from domestic and foreign sources, (iv) creating employment opportunities, and (v) developing the infrastructure facilities. The incentives and facilities provided to the units in SEZs include:

- Duty free import/domestic procurement of goods for the development, operation and maintenance of SEZ units.
- 100% income tax exemption on export income for the first 5 years, 50% for the next 5 years and 50% of the ploughed back export profit for the next 5 years.
- Exemption from Minimum Alternate Tax[4], Central Sales Tax, Service Tax and State Sales Tax.
- Single window clearance for central and state level approvals.

EPZs established prior to the 2005 Act were notified and converted into SEZs, continuing their operation under the new policy.[5] Any individual, cooperative society, company or partnership firm, including foreign firms, can submit a proposal for setting up an SEZ. They are referred to as developers of SEZs. Compared to SEZs in other countries, SEZs in India are not spatial units designated by the government. Rather, firms must apply for permission to develop an SEZ and customs boundaries are redrawn around the existing location. Therefore, an SEZ status can be assigned even to a single firm. This is a distinctive feature of India's SEZ policy. Another particular feature of Indian SEZs is that the policy provides equal opportunities to develop an SEZ for government, private or joint sectors. We will return to both of these issues in our empirical analysis below.

[4] This exemption was withdrawn on 01.04.2012, however, other incentives remain in place.

[5] In the analysis, only SEZs notified under the 2005 Act are used. That is, we exclude 19 converted SEZs to eliminate the concern that the initial incentives and goals of converted and newly notified SEZs are different. Table A.1 in Appendix provides summary statistics for SEZs notified under the 2005 Act and converted SEZs established before the 2005 Act. On average, converted SEZs have a bigger area compared to newly established SEZs which can be explained by the export-oriented policy of initially designed EPZs.

The establishment of an SEZ proceeds in three steps: approval, notification and operation. The most crucial criterion for approval is the possession of land. When a developer is in the process of acquiring land, only in-principal approval can be granted. Furthermore, the formal approval can be issued only after (i) the state government has signed the project, (ii) the developer can prove the possession of land, and (iii) the state government has provided exemptions from taxes, ensured adequate infrastructure and issued clearance from the state regulatory bodies. After approval, the board provides notification for the authorization to begin the operation, at which point the investment and construction can be initiated (Alkon, 2018). However, not all approved or notified SEZs become finally operational.

For our analysis, we obtained a list of notified SEZs under the 2005 Act from the Ministry of Commerce and Industry, Department of Commerce.[6] The dataset contains information on the name of the developer (which is a private or public company or organization that received notification of approval for developing an SEZ), the village and state names where an SEZ is located, the industry in which the SEZ specializes, the area, and the date of notification. There is no information on the number of units operating in each SEZ nor the amount of attracted investment or people employed in each SEZ. Overall, there are 354 notified SEZs reported by 2020 with the first zone being notified in 2006.[7] Though the type of ownership is not indicated in the list of notified SEZs, 35 out of 354 zones are classified as state-owned according to the list of the Council of State Industrial Development and Investment Corporations of India.[8]

Regarding the location choice, the SEZs Act provides no limitation on the geographic location of the zones. However, it is not surprising to observe the concentration of zones in areas with developed infrastructure, targeting primarily big cities in the most industrialized regions (Kennedy and Rundell, 2014; Palit, 2009; Jenkins et al., 2015). Our data show that 84% of notified SEZs are located in India's eight most industrialized states (Andhra Pradesh, Gujarat, Karnataka, Kerala, Maharashtra, Tamil Nadu, Telangana and Uttar Pradesh). Furthermore, Figure 1 illustrates that there is great heterogeneity with respect to the number of established zones across districts, with some districts receiving up to 44 zones compared to no SEZs in the northern and eastern parts of India. The non-random assignment of zones poses a potential threat to the causal identification of the effect of zones due to a positive selection bias, which will be addressed in our analysis using inverse probability weighting.

There is also substantial heterogeneity in terms of the area size of the SEZs.[9] Figure 2

[6] The list is available under: `http://sezindia.nic.in/upload/uploadfiles/files/notify.pdf`. Last update 29/02/2020.

[7] Table A.2 in the Appendix shows the number of notified SEZs over time.

[8] Available under: `https://www.cosidici.com/`

[9] To facilitate the expansion of large-sized SEZs, the Indian Government introduced a sector-wise minimum land area requirement for establishing a zone. SEZs in sectors other than IT, Biotech and health services have a minimum requirement of land area of 50 hectares, whereas for the latter there is no minimum land area requirement. Given that 67% of SEZs are in the IT sector, the distribution of the area is right-skewed with the median area being 19.55 hectares, mean area - 107.8 ha and standard deviation of 411.82 ha. All of the outliers are multi-product SEZs with

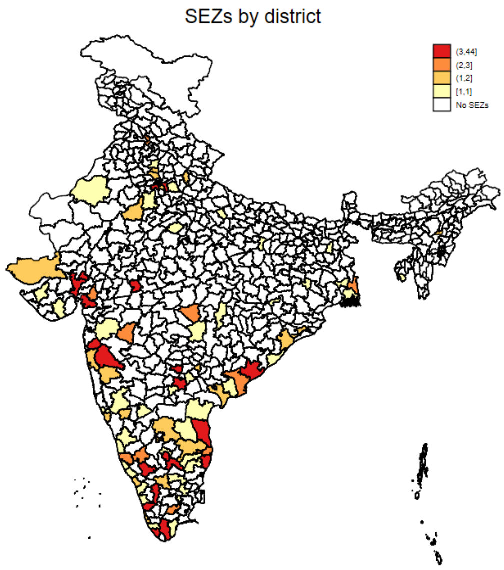

Figure 1: Number of SEZs by district.

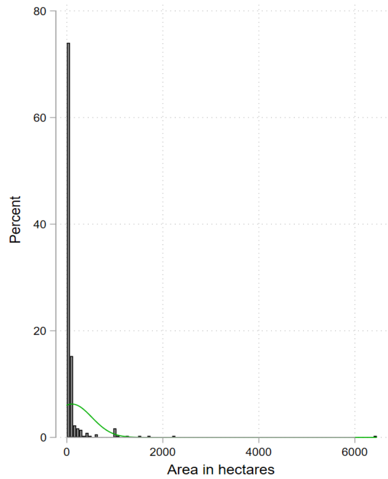

Figure 2: Histogram of the area and the radius of SEZs.

depicts the histogram of the area, while Table 1 presents summary statistics of the area by SEZ-sector, showing substantial differences in average area size across sectors. Additionally, Figure A.1 in the Appendix illustrates the distribution of SEZs by sector and time.

To summarise, comparing Indian SEZs to China, where SEZ have also been widely used, shows that India has several distinctive features. First, initial waves of China's SEZs targeted coastal regions with easy access to port and transportation networks, whereas in India there are no imposed restrictions on the geographic location of SEZs. Secondly, unlike China's SEZs which are large open territories covering whole cities and spanning over hundreds of thousands of hectares, SEZs in India are fenced-in zones, the smallest of which is one hectare. Thirdly, in India developers submit the proposal for establishing an

the largest being Adani Port and SEZs (6.456 ha) and Andhra Pradesh Industrial Infrastructural Corporation Ltd. (2.206 ha).

Table 1: Summary statistics of area by sector.

	Area in hectares				
	Mean	SD	Min	Max	N
Aviation	101.69	0.45	101.17	101.98	3
Biotech	18.97	9.79	10.00	40.47	15
Construction	106.46	.	106.46	106.46	1
Energy	76.49	84.15	10.00	222.67	6
Engineering	124.23	69.24	36.42	317.71	16
Food processing	48.20	44.63	11.88	119.14	7
Free Trade and Warehousing Zones	109.55	144.91	40.63	434.86	7
Gems and Jewellery	68.80	.	68.80	68.80	1
Handicrafts	10.49	.	10.49	10.49	1
IT	24.12	33.39	1.05	223.00	237
Minerals	119.86	41.17	50.75	166.91	6
Multi-product	1,165.60	1,355.21	105.44	6,456.33	20
Paper products	109.81	.	109.81	109.81	1
Pharmaceuticals	94.47	54.50	11.47	247.39	18
Port	224.57	98.90	110.47	285.84	3
Textile	133.69	107.25	20.41	404.70	12
Total	107.80	411.82	1.05	6,456.33	354

SEZ, which is then reviewed by state and central governments and finally approved by the Board of Approval. In China, on the contrary, the government assigns a particular area an SEZ status which consequently attracts foreign and domestic firms due to stimulating fiscal regulations (Zheng and Aggarwal, 2020).

3 Compiling the data set

One of the main challenges in assessing the impact of SEZs, particularly in developing countries, is the unavailability of data on firms operating inside SEZs. Therefore, in an effort to overcome this limitation, we merge firm-level data from Prowess with a list of notified SEZs obtained from the Ministry of Commerce and Industry using a spatial approach.

Firm-level data used in the analysis are obtained from Prowess, a database of financial performance of Indian companies, collected by the Centre for Monitoring the Indian Economy (CMIE). This data has been used extensively in other strands of research, see, e.g., Goldberg et al. (2010); De Loecker et al. (2016). Prowess includes relatively large firms and accounts for 60-70% of the economic activity in the industrial sector (Goldberg et al., 2010). Our data set is an unbalanced panel of firms covering the period from 1988 to 2020 and consists of 18.516 firms. The dataset provides information on the financial statements of firms, including sales, assets, raw materials, compensation to employees, exports, industry, and most importantly, the address of the registered office of the firm.[10]

[10] An advantage of using Prowess compared to the Annual Survey of Industries (ASI) is that the precise location information is available in Prowess, whereas only the administrative territorial unit such as state or district is reported in ASI. This allows us to focus on firms as treatment units and determine whether a firm is inside, in the vicinity, or far away from an SEZ. Using district level information in ASI would only allow to classify a firm as within a district that has an SEZ, resulting in potential aggregation bias. Another advantage is that Prowess contains firms operating in both the manufacturing and service sectors. The majority of firms (27.5%) operate in financial service activities, followed by wholesale and retail trade (15%) and chemicals (3%) as presented in Table A.3. However, because firms are under no legal obligation to report the data, only less

Since Prowess does not directly report information on the SEZ status of the firm, the address is used to identify the geographic coordinates of the firm.[11] The latitude and longitude of each firm, together with spatial rings of different radii around the centroid of SEZs, are plotted on a map using ArcGIS to identify the location of a firm in relation to SEZs.

The primary difficulty in pinning down the exact location of the SEZ is imprecise location information, which is available at the village level in the most disaggregated form. Thus, to pinpoint the accurate address of the SEZ, we manually identify the latitude and longitude of the zone using the name of the developer combined with the village and state names. Since SEZs are not points on a map but rather geographic zones, we use the information on the area of the SEZ to create spatial rings around the centroid of the zones. Because we do not know the actual boundaries of SEZs, we assume that they have a circular shape.[12]

We then create the first spatial ring using the information on the original radius of SEZs presented in Table 1. We assume that all firms within this circle are inside the SEZ and we therefore refer to this circle as *inside*. Subsequently, the original radius is increased by 5 kilometers to create the second spatial ring with the radius $5km + r$, where r is the original radius. This allows us to identify firms that are not in the SEZ but in the vicinity, not further than 5 km away from the SEZ. The aforementioned procedure is repeated to increase the radius by $10km + r$ and $15km + r$ as is shown in Figure A.2. We expect that firms in the vicinity of the SEZ may be affected by spillovers. Given that spillovers have been shown to be localized, they are expected to be most pronounced for firms within a 10 km band around the SEZ (Barrios et al., 2012; Baldwin et al., 2010). Our control group consists of firms more than 40km away from the centre of an SEZ.

To merge the area of SEZs with firm data, we plot the geocoded firms on a map. Firms that fall inside the created spatial rings are defined as treated. To avoid an additive effect, we exclude all firms falling inside the previous spatial ring from each subsequent ring, e.g. $10km + r$ contains firms that are located between $5km + r$ and $10km + r$. Hence, distance bands for *inside*, $0 - 5km$, $5 - 10km$ and $10 - 15km$ are formed. Table A.4 in the Appendix depicts the number of treated firms inside each distance band. We observe that only a small number of firms fall inside the original radius of SEZs. As we increase the radius and keeping in mind that SEZs tend to be spatially concentrated, the rings overlap and firms fall inside multiple rings, which is referred to as treatment intensity. In this case, a treatment year is assigned as the earliest year among all SEZs and one observation per firm

than 10% of firms (mostly public sector and large IT companies) disclose employment information, which makes Prowess unsuitable to analyze the labor market implications of SEZs. Prowess does provide wage bill information which is used later on for the TFP estimation.

[11] Geocoding is done using ArcGIS Online Geocoding Service. We exclude from the analysis the following states due to the small number of observations: Andaman and Nicobar, Arunachal Pradesh, Chandigarh, Goa, Nagaland and Manipur.

[12] This assumption will lead to some firms that are actually located in SEZs being classified as non-SEZ firms and vice versa. However, observing estimated effects that go into different directions for inside SEZ and firms in the 0-5 kilometers distance band give us confidence that the potential mis-classification does not seem to be too much of a threat.

is kept. To not further complicate the analysis, we omit a time aspect of SEZs opening, meaning that the presented numbers are time-invariant as of 2020.

An important point to be made is that the list of notified SEZs is used for the analysis. As described above, the establishment process consists of three stages: approval, notification and operational stage. Not all approved SEZs become eventually notified or operational. As of 2020, there are 421 formal approvals, 84% of which are notified and only 57% are operational. However, at the time of formal notification, investments and construction can begin, which may already affect the performance of firms. Following this reasoning, we chose the notification stage as our treatment. Moreover, we do not consider in our analysis SEZs notified prior to the enactment of the 2005 Act. These are 19 EPZs that were established before the SEZs policy and were converted into SEZs with the enforcement of the 2005 Act. Since the initial goal of EPZs was primarily to promote exports, whereas SEZs' focus is turned into developmental effects, the provided incentives may be different, which leads us to focus solely on SEZs notified under the SEZs Act.[13]

The analysis focuses on TFP growth as the main variable of interest. All variables are deflated using industry-specific Wholesale Price Index (WPI) for manufacturing firms and yearly WPI for service firms and transformed into logarithms. Total factor productivity is estimated using Ackerberg et al. (2015) approach (a detailed explanation on the estimation is presented in Appendix A.3). Other variables used as baseline controls include age, a foreign ownership dummy, dummy variables for manufacturing and service sectors measured in 2005 and time-invariant state dummies.[14] We classify a firm as foreign-owned if the percentage of equity shares held by foreign individuals, corporate bodies or institutions exceeds 25%. Time-varying covariates include, depending on the specification, total assets, sales, TFP and exporter dummy.

Table 2 presents summary statistics of treated firms in each respective distance band (Columns 1 through 4) and a fixed control group (Column 5). Panel A depicts the mean of the variables for all years, Panel B presents the summary statistics for pre-treatment 2005 year. Looking at the initial level of productivity for inside SEZ firms and the control group, we observe that the TFP level is higher for the treated group compared to the control group, which indicates that initially more productive firms self-selected into SEZs. Those firms are also more likely to export. However, mean pre-treatment assets and age are larger for the control group than for SEZ firms, which indicates that SEZ firms are initially smaller and younger. Overall, this pre-treatment mean comparison indicates a positive selection bias. To overcome this problem, we control for the pre-treatment variables so that the results can be interpreted accounting for this selection.

Further examining trends for the whole sample, we observe that TFP growth is smaller for the treated group than the control group. Sales and wage growth are greater for inside

[13]None of the firms in the control group falls inside SEZs established prior to the 2005 Act. We further provide a robustness check excluding firms in the treated group that are located in those converted SEZs.

[14]We do not observe the change in the registered address of the firms, hence the location information is time-invariant as of the latest financial report.

SEZ firms relative to the control group. Moreover, SEZ firms have a higher share of foreign ownership and operate mostly in the service sector compared to the control group. The share of manufacturing firms is also higher in the control group.

Looking at the firms in the distance bands, they also exhibit higher productivity levels but lower growth rates compared to the control group. This is accompanied by higher sales and wage growth for all firms except those in the 10-15km distance band. They are also less likely to be importers and manufacturing firms.

Table 2: Summary statistics of firms.

	(1) Inside Mean	(2) 0-5km Mean	(3) 5-10km Mean	(4) 10-15km Mean	(5) Control Mean
Panel A: Whole sample					
TFP	1.295	1.210	1.202	1.253	1.139
Exporter dummy	0.288	0.298	0.266	0.269	0.269
Log of assets	4.156	4.266	4.117	3.970	4.412
TFP growth	-0.019	-0.007	-0.007	-0.004	-0.009
Sales growth	0.026	0.036	0.028	0.012	0.019
Wage growth	0.056	0.069	0.061	0.046	0.053
Material expenses/sales growth	-0.006	-0.012	-0.013	-0.017	-0.013
Age	37.205	37.844	39.793	41.619	39.659
Foreign ownership	0.014	0.019	0.009	0.009	0.011
Manufacturing dummy	0.352	0.364	0.364	0.349	0.602
Services dummy	0.511	0.521	0.536	0.568	0.297
Importer dummy	0.299	0.305	0.278	0.269	0.333
Panel B: Pre-treatment variables in 2005					
TFP	1.288	1.231	1.229	1.281	1.077
Exporter dummy	0.265	0.267	0.215	0.228	0.211
Log of assets	3.498	3.554	3.375	3.251	3.712
TFP growth	-0.063	-0.023	0.016	0.047	0.030
Sales growth	-0.014	0.072	0.056	0.009	0.060
Wage growth	-0.014	0.032	0.021	0.009	0.005
Material expenses/sales growth	0.015	0.045	0.037	0.023	0.006
Age	36.000	37.007	39.169	40.612	38.943
Foreign ownership	0.009	0.022	0.010	0.008	0.013
Manufacturing dummy	0.341	0.330	0.329	0.310	0.576
Services dummy	0.538	0.560	0.577	0.603	0.325
Importer dummy	0.247	0.266	0.237	0.228	0.274

4 Evidence from event studies

Recall that the aim of the paper is to estimate the effects of the establishment of an SEZ on firm performance, distinguishing those located within the SEZ, and those in the vicinity. As pointed out above, the main purpose of establishing SEZs was to improve the economic development of the regions - and not, as e.g., in China, to boost exports. We, therefore, focus in our analysis on firm productivity and compare the productivity growth of firms before and after the establishment of SEZs relative to the firms that are not exposed to the program. We consider growth rather than levels as this differences out differences in productivity levels across firms. The key assumption then is that treated and control groups would have evolved in the same way in the absence of treatment, in other words, the conditional mean independence (CIA) assumption should be satisfied.

Before proceeding to a more formal econometric approach, we start by identifying a within-firm estimator using an event-study design. This illustrates the development of the variables of interest in the years preceding and following the establishment of SEZs for each firm i. Accounting for differential timing of treatment, the approach thus handles pre-trends and post-treatment dynamics.[15]

As suggested by De Chaisemartin and d'Haultfoeuille (2020), we employ a newly proposed estimator which estimates the average treatment effect in the groups that switch treatment at the time of switching. Thus, at each time a difference-in-differences is estimated based on groups that change their treatment status at time t relative to groups with stable treatment. Subsequently, these difference-in-differences are averaged over the whole observation period with weights depending on the number of switchers at each time. As a robustness check, we further present results using an alternative estimator proposed by Borusyak et al. (2021) and a two-way fixed effects specification.

We restrict the choice of the control group to a sub-sample of firms located further than 40 kilometers away from the zones to alleviate the concern that the control group is affected by the treatment. In the choice of the control group, we relied on two primary factors: the control group should not be affected by the treatment and it should be comparable to the treated group. We also provide a robustness check using two alternative control groups - firms located further than 30km and 45km away from the zones in Figures A.3 and A.4 in Appendix.[16]

The event window is restricted to 10 years before and after treatment. This requires assumptions about the nature of the effect outside of the window. The model which excludes all periods outside of the event window makes an implicit assumption that treatment effects drop to zero outside of the event window. Following the suggestion of Schmidheiny

[15]Recent econometric literature has raised concerns regarding the unbiasedness of the two-way fixed effects estimator in the presence of treatment heterogeneity and staggered treatment adoption (Borusyak et al., 2021; Callaway and Sant'Anna, 2021; De Chaisemartin and d'Haultfoeuille, 2020; Goodman-Bacon, 2021; Sun and Abraham, 2021). When the treatment effects are heterogeneous, the average treatment effect in the two-way fixed effects estimator is a weighted average of all the heterogeneous treatment effects. In the presence of both the variation in treatment timing and treatment heterogeneity, the average treatment effect is identified in part through the changes over time within already treated units. Consequently, some of the weights on the heterogeneous treatment effect underlying the average treatment effect can be negative, particularly for groups treated for many periods. Specifically, De Chaisemartin and d'Haultfoeuille (2020) show that even though all the average treatment effects are positive, the linear regression coefficient may be negative. This leads to difficulties in interpreting a two-way fixed effects estimator.

[16]Another possibility would be to form a control group consisting of firms that applied for SEZs but were rejected, in line with Kline and Moretti (2014) and Helmers and Overman (2017). Examining the minutes of the meeting of the Board of Approval reveals that the majority of applications are approved, and, when the required documents are lacking, "in-principal" approval is granted or the application is deferred until the developer is able to present the required clearances or satisfies the minimum land requirement. Thus, due to the limited number of observations, this approach is not possible here.

and Siegloch (2019), we bin the endpoints assuming constant treatment effects before and after the event. Binning introduces important parameter restrictions which ensure that the model is identified econometrically.

We estimate the following regression equation:

$$\Delta Y_{it}(db) = \alpha + \sum_{k \geq -10}^{10} \beta_k \times D_{it}^k(db) + \gamma X_{it} + \phi_i + \lambda_t + \epsilon_{it}, \tag{1}$$

where event dummies for the window $-10 \leq k \leq 10$ are created. $D_{it}^k(db)$ represents the SEZ program establishment event. $D_{it}^k(db) = 1$ if the observations' period of firm i at time t relative to the first period when firm i is treated by an SEZ equals the value of k for each distance band db. $D_{it}^k(db)$ is always 0 for never-treated firms. ΔY_{it} is the outcome variable defined as TFP growth for firm i at time t. ϕ_i are firm fixed effects that control for time-invariant differences between firms. λ_t represents year fixed effects that control for business cycle trends common across all firms in India. Standard errors are clustered at the firm level to account for the fact that observations within a firm correlate across time. X_{it} is a vector of controls including exporter dummy, manufacturing and services dummy and foreign ownership dummy which change over time. Since the aim of the analysis is to estimate potential spillovers from SEZs, Equation 1 is estimated separately for each distance band. The coefficient of interest, β_k, identifies the effect of SEZs program k years following its implementation.

To visualize the dynamic effects, the point estimates together with 95% confidence intervals are plotted in Figure 3. Importantly, looking at pre-treatment trends, we cannot reject the null hypothesis of no significant differences between treated and control groups prior to treatment. Looking at the post-treatment periods, we observe a significant decline in productivity growth for firms inside the zones right after the program implementation. The effect remains significant and negative also at later periods, e.g. 7 or 9 years after the event. However, there is no clear indication of spillovers; firms in the 5-10km distance band experience a negative productivity trend, but the coefficients are not significant at 5%.[17]

As discussed above, we also provide a robustness check using two-way fixed effects and Borusyak et al. (2021) estimators in Figure A.5 in the Appendix. The point estimates of a newly proposed estimator closely parallel two-way fixed effects results, with the latter having wider confidence intervals. The broader picture, however, remains the same.

Since the sample used for the analysis is unbalanced, it may create a concern that the attrition of firms is non-random. As an additional robustness check, we keep only those treated firms that are observed for consecutive ten years before and after the treatment for each distance band and re-estimate Equation 1 for the sample of balanced treated firms. We keep the control firms as before not to lose observations. Figure A.6 in the Appendix depicts the results. We still observe a significant productivity decline for inside

[17] Wider confidence intervals are observed in some graphs compared to others which is attributed to the smaller number of treated firms for particular distance bands as is reported in Table A.4.

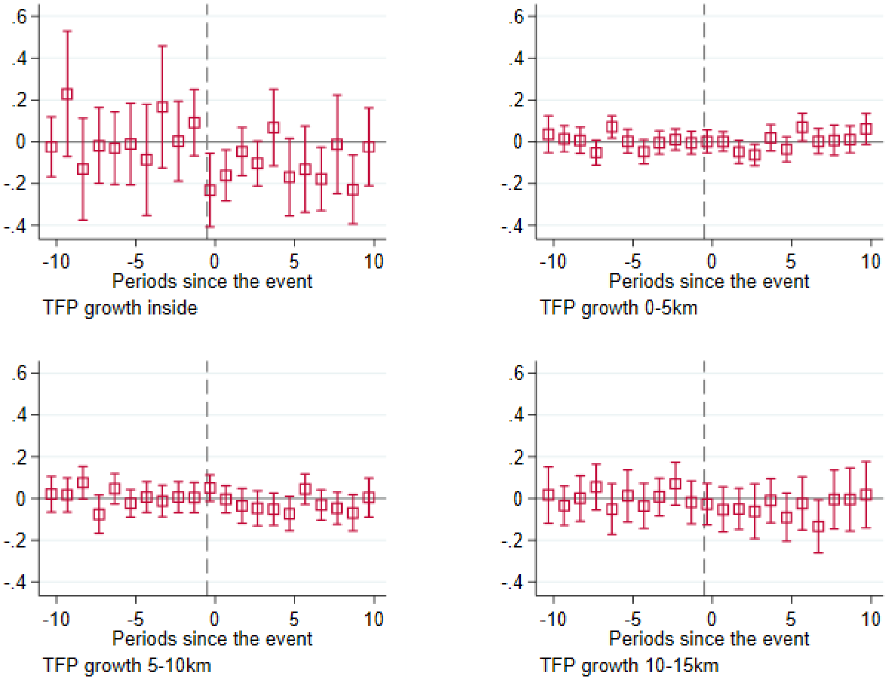

Figure 3: Event study graph for TFP growth using De Chaisemartin and d'Haultfoeuille (2020). 95% confidence interval is reported.

SEZ firms, which is more pronounced in the initial years following the event. Interestingly, the downward trend in TFP growth observed for firms in the 5-10km distance band now turns statistically significant and is also more pronounced for the initial 5 years after the event.[18]

While the identified within-firm estimates may bring us one step closer to a plausible treatment effect, the identification relies primarily on assumption that the event is exogenous. This may not necessarily hold given that firms self-select into treatment. To account for selection and ensure that treated and control groups are comparable, we now proceed with our preferred specification of determining the average treatment effect of SEZs on firm performance using a combined difference-in-differences and propensity-score re-weighting approach, where weights are created at each point in time.

5 Time-Varying Treatment Approach

5.1 Methodology

Standard propensity score methods applied to longitudinal data may be misleading when the treatment and the variable of interest are observed at multiple points in time (Girma and Görg, 2022). To illustrate, firm variables change over time depending on previous confounders, the treatment history, and the development of the outcome variable in the

[18]Due to a sharp decrease in the number of treated firms in a balanced 20-year sample, as a robustness check, we reduce the event window to five years before and after the treatment. Figure A.7 in the Appendix presents the results, which appear similar.

preceding periods (Thoemmes and Ong, 2016). Moreover, pre-treatment covariates used for deriving conditional probabilities vary over time in a way that is possibly influenced by previous outcome variables. Therefore, the longitudinal structure of panel data and the rolling introduction of the treatment make it difficult to use the standard inverse probability weighing technique, which may lead to biased estimates.

To overcome this issue, we follow a growing literature on time-varying treatments and calculate weights at each point in time (Fitzmaurice et al., 2008; Thoemmes and Ong, 2016; Girma and Görg, 2022). To illustrate, at the first treatment occurrence, we predict treatment assignment given the observed history of the covariates. At the next time point, a different set of weights is constructed that makes the treatment selection at time two orthogonal of all observed covariates prior to this treatment selection. Repeating this procedure for each year following the first treatment introduction results in a set of weights, which are eventually cross-multiplied to form a unique final weight for firm i at time t.[19]

Taking all together the stabilized weight is estimated as follows:

$$SW_{it}(db) = \prod_{t=1}^{T} \frac{Pr(SEZs_{it}(db) = 1 \mid X_i^0)}{Pr(SEZs_{it}(db) = 1 \mid \bar{X}_{it-1}, X_i^0)}, \tag{2}$$

where $SEZs_{it}(db)$ is an indicator for a post-SEZ period for firm i, time t and each distance band db. It is always zero for never-treated firms. X_i^0 are time-invariant covariates which include age, a foreign ownership dummy, dummies for manufacturing and service industries in 2005 and state dummies. $\bar{X}_{it-1}$ are time-varying covariates up until $t-1$, including the log of total assets, log of sales, exporter dummy and the history of the outcome variable. To incorporate information on the values of time-varying covariates before the start of the treatment, the value for 2006 is replaced by the mean value for 2004-2006. Thus, the stabilized weights are defined for each firm i at time t and each distance band db. Since $SEZs_{it}$ is distance band-specific, weighting is done for each distance band and fixed control group separately.

The intuitive interpretation is similar to standard propensity score methods. Firms that exhibit a high propensity to be treated and are ultimately treated are down-weighted in the pseudo-population because they are over-represented relative to the control group, which exhibits high treatment probability but is not treated. It is worth noting that propensity score weighting helps get the treatment independent of observable covariates but there may still be some unobserved factors inducing the selection decision, which cannot be controlled for.

The propensity scores are estimated using covariate balancing propensity score (CBPS) following Imai and Ratkovic (2014). This methodology makes use of the dual characteristics of the propensity score as covariate balancing score and the conditional likelihood of treatment assignment. Thus, CBPS models treatment probability while at the same

[19] For instance, if firm i is treated in 2006 and observed throughout the whole sample period up until 2020, a set of 14 weights will be estimated. Taking the product of all the weights results in a unique weight for firm i and time t.

time optimizing the covariate balance, which eliminates the need for separate covariance balancing checks.[20]

Once the weights are formed, they can be included in the final regression. The estimated weighted difference-in-differences regression equation takes the following form:

$$\Delta Y_{it}(db) = \alpha + \beta SEZs_{it}(db) + \theta X_i^0 + \lambda_t + \mu_s + \epsilon_{it} \tag{3}$$

where $\Delta Y_{it}(db)$ is the dependent variable (defined as productivity growth), $SEZs_{it}(db) = 1$ for SEZ firm in post-SEZ period and zero otherwise, X_i^0 include baseline controls such as age, a dummy variable for foreign ownership and dummies for manufacturing and service sectors in 2005. λ_t are year fixed effects that control for time trends common to all firms. μ_s are state fixed effects which absorb differences in the geographic location of the zones. Standard errors are clustered by firm to allow for within-firm correlation of the dependent variable over time.To eliminate the unobserved firm-specific effects, the variables are log differenced. β is the coefficient of interest that shows whether the expected change in the outcome from pre-SEZs to post-SEZs is different in the treated group relative to the control group. The regression is distance band-specific, e.g. weighted firms in each distance band are compared to a fixed set of control group firms further than 40 kilometers away.

To contrast the results with a more conventional approach, we also estimate a standard propensity score re-weighting based on a single pre-treatment year. First, we estimate nearest-neighbour propensity score matching as follows: $P(db) = Pr(D_i = 1|X_i^0)$, where $D_i = 1$ if firm i is in an SEZ for each distance band db and zero for never-treated firms. X_i^0 is a vector of pre-treatment covariates including age, dummies for manufacturing and service sectors, a foreign ownership dummy all measured in 2005, time-invariant state dummies, and mean of log of sales and mean of log of assets for 2004-2006. Once the probabilities are estimated, they are transformed into weights. The treatment group receives a weight of $\frac{1}{Pr(D_i=1|X_i^0)}$ and the control group is weighted by $\frac{1}{1-Pr(D_i=1|X_i^0)}$. The main difference is that this approach predicts treatment assignment based on a fixed set of pre-treatment 2005 covariates for firms treated in earlier or later periods equally. In contrast, we account for the development of time-varying variables up until the treatment takes place.

5.2 Results

The results of the time-varying treatment estimation approach are presented in this section. We start by estimating the direct and spillover effects of SEZs on TFP growth in Table 3. Results show that conditional on controlling for initial selection, the productivity growth of firms inside the zones decreased significantly by $(\exp^{-0.167} - 1) \times 100 = -15.4\%$ after the establishment of the zones compared to firms that are never treated, on average. There is no indication of significant spillovers for neighbouring firms in the 0-5km distance band. However, firms in the 5-10km distance band experience a significant decrease in their productivity growth. This result on spillovers is not robust in alternative specifications

[20]For illustrative purposes, we report some balancing tests in Table A.5 in the Appendix.

below, however.

Table 3: Time-varying treatment effect of SEZs on the growth rate of TFP.

	(inside) TFP growth	(0-5km) TFP growth	(5-10km) TFP growth	(10-15km) TFP growth
SEZs	-0.167**	0.0428	-0.146**	0.0116
	(0.0808)	(0.0947)	(0.0595)	(0.107)
Age 2005	-0.00644	-0.00325**	-0.00759**	0.00445***
	(0.00496)	(0.00145)	(0.00362)	(0.00170)
Service 2005	-0.112	-0.234**	-0.0358	-0.0542
	(0.140)	(0.113)	(0.144)	(0.156)
Manufacturing 2005	-0.0535	0.0365	-0.168	-0.0768
	(0.105)	(0.0543)	(0.111)	(0.143)
Foreign ownership 2005	0.0335	0.0143	0.0693	-0.0209
	(0.0415)	(0.0576)	(0.152)	(0.0924)
Constant	0.371	0.0713	0.774***	0.252
	(0.218)	(0.121)	(0.243)	(0.153)
N	1813	12808	8256	5404
Year Fixed Effects	Yes	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

Note: The outcome variable is measured as the log difference of TFP. TFP is measured using Ackerberg et al. (2015) approach. Time-varying covariates for creating the propensity scores include log of assets, log of sales, and the history of the outcome variable. Time-invariant covariates include age, a foreign ownership dummy, dummies for manufacturing and service industries in 2005 and state dummies. The weights are derived using CBPS and Equation 2. Standard errors are clustered at the firm level.

To make sure that our results are not due to the specific measure of TFP, we now look at three alternative proxies of firm level productivity. Firstly, we use the well-established TFP measurement due to Levinsohn and Petrin (2003). Table 4 reports the estimation results. There is a significant negative effect on the TFP growth for inside SEZ firms, whose productivity growth declined by 11.3%. Neighbouring firms, on the other hand, are not impacted by the establishment of SEZs.

Secondly, we consider total sales growth as an alternative outcome variable in Table 5. We can see that the sales growth of firms inside the zones decreased significantly by 25.2% compared to similar firms not exposed to treatment. Neighbouring firms are not influenced in any discernible (statistically significant) way. Thus, in line with the negative implications for TFP growth, the establishment of SEZs also affects negatively the growth rate of sales for firms inside the zones.

A third alternative measure is wage intensity growth measured as the ratio of total wage bill to sales, as changes in wages may also reflect changes in (labour) productivity (the same measure of the share of wages in total revenue is used by Ahsan and Mitra (2014)) . The results using the growth of wage intensity as a dependent variable are reported in Table 6. Similar to previous results, we find that firms inside the zones experience a significant

Table 4: Alternative TFP measure: Levinsohn and Petrin (2003) approach.

	(inside) TFP growth LP	(0-5km) TFP growth LP	(5-10km) TFP growth LP	(10-15km) TFP growth LP
SEZs	-0.120***	0.0817	0.262	-0.0637
	(0.0407)	(0.0750)	(0.206)	(0.0957)
Age 2005	-0.000853	-0.00230**	-0.0127**	0.00379**
	(0.00103)	(0.000907)	(0.00636)	(0.00157)
Service 2005	-0.141	-0.0529	-0.200	0.109
	(0.110)	(0.0649)	(0.136)	(0.165)
Manufacturing 2005	-0.0843	0.154***	-0.365**	0.0826
	(0.0736)	(0.0346)	(0.175)	(0.147)
Foreign ownership 2005	0.0232	0.0201	0.159	-0.112
	(0.0460)	(0.0481)	(0.128)	(0.147)
Constant	0.188	-0.0998	0.999***	0.101
	(0.111)	(0.101)	(0.253)	(0.148)
N	1813	12808	8256	5404
Year Fixed Effects	Yes	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes	Yes

Standard errors in parentheses
** $p < 0.05$, *** $p < 0.01$

Note: Time-varying covariates for creating the propensity scores include log of assets, log of sales and the history of the outcome variable. TFP is measured using Levinsohn and Petrin (2003) approach. Time-invariant covariates include age, a foreign ownership dummy, dummies for manufacturing and service industries in 2005 and state dummies. The weights are derived using CBPS and Equation 2. Standard errors are clustered at the firm level.

decrease in wage intensity growth relative to, ceteris paribus, matched never treated firms. This is consistent with firms passing negative TFP growth effects onto their workers in terms of reduced wages. As before, we do not find any statistically significant effects on neighbouring firms.

An advantage of using firm level data is, of course, that we can allow for heterogeneous effects of SEZs establishment across different firm types. Ownership, size, or industry of the firm can potentially affect a firm's ability to benefit from the incentives and infrastructure brought by the SEZ implementation. We therefore now check effect heterogeneity across these different categories. To do so, we proceed as follows. First, we calculate the weights for the whole sample using Equation 2. Second, we split the sample based on the following characteristics: (i) the dummy variable takes the value of one if the firm is foreign-owned and zero otherwise, (ii) the second dummy variable is equal to one for firms with log sales above the sample median in 2005 as a proxy for firm size and zero otherwise, (iii) separate dummy variables for firms operating in manufacturing or service sectors.[21] Then, Equation 3 is estimated for each sub-sample including weights.

Table 7 depicts the results. To save space, we focus only on TFP growth measured as in Table 3. It appears that domestic firms, large firms, and manufacturing firms are negatively affected by the SEZs program. As regards the impact on neighbouring firms,

[21] Though the percentage of firms with certain characteristics is different depending on the distance band, the example is provided for firms located *inside*. 1% of firms located inside the original SEZ radius are foreign-owned, 40% are small-size firms, 56% operate in the manufacturing sector and 33% in the service sector.

Table 5: Time-varying treatment effect of SEZs on the growth rate of sales.

	(inside) Sales growth	(0-5km) Sales growth	(5-10km) Sales growth	(10-15km) Sales growth
SEZs	-0.291**	-0.188	0.255	0.178
	(0.117)	(0.246)	(0.171)	(0.151)
Age 2005	-0.00516***	0.00190	-0.00819***	0.00272
	(0.00170)	(0.00359)	(0.00311)	(0.00170)
Service 2005	-0.0263	0.0728	-0.386	0.376***
	(0.185)	(0.0678)	(0.318)	(0.0626)
Manufacturing 2005	-0.128	0.159	-0.368	0.312***
	(0.0977)	(0.0907)	(0.364)	(0.113)
Foreign ownership 2005	0.388***	0.113	-0.0153	0.0680
	(0.0908)	(0.131)	(0.146)	(0.172)
Constant	-0.104	-0.0207	0.717**	-0.426**
	(0.119)	(0.288)	(0.344)	(0.203)
N	1969	13301	8570	5653
Year Fixed Effects	Yes	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes	Yes

Standard errors in parentheses
** $p < 0.05$, *** $p < 0.01$

Note: The outcome variable is measured as the log difference of sales. Time-varying covariates for creating the propensity scores include log of assets, TFP and the history of the outcome variable. Time-invariant covariates include age, a foreign ownership dummy, dummies for manufacturing and service industries in 2005 and state dummies. The weights are derived using CBPS and Equation 2. Standard errors are clustered at the firm level.

again no robust picture emerges.[22]

As we increase the radius of the spatial circles around SEZs, one firm may fall within multiple SEZs because they are located close to each other. Such firms may experience differential effects as they are able to absorb the benefits of (or be hurt by) more than one SEZ. To check this, we now allow for such differential *treatment intensity*. We define a new variable "SEZ intensity" that counts the number of other than own SEZs within a 15 kilometer radius. If there is no other SEZ in the vicinity, then "SEZ intensity" is zero.[23] We re-estimate our baseline specification adding this additional explanatory variable which captures the average additional effect from each of the other SEZs.

Results are presented in Table 8. Whereas the negative effects from the establishment

[22]In another way of looking at heterogeneity, we estimate the effects for solely Hi-tech SEZs, given the predominant number of IT and other technologically intensive SEZs. Thus, we keep only SEZs that operate in IT, electronics, biotechnology, pharmaceuticals, engineering, aviation, and aerospace industry and non-conventional energy sectors and estimate our baseline specification. Results in Table A.6 in the Appendix indicate that conditional on controlling for initial selection, there are no further productivity gains from the establishment of Hi-tech SEZs for directly affected firms. Because a significant negative effect for inside SEZ firms disappears, it indicates that Hi-tech SEZs are, on average, more productive compared to other zones. Furthermore, for firms in the 10-15km distance band, we indeed find a significant positive effect on TFP growth. Thus, the evidence points to heterogeneous effects of SEZs with Hi-tech zones being relatively more productive compared to zones specializing in other industries.

[23]We limit the distance to 15 kilometers following the argument advanced in the literature that spillovers tend to be very localized in space.

Table 6: Time-varying treatment effect of SEZs on wage intensity growth.

	(inside) Wage intensity growth	(0-5km) Wage intensity growth	(5-10km) Wage intensity growth	(10-15km) Wage intensity growth
SEZs	-0.227**	-0.00466	-0.407	-0.125
	(0.0932)	(0.0467)	(0.255)	(0.0885)
Age 2005	0.00415	0.00328	0.00842***	-0.00325**
	(0.00252)	(0.00227)	(0.00239)	(0.00149)
Service 2005	0.0860	-0.0380	-0.0629	-0.353***
	(0.182)	(0.0660)	(0.191)	(0.124)
Manufacturing 2005	-0.0589	0.00906	-0.148	-0.433***
	(0.0885)	(0.0607)	(0.235)	(0.116)
Foreign ownership 2005	-0.180	-0.0289	0.0709	-0.304
	(0.118)	(0.0763)	(0.103)	(0.169)
Constant	0.0151	-0.0814	0.00379	0.520***
	(0.142)	(0.129)	(0.226)	(0.135)
N	1960	13247	8546	5618
Year Fixed Effects	Yes	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

Note: The outcome variable is measured as the log difference of wages. Time-varying covariates for creating the propensity scores include log of assets, TFP and the history of the outcome variable. Time-invariant covariates include age, a foreign ownership dummy, dummies for manufacturing and service industries in 2005 and state dummies. The weights are derived using CBPS and Equation 2. Standard errors are clustered at the firm level.

of SEZs remain, there is no indication of additional negative effects from other than own SEZs in the vicinity for firms located inside the zones. Interestingly, there is some evidence for positive spillovers for firms in the 0-5km distance band if there is more than one SEZ in the vicinity. Though the coefficient is small in magnitude, this can be interpreted as an indication of the potential benefits of SEZs if they form a cluster and congest in one region rather than being spread across the country.

5.3 Further robustness checks

In order to further strengthen the robustness of our results thus far, we perform a number of additional checks: (i) use alternative distance bands, (ii) exclude firms in SEZs established prior to the 2005 SEZs Act and (iii) use more conventional nearest-neighbour propensity score matching.

To check that the estimation results are not sensitive to the somewhat arbitrary chosen width of the distance bands, we provide results using alternatively defined distance bands. As argued by Rosenthal and Strange (2003), any arbitrary number of concentric rings can be used to define the distance bands. Thus, we increase the radius to 7 kilometers instead of 5 kilometers and construct the following distance bands: *inside*, $0 - 7km$, $7 - 14km$, and $14 - 21km$ with the control group being firms located further than 40 kilometers away. Table A.7 illustrates the number of treated firms inside each distance band. The results presented in Table A.8 in the Appendix are robust to an alternative definition of distance bands.

In our analysis, we do not consider EPZs established prior to the SEZs Act and later converted to SEZs with the enactment of the 2005 Act. However, firms that are located in

Table 7: Heterogeneous effects of SEZs on TFP growth.

	(inside) TFP growth	(0-5km) TFP growth	(5-10km) TFP growth	(10-15km) TFP growth
Panel A: Ownership type				
SEZs - foreign-owned private firms	-0.0845	0.0290	-0.0454	0.0374
	(0.122)	(0.0327)	(0.0864)	(0.0725)
SEZs - other domestic firms	-0.170**	0.0427	-0.147**	0.00633
	(0.0822)	(0.0971)	(0.0596)	(0.110)
Panel B: Firm size				
SEZs - large-size firms	-0.192**	-0.0224	-0.145**	0.0988
	(0.0909)	(0.0615)	(0.0615)	(0.170)
SEZs - small-size firms	0.0236	0.529	0.0406	-0.451
	(0.0720)	(0.288)	(0.179)	(0.248)
Panel C: Sector				
SEZs - manufacturing	-0.194**	-0.00221	-0.132	-0.0399
	(0.0974)	(0.0325)	(0.0840)	(0.0503)
SEZs - services	-0.437	-0.233**	-0.152	-0.0427
	(0.302)	(0.109)	(0.117)	(0.156)

Standard errors in parentheses

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

Note: Time-varying covariates for creating the propensity scores include log of asset, log of sales and the history of the outcome variable. Time-invariant covariates include age, a foreign ownership dummy, dummies for manufacturing and service industries in 2005 and state dummies. The weights are derived using CBPS and Equation 2. Standard errors are clustered at the firm level. After the weights are estimated, the Equation 3 is re-estimated for different samples. Large-size firms are firms with log of sales above the sample median in 2005. All specifications control for year and state fixed effects.

converted SEZs and appear in our treated or control groups may be affected by a different type of incentives and therefore may lead to biased results. Geocoding EPZs and identifying firms located inside them shows that none of the firms in the control group is located in converted SEZs. However, some treated firms are indeed located in those converted SEZs. In Table A.9 in the Appendix we exclude those firms and observe that the results are not affected.

In order to contrast our main results with a more conventional PSM approach, we also present estimation results using nearest-neighbour matching, where the probability of being treated is derived based on a single set of pre-treatment 2005 covariates for all firms equally. Table A.10 in the Appendix shows that the coefficient for inside is not statistically significant, which may suggest that the results are overestimated when we do not control for selection bias using the longitudinal structure of the data.

5.4 Potential explanations

To sum up, our results thus far indicate that, conditional on controlling for initial selection, the evidence on spillovers on neighbouring firms appears not very robust and depends on the specification taken. By contrast, the establishment of SEZs resulted in a significant direct productivity growth decline for firms located inside the zones, on average. How can we explain such a negative direct effect?

Alkon (2018) demonstrates that SEZs in India did not bring local socioeconomic development and argues that the mechanisms underlying the inefficiency of SEZs may be exces-

Table 8: Treatment intensity.

	(inside) TFP growth	(0-5km) TFP growth	(5-10km) TFP growth	(10-15km) TFP growth
SEZs	-0.205***	-0.0357	-0.268**	-0.0934
	(0.0682)	(0.102)	(0.116)	(0.132)
SEZs intensity	0.00400	0.0383**	0.102	0.0718
	(0.00461)	(0.0168)	(0.0744)	(0.0439)
Age 2005	-0.00637	-0.00247**	-0.00766**	0.00406**
	(0.00502)	(0.00125)	(0.00364)	(0.00180)
Service 2005	-0.108	-0.272***	0.00509	-0.0629
	(0.138)	(0.103)	(0.133)	(0.155)
Manufacturing 2005	-0.0508	-0.0452	-0.112	-0.0756
	(0.103)	(0.0538)	(0.102)	(0.144)
Foreign ownership 2005	0.0267	-0.00267	0.0504	-0.0404
	(0.0434)	(0.0569)	(0.138)	(0.0845)
Constant	0.373	0.171	0.734***	0.288
	(0.216)	(0.120)	(0.228)	(0.154)
N	1813	12808	8256	5404
Year Fixed Effects	Yes	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

Note: SEZ intensity counts the number of other than own SEZs within 15 kilometers. Time-varying covariates for creating the propensity scores include log of assets, log of sales, and the history of the outcome variable. Time-invariant covariates include age, a foreign ownership dummy, dummies for manufacturing and service industries in 2005 and state dummies. The weights are derived using CBPS and Equation 2. Standard errors are clustered at the firm level.

sive governmental involvement and rent-seeking. The presence of state-owned industrial development corporations is one of the key drivers of SEZ location. These state development corporations facilitate land acquisition; however, government intervention may fail to account for market conditions, infrastructure, labor availability, and other necessary inputs, thus making SEZs large developer's projects with little productivity gains. To investigate this hypothesis, we exclude 35 state-owned industrial development corporations listed on the website of the Council of State Industrial Development and Investment Corporations of India from the list of notified SEZs.[24] Then, the empirical model is re-estimated and the results are reported in Table 9. We observe that the significant negative effect for inside SEZ firms disappears, suggesting that firms in SEZs located or being developed by state-owned industrial development corporations may be less productive compared to firms in privately developed zones. There is no indication of any significant effect on neighbouring firms. Thus, these results are in line with the hypothesis of Alkon (2018) that excessive governmental intervention in the development of SEZs brings in practice little productivity

[24]The observations are dropped by the name of the developer of SEZs since the information on the ownership is not directly reported.

gains.

Table 9: SEZs effect on TFP growth excluding the zones located in state-owned industrial development corporations.

	(inside) TFP growth	(0-5km) TFP growth	(5-10km) TFP growth	(10-15km) TFP growth
SEZs	-0.0789	0.00830	0.0218	-0.0587
	(0.0489)	(0.0682)	(0.0614)	(0.0558)
Age 2005	-0.00292	-0.00583***	-0.0157***	0.000831
	(0.00184)	(0.00139)	(0.00283)	(0.00112)
Service 2005	-0.138	-0.209**	-0.218	-0.163**
	(0.187)	(0.0815)	(0.175)	(0.0733)
Manufacturing 2005	-0.0857	-0.0918**	-0.337	-0.111
	(0.104)	(0.0419)	(0.199)	(0.0633)
Foreign ownership 2005	0.123***	-0.0959	0.223	-0.116
	(0.0460)	(0.0678)	(0.179)	(0.0933)
Constant	0.169	0.423***	0.987***	0.123
	(0.147)	(0.0653)	(0.282)	(0.108)
N	1619	11457	7716	5199
Year Fixed Effects	Yes	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

Note: Time-varying covariates for creating the propensity scores include log of sales, assets growth and the history of the outcome variable. Time-invariant covariates include age, a foreign ownership dummy, dummies for manufacturing and service industries in 2005 and state dummies. 35 SEZs located in state-owned industrial development corporations are excluded. The weights are derived using CBPS and Equation 2. Standard errors are clustered at the firm level.

In addition to government interference, another potential explanation is related to rent-seeking and may be attributed to one of the peculiarities of the Indian SEZ policy: The possibility of single-firm zones. In India, firms must apply for permission to develop and subsequently operate in an SEZ. Thus, if a firm decides that benefits provided by the SEZ program are higher than potential costs (e.g. expenses on infrastructure development in the processing area or social amenities in the non-processing area), it may file an application and if approved, customs boundaries are redrawn around the existing location. There is no locational component in the program and therefore even a single firm may have an SEZ status. Arguably, the possibility of a single-firm SEZ may exacerbate the potential for rent-seeking. Blending the role of SEZ developer and firm director in one person may give particular incentives in the interaction with policy makers deciding on the SEZ status. In particular, the developer-cum-director may have an incentive for rent extraction even at the disadvantage of company performance. Thus, the director may strive to maximise personal income rather than firm performance. This is less likely to be an issue in a "regular" SEZ where a developer, who interacts with local policy makers, is distinct from a number of firm managers / directors representing firm interests.

Unfortunately, we do not have information on whether an SEZ is a single or multi-firm entity. We do, however, have information on the area size of the SEZ and it is a likely assumption that one firm SEZs are on average smaller than multi-firm zones. In order to look into this issue, we estimate the effect of SEZs on productivity keeping only SEZs that have size above the mean SEZ area. Results depicted in Table 10 indicate that firms located in relatively large SEZs increased their TFP growth by more than 40 percent. This result, in combination with previous findings, suggests that the distortions leading to negative direct productivity effects come mainly from the large number of small zones, which likely include many single firm SEZs. However, if SEZs are large territories hosting many units, as is the case in China, potential benefits, likely also reflecting agglomeration effects, are higher. Interestingly, we now also find positive spillover effects on neighbouring firms in the 5-10km band.

Table 10: Heterogeneity effect from relatively big SEZs.

	(inside) TFP growth	(0-5km) TFP growth	(5-10km) TFP growth	(10-15km) TFP growth
SEZs	0.416**	0.0481	0.323***	-0.195
	(0.166)	(0.0780)	(0.0186)	(0.147)
Age 2005	0.00208	-0.00648	-0.00234	0.000599
	(0.00157)	(0.00369)	(0.00130)	(0.000336)
Service 2005	-0.246	-0.366***	-0.00608	-0.145**
	(0.159)	(0.124)	(0.0615)	(0.0670)
Manufacturing 2005	-0.0861	-0.202***	0.0811	-0.0248
	(0.0584)	(0.0719)	(0.0567)	(0.0389)
Foreign ownership 2005	0	0.199***	0	0
	(.)	(0.0746)	(.)	(.)
Constant	-0.0170	0.423***	0.00305	0.0498
	(0.0592)	(0.0867)	(0.0852)	(0.0961)
N	472	3349	2622	2954
Year Fixed Effects	Yes	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

Note: We keep only SEZs with above mean area. Time-varying covariates for creating the propensity scores include log of asset, log of sales and the history of the outcome variable. Time-invariant covariates include age, a foreign ownership dummy, dummies for manufacturing and service industries in 2005 and state dummies. The weights are derived using CBPS and Equation 2. Standard errors are clustered at the firm level.

Our data base provides detailed information on directors' pay, which we can use to probe a bit further into this potential misallocation effect. Is it the case that directors' increase their salaries even if, as we have shown, average productivity effects are negative for firms in SEZs? To investigate this, we use the growth of total remuneration to a company's directors as the outcome variable in Table 11.[25] Results are in line with our hypothesis.

[25] Unfortunately, the reporting of directors' remuneration is not obligatory, which leads to a

The total remuneration of directors in SEZ-firms increased, on average, significantly by around 15%. At the same time, the compensation of directors in the neighbouring firms did not experience any significant change. In an extension to this analysis, we find that this effect does not hold when only considering large SEZs (as in Table 10), suggesting that it is driven by SEZs of small scale, including single firm SEZs. Results for this are available on request.

Table 11: Time-varying treatment effect on directors' remuneration growth.

	(inside) Dir remuneration	(0-5km) Dir remuneration	(5-10km) Dir remuneration	(10-15km) Dir remuneration
SEZs	0.152**	-0.0373	0.132	-0.298**
	(0.0572)	(0.182)	(0.0682)	(0.123)
N	293	3438	2092	1712
Panel A: Share of executive directors less than half				
SEZs	0.150**	-0.0620	0.140	-0.376***
	(0.0703)	(0.174)	(0.0755)	(0.119)
N	243	2885	1720	1467
Panel B: Share of independent directors greater than half				
SEZs	0.459	0.223	0.198	-0.311
	(0.235)	(0.273)	(0.111)	(0.179)
N	117	1587	1045	806

Standard errors in parentheses
** $p < 0.05$, *** $p < 0.01$

Note: Time-varying covariates for creating the propensity scores include log of asset, log of sales and the history of the outcome variable. Time-invariant covariates include age, a foreign ownership dummy, dummies for manufacturing and service industries in 2005 and state dummies. The weights are derived using CBPS and Equation 2. Standard errors are clustered at the firm level. All regressions include year and state fixed effects.

In panels A and B, we try to understand which type of directors are the ones benefiting. Panel A tells us that firms, where the share of non-executive directors is above one-half, are the ones experiencing significant increases in total remuneration. This is not surprising as non-executive directors play a crucial role in the formulation of the company's policies, whereas executive directors manage the routine of the company and may be considered more akin to high-skilled employees. Hence, their salary directly depends on the performance of firms. If, however, non-executives received financial motivation to attract the company into a single-firm SEZ program, their total compensation will increase without necessarily increasing the performance of firms. However, whether there are more independent directors in the company does not seem to play any role.[26]

6 Conclusion

Incentives brought by the SEZs Act pose a natural question of whether firms directly affected by the program experience significant improvements in their performance. Moreover, we evaluate whether these clusters of economic activity create spillovers to the neighbouring

substantial drop in observations. To make sure that the smaller sample does not introduce a bias, we re-estimate the TFP growth model in Table 3 also on this smaller sample used here. We still find statistically significant negative direct effects on TFP growth for firms inside SEZs. Results are available on request.

[26]Independent directors provide external guidance and are appointed for a maximum term of five years by the Board of Directors.

firms as a result of agglomeration economies. Our analysis is based on a unique representative geocoded dataset of firms and their assigned SEZ status covering all of India.

Our findings demonstrate that conditional on controlling for initial selection, India's SEZs program induced a significant drop in productivity growth for firms located inside the zones, on average. The result is robust to different specifications. When looking at possible spillovers to neighbouring firms, we do not find any robust evidence for such externalities to firms within 15km of the SEZ.

These findings go against what policy makers generally expect, and they also differ from results found for SEZs in China. They do, however, mirror earlier findings by Alkon (2018) based on a more aggregate analysis. In an attempt to explain this, we focus on the possibility of distortions through political interference and rent-seeking, aided by the fact that India allows for single-firm SEZs - a marked distinction to policy in China and other countries.

In this regard, our findings show that the negative direct effect on firms inside SEZ does not hold when only considering SEZ operated by non-state entities, suggesting that state interference may be an issue. Furthermore, we show that rent-seeking on the part of companies' directors may contribute to the negative productivity effects, as we find that directors' pay growth increases as a result of establishing an SEZ. We also show that these results are particularly pertinent in small SEZs. By contrast, our findings suggest that the productivity effect of SEZs on firms is positive in relatively large, i.e. above mean area, SEZs. Here, we find a strong positive and sizable in magnitude productivity growth increase for inside SEZ firms. These results are in line with the idea that the inefficiency of the program may be driven by single-firm SEZs, which may make political interference and rent-seeking more likely than in a large SEZ with multiple firms.

From a policy perspective, our findings suggest that the design of the SEZ policy in India may need to be re-evaluated. In particular, the possibility of single-firm SEZs may provide too much potential for political interference and rent-seeking, thus stifling otherwise positive development effects of SEZs.

Due to the growing popularity of SEZs as policy tools in developing countries, further efforts should be carried out to analyze the effectiveness of the program in India as well as other countries. For example, our focus is on the productivity growth of firms, leaving aside considerations about labour market implications at the firm level. Only with good data and adequate identification strategies can one provide constructive advice for policymakers on the local developmental implications of the program.

A Appendix

A.1 Figures

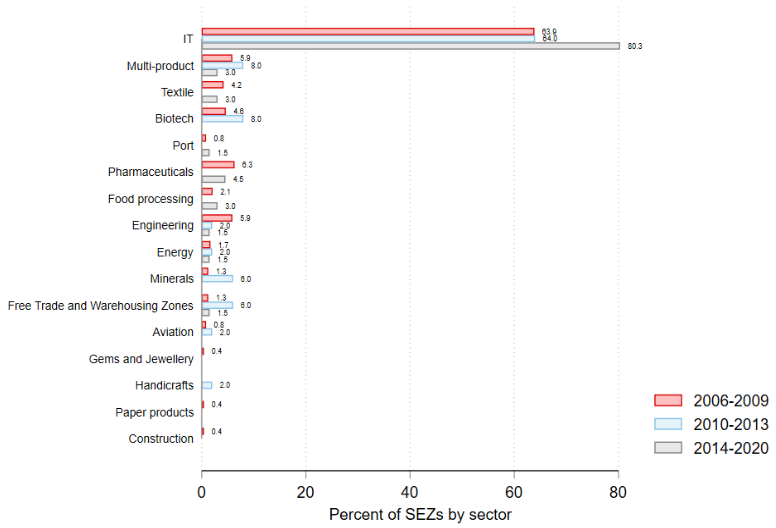

Figure A.1: Sector-wise distribution of SEZs over time.

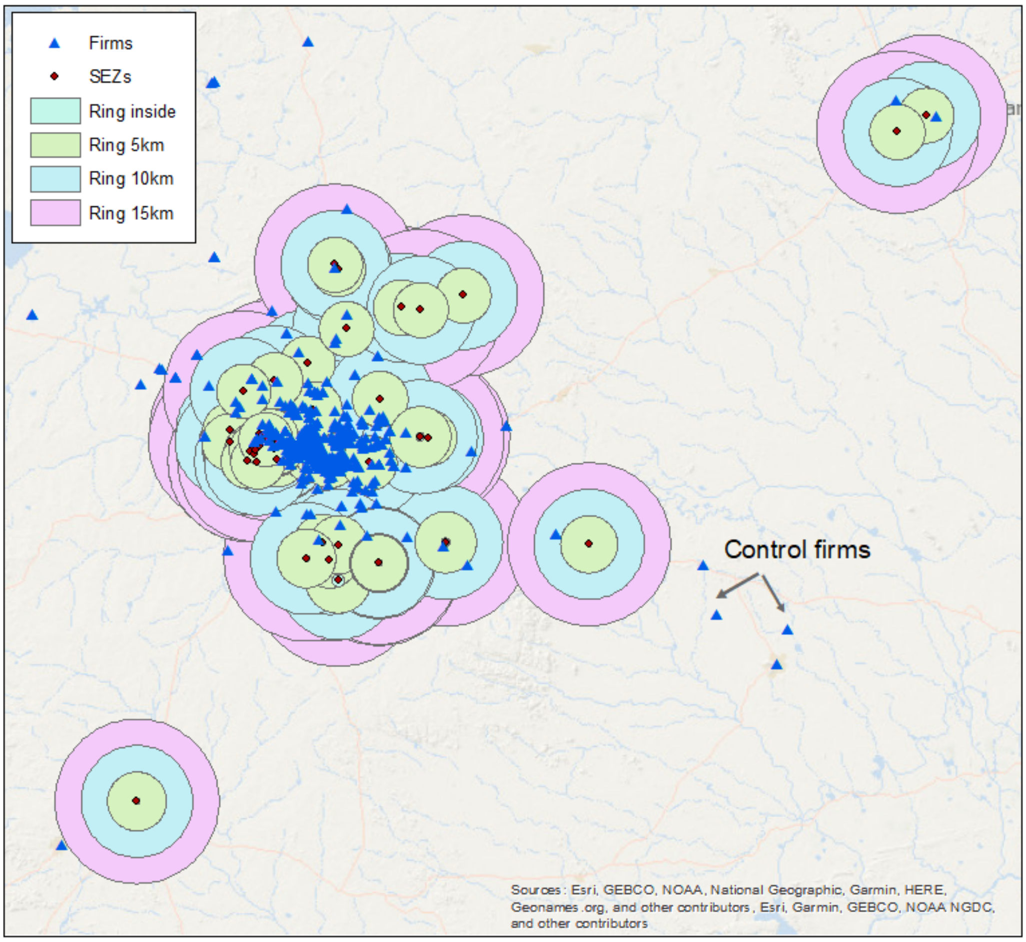

Figure A.2: The mapping of firms in the SEZs. The blue triangles represent geocoded firms. The red dots represent geocoded SEZs. Using the information on the zones' area, a radius is created and subsequently increased by 5 km. Buffers of various sizes are created around the centroid of SEZs using ArcGIS.

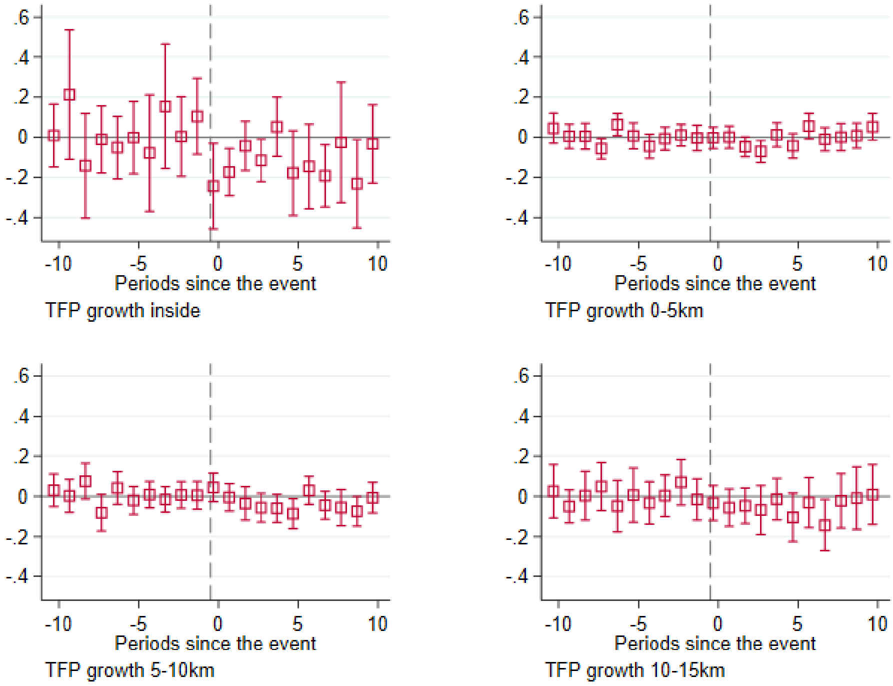

Figure A.3: Robustness check using De Chaisemartin and d'Haultfoeuille (2020) estimator and firm further than 30 km away as a control group. 95% confidence interval is reported.

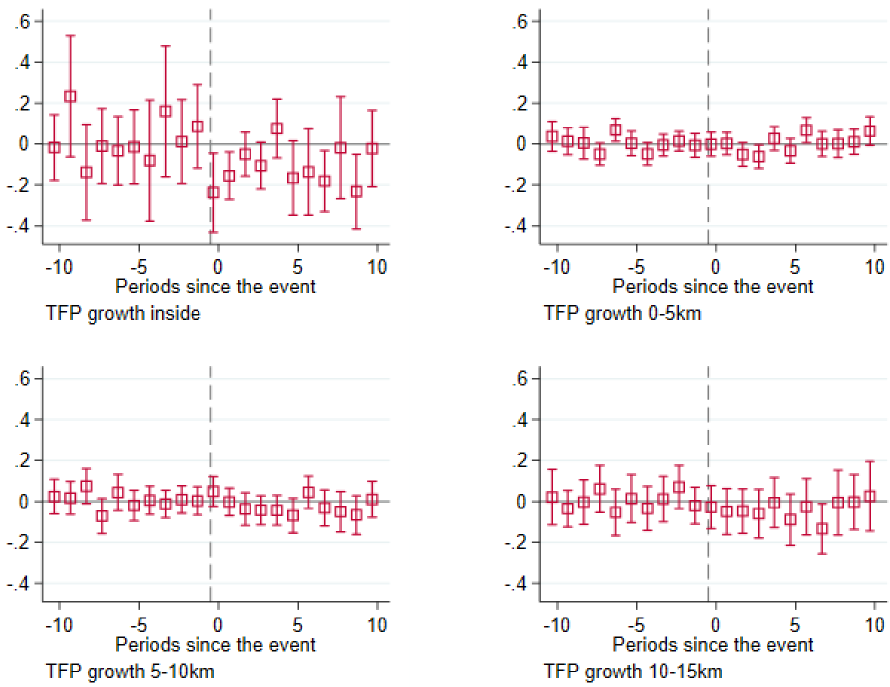

Figure A.4: Robustness check using De Chaisemartin and d'Haultfoeuille (2020) estimator and firm further than 45 km away as a control group. 95% confidence interval is reported.

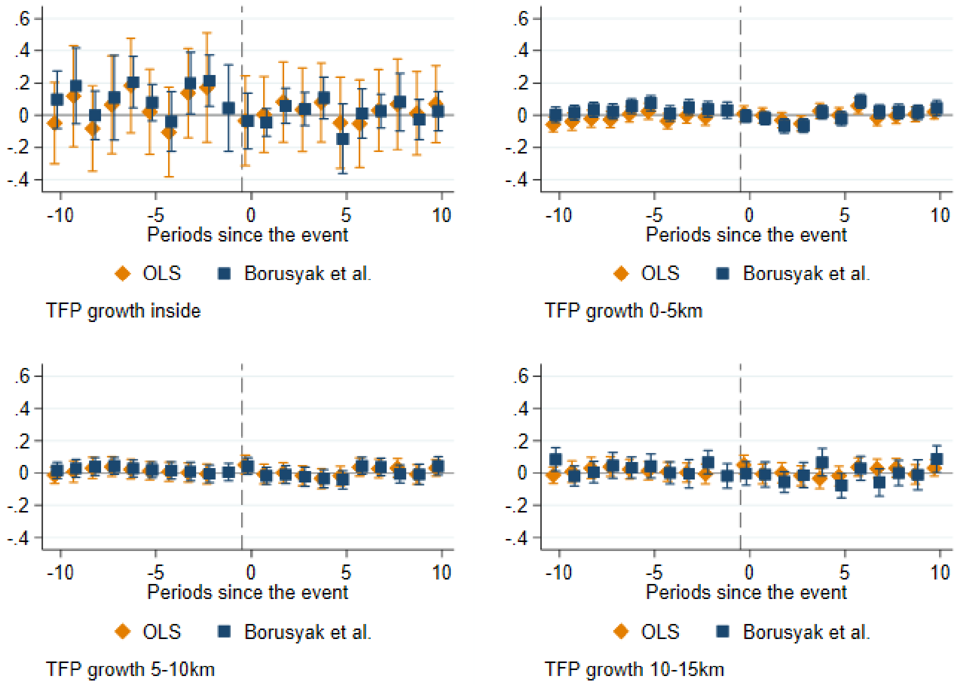

Figure A.5: Robustness check using two-way fixed effects and Borusyak et al. (2021). 95% confidence interval is reported.

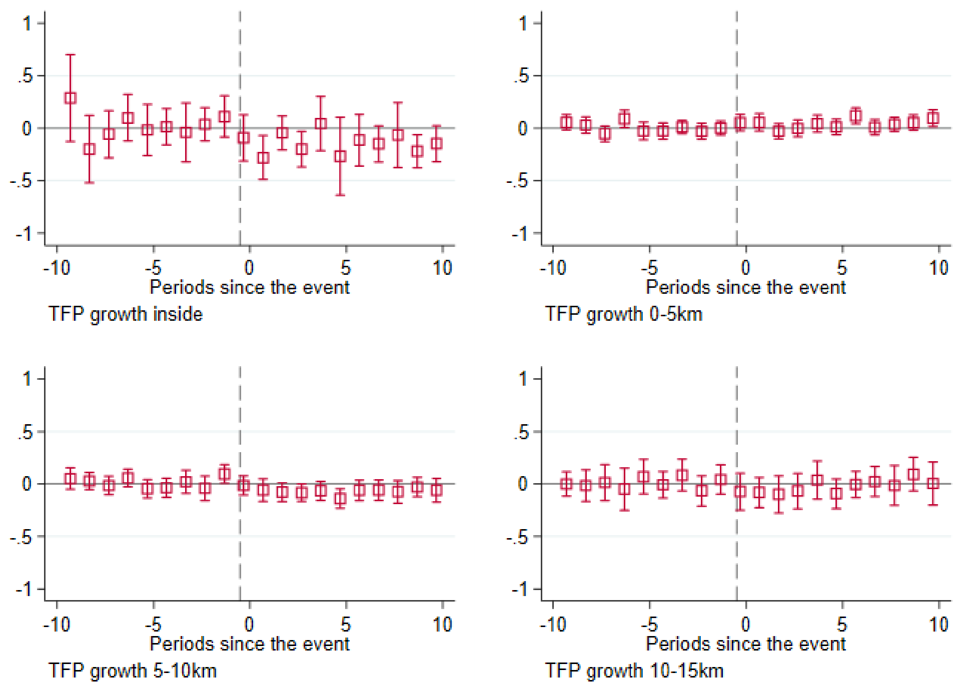

Figure A.6: Event study graph for TFP growth. 95% confidence interval is reported. The sample of treated firms is balanced for the event window [-10, 10].

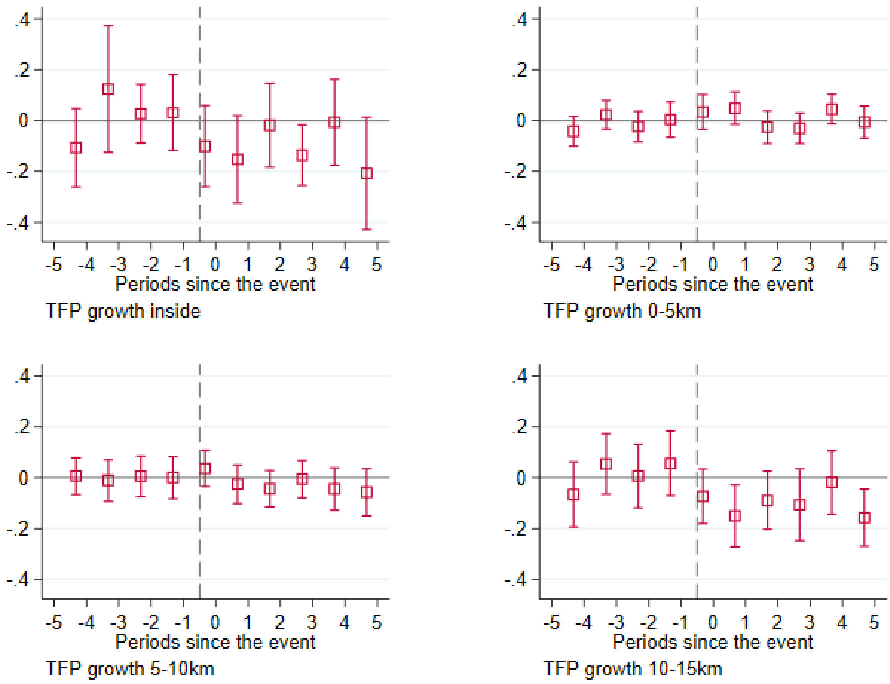

Figure A.7: Event study graph for TFP growth. 95% confidence interval is reported. The sample of treated firms is balanced for the event window [-5, 5].

A.2 Tables

Table A.1: Comparison of SEZs notified under the 2005 Act and converted SEZs established prior to the 2005 Act.

	(1) SEZs					(2) SEZs established before 2005 Act				
	N	Mean	SD	Min	Max	N	Mean	SD	Min	Max
Notification date	354	2009.69	3.92	2006	2020	19	2002.89	1.70	2000	2005
Area in ha	354	107.80	411.82	1.05	6456.33	19	150.58	239.12	2.02	1052.18
Radius in meters	354	396.68	431.65	57.82	4533.34	19	576.64	393.63	80.19	1830.08
Commencement of operation						19	1996.05	12.19	1965	2006

Table A.2: The establishment of SEZs over the 2006-2020 period.

	Frequency	Percent	Cum. percent
2006	54	15.25	15.25
2007	89	25.14	40.40
2008	50	14.12	54.52
2009	45	12.71	67.23
2010	20	5.65	72.88
2011	14	3.95	76.84
2012	6	1.69	78.53
2013	10	2.82	81.36
2014	5	1.41	82.77
2015	3	0.85	83.62
2016	11	3.11	86.72
2017	30	8.47	95.20
2018	4	1.13	96.33
2019	9	2.54	98.87
2020	4	1.13	100.00
Total	354	100.00	

Table A.3: Summary statistics of firms by industry. Pooled sample for 1988-2020.

	Frequency	Percent	Cum. percent
Crop & animal production	5795	2.080	2.080
Forestry & logging	6550	2.351	4.431
Fishing & aquaculture	83	0.0298	4.461
Mining of coal & lignite	485	0.174	4.635
Extraction of crude petroleum & natural gas	284	0.102	4.737
Mining of metal ores	505	0.181	4.919
Other Mining & quarrying	1937	0.695	5.614
Food	11479	4.121	9.735
Beverages	2176	0.781	10.52
Tobacco	322	0.116	10.63
Textiles	10792	3.874	14.51
Wearing apparel	646	0.232	14.74
Leather	1040	0.373	15.11
Wood	762	0.274	15.38
Paper	3309	1.188	16.57
Printing & reproduction of recorded media	191	0.0686	16.64
Coke & refined petroleum products	1182	0.424	17.06
Chemicals	14919	5.355	22.42
Pharmaceuticals	5150	1.849	24.27
Rubber & plastics products	7548	2.710	26.98
Other non-metallic mineral	4447	1.596	28.57
Basic metals	10423	3.742	32.32
Fabricated metal products	3849	1.382	33.70
Computer, electronic & optical products	4330	1.554	35.25
Electrical equipment	6139	2.204	37.46
Machinery & equipment	7518	2.699	40.15
Motor vehicles	4001	1.436	41.59
Other transport equipment	1018	0.365	41.96
Furniture	191	0.0686	42.02
Other manufacturing	8977	3.222	45.25
Electricity, gas etc. supply	2375	0.853	46.10
Water collection	63	0.0226	46.12
Construction of buildings	6040	2.168	48.29
Civil engineering	4675	1.678	49.97
Specialized construction activities	327	0.117	50.09
Wholesale & retail trade & repair of motor-vehicles and motorcycles	1114	0.400	50.49
Other wholesale & retail trade	33816	12.14	62.63
Retail trade	1325	0.476	63.10
Land transport & transport via pipelines	988	0.355	63.46
Water transport	617	0.221	63.68
Air transport	353	0.127	63.80
Warehousing & support activities for transportation	2581	0.927	64.73
Postal and courier activities	182	0.0653	64.80
Accommodation	3554	1.276	66.07
Food and beverage service activities	21	0.00754	66.08
Publishing activities	965	0.346	66.43
Music publishing activities	1046	0.375	66.80
Programming and broadcasting activities	65	0.0233	66.82
Telecommunications	1481	0.532	67.36
Computer programming	6661	2.391	69.75
Information service activities	822	0.295	70.04
Financial service activities	62925	22.59	92.63
Insurance, reinsurance and pension funding	25	0.00897	92.64
Other financial activities	4884	1.753	94.39
Real estate activities	24	0.00862	94.40
Legal & accounting activities	7	0.00251	94.40
Activities of head offices	2706	0.971	95.37
Architecture & engineering activities	1207	0.433	95.81
Scientific research & development	160	0.0574	95.87
Advertising & market research	801	0.288	96.15
Other scientific activities	136	0.0488	96.20
Rental and leasing activities	4592	1.648	97.85
Employment activities	185	0.0664	97.92
Travel agency etc. activities	481	0.173	98.09
Security & investigation activities	184	0.0661	98.16
Office administrative etc. activities	1316	0.472	98.63
Public administration & defence	87	0.0312	98.66
Education	549	0.197	98.86
Residential care activities	1725	0.619	99.48
Creative, arts & entertainment activities	356	0.128	99.60
Sports activities	335	0.120	99.72
Activities of membership organizations	562	0.202	99.92
Repair of computers	175	0.0628	99.99
Other personal service activities	34	0.0122	100
Total	278575	100	

Table A.4: Number of treated firms in each distance band.

	inside	0-5km	5-10km	10-15km
Number of firms	365	7864	4868	2475

Table A.5: Balancing tests for Table 3 for inside.

	Mean_T	Mean_C	Diff	Std_diff	SD_pool	Var_ratio
TFP growth 2007	-.0810941	.0409531	-.1220472	-.3573565	.3415279	2.323236
TFP growth 2008	-.0768715	-.0118758	-.0649957	-.2719998	.2389548	.7663949
TFP growth 2009	.1877508	.0058396	.1819112	.5271367	.345093	3.866414
TFP growth 2010	.0206099	-.0121002	.0327101	.1292636	.2530494	.409122
TFP growth 2011	.0563886	.0985461	-.0421575	-.1450173	.2907066	2.29387
TFP growth 2012	-.0103206	-.0312178	.0208972	.0695467	.3004779	1.368705
TFP growth 2013	-.0732748	.003007	-.0762819	-.2696978	.282842	1.454788
TFP growth 2014	.0339198	.0225921	.0113276	.051969	.2179693	1.097159
TFP growth 2015	-.1193401	.0234242	-.1427643	-.3518834	.4057148	2.817334
TFP growth 2016	-.0219982	-.0210187	-.0009794	-.0023536	.4161473	3.485207
TFP growth 2017	.0638089	.0239686	.0398402	.1404902	.2835801	1.913899
TFP growth 2018	.037599	-.0699055	.1075045	.3348593	.3210437	.2755607
TFP growth 2019	-.0423475	.0598461	-.1021936	-.2747312	.3719766	.1931344
TFP growth 2020	-.006378	-.0867667	.0803887	.5219441	.1540178	.852193
Sales, log 2007	5.990288	6.070082	-.0797941	-.0423318	1.884968	.4250121
Sales, log 2008	5.98637	6.032049	-.045679	-.024817	1.840634	.5928943
Sales, log 2009	6.053806	5.951294	.1025118	.0527945	1.941713	.5539221
Sales, log 2010	5.822902	5.950248	-.1273465	-.0634976	2.005533	.4637786
Sales, log 2011	5.752333	5.981756	-.2294228	-.1158671	1.980052	.430202
Sales, log 2012	5.650609	5.867165	-.216556	-.1069767	2.024329	.4433153
Sales, log 2013	5.894158	5.878824	.0153338	.0083772	1.830421	.3400056
Sales, log 2014	5.82926	5.779319	.049941	.0277566	1.79925	.3398889
Sales, log 2015	5.792799	5.724564	.0682356	.0387607	1.760434	.3542585
Sales, log 2016	5.66546	5.60017	.0652906	.0375287	1.739749	.4188397
Sales, log 2017	5.652296	5.550956	.1013401	.0583483	1.736815	.3519593
Sales, log 2018	5.550934	5.461365	.0895695	.0512048	1.749239	.3560709
Sales, log 2019	5.423936	5.467157	-.0432215	-.0261288	1.654169	.3613243
Sales, log 2020	5.23627	5.307611	-.0713411	-.0458795	1.554967	.3694271
Assets, log 2007	6.173651	6.085292	.0883584	.0468496	1.886	.7002241
Assets, log 2008	6.115835	6.070371	.0454638	.0246542	1.844062	.7392428
Assets, log 2009	6.080168	6.021612	.0585564	.0316067	1.852658	.7394295
Assets, log 2010	5.962739	6.01307	-.0503305	-.0268505	1.874468	.6883179
Assets, log 2011	5.910708	5.965241	-.0545329	-.0293545	1.857732	.6780151
Assets, log 2012	5.818966	5.899089	-.0801233	-.0436185	1.836911	.669197
Assets, log 2013	5.949503	5.904897	.0446054	.0256739	1.737388	.553699
Assets, log 2014	5.844499	5.853594	-.0090948	-.0053725	1.692841	.6259351
Assets, log 2015	5.86369	5.805671	.0580198	.0346838	1.672822	.6415935
Assets, log 2016	5.808872	5.746177	.0626945	.0386667	1.62141	.6423504
Assets, log 2017	5.692057	5.697986	-.0059281	-.0037233	1.59219	.6088976
Assets, log 2018	5.485865	5.662764	-.1768995	-.1140926	1.55049	.6072934
Assets, log 2019	5.356044	5.592245	-.2362008	-.1573484	1.501133	.6230818
Assets, log 2020	5.199145	5.41042	-.2112748	-.1454554	1.452505	.6052089
Age 2005	47.0625	41.61765	5.444853	.3231678	16.84838	1.852496
Services 2005	0	0	0	.	0	.
Manufacturing 2005	.75	.8529412	-.1029412	-.2652522	.3880879	1.547586
Foreign ownership 2005	.0625	.0294118	.0330882	.1671556	.1979487	2.125
Andhra Pradesh	0	0	0	.	0	.
Assam	0	0	0	.	0	.
Bihar	0	0	0	.	0	.
Chhattisgarh	0	0	0	.	0	.
Dadra & Nagar Haveli	0	0	0	.	0	.
Daman & Diu	0	0	0	.	0	.
Gujarat	.125	.2941176	-.1691176	-.3920031	.4314191	.5454167
Haryana	0	0	0	.	0	.
Himachal Pradesh	0	0	0	.	0	.
Jammu & Kashmir	0	0	0	.	0	.
Jharkhand	0	0	0	.	0	.
Karnataka	0	0	0	.	0	.
Kerala	0	0	0	.	0	.
Madhya Pradesh	0	0	0	.	0	.
Maharashtra	.1875	.5588235	-.3713235	-.7405343	.5014265	.6397368
Meghalaya	0	0	0	.	0	.
NCT of Delhi	0	0	0	.	0	.
Odisha	0	0	0	.	0	.
Puducherry	0	0	0	.	0	.
Punjab	0	0	0	.	0	.
Rajasthan	0	0	0	.	0	.
Tamil Nadu	.625	.0882353	.5367647	1.211419	.4430875	3.016129
Telangana	0	0	0	.	0	.
Uttar Pradesh	0	0	0	.	0	.
Uttarakhand	0	0	0	.	0	.
West Bengal	0	0	0	.	0	.

Note: The table is based on estimation for the last year in the sample. As rules of thumb, the balancing is considered to be achieved if the variance ratio is between 0.5 and 2, and standardized difference < 0.2 for key variables.

Table A.6: Heterogeneity effect from Hi-tech SEZs.

	(inside) TFP growth	(0-5km) TFP growth	(5-10km) TFP growth	(10-15km) TFP growth
SEZs	-0.189	0.0428	0.134	0.131**
	(0.108)	(0.0947)	(0.0969)	(0.0659)
Age 2005	-0.000209	-0.00325**	-0.0145***	-0.00647***
	(0.00194)	(0.00145)	(0.00155)	(0.000578)
Service 2005	0.0834	-0.234**	-0.148	0.157
	(0.135)	(0.113)	(0.107)	(0.154)
Manufacturing 2005	-0.0277	0.0365	-0.149	0.0262
	(0.0393)	(0.0543)	(0.107)	(0.0588)
Foreign ownership 2005	0	0.0143	0.204**	0.162
	(.)	(0.0576)	(0.100)	(0.0851)
Constant	-0.125***	0.0713	0.771***	0.343***
	(0.0479)	(0.121)	(0.162)	(0.0617)
N	1103	12808	6736	3321
Year Fixed Effects	Yes	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes	Yes

Standard errors in parentheses
** $p < 0.05$, *** $p < 0.01$

Note: Time-varying covariates for creating the propensity scores include log of asset, log of sales and the history of the outcome variable. Time-invariant covariates include age, a foreign ownership dummy, dummies for manufacturing and service industries in 2005 and state dummies. The weights are derived using CBPS and Equation 2. Standard errors are clustered at the firm level.

Table A.7: Number of treated firms in the alternative distance bands.

	inside	0-7km	7-14km	14-21km
Number of firms	365	10584	4549	391

Table A.8: Robustness check using alternative distance bands.

	(inside) TFP growth	(0-7km) TFP growth	(7-14km) TFP growth	(14-21km) TFP growth
SEZs	-0.167**	-0.201	-0.160***	0.0615
	(0.0808)	(0.110)	(0.0544)	(0.123)
Age 2005	-0.00644	-0.00567**	-0.00285	-0.00136
	(0.00496)	(0.00228)	(0.00209)	(0.00119)
Service 2005	-0.112	0.0296	-0.298***	-0.343**
	(0.140)	(0.146)	(0.0962)	(0.153)
Manufacturing 2005	-0.0535	0.0491	-0.232***	0.0114
	(0.105)	(0.145)	(0.0788)	(0.0758)
Foreign ownership 2005	0.0335	-0.306	0.0133	-0.0606
	(0.0415)	(0.256)	(0.0720)	(0.104)
Constant	0.371	0.360***	0.313***	0.186
	(0.218)	(0.128)	(0.100)	(0.124)
N	1813	15692	7876	2919
Year Fixed Effects	Yes	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes	Yes

Standard errors in parentheses
** $p < 0.05$, *** $p < 0.01$

Note: Time-varying covariates for creating the propensity scores include log of assets, log of sales and the history of the outcome variable. Time-invariant covariates include age, a foreign ownership dummy, dummies for manufacturing and service industries in 2005 and state dummies. The weights are derived using CBPS and Equation 2. Standard errors are clustered at the firm level.

Table A.9: Robustness check excluding firms located in EPZs established prior to the enactment of the SEZs Act and later converted to SEZs.

	(inside) TFP growth	(0-5km) TFP growth	(5-10km) TFP growth	(10-15km) TFP growth
SEZs	-0.167**	0.0685	-0.125**	0.0123
	(0.0808)	(0.0920)	(0.0601)	(0.107)
Age 2005	-0.00644	-0.00317**	-0.00724**	0.00440***
	(0.00496)	(0.00134)	(0.00361)	(0.00169)
Service 2005	-0.112	-0.233**	-0.0133	-0.0544
	(0.140)	(0.108)	(0.149)	(0.156)
Manufacturing 2005	-0.0535	0.0318	-0.153	-0.0769
	(0.105)	(0.0545)	(0.114)	(0.143)
Foreign ownership 2005	0.0335	0.0363	0.0531	-0.0184
	(0.0415)	(0.0554)	(0.151)	(0.0909)
Constant	0.371	0.0500	0.746***	0.255
	(0.218)	(0.122)	(0.245)	(0.153)
N	1813	12696	8240	5404
Year Fixed Effects	Yes	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes	Yes

Standard errors in parentheses
** $p < 0.05$, *** $p < 0.01$

Note: Time-varying covariates for creating the propensity scores include log of assets, log of sales and the history of the outcome variable. Time-invariant covariates include age, a foreign ownership dummy, dummies for manufacturing and service industries in 2005 and state dummies. The weights are derived using CBPS and Equation 2. Standard errors are clustered at the firm level.

Table A.10: Estimation results using standard propensity score re-weighting.

	(inside) TFP growth	(0-5km) TFP growth	(5-10km) TFP growth	(10-15km) TFP growth
SEZ	0.00573	0.0100	0.00586	-0.0200
	(0.0197)	(0.00936)	(0.0111)	(0.0164)
Age 2005	0.0000551	-0.0000330	-0.000197	-0.000480**
	(0.000332)	(0.000181)	(0.000182)	(0.000235)
Service 2005	-0.00871	-0.0203	-0.0106	-0.0351
	(0.0303)	(0.0229)	(0.0213)	(0.0220)
Manufacturing 2005	0.0130	0.00358	-0.00566	-0.0146
	(0.0157)	(0.00859)	(0.00771)	(0.00968)
Foreign ownership 2005	0.0121	0.0158	0.00273	0.0754***
	(0.0404)	(0.0100)	(0.0166)	(0.0267)
Assets, log 2005	-0.00608	-0.00643	-0.00210	-0.00786
	(0.00799)	(0.00534)	(0.00698)	(0.00638)
Sales, log 2005	0.0137	0.00808	0.00621	0.0106
	(0.00899)	(0.00531)	(0.00660)	(0.00628)
Constant	-0.0612***	-0.0228	-0.0162	0.00406
	(0.0217)	(0.0210)	(0.0203)	(0.0275)
N	5270	32409	22027	14820

Standard errors in parentheses
** $p < 0.05$, *** $p < 0.01$

Note: First, we estimate nearest-neighbour propensity score matching as follows: $P(db) = Pr(D_i = 1|X_i^0)$, where $D_i = 1$ if firm i is in an SEZ for each distance band db and zero for never-treated firms. X_i^0 is a vector of pre-treatment covariates including age, dummies for manufacturing and service sectors, a foreign ownership dummy all measured in 2005, a time-invariant state dummies, and mean of log of sales and mean of log of assets for 2004-2006. Once the probabilities are estimated, they are transformed into weights. The treatment group receives a weight of $\frac{1}{Pr(D_i=1|X_i^0)}$ and the control group is weighted by $\frac{1}{1-Pr(D_i=1|X_i^0)}$.

A.3 TFP estimation

Consider the following Cobb-Douglas production technology with Hicks-neutral productivity in logarithmic form:

$$q_{it} \;=\; \beta_l\, l_{it} \;+\; \beta_k\, k_{it} \;+\; \omega_{it} \;+\; \epsilon_{it}, \tag{4}$$

where q_{it} is the logarithm of value added, l_{it} and k_{it} denote the log of labor and capital inputs, respectively, all of which are observed. There are two econometrically unobserved terms: ω_{it} and ϵ_{it}. The latter term represents shocks to the production that are not observed by the firm before making the input decision at time t. In contrast, ω_{it} represents productivity shocks that are potentially observed by the firm while making the input decision. To illustrate, the examples of ω_{it} might be the managerial ability of a firm, the expected delays and down-time due to a machine breakdown, the expected amount of rainfall at a farm, etc. On the other hand, ϵ_{it} represents the deviation from the predicted rainfall or the expected delay time, a sudden breakage of a machinery and other unexpected shocks or a measurement error.

The challenge in obtaining consistent production function estimates lies in the correlation between the unobserved productivity shocks and the input decision. The decision of a firm on the production inputs (l_{it}, k_{it}) will most likely depend on the observed by the firm ω_{it}, which makes OLS estimates of β_l and β_k inconsistent.

The control function approach proposed by Olley and Pakes (1996) and extended by Levinsohn and Petrin (2003) and the Ackerberg et al. (2015) technique are applied. The unobserved productivity shocks are proxied by the following material demand function:

$$m_{it} \;=\; m_t(l_{it}, k_{it}, \omega_{it}) \tag{5}$$

By inverting (5), productivity is expressed as:

$$\omega_{it} \;=\; h_t(l_{it}, k_{it}, m_{it}) \tag{6}$$

The estimation then proceeds in two steps. In the first stage, Equation (4) is estimated, where ω_{it} is substituted with its proxy from Equation (6). Thus, the estimation equation is as follows:

$$q_{it} \;=\; \Phi_t(l_{it}, k_{it}, m_{it}) + \epsilon_{it}, \tag{7}$$

where $\Phi_t(l_{it}, k_{it}, m_{it}) \;=\; \beta_l\, l_{it} \;+\; \beta_k\, k_{it} \;+\; h_t(l_{it}, k_{it}, m_{it})$. Important to notice that none of the coefficients $\beta \;=\; (\beta_l, \beta_k)$ are estimated in the first stage due to perfect collinearity, however, the predicted output is used to express the productivity:

$$\omega_{it}(\beta) \;=\; \hat{\Phi}_{it} - \beta_l\, l_{it} - \beta_k\, k_{it}. \tag{8}$$

In the second stage, moment conditions are formed to identify the production function coefficients. Thus, the law of motion for productivity explains the current level productivity as a function of productivity in the previous period and the innovation term ξ_{it} in the productivity shock ω_{it}:

$$\omega_{it} \;=\; g_t(\omega_{it-1}) + \xi_{it}. \tag{9}$$

Non-parametrically regressing $\omega_{it}(\beta)$ on $g(\omega_{it-1})$, the innovation term $\xi_{it}(\beta) \;=\; \omega_{it}(\beta) - E\left[\,\omega_{it}(\beta)\,|\,\omega_{it-1}(\beta)\,\right]$ is obtained from the residuals of the regression.

Given the timing assumptions that k_{it} was decided at $t-1$ and that lagged labor, l_{it-1}, is chosen at $t-b-1$, prior to m_{it} being chosen at t, where $0 < b < 1$, implies that the innovation term in productivity shocks is uncorrelated with all input choices prior to t. Thus, the moment conditions are:

$$E\left(\xi_{it}(\beta)\begin{pmatrix} l_{it-1} \\ k_{it} \end{pmatrix}\right) \;=\; 0 \tag{10}$$

Once the production function coefficients have been estimated, a firm-level total factor productivity is calculated as:

$$\hat{\omega}_{it} \;=\; \hat{\Phi}_{it} - \hat{\beta}_l\, l_{it} - \hat{\beta}_k\, k_{it}. \tag{11}$$

To account for industry differences in the production technology, the elasticities are estimated by industry. Some industries are combined to ensure enough observations in each group.

Value added is measured as firm revenue less expenditures on material inputs. Material inputs are defined as the sum of expenditures on raw material expenses and consumption of stores and spares. Labor input is measured by the total wage bill which comprises wages, social security contributions, bonuses, paid-leaves, etc. Capital input is represented by the gross fixed assets which include the movable, immovable and intangible assets of a firm.

Wages, value added, capital and intermediate materials are deflated by the 2-digit NIC-Industry Wholesale Price Index. Variables of firms in the service sector are deflated by the yearly WPI. All variables are monotonically transformed using the inverse hyperbolic sine (asinh). The inverse hyperbolic sine closely parallels log transformation but is defined at zero.[27] The interpretation of the regression coefficients is similar to log-transformed variables (Card and DellaVigna, 2020; Bahar et al., 2019).

Additionally, alternative measures of TFP, namely the approaches of Wooldridge (2009) and Levinsohn and Petrin (2003) are calculated and presented in the correlation Table A.11.

[27]The inverse hyperbolic sine (asinh) is defined as $ln(\alpha + \sqrt{(\alpha^2+1)}$. For $\alpha \geq 2$, $asinh(\alpha) = ln(2) + ln(\alpha)$ and $asinh(0) = 0$.

Table A.11: Correlation table for different TFP measures.

	(1)		
	ACF	Wooldridge	Levpet
ACF	1		
Wooldridge	0.799***	1	
Levpet	0.941***	0.863***	1

* $p < 0.1$, ** $p < .05$, *** $p < 0.01$

References

Ackerberg, D. A., Caves, K., and Frazer, G. (2015). "Identification Properties of Recent Production Function Estimators". *Econometrica*, 83(6):2411–2451.

Aggarwal, A. (2007). "Impact of Special Economic Zones on Employment, Poverty and Human Development". Development economics working papers, East Asian Bureau of Economic Research.

Aggarwal, A. (2012). "Social and Economic Impact of SEZs in India".

Ahsan, R. N. and Mitra, D. (2014). "Trade Liberalization and Labor's Slice of the Pie: Evidence from Indian Firms". *Journal of Development Economics*, 108:1–16.

Alkon, M. (2018). "Do Special Economic Zones Induce Developmental Spillovers? Evidence from India's States". *World Development*, 107:396–409.

Bahar, D., Özgüzel, C., Hauptmann, A., and Rapoport, H. (2019). "Migration and Post-Conflict Reconstruction: The Effect of Returning Refugees on Export Performance in the Former Yugoslavia". IZA Discussion Papers 12412, Bonn.

Baldwin, J. R., Brown, W. M., and Rigby, D. L. (2010). "Agglomeration Economies: Microdata Panel Estimates from Canadian Manufacturing". *Journal of Regional Science*, 50(5):915–934.

Barrios, S., Bertinelli, L., Heinen, A., and Strobl, E. (2012). "Exploring the Existence of Local and Global Knowledge Spillovers: Evidence from Plant-Level Data". *The Scandinavian Journal of Economics*, 114(3):856–880.

Blakeslee, D., Chaurey, R., Fishman, R., and Malik, S. (2022). "Land Rezoning and Structural Transformation in Rural India: Evidence from the Industrial Areas Program". *The World Bank Economic Review*, 36(2):488–513.

Borusyak, K., Jaravel, X., and Spiess, J. (2021). "Revisiting Event Study Designs: Robust and Efficient Estimation". *arXiv preprint arXiv:2108.12419*.

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

Busso, M., Gregory, J., and Kline, P. (2013). "Assessing the Incidence and Efficiency of a Prominent Place-Based Policy". *American Economic Review*, 103(2):897–947.

Callaway, B. and Sant'Anna, P. H. (2021). "Difference-in-Differences with Multiple Time Periods". *Journal of Econometrics*, 225(2):200–230.

Card, D. and DellaVigna, S. (2020). "What Do Editors Maximize? Evidence from Four Economics Journals". *Review of Economics and Statistics*, 102(1):195–217.

Criscuolo, C., Martin, R., Overman, H. G., and Van Reenen, J. (2019). "Some Causal Effects of an Industrial Policy". *American Economic Review*, 109(1):48–85.

De Chaisemartin, C. and d'Haultfoeuille, X. (2020). "Two-way Fixed Effects Estimators with Heterogeneous Treatment Effects". *American Economic Review*, 110(9):2964–96.

De Loecker, J., Goldberg, P. K., Khandelwal, A. K., and Pavcnik, N. (2016). "Prices, Markups, and Trade Reform". *Econometrica*, 84(2):445–510.

Fitzmaurice, G., Davidian, M., Verbeke, G., and Molenberghs, G. (2008). *"Longitudinal Data Analysis"*. CRC press.

Girma, S. and Görg, H. (2022). "Productivity Effects of Processing and Ordinary Export Market Entry: A Time-varying Treatments Approach". *Review of International Economics*, 30(3):836–853.

Goldberg, P. K., Khandelwal, A. K., Pavcnik, N., and Topalova, P. (2010). "Imported Intermediate Inputs and Domestic Product Growth: Evidence from India". *The Quarterly Journal of Economics*, 125(4):1727–1767.

Goodman-Bacon, A. (2021). "Difference-in-Differences with Variation in Treatment Timing". *Journal of Econometrics*, 225(2):254–277.

Ham, J. C., Swenson, C., İmrohoroğlu, A., and Song, H. (2011). "Government Programs Can Improve Local Labor Markets: Evidence from State Enterprise Zones, Federal Empowerment Zones and Federal Enterprise Community". *Journal of Public Economics*, 95(7-8):779–797.

Hasan, R., Jiang, Y., and Rafols, R. M. (2021). "Place-based Preferential Tax Policy and Industrial Development: Evidence from India's Program on Industrially Backward Districts". *Journal of Development Economics*, 150:102621.

Helmers, C. and Overman, H. G. (2017). "My Precious! The Location and Diffusion of Scientific Research: Evidence From the Synchrotron Diamond Light Source". *The Economic Journal*, 127(604):2006–2040.

Imai, K. and Ratkovic, M. (2014). "Covariate Balancing Propensity Score". *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 76(1):243–263.

Jenkins, R., Kennedy, L., Mukhopadhyay, P., and Pradhan, K. C. (2015). "Special Economic Zones in India: Interrogating the Nexus of Land, Development and Urbanization". *Environment and Urbanization ASIA*, 6(1):1–17.

Kennedy, L. and Rundell, E. (2014). "The State and Industrial Development in India: From Small Industry to Special Economic Zones". *Critique internationale*, 63(2):77–93.

Kline, P. and Moretti, E. (2014). "Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority". *The Quarterly Journal of Economics*, 129(1):275–331.

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

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

Nazarczuk, J. (2018). "Do Operations in SEZs Improve Firm's Productivity? Evidence from Poland". *New Trends and Issues Proceedings on Humanities and Social Sciences*, 4:256–264.

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

Palit, A. (2009). "Growth of Special Economic Zones (SEZs) in India: Issues and Perspectives". *Journal of Infrastructure Development*, 1(2):133–152.

Rosenthal, S. S. and Strange, W. C. (2003). "Geography, Industrial Organization, and Agglomeration". *Review of Economics and Statistics*, 85(2):377–393.

Schmidheiny, K. and Siegloch, S. (2019). "On Event Studies and Distributed-lags in Two-way Fixed Effects Models: Identification, Equivalence, and Generalization". *Equivalence, and Generalization (January 2019).*

Shenoy, A. (2018). "Regional Development Through Place-based Policies: Evidence from a Spatial Discontinuity". *Journal of Development Economics*, 130:173–189.

Steenbergen, V. and Javorcik, B. (2017). "Analyzing the Impact of the Kigali Special Economic Zone on Firm Behavior". *International Growth Centre Policy Working Paper.*

Sun, L. and Abraham, S. (2021). "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects". *Journal of Econometrics*, 225(2):175–199.

Thoemmes, F. and Ong, A. D. (2016). "A Primer on Inverse Probability of Treatment Weighting and Marginal Structural Models". *Emerging Adulthood*, 4(1):40–59.

Wang, J. (2013). "The Economic Impact of Special Economic Zones: Evidence from Chinese Municipalities". *Journal of Development Economics*, 101:133–147.

Wooldridge, J. M. (2009). "On Estimating Firm-Level Production Functions Using Proxy Variables to Control for Unobservables". *Economics Letters*, 104(3):112–114.

World Bank, G. (2017). "Special Economic Zones: An Operational Review of Their Impacts".

Zheng, Y. and Aggarwal, A. (2020). "Special Economic Zones in China and India: A Comparative Analysis". In *The Oxford Handbook of Industrial Hubs and Economic Development*, pages 607–622. Oxford University Press.