I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15123

Place-Based Policies and Agglomeration Economies: Firm-Level Evidence from Special Economic Zones in India

Holger Görg
Alina Mulyukova

MARCH 2022

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15123

Place-Based Policies and Agglomeration Economies: Firm-Level Evidence from Special Economic Zones in India

Holger Görg
Kiel Institute for the World Economy and IZA

Alina Mulyukova
Kiel Institute for the World Economy

MARCH 2022

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany
Phone: +49-228-3894-0
Email: publications@iza.org
www.iza.org

ABSTRACT

Place-Based Policies and Agglomeration Economies: Firm-Level Evidence from Special Economic Zones in India*

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, thus creating a novel dataset of firms with their assigned SEZ status. To overcome the selection bias we employ inverse probability weighting with time-varying covariates in a difference-in-differences framework. Our analysis yields that conditional on controlling for initial selection, the establishment of SEZs induced no further productivity gains for within SEZ firms, on average. This effect is predominantly driven by relatively less productive firms, whereas more productive firms experienced significant productivity gains. However, SEZs created negative externalities for firms in the vicinity which attenuate with distance. Neighbouring domestic firms, large firms, manufacturing firms and non-importer firms are the main losers of the program. Evidence points at the diversion of inputs from non-SEZ to SEZ- firms as a potential mechanism.

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

Keywords: Special Economic Zones, agglomerations, firm performance, India

Corresponding author:
Alina Mulyukova
Kiel Institute for the World Economy
Kiellinie 66
24105 Kiel
Germany

E-mail: alina.mulyukova@ifw-kiel.de

* The authors are very grateful to 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). However, there are only a few studies that evaluate such programs in developing countries primarily due to lack of data and endogeneity issues related to the non-random location of such policies. Furthermore, 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 economic growth of the region (World Bank, 2017).

This paper contributes to the literature by providing 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 to 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. However, this observational evidence is insufficient and further investigation using micro-level data, such as on firms or individuals, is required.

Apart from looking at this highly relevant country, a key novelty of this paper is that firms are used as a unit of analysis. This allows more precise estimation of the impact of SEZs, taking into account firm heterogeneity, compared to aggregating the data to the administrative unit level (as in, e.g., Wang (2013) and Alkon (2018)). Specifically, we use firm-level data to estimate the impact of SEZs on productivity growth, and also consider other performance measures such as export activity, sales growth, labour and intermediate input use 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.[1]

The main challenge confronting us in assessing the SEZs impact is the unavailability of data on firms operating inside the zones. While detailed information on the actual SEZ (e.g. location, size, establishment year, etc) is publicly available, 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 the information on the size of the zone. Subsequently, the radius is increased by 5 kilometers to create the following distance bands: $inside$, $0-5km$, $5-10km$ and $10-15km$. This approach enables the estimation of the spatial extent of any potential spillovers stemming from SEZs. Only with such fine-grained spatial data, it is possible to empirically identify any potential effect on firms inside SEZ and adjacent non-SEZ firms.

To the best of our knowledge, this study is the first to assemble a representative

[1] With data aggregated at some geographical level, such a distinction is not possible, of course.

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 financial statements, industry, ownership, products produced, age of a firm, its 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).[2] 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. The key challenge in identifying a causal effect is to find a proper counterfactual. 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, 2021). 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 find no further productivity gains for firms located inside the zones as a result of SEZs implementation. This effect is primarily driven by relatively less productive firms, whereas more productive firms experienced significant productivity gains. Looking at the spillovers, we find that in contrast to expected positive spillovers, the establishment of SEZs in India induced negative externalities on adjacent firms, whose productivity growth decreased significantly by 17 percent, on average. While this effect predominantly holds for more productive firms, less productive firms significantly increased their TFP growth, which can be explained by higher catch-up rates for firms that are further away from the technological frontier. We also find results pointing at negative spillovers when looking at the export activity of firms outside SEZs. In line with the spillovers literature, we find that the negative effects attenuate with distance and are more pronounced for firms in the immediate proximity and within 5 kilometers of SEZs.

We furthermore find that firms inside SEZs are able to increase their labour and intermediate input use as a result, while firms in the vicinity experience reductions in the use of these two inputs. These results are in line with a potential mechanism for the observed

[2] Prowess has been used extensively in other strands of research, see, e.g., Goldberg et al. (2010). 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. Moreover, Prowess contains firms in both the manufacturing and service sectors.

productivity effects. Given the incentives SEZ firms receive, they may be able to attract workers and intermediate good suppliers away from non-SEZ firms.

Further, we find that the adverse effects amplify with the number of other SEZs in the vicinity. Turning to effect heterogeneity, neighbouring domestic firms, large firms, manufacturing firms and non-importer firms seem to be the main losers of the program, showing a significant downward productivity trend. Interestingly, manufacturing firms inside SEZs experience a significant decline in TFP growth following the SEZ implementation. This may be attributed to the fact that goods from SEZs sold in the domestic market are considered as imports, which decreases the comparative advantage relative to other domestic manufacturing firms. Further, we find that most of the adverse effect comes from the technologically intensive zones.

Our paper contributes to a broader literature examining the effects of place-based policies in the presence of agglomeration economies (Ham et al., 2011; Busso et al., 2013; Kline and Moretti, 2014; Chaurey, 2017). Particularly, 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. While the need to understand the welfare and developmental implications of SEZs is growing, the literature remains scarce primarily due to the lack of data and the endogeneity concerns related to the non-random assignment of SEZs. Two papers that closely relate to this work are the evaluation of SEZs in China by Wang (2013) and Lu et al. (2019), which showed that the SEZs establishment increased capital investment, employment, wages and productivity of firms by achieving agglomeration economies. Other papers examining the implications of SEZs on firm-level outcomes include Steenbergen and Javorcik (2017), Nazarczuk (2018) and Nazarczuk and Umiński (2018) showing that firms moving into the zones benefited in terms of their sales, value-added and export performance. Regarding the SEZs program in India, several works provide descriptive evidence evaluating the efficiency of the program (Aggarwal, 2007, 2012), however, a paper by Alkon (2018) is methodologically mostly comparable to this one. Alkon (2018) uses census subdistrict-level data and finds that SEZs failed to induce developmental spillovers.

The rest of the paper is structured as follows. Section 2 describes the background information on the SEZs program in India. Section 3 presents theoretical evidence on agglomeration economies stemming from place-based policies and formulates the key hypotheses. In Section 4 the data are introduced. Section 5 presents the event study. Section 6 proceeds with describing the methodology of time-varying treatment, estimates the effect and presents the results, heterogeneity analysis and robustness checks. Section 7 concludes.

2 Background of Indian SEZs

India was one of the first countries in Asia to recognize the importance of Export Processing Zones (EPZs) in 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 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 the 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 SEZs were expected to attract FDI of US$ 6 billion and create 50.000 jobs by the end of 2007 (Aggarwal, 2012). 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[3], 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.[4] 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 apply for creating an SEZ and thus SEZ status can be assigned to one firm.

The Board of Approval is the single-clearance window mechanism for establishing the SEZs units. The establishment process 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. According to a representative survey conducted by Mukherjee and Bhardwaj (2016), 84% of interviewed units across 32 SEZs indicated that income tax holidays were the most crucial factor in their decision to begin operation inside the SEZs, pointing to the importance of fiscal incentives.

One distinctive feature of Indian SEZs is that the policy provides equal opportunities to develop an SEZ for government, private or joint sectors. In India, state governments are responsible for property rights and are entitled to acquire private land if it serves a "public purpose". Alkon (2018) shows that there is a positive correlation between the presence of state-owned industrial development corporations and the number of approved SEZs. The main purpose of the industrial corporations is to facilitate the development, to promote

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

[4]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 B.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.

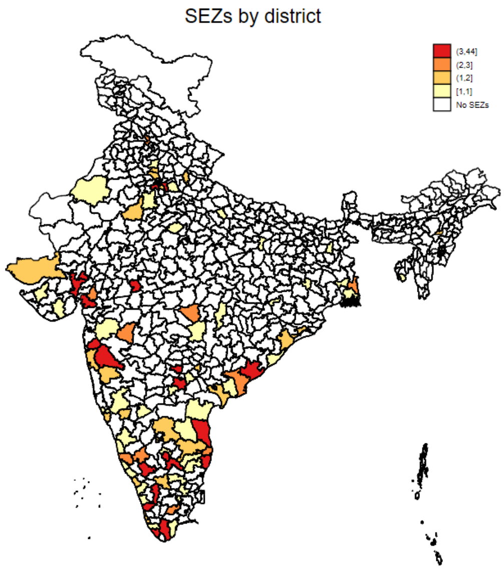

Figure 1: Number of SEZs by district.

industrialization and growth of the region by providing suitable infrastructure, mediating the land acquisition as well as the clearance attainment (e.g. environmental clearance) for private developers. Thus, state-owned development corporations serve as primary developer or co-developer for many SEZs. 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.[5]

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). 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 threat to the causal identification of the effect of zones due to a positive selection bias, which will be addressed using inverse probability weighting.

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 the largest being Adani Port and SEZs (6.456 ha) and Andhra Pradesh Industrial Infrastructural Corporation Ltd. (2.206 ha). In the empirical analysis below, we use concentric rings drawn around SEZs using the information on the area of the zones. Figure 2 depicts the histograms of the area and the radius. Table 1 presents summary statistics of the area and radius by SEZ-sector, showing great sector-wise heterogeneity in terms of the size of the zones. Additionally, Figure A.1 illustrates the distribution of SEZs by sector and time. We define the first wave years from 2006-2009, the second wave are years 2010-2013 and the third wave is 2014-2020. 64%

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

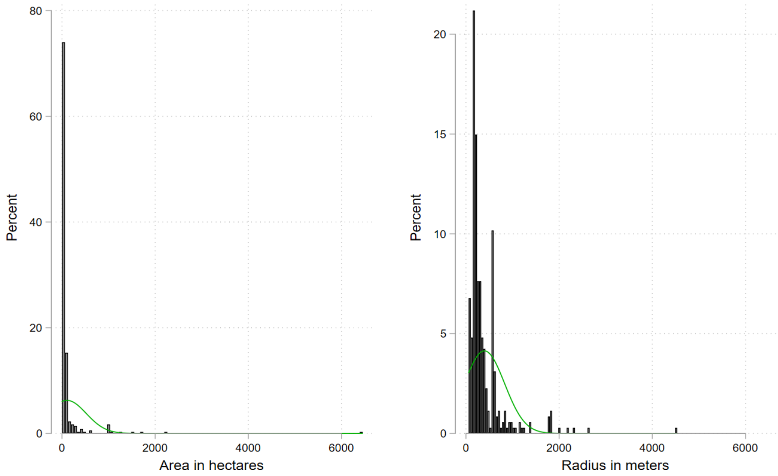

Figure 2: Histogram of the area and the radius of SEZs.

of the zones established in the first wave are IT-based, with the number increasing and reaching 80% in the last wave.

Table 1: Summary statistics of area and radius by sector.

	Area in hectares				Radius in meters				
	Mean	SD	Min	Max	Mean	SD	Min	Max	N
Aviation	101.69	0.45	101.17	101.98	568.94	1.26	567.48	569.75	3
Biotech	18.97	9.79	10.00	40.47	239.01	59.18	178.41	358.92	15
Construction	106.46	.	106.46	106.46	582.13	.	582.13	582.13	1
Energy	76.49	84.15	10.00	222.67	423.91	276.65	178.41	841.89	6
Engineering	124.23	69.24	36.42	317.71	610.18	157.07	340.49	1,005.64	16
Food processing	48.20	44.63	11.88	119.14	356.15	176.10	194.45	615.83	7
Free Trade and Warehousing Zones	109.55	144.91	40.63	434.86	522.83	296.53	359.60	1,176.52	7
Gems and Jewellery	68.80	.	68.80	68.80	467.97	.	467.97	467.97	1
Handicrafts	10.49	.	10.49	10.49	182.71	.	182.71	182.71	1
IT	24.12	33.39	1.05	223.00	241.22	136.61	57.82	842.51	237
Minerals	119.86	41.17	50.75	166.91	608.52	116.14	401.94	728.90	6
Multi-product	1,165.60	1,355.21	105.44	6,456.33	1,738.86	850.14	579.33	4,533.34	20
Paper products	109.81	.	109.81	109.81	591.22	.	591.22	591.22	1
Pharmaceuticals	94.47	54.50	11.47	247.39	524.37	165.14	191.10	887.39	18
Port	224.57	98.90	110.47	285.84	828.84	204.37	592.99	953.87	3
Textile	133.69	107.25	20.41	404.70	610.53	240.00	254.89	1,134.99	12
Total	107.80	411.82	1.05	6,456.33	396.68	431.65	57.82	4,533.34	354

Overall, compared to place-based policies in China, India's SEZs have 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 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 to a particular area an SEZ status which consequently attracts foreign and domestic firms due to stimulating fiscal regulations.

3 Conceptual framework

To fix ideas and motivate our empirical analysis, this section briefly sketches a conceptual background to outline expected effects. Recall that the aim of the paper is to estimate the impact of the new establishment of SEZs on firms, distinguishing firms located within a newly established SEZ, and those in the geographic vicinity. The main threat to the identification strategy is selection bias (Allcott, 2015; Lu et al., 2019). Since place-based policies in developing countries target predominantly more prosperous regions with advanced infrastructure and network, one may expect that this positive selection may systematically bias the results upwards. In the case of Indian SEZs, it could be that initially more productive firms apply for creating an SEZ since they are able to bear the investment costs. We show evidence of positive selection and address this issue in the following sections.

Turning to the SEZ effects on firm performance, we would expect different mechanisms impacting these firms. As for firms within SEZs, one would expect that the provided tax incentives would result in a reduced marginal cost for firms within SEZs. Given the lower costs, they are therefore able to charge lower prices, ceteris paribus, and hence increase demand for their output compared to firms located outside of the SEZs. The lower tax rates will, thus, generate a surplus in profits, which firms can invest in innovation activities, further boosting their innovativeness and productivity.

However, the inefficient use of incentives could also result in no productivity gains. Firms might use the given benefits without expanding their production, substituting public benefits for privately funded investment. In other words, the public incentives may crowd out private investment, which will lead to no improvement in firms' performance. Also, depending on whether a firm is in the upper or lower tail of productivity distribution, its absorptive capacity to assimilate knowledge and technology brought by SEZ incentives may vary. One may expect that firms in the upper tail of productivity distribution may be the main beneficiaries of the program, whereas less productive firms will be left behind (Bos and Vannoorenberghe, 2018).

Apart from these direct effects on firms located in SEZs, we are also interested in estimating the impact on firms located in the geographic vicinity of the zones. Here, existing empirical evidence shows that place-based policies may achieve agglomeration economies (externalities or spillovers) due to the clustering of economic activity (Rosenthal and Strange, 2004; Wang, 2013; Kline and Moretti, 2014; Neumark and Simpson, 2015; Zheng et al., 2015; Lu et al., 2019). Firms that are located in the vicinity may benefit from industrial clusters in various ways, depending upon the nature of externalities. Geographic proximity is considered to be important because if agents are physically closer, there is more potential for interaction and thus spillovers (Rosenthal and Strange, 2004). Conversely, agglomeration economies attenuate with distance, with the first few miles experiencing a rapid decline, slowing down thereafter (Rosenthal and Strange, 2003; Combes and Gobillon, 2015; Helmers and Overman, 2017).

Spillovers can be positive or negative. The general mechanism underlying positive spillovers is the diffusion of knowledge and ideas. If firms located in SEZs are more technologically advanced and innovative compared to their non-SEZs counterparts, then the accumulated knowledge can spread across neighbouring firms. However, for this to happen, non-SEZs firms should have the absorptive capacity to assimilate and adopt these new technologies. On the one hand, firms with a large technological gap possibly have the highest marginal return to innovation, however, they are left behind the best practices and potentially lack the technical competency to catch up. On the other hand, firms that are at the frontier of innovations likely do not have the incentives to alter the existing practices, but could easily adopt new technologies (Blalock and Gertler, 2009). Griffith et al. (2002), for instance, show that firms further away from the technological frontier have higher catch-up rates.[6]

Not all externalities may be positive, however. SEZ firms may negatively impact on within-firm performance of non-SEZ firms in a variety of ways. Enjoying the fiscal benefits, SEZs firms may be able to pay a higher marginal wage and attract the best workers from the neighbouring firms, resulting in productivity declines of adjacent non-SEZs firms. SEZs firms may also attract better quality suppliers compared to their non-SEZs competitors, thus "stealing" best quality inputs from other non-SEZ firms. As a "stealing effect", whether it refers to labor or material inputs, may lead to negative impacts on the performance of firms outside of SEZs.

Given that the spillovers are localized, the effects are expected to be most pronounced for firms within 5 to 10 kilometers of SEZs, steadily losing their significance thereafter. Barrios et al. (2012), for instance, show that R&D spillovers are bounded within 10 kilometers. The same results apply to knowledge spillovers that extend at most 10 kilometers away (Baldwin et al., 2010). Furthermore, Rosenthal and Strange (2003) show that agglomeration economies attenuate rapidly with distance, with the initial effect in the first mile being up to 10 to 1000 times larger than the effect two to five miles away. After five miles, the attenuation is less pronounced. We also look into such a distance decay in our empirical analysis.

To investigate such geographic spillovers, we apply a concentric ring approach, where rings of different radii are created around the centroid of SEZs. In principle, as argued by Rosenthal and Strange (2003), any arbitrarily large number of concentric rings can be used, however, to maintain a parsimonious specification, it is recommended to aggregate the geographic details into distance bands. In this paper, the actual radius of SEZs is subsequently increased by 5 kilometers to construct the following distance bands: *inside*, $0-5km$, $5-10km$ and $10-15km$ (similar to Rosenthal and Strange (2003) and De Silva and McComb (2012)). Firms located further than 40 km away from the centroid of SEZs constitute the control group. Figure A.2 illustrates the mapping of firms into the circles of various radii around SEZs. Due to the clustering of economic activity, an increase in the radius leads firms to be treated by multiple SEZs, which is denoted as treatment intensity[7]. Thus, the estimation results for *inside* indicate the direct effect of SEZs opening, whereas distance bands show potential spillover effects stemming from the zones.

When considering firm heterogeneity, one may expect that depending on firms' characteristics, the SEZ effects may differ due to the particular structure and primary goals of the program. Intuitively, since SEZs require no minimal land area for firms in the IT and Biotech sectors, serving as an incentive for creating start-ups, one may expect that service-sector firms and young firms will be the main beneficiaries of the program. Regarding the ownership status, the SEZs aim at attracting not only foreign investors but also encourage domestic producers to expand their production and enter international markets. Therefore, the heterogeneous effects by ownership type are ambiguous and should be tested empirically. Further, given the duty free procurement of goods, one may expect that firms that source intermediate inputs from the international market would benefit more compared to non-importer firms.

Finally, given the observed clustering of SEZ activity in certain regions, a given firm may be affected by more than one SEZ as the radius of the circle increases. Thus, we also

[6]Another potential channel through which positive spillovers may happen is through input-output linkages between neighbouring firms. SEZs firms may experience a higher demand for their output through increased competitiveness resulting from the reduced marginal cost of production. To better respond to the increased demand, a firm, in turn, will increase its production and, therefore, the demand for intermediate inputs, which may be supplied by firms in the vicinity. As we focus on within-industry effects this argument does not apply in our analysis.

[7]When firms are treated by multiple SEZs, treatment year is defined as the minimum year among all SEZs.

estimate whether the presence of other than own SEZs in the vicinity amplifies the SEZ effects.

4 Data

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, two different data sources, namely firm-level data and the list of notified SEZs, were merged using a spatial approach.

SEZs Data. - First, the list of notified SEZs under the 2005 Act was obtained from the Ministry of Commerce and Industry, Department of Commerce.[8] 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.

An important point to be made here 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.[9]

Firm Data. - 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).[10] This is an unbalanced panel of firms covering the period from

[8] The list is available under: `http://sezindia.nic.in/upload/uploadfiles/files/notify.pdf`. Last update 29.02.2020.

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

[10] Prowess accounts for 60-70% of the economic activity in the industrial sector (Goldberg et al., 2010). Since firms are not obliged to report their financial performance to CMIE, the entry and exit of the firms in the sample may not represent their true status. Moreover, Prowess represents relatively large Indian firms. One of the main advantages of using Prowess compared to the Annual Survey of Industries (ASI) is that the precise address information is available in Prowess, whereas ASI reports only the administrative territorial unit (e.g. district or state). Therefore, it would not be possible to identify precisely firms operating in SEZs. Using a district as a unit of analysis would lead to a downward bias due to the aggregation and inclusion of non-SEZ firms. Secondly, 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 B.3. However, because firms are under no legal obligation to report the data, only less than 10 percent 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

1988 to 2020. Overall, the sample 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. 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 treatment status of firms as described below.

The analysis focuses on TFP growth as the main variable of interest. We further consider sales growth, export probability, wage growth and material expenses growth to better understand what drives the TFP results. 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 C). 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 code.[12] We classify a firm as foreign-owned if the percent of equity shares held by foreign individuals, corporate bodies or institutions exceeds 25%.[13] Time-varying covariates include, depending on the specification, total assets, sales, TFP and export sales.

Table 2 presents the 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 in percent, 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. Moreover, pre-treatment mean assets is higher for within SEZ firms compared to the control group, indicating that SEZ firms are initially bigger. Further, within SEZ firms have a higher initial propensity to export and higher export share relative to the control group. Overall, this pre-treatment mean comparison indicates positive selection bias. To overcome this problem, we control for the pre-treatment variables so that the results can be interpreted accounting for the site selection.

Further examining trends for the whole sample, we observe that TFP grows less in each treated group compared to the growth rate in the control group. The sales growth, asset growth and wage growth of firms in the 0-5km distance band are higher than the respective values of SEZ firms. The exports growth is lower for inside SEZ firms compared to the control group. Moreover, SEZ firms are younger, have a higher share of foreign ownership and operate mostly in the service sector compared to the control group.

Looking at pre-treatment trends, the TFP growth for inside SEZ firms is significantly lower than that of the control group, accompanied by lower growth rates of sales, assets and wages. The same pattern holds for TFP growth for firms in the 0-5 kilometers distance band, however, the growth rate of sales, assets and wages is higher than that of the never-treated firms. Interestingly, the proportion of SEZ firms in the manufacturing sector is significantly lower than that in the control group, which can potentially be explained by the program design. Any goods from SEZs to the Domestic Tariff Area - an area within India that is outside the specially designated zones - are treated as imports and are subjected to duties of customs, which may reduce the incentives of manufacturing firms oriented for the

TFP estimation.

[11]Geocoding is done using ArcGIS Online Geocoding Service.

[12]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.

[13]We provide a robustness check using a 50% threshold.

domestic market to locate in an SEZ. This is also reflected in the growth of export sales, which is significantly higher for firms in the control group (2.49% for within SEZ firms vs. 17.08% for the control group).

Table 2: Summary statistics of firms.

		(1) Inside			(2) 0-5km			(3) 5-10km			(4) 10-15km			(5) Control	
	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
Panel A: Whole sample															
TFP	1954	32.437%	0.64	39833	32.932%	0.73	23517	31.031%	0.71	11306	31.759%	0.73	15175	23.792%	0.65
Assets	4678	50.27778	454.86	97024	46.19758	459.63	58529	47.12143	595.94	31064	47.59398	413.14	25274	28.19869	126.90
Exporter dummy	4692	28.943%	0.45	97526	29.922%	0.46	58719	26.804%	0.44	31223	27.259%	0.45	25360	27.129%	0.44
Export share	1347	39.485%	2.30	28710	503.221%	555.94	15503	212.632%	116.48	8275	205.107%	70.36	6862	25.173%	0.83
TFP growth	1636	-0.655%	0.31	33753	-0.136%	0.32	19811	-0.062%	0.33	9574	-0.390%	0.33	12847	0.069%	0.37
Sales growth	3007	-0.785%	0.84	59525	0.151%	0.80	34542	-0.654%	0.79	17420	-2.475%	0.83	17267	-1.883%	0.76
Asset growth	4053	0.535%	0.46	84249	2.370%	0.49	50726	2.315%	0.47	26970	0.317%	0.51	21861	0.535%	0.41
Wage growth	3361	2.359%	0.54	69016	3.502%	0.54	41439	2.708%	0.54	20943	1.053%	0.53	19060	1.668%	0.49
Material expenses/sales growth	1763	-0.586%	0.59	37010	-1.227%	0.58	21954	-1.351%	0.57	10669	-1.671%	0.59	14003	-1.291%	0.53
Exports growth	1089	2.547%	1.02	23381	4.673%	1.00	12545	1.989%	1.01	6809	2.261%	1.03	5564	3.280%	0.98
Age	4690	37.17932	16.62	97515	37.79740	17.52	58715	39.72385	19.55	31210	41.56530	20.24	25358	39.66488	17.35
Foreign ownership dummy	4692	1.428%	0.12	97526	1.924%	0.14	58719	0.935%	0.10	31223	0.922%	0.10	25360	1.104%	0.10
Manufacturing dummy	4692	35.102%	0.48	97526	36.482%	0.48	58719	36.496%	0.48	31223	35.115%	0.48	25360	60.193%	0.49
Services dummy	4692	51.002%	0.50	97526	52.020%	0.50	58719	53.543%	0.50	31223	56.494%	0.50	25360	29.700%	0.46
Importer dummy	4692	29.838%	0.46	97526	30.420%	0.46	58719	27.890%	0.45	31223	27.153%	0.44	25360	33.253%	0.47
Panel B: Pre-treatment variables in 2005															
TFP	81	31.110%	0.61	1601	34.479%	0.75	954	32.659%	0.70	458	26.625%	0.72	618	22.174%	0.59
Assets	224	22.62019	242.91	4515	25.93660	270.31	2804	27.22806	343.88	1489	20.88516	181.98	1177	16.61342	81.67
Exporter dummy	224	26.339%	0.44	4539	26.636%	0.44	2811	21.558%	0.41	1499	23.215%	0.42	1180	21.271%	0.41
Export share	59	37.116%	0.36	1200	38.710%	1.83	603	26.598%	0.33	343	79.383%	8.33	251	24.633%	0.27
TFP growth	64	0.890%	0.28	1285	0.665%	0.35	767	2.000%	0.32	399	1.214%	0.30	497	3.070%	0.32
Sales growth	125	1.366%	1.01	2376	9.941%	0.82	1424	8.715%	0.86	789	3.950%	0.92	694	8.748%	0.69
Asset growth	181	5.189%	0.36	3607	8.670%	0.47	2243	8.153%	0.54	1316	4.355%	0.53	947	6.484%	0.38
Wage growth	147	1.806%	0.49	2818	6.176%	0.58	1692	5.262%	0.55	962	4.076%	0.59	783	3.439%	0.50
Material expenses/sales growth	70	1.522%	0.46	1398	4.453%	0.56	834	4.455%	0.55	439	2.220%	0.62	543	0.643%	0.45
Exports growth	51	2.487%	1.20	952	27.165%	1.04	480	28.378%	1.01	296	31.078%	1.00	208	17.080%	0.88
Age	224	35.95982	15.23	4539	36.93325	17.19	2811	39.08573	18.92	1499	40.55904	19.26	1180	38.94407	17.81
Foreign ownership dummy	224	0.893%	0.09	4539	2.137%	0.14	2811	0.961%	0.10	1499	0.801%	0.09	1180	1.271%	0.11
Manufacturing dummy	224	33.929%	0.47	4539	33.069%	0.47	2811	33.155%	0.47	1499	31.154%	0.46	1180	57.627%	0.49
Services dummy	224	53.571%	0.50	4539	55.849%	0.50	2811	57.560%	0.49	1499	60.173%	0.49	1180	32.542%	0.47
Importer dummy	224	24.554%	0.43	4539	26.393%	0.44	2811	23.693%	0.43	1499	23.282%	0.42	1180	27.373%	0.45

Treatment Definition. - Since the information on the SEZ status of the firm is not directly reported by Prowess, we merge two geocoded datasets using a spatial ring approach. The primary difficulty in pinning down the exact location of the zone is the imprecise location information, which is available at the village level in the most disaggregated form. Thus, to pinpoint the accurate address of the zone, 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.[14]

We create the first spatial ring using the information on the original radius of SEZs presented in Table 1, which is from now on referred to 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. The aforementioned procedure is repeated to increase the radius by $10km + r$ and $15km + r$ as is shown in Figure A.2. 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 the 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 B.4 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

[14]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 miss-classification does not seem to be a threat.

year among all SEZs and one observation per firm 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.

5 Event Study

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 we compare productivity growth of firms before and after the establishment of SEZs relative to the firms that are not exposed to the program. The key assumption 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.

One of the identification problems is the non-random treatment assignment.[15] To deal with the potential self-selection problem, we first apply nearest-neighbour propensity score matching and estimate the following equation:

$$P(db) = Pr(D_i = 1|X_i^0), \tag{1}$$

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 code, and mean of log of sales, log of assets, TFP growth and exporter status 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 $\frac{1}{1-Pr(D_i=1|X_i^0)}$. In this way, we create a pseudo-population in which the treatment assignment is independent of the covariates. Furthermore, 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.[16] The propensity score is thus performed on a restricted sub-sample.[17]

[15] In developed countries, place-based policies aim at reducing regional inequality and hence target disadvantaged geographic areas, which may bias the estimates downwards because treated units are already located on a downward trend (Kline and Moretti, 2014; Criscuolo et al., 2019). In developing countries, place-based policies are used primarily to attract foreign direct investment and are therefore established based on favourable geographic location and advanced infrastructure of the region, leading to an upward bias of the estimates due to exhibiting increasing productivity trends of firms (Aggarwal, 2012; Lu et al., 2019).

[16] In the choice of the control group we relied on two primary factors: the control group should not be affected by treatment and it should be comparable to the treated group. We also provide a robustness check for the time-varying treatment approach using an alternative control group.

[17] 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

Once the weights are formed, they are included in the following estimation equation:

$$y_{it}(db) = \alpha + \sum_{k \geq -10, k \neq -1}^{10} \beta_k \times D_{it}^k(db) + \gamma X_{it} + \phi_i + \lambda_t + \epsilon_{it}, \tag{2}$$

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.[18] $D_{it}^k(db)$ is always 0 for never-treated firms. y_{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 are not independently and identically distributed but rather correlate across time. X_{it} includes as controls the log of sales, log of total assets, and exporter status. The year preceding the treatment ($k = -1$) is chosen as a benchmark so that the post-treatment effects are relative to the year immediately prior to the program implementation. Since the aim of the analysis is to estimate potential spillovers from SEZs, Equation 2 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 an increasing TFP growth trend for firms inside SEZs, reaching its peak four years after treatment. The effect though is not significant at 5%. For firms in the 0-5 kilometers distance band, we observe a significant negative effect on the productivity growth, which is most pronounced two to three years after SEZs establishment. A declining trend is also observed for firms in the 10-15 kilometers distance band, however, the effect is not significant.[19]

Since the sample used for the analysis is unbalanced, it may create a concern that the attrition of firms is non-random. As a 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 2 for the sample of balanced treated firms. We keep the control firms as before not to lose observations. Figure A.3 in the Appendix depict the results. We still observe a positive trend for inside firms up to the fourth period after SEZs opening and negative trends for the neighbouring firms.[20]

While the identified within-firm estimates may bring us one step closer to a plausible treatment effect, this approach raises one important question. Since the applications for setting up an SEZ are considered on a rolling basis, the treatment for each firm sets in at different points in time. Therefore, the relative importance of the covariates that determine initial selection into treatment changes over time. To account for that, we create time-varying weights. Thus, we now turn to our preferred specification of determining

---

or satisfies the minimum land requirement. Thus, due to the limited number of observations, this approach is not possible here.

[18] Observations outside of the event window are dropped. However, the results do not change when these observations are replaced to be inside the event window.

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

[20] 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.4 in the Appendix presents the results, which appear similar.

the average treatment effect of SEZs on firm performance using a combined difference-in-differences and propensity-score reweighting approach, where weights are created at each point in time.

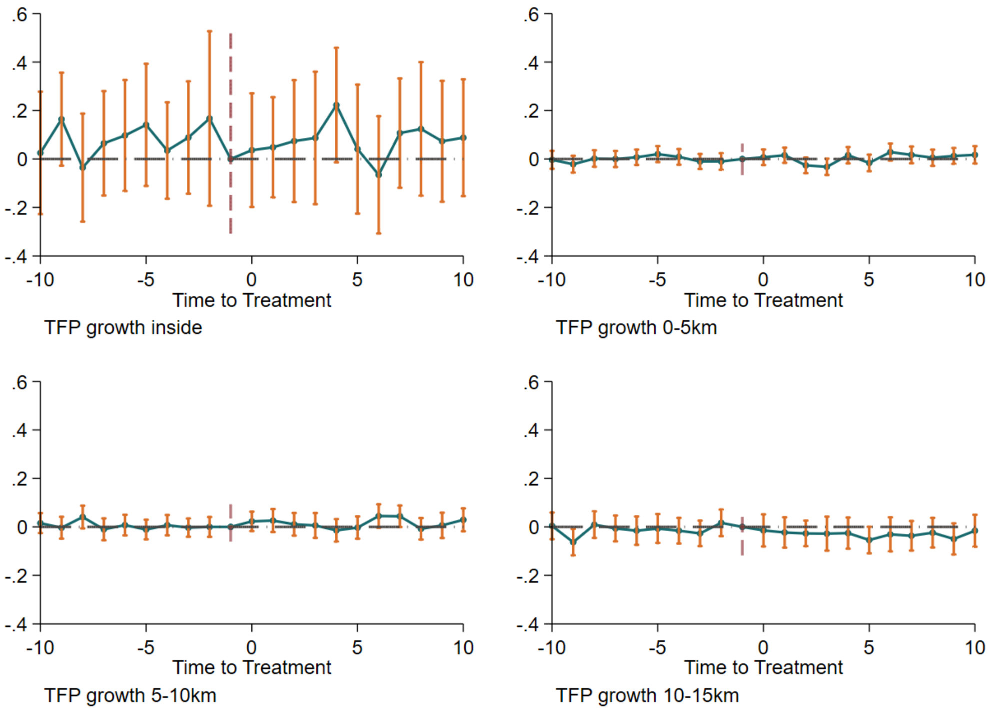

Figure 3: Event study graph for TFP growth. 95% confidence interval is reported. The sample is matched.

6 Time-Varying Treatment Approach

6.1 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, 2021). To illustrate, firm variables change over time depending on previous confounders, the treatment history, and the development of the outcome variable in the 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 would lead to biased estimates.

To overcome the aforementioned problem, we follow a growing literature on the time-varying treatment approach and calculate weights at each point in time (Fitzmaurice et al., 2008; Thoemmes and Ong, 2016; Girma and Görg, 2021). To illustrate, at the first treatment occurrence, we estimate a logistic regression to predict a binary treatment selection given the observed history of the covariates. At the next time point, a different set of weights from the logistic regression is constructed that makes the treatment selection at time two orthogonal of all observed covariates prior to this treatment selection. Since the first treatment occurred in 2006 and the last one in 2020, repeating this procedure for each year following the first treatment introduction results in a set of 14 weights, which are eventually cross-multiplied to form a unique final weight for each firm i at time t.

Taking all together, the stabilized weight is estimated as follows:

$$SW_{it}(db) = \prod_{t=1}^{T} \frac{P(SEZs_{it}(db) = 1 \mid X_i^0)}{P(SEZs_{it}(db) = 1 \mid \bar{X}_{it-1}, X_i^0)}, \tag{3}$$

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 code. $\bar{X}_{it-1}$ are time-varying covariates up until $t-1$, including the log of total assets, log of sales, TFP, exporter dummy and the history of the outcome variable depending on the specification. 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.

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:

$$Y_{it}(db) = \alpha + \beta SEZs_{it}(db) + \theta X_i^0 + \lambda_t + \epsilon_{it} \tag{4}$$

where $Y_{it}(db)$ is the dependent variable (alternatively defined as TFP growth, sales growth, or change in export probability). $SEZs_{it}(db) = 1$ 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. Standard errors are clustered by firm to allow for within-firm correlation of the dependent variable over time.[21] 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 firms further than 40 kilometers away.

6.2 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 selection, we do not find further positive effects of the establishment of SEZs on the productivity growth of within SEZ firms, on average. However, we observe significant negative externalities for firms located in the spatial proximity to SEZs. The productivity growth of firms within 5 kilometers decreased significantly by $(\exp^{-0.157} -1) \times 100 = 17\%$ compared to, ceteris paribus, similar firms in the control group. Firms in the 5-10 kilometers distance band are also affected negatively by the presence of SEZs, decreasing their productivity growth by 16%. However, the adverse effect is less pronounced for firms in the 10-15 kilometers distance band, accounting for a decline by 7%. Thus, we observe a continuous distance decay effect - the further the firms are from SEZs, the less pronounced are the negative externalities stemming from SEZs.

However, splitting the sample on relatively more or less productive firms shows a slightly different pattern. Table 4 illustrates that conditional on controlling for selection, within SEZ firms with above median TFP growth significantly increase their productivity growth as a result of SEZs establishment. Their productivity growth increased by 41% compared to, ceteris paribus, similar firms in the control group. More productive firms in the 0-5 kilometers distance band, however, show a significant negative TFP growth decline, accounting for a drop of 9%. The negative externalities disappear after 5 kilometers. Looking at firms in the lower tail of productivity distribution, we observe that whereas SEZ firms experienced no significant effect as a result of policy implementation, neighbouring firms within 15 kilometers see significant productivity gains. As discussed above, the intuitive explanation is that firms that are further away from the technological frontier potentially have a higher marginal return to innovation and thus will have higher catch-up rates.

In order to get a better understanding of what may be driving these TFP effects, we now look at some further variables that may contribute to productivity growth: total sales, exports, total labour and intermediate inputs. We start with total sales growth in Table 5. Conditional on controlling for positive selection, we find no further discernible (statistically significant) impact on the sales growth of within SEZ firms. The impact on the neighbouring firms is also not statistically significant, so reductions in output are unlikely to be able to explain the negative spillovers found in the TFP estimations.

Table 6 considers changes in export activity, as export promotion is generally a goal of the implementation of SEZs. Specifically, we look at total export sales. Conditional on controlling for initial selection, results show no significant further gains of export sales as a result of SEZs establishment. Neighbouring firms, however, experience a significant drop in the growth rate of export sales. It thus appears that, in line with negative implications for TFP growth, the establishment of an SEZ also affects negatively the growth rate of exports, for firms located in the vicinity of the zone.[22]

[21] The specification does not control for firm fixed effects because a firm i has different weights at each point in time depending on the development of previous confounders. To eliminate the unobserved firm-specific effects, the variables are hence transformed to log differences.

[22] We also look at the probability of exporting, see Table B.8 in Appendix. Results show that the

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	-0.0698	-0.157***	-0.148***	-0.0712***
	(0.124)	(0.0393)	(0.0428)	(0.0177)
Age 2005	0.00570	-0.00514***	-0.00893***	-0.000786
	(0.00334)	(0.000768)	(0.00178)	(0.00108)
Service 2005	0.683***	-0.330***	0.129	0.0752
	(0.107)	(0.0761)	(0.0899)	(0.0729)
Manufacturing 2005	0.792***	-0.260***	0.158	-0.0651
	(0.117)	(0.0630)	(0.0972)	(0.0517)
Foreign ownership 2005	-0.135**	0.0789***	0.0586	-0.000795
	(0.0662)	(0.0244)	(0.0350)	(0.0163)
Constant	0.0468	0.689***	0.455***	0.199***
	(0.0865)	(0.0675)	(0.128)	(0.0718)
N	4395	14808	10090	6727
Year 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 code. The weights are derived from the logistic regression using Equation 3. Standard errors are clustered at the firm level.

As pointed out above, a potential mechanism that would explain negative externalities is the movement of workers from firms outside of SEZs to those located inside. Unfortunately, Prowess does not report comprehensive data on the amount of labour input used in firms, let alone movement of workers between firms. Still, in order to supplement our results thus far, we use the total wage bill as a measure of labour input and investigate how this changes in response to the establishment of SEZs. The results for using changes in the total wage bill as a dependent variable are reported in Table 7. We find that conditional on controlling for selection, firms inside the zones increase their total wage bill, however, firms in the vicinity experience a decline in the growth of paid wages. Most affected are firms in the 10-15 kilometers distance band, whose wage growth decreased by 29%. This finding is thus in line with the labour reallocation argument advanced above: due to received benefits, more productive firms in SEZs hire new workers, increasing their total wage bill. Some of these new hires may come from firms in the vicinity, reducing their total labour input. Unfortunately, with the data at hand, we are not able to dig deeper into this to actually see whether workers move between firms.[23]

Finally, we consider the growth of spending on intermediate inputs as the outcome variable. Similar to the argument on labour reallocation, SEZ-firms may increase their

likelihood of exporting for within SEZ firms declined as a result of SEZs. Neighbouring firms are also less likely to export, though the coefficient is only significant for firms in the 5-10 kilometers distance band.

[23]The results are also consistent with firms merely passing negative TFP effects onto their workers in terms of reduced wages. Given the data at hand, we are unfortunately not able to discriminate between these two intuitively equally plausible interpretations.

Table 4: Time-varying treatment effect of SEZs on TFP growth of more and less productive firms.

	Above median TFP growth				Below median TFP growth			
	(inside)	(0-5km)	(5-10km)	(10-15km)	(inside)	(0-5km)	(5-10km)	(10-15km)
SEZs	0.347***	-0.0835**	0.00166	-0.0261	-0.160	0.114**	0.218***	0.128***
	(0.119)	(0.0392)	(0.0115)	(0.0221)	(0.0900)	(0.0471)	(0.0711)	(0.0239)
Age 2005	-0.00371	-0.00210	-0.000771	0.00111	0.00806**	0.00543**	0.00766***	-0.000301
	(0.00532)	(0.00127)	(0.000402)	(0.000733)	(0.00354)	(0.00245)	(0.00273)	(0.00121)
Service 2005	0.744***	-0.139**	-0.0401	-0.0608	0.181	-0.00132	-0.0976	-0.121
	(0.118)	(0.0579)	(0.0255)	(0.0519)	(0.125)	(0.0379)	(0.0625)	(0.0677)
Manufacturing 2005	0.883***	-0.196***	-0.0284	-0.0682	0.203**	-0.00447	-0.0671	-0.0365
	(0.106)	(0.0692)	(0.0214)	(0.0442)	(0.0855)	(0.0676)	(0.0587)	(0.0447)
Foreign ownership 2005	-0.165***	0.0541**	-0.00348	0.0126	-0.0386	-0.0548**	-0.0229	-0.0891
	(0.0447)	(0.0242)	(0.0211)	(0.0293)	(0.0781)	(0.0234)	(0.0341)	(0.117)
N	2130	7109	4915	3252	2265	7699	5175	3475
Year Fixed Effects	Yes	Yes	Yes	Yes	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 code. The weights are derived from the logistic regression using Equation 3. Standard errors are clustered at the firm level.

material spending as a result of the provided incentives, particularly duty free import and domestic procurement of goods, in an attempt to increase the output. This additional demand for intermediate inputs, if satisfied by local suppliers, may imply reductions in intermediate input purchases by firms in the vicinity. Results of the estimation are reported in Table 8. Conditional on initial selection, there is a statistically significant increase in adjusted intermediate input spending for firms inside SEZs. Neighbouring firms exhibit a negative effect, however, it is significant only for firms in the 10-15 kilometers distance band, results which closely parallel the negative effects on the growth of wage bill. While this is in line with the idea of attracting intermediate inputs away from firms in the vicinity, our data do not allow us to be thorough enough to be able to clearly conclude on this.

To sum up, our results thus far indicate that conditional on controlling for initial selection, there is no further significant increase in productivity growth of within SEZ firms as a result of SEZs establishment, on average. However, this effect is mostly driven by relatively less productive firms, whose TFP growth declined compared to more productive firms, which enjoyed significant productivity gains. Firms in the vicinity of the SEZs experience, on average, reductions in TFP growth and export activity, however. These negative externalities attenuate with distance. The negative TFP effects are not accompanied by negative effects on sales, but go together with lower growth in labour and intermediate input use. While these latter findings are consistent with the idea that workers and suppliers of intermediate inputs move from non-SEZ to SEZ firms - implying a deteriorating productivity performance - our data are not exhaustive enough to come to a final conclusion.

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	0.427	0.0576	-0.0979	0.0952
	(0.251)	(0.134)	(0.0911)	(0.0832)
Age 2005	0.00699	0.00159	0.00335**	-0.00442
	(0.00675)	(0.00366)	(0.00140)	(0.00326)
Service 2005	1.221***	-0.0324	0.00512	0.302***
	(0.304)	(0.366)	(0.122)	(0.0970)
Manufacturing 2005	1.274***	0.195	-0.0407	0.263***
	(0.170)	(0.405)	(0.151)	(0.0940)
Foreign ownership 2005	-0.0444	-0.0143	0.0568	-0.0411
	(0.109)	(0.0779)	(0.0431)	(0.0751)
Constant	-1.677***	-0.439	-0.272	-0.206
	(0.177)	(0.476)	(0.189)	(0.175)
N	4668	15316	10441	6976
Year 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 code. The weights are derived from the logistic regression using Equation 3. Standard errors are clustered at the firm level.

Table 6: Time-varying treatment effect of SEZs on the growth rate of export earnings.

	(inside) Exports growth	(0-5km) Exports growth	(5-10km) Exports growth	(10-15km) Exports growth
SEZs	0.0464	-0.0832**	-0.0713	-0.460**
	(0.0701)	(0.0413)	(0.0860)	(0.228)
Age 2005	0.000817	-0.00410	-0.00255	0.00471
	(0.00143)	(0.00228)	(0.00519)	(0.00672)
Service 2005	-0.279	-0.275**	-0.00727	-0.563***
	(0.277)	(0.127)	(0.106)	(0.201)
Manufacturing 2005	-0.0868	0.0751	0.0685	-0.219
	(0.0586)	(0.0642)	(0.105)	(0.203)
Foreign ownership 2005	0.0436	0.109**	0.0963	0.156
	(0.0827)	(0.0447)	(0.0577)	(0.129)
Constant	0.0339	0.175	0.139	-0.0782
	(0.0705)	(0.132)	(0.271)	(0.189)
N	1507	6241	3923	2561
Year 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 the sum of earnings from exports of goods and services. 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 code. The weights are derived from the logistic regression using Equation 3. Standard errors are clustered at the firm level.

Table 7: Time-varying treatment effect of SEZs on wage growth.

	(inside) Wage growth	(0-5km) Wage growth	(5-10km) Wage growth	(10-15km) Wage growth
SEZs	0.0812**	-0.0172	0.0314	-0.251***
	(0.0334)	(0.0211)	(0.0255)	(0.0596)
Age 2005	0.00597	0.000956	0.000766	-0.000327
	(0.00310)	(0.000776)	(0.000609)	(0.00158)
Service 2005	-0.0111	-0.0840***	-0.0198	-0.426***
	(0.0660)	(0.0314)	(0.0357)	(0.0800)
Manufacturing 2005	0.00244	-0.0184	-0.00433	-0.0292
	(0.0555)	(0.0200)	(0.0334)	(0.0472)
Foreign ownership 2005	-0.00195	0.0138	0.118	0.346**
	(0.0155)	(0.0231)	(0.0941)	(0.146)
Constant	-0.243**	-0.0454	-0.0640**	0.0263
	(0.102)	(0.0269)	(0.0293)	(0.0812)
N	6500	24569	16146	10502
Year 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, 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 code. The weights are derived from the logistic regression using Equation 3. Standard errors are clustered at the firm level.

Table 8: Time-varying treatment effect of SEZs on scaled material spending growth.

	(inside) Material expenses/sales growth	(0-5km) Material expenses/sales growth	(5-10km) Material expenses/sales growth	(10-15km) Material expenses/sales growth
SEZs	0.226***	-0.0162	0.00908	-0.267***
	(0.0855)	(0.0433)	(0.0270)	(0.0898)
Age 2005	0.00334	-0.000114	-0.000317	-0.00912***
	(0.00351)	(0.000887)	(0.000501)	(0.00141)
Service 2005	-0.0863	0.285	-0.0543	0.408
	(0.130)	(0.262)	(0.0574)	(0.238)
Manufacturing 2005	-0.175**	0.334***	0.00510	0.378
	(0.0870)	(0.0588)	(0.0233)	(0.234)
Foreign ownership 2005	0.112**	-0.0110	-0.0116	-0.142**
	(0.0521)	(0.0329)	(0.0241)	(0.0685)
Constant	-0.749***	-0.786***	0.0368	0.337
	(0.127)	(0.0310)	(0.0455)	(0.259)
N	4393	14760	10089	6681
Year Fixed Effects	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

Note: Material spending is defined as the sum of raw material and stores & spares expenditure. The outcome variable is measured as the log difference of the growth of scaled material spending to 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 code. The weights are derived from the logistic regression using Equation 3. Standard errors are clustered at the firm level.

6.3 Treatment intensity

As we increase the radius of the spatial circles around SEZs, one firm falls within multiple SEZs and, therefore, may have differential effects because it is able to absorb the benefits of (or be hurt by) more than one SEZ. To check the intensive margin of the effect, 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 15 kilometer radius. If there is no other SEZ in the vicinity, then "SEZ intensity" is zero. We limit the distance to 15 kilometers since there are no spillover effects of SEZ beyond the 10-15 kilometers distance band. 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 9. Whereas the negative effects from the establishment of SEZs remain, there is an additional negative effect of other SEZs within 15 kilometers. Intuitively, given the found negative effects of SEZs, it is expected that the presence of multiple zones will intensify the adverse effects. Thus, we also observe a negative sign of SEZ intensity, though the coefficients are small in magnitude.

Table 9: Treatment intensity results.

	(inside) TFP growth	(0-5km) TFP growth	(5-10km) TFP growth	(10-15km) TFP growth
SEZs	-0.255	-0.149***	-0.135***	0.0288
	(0.154)	(0.0421)	(0.0434)	(0.0320)
SEZ intensity	0.0129	-0.00406	-0.00766	-0.0532***
	(0.00976)	(0.00396)	(0.0155)	(0.0161)
Age 2005	0.00725***	-0.00515***	-0.00893***	-0.000219
	(0.00274)	(0.000767)	(0.00178)	(0.000837)
Service 2005	0.650***	-0.332***	0.130	0.0661
	(0.121)	(0.0768)	(0.0901)	(0.0693)
Manufacturing 2005	0.750***	-0.260***	0.155	-0.0876
	(0.129)	(0.0629)	(0.0991)	(0.0463)
Foreign ownership 2005	-0.112	0.0805***	0.0602	-0.00140
	(0.0693)	(0.0240)	(0.0353)	(0.0156)
Constant	0.00401	0.689***	0.457***	0.191***
	(0.0812)	(0.0675)	(0.129)	(0.0640)
N	4395	14808	10090	6727
Year 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 and state code in 2005. The weights are derived from the logistic regression using Equation 3. Standard errors are clustered at the firm level.

6.4 Heterogeneity analysis

The differences in absorptive capacity of neighbouring firms may result in heterogeneous effects of SEZs establishment. Ownership, size, age, industry or importer status 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 3. 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) the third dummy variable equals one for firms with age above the sample median in 2005 and zero otherwise, (v) separate dummy variables for firms operating in manufacturing or service sectors, and (vi) the dummy variable for importer firm.[24] Then, Equation 4 is estimated for each sub-sample including weights. To save space, we focus only on the productivity growth of firms to shed some light on which type of firm is the most affected.

Table 10 depicts the results. As regards externalities, it appears that domestic firms, large firms, manufacturing firms and non-importer firms seem to be the main losers of the program implementation, showing significant downward productivity growth effects. Interestingly, looking at firms inside SEZs we find that manufacturing firms experience a significant decline in TFP growth following the SEZ implementation. This may be attributed to the fact that sales from SEZ firms to the domestic market are subjected to import tariffs, which decreases the comparative advantage relative to other domestic manufacturing firms. Additionally, as we have seen above, export sales also did not increase but rather decreased for inside SEZ firms, which hence is reflected in declining productivity growth. The effects do not appear to be influenced in any clear way by firm age.

Secondly, given the predominant number of IT and other technologically intensive SEZs, we estimate the effect separately from solely Hi-tech 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 presented in Table 11 indicate that conditional on controlling for initial selection, the establishment of Hi-tech SEZs led to a significant decline in TFP growth for directly affected firms, whose productivity growth decreased by 8%. Similar to the baseline results, the induced negative externalities are more pronounced for firms in the 0-5 kilometers distance band, whose productivity growth declined by 18%. The effects decay with distance but still remain negative and significant. The same results hold when estimating the effect solely from IT SEZs.

Alkon (2018) demonstrates that SEZs in India did not bring local socioeconomic development and argues that the mechanism underlying the inefficiency of SEZs is excessive 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 the 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 test this assumption, 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.[25] Then, the empirical model is re-estimated for each distance band and the results are reported in Table 12. The productivity growth of firms inside the zones has a positive sign but still remains insignificant. The negative

[24]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, 41% are small-size firms, 46% are young firms, 56% operate in the manufacturing sector and 33% in the service sector, and 33% are importers.

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

externalities on the neighbouring firms are still present, though significant only for the 0-5 kilometers distance band. Thus, we do not find evidence that the presence of state-owned industrial development corporations is the main driver underlying the inefficiency of SEZs.

Table 10: 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	0.00261	0.0102	-0.0307	0.0168
	(0.0299)	(0.0176)	(0.0190)	(0.0216)
SEZs - other domestic firms	-0.0722	-0.159***	-0.150***	-0.0717***
	(0.126)	(0.0396)	(0.0430)	(0.0182)
Panel B: Firm size				
SEZs - large-size firms	-0.186	-0.172***	-0.146***	-0.0737***
	(0.112)	(0.0444)	(0.0464)	(0.0204)
SEZs - small-size firms	0.0291	0.0550	0.0576	-0.0264
	(0.0283)	(0.0458)	(0.0321)	(0.0209)
Panel C: Firm age				
SEZs - old firms	0.0137	-0.0516***	-0.119	-0.0494
	(0.0178)	(0.0177)	(0.0642)	(0.0263)
SEZs - young firms	0.0992	-0.0303	0.0559	0.0143
	(0.0508)	(0.0342)	(0.0384)	(0.0749)
Panel D: Sector				
SEZs - manufacturing	-0.232**	-0.0732***	-0.185***	-0.0735***
	(0.116)	(0.0247)	(0.0377)	(0.0207)
SEZs - services	0.162	-0.0124	0.0155	-0.0603
	(0.0983)	(0.0259)	(0.0402)	(0.0353)
Panel E: Importer				
SEZs - importer	0.0426	0.00249	0.0131	0.0189
	(0.0260)	(0.0390)	(0.0128)	(0.0186)
SEZs - non-importer	-0.203	-0.136***	-0.115**	-0.0851***
	(0.151)	(0.0501)	(0.0447)	(0.0321)

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 code. The weights are derived from the logistic regression using Equation 3. Standard errors are clustered at the firm level. After the weights are estimated, the Equation 4 is re-estimated for different samples. Large-size firms are firms with log of sales above the sample median in 2005. Young firms are firms with the age below the sample median in 2005. All specifications control for year fixed effects.

6.5 Sensitivity analysis

In order to establish the robustness of our results thus far, we perform a number of checks: (i) use a winsorized sample at the 1st and the 99th percentile, (ii) use an alternative measure of TFP growth, (iii) use an alternative control group, (iv) use alternative distance bands, (v) exclude firms in SEZs established prior to the 2005 SEZs Act, and (vi) use an alternative definition of foreign ownership.

Winsorized sample. - To verify that the results are not driven by some outliers, we re-estimate our baseline specification but winsorize the sample at the 1^{st} and the 99^{th} percentile. The results presented in Table B.9 are comparable to the main results, showing no significant impact on the directly affected firms, conditional on selection, and induced negative externalities on the neighbouring firms. The adverse effect is significant for firms in the 10-15 kilometers distance band.

Table 11: Heterogeneity effect from Hi-tech SEZs.

	(inside) TFP growth	(0-5km) TFP growth	(5-10km) TFP growth	(10-15km) TFP growth
SEZs	-0.0779**	-0.162***	-0.142**	-0.0455**
	(0.0309)	(0.0342)	(0.0581)	(0.0205)
Age 2005	0.000401	-0.00460***	-0.00528	-0.00281***
	(0.000858)	(0.000861)	(0.00284)	(0.000765)
Service 2005	0.00775	-0.439***	0.446***	-0.0601
	(0.0480)	(0.0663)	(0.136)	(0.0732)
Manufacturing 2005	0.0477	-0.316***	0.478***	0.111
	(0.0343)	(0.0366)	(0.140)	(0.0847)
Foreign ownership 2005	0.119***	0.0607**	0.0504	-0.0192
	(0.0241)	(0.0246)	(0.0368)	(0.0232)
Constant	0.0132	0.722***	-0.0585	0.0526
	(0.0528)	(0.0322)	(0.164)	(0.0732)
N	4248	13187	8575	4953
Year 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 code. The weights are derived from the logistic regression using Equation 3. Standard errors are clustered at the firm level.

Alternative TFP measures. - For the sake of completeness and to corroborate the robustness of the results to alternative TFP measure, we use an alternative approach for productivity estimation, namely Wooldridge (2009). Table B.10 reports the estimation results. There is a significant negative effect on the TFP growth of inside SEZ firms, whose productivity growth declined by 33%. Neighbouring firms within 5 kilometers also experience a significant drop in their performance, accounting for 8% decline. The sign of the coefficient for the 5-10 kilometers distance band remains negative but turns insignificant.

Alternative control group. - Since the choice of the control group - firms located further than 40 kilometers away from the zones - is somehow arbitrary, we, in addition, present the estimation results using an alternative control group, namely firms that are located further than 30 kilometers away from the zones.[26] The rest of the analysis is identical to the one used for estimating Table 3. The results presented in Table B.11 are consistent with previous findings, indicating no significant impact on within SEZ firms, conditional on controlling for initial selection, and negative externalities for the neighbouring firms.

Alternative distance bands. - To verify 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

[26] 1974 firms are located further than 40 kilometers away and 2315 firms are located further than 30 kilometers away, and 1876 firms are located further than 45 kilometers from SEZs. Given the decreasing number of observations as we increase the threshold, we choose to use 30 kilometers.

Table 12: 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	0.0566	-0.0973***	-0.0201	0.0188
	(0.0525)	(0.0323)	(0.0272)	(0.0340)
Age 2005	0.00338***	-0.00413***	-0.000759	-0.00128***
	(0.00115)	(0.00104)	(0.00107)	(0.000476)
Service 2005	0.392***	-0.385***	0.0658	-0.0121
	(0.116)	(0.0847)	(0.0820)	(0.0273)
Manufacturing 2005	0.473***	-0.216**	0.0949	0.000151
	(0.125)	(0.0944)	(0.0926)	(0.0261)
Foreign ownership 2005	0.0264	0.0676***	-0.00324	0.00604
	(0.0275)	(0.0200)	(0.0157)	(0.0185)
Constant	0.536***	0.555***	0.0574	0.122***
	(0.164)	(0.106)	(0.126)	(0.0400)
N	4187	14061	9700	6658
Year 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 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 code. 35 SEZs located in state-owned industrial development corporations are excluded. The weights are derived from the logistic regression using Equation 3. Standard errors are clustered at the firm level.

distance bands: $inside$, $0-7km$, $7-14km$, and $14-21km$ with the control group being firms located further than 40 kilometers away. Table B.7 illustrates the number of treated firms inside each distance band. The results presented in Table B.12 are robust to an alternative definition of distance bands. Interestingly, we observe that the negative externalities are very localized and extend up to 14 kilometers from SEZs since the coefficient loses its significance for firms in the 14-21 kilometers distance band.

Excluding firms in SEZs established before the 2005 Act. - 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 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 B.13 we exclude those firms and observe that the results are not affected.

Alternative definition of foreign ownership. - In our baseline regression, we defined the firm as foreign-owned if the proportion of shares held by foreign individuals, corporate bodies or institutions is greater than 25%. As a robustness check, we define the company as foreign-owned if the majority of equity shares are held by foreigners, that is, we move the threshold is 50%. Results presented in Table B.14 are robust to an alternative definition.

7 Conclusion

The incentives brought by the SEZs Act pose a natural question on 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 firms as a result of agglomeration economies. The main novelty of this paper is that firms are used as a unit of analysis which enables the granular estimation of the effects and identification of the distance threshold of spillovers. Our contribution is therefore the creation of a 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, on average, no further productivity gains for within SEZ firms. However, this effect is predominantly driven by less productive firms, whereas relatively more productive firms experienced significant productivity gains. Additionally, the establishment of SEZs appears to have created negative externalities on the neighbouring non-SEZ firms. These negative effects attenuate with distance. A potential underlying mechanism is that more productive SEZ firms divert raw material inputs and labor from firms in the neighbourhood. The adverse effect is more pronounced for neighbouring domestic firms, large firms, manufacturing firms and non-importer firms.

Though these results are striking and opposite to the generated agglomeration economies in China, it would be interesting to delve deeper into the incentive scheme of the program to be able to identify the reasons why the provided benefits, on average, resulted in no further productivity gains for within SEZ firms and moreover, induced negative effects on the firms in the neighbourhood. The taxation scheme, the lack of absorptive capacity of the adjacent firms as well as the insufficiency of linkages between SEZ and non-SEZ firms can potentially explain the negative externalities. Secondly, it would be interesting to perform a cost-benefit analysis given the adverse effects of SEZs to evaluate the general welfare benefits of the policy. Lastly, it would be interesting to investigate worker mobility across SEZ and non-SEZ firms as a channel explaining the spillover effects. Given the low levels of across-state migration in India, it would be necessary to collect linked employer-employee data to uncover the linkages between firms inside the zones and in the immediate proximity.

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 other countries. 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: Figures

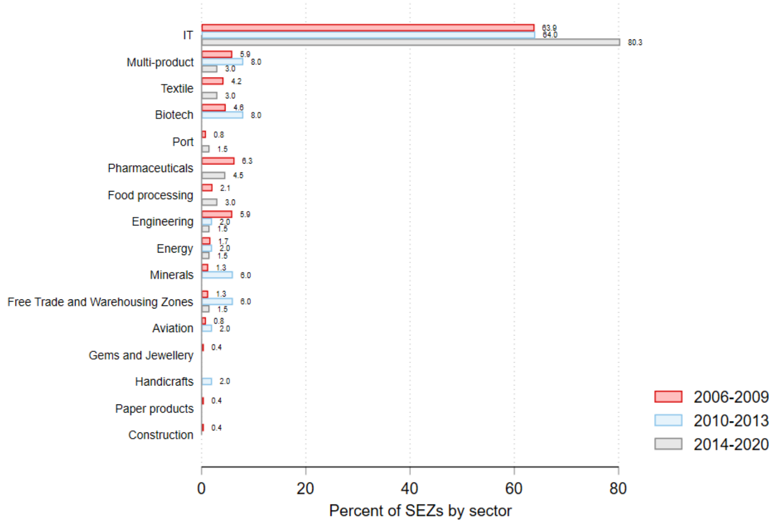

Figure A.1: Sector-wise distribution of SEZs over time.

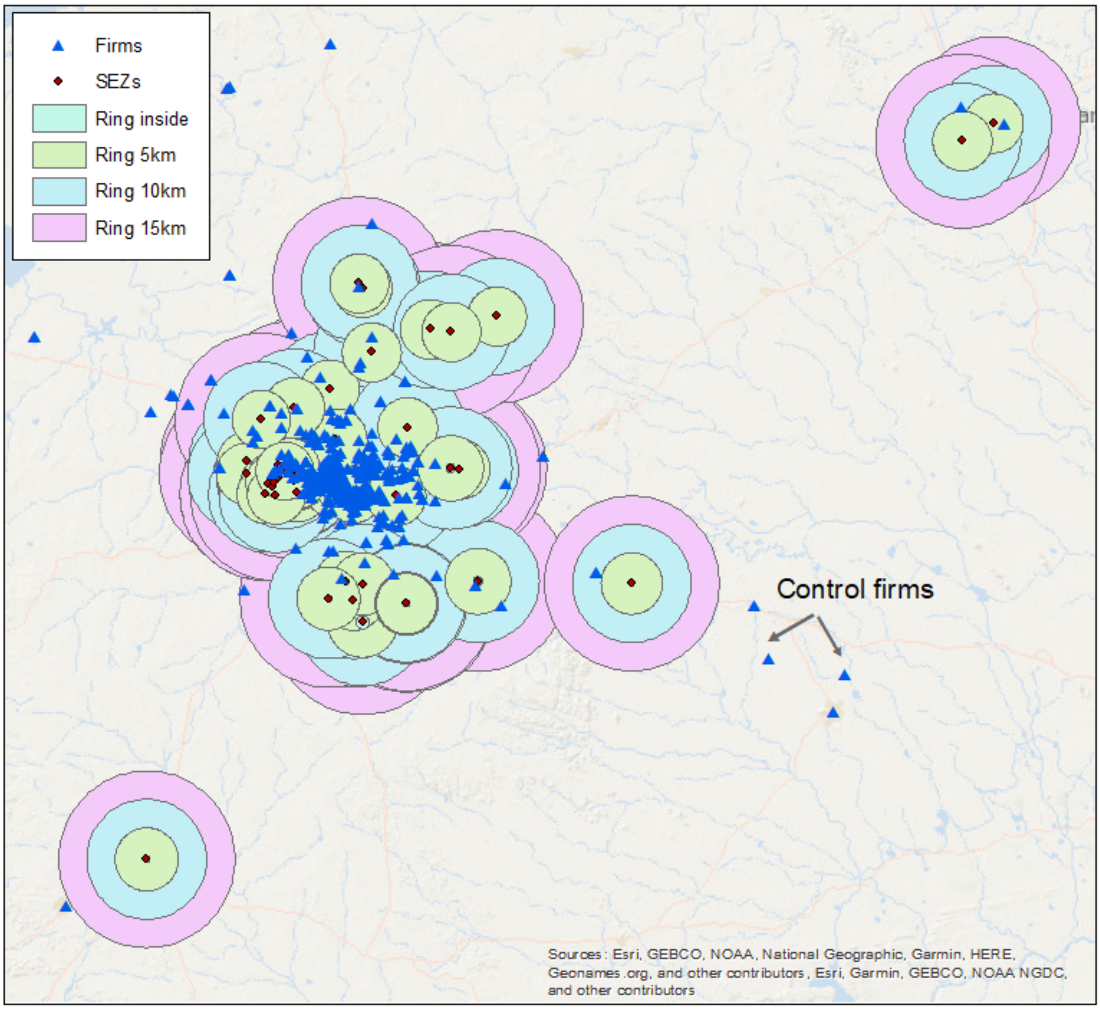

Figure A.2: This figure illustrates the mapping of firms in the SEZs. The blue triangles represent geocoded firms. The red dots represent geocoded SEZs. Using the information on the zones' area, a radius is created and subsequently increased by 5 km. Buffers of various sizes are created around the centroid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

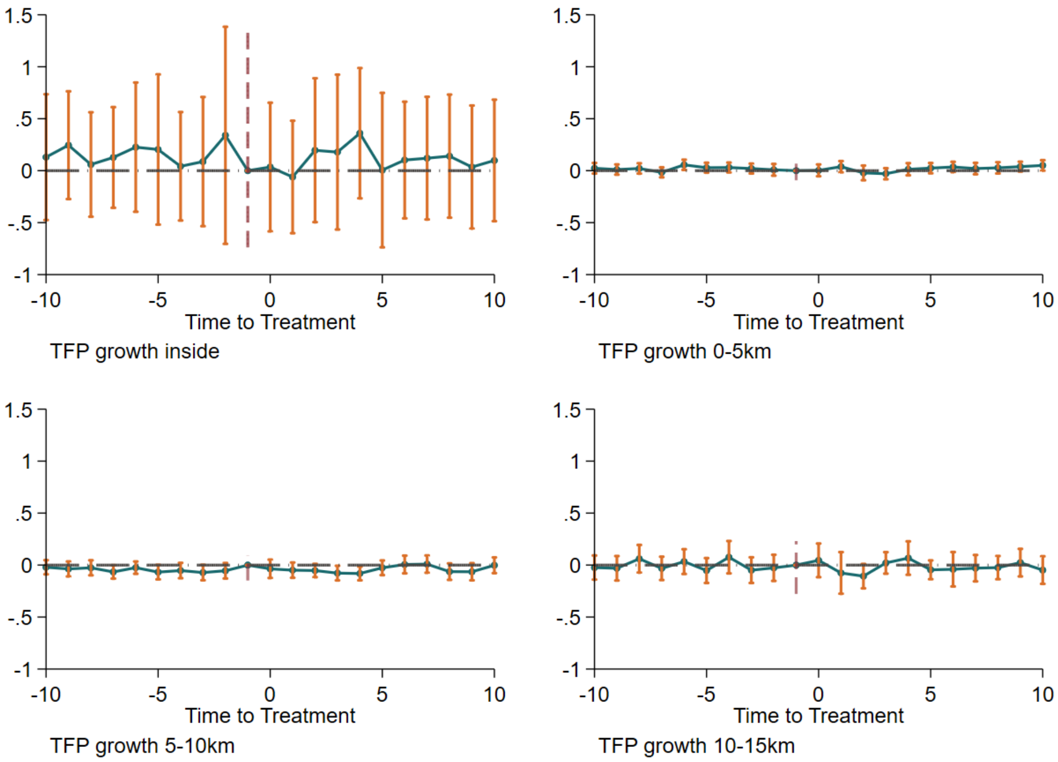

Figure A.3: Event study graph for TFP growth. 95% confidence interval is reported. The sample of treated firms is balanced for the event window [-10, 10].

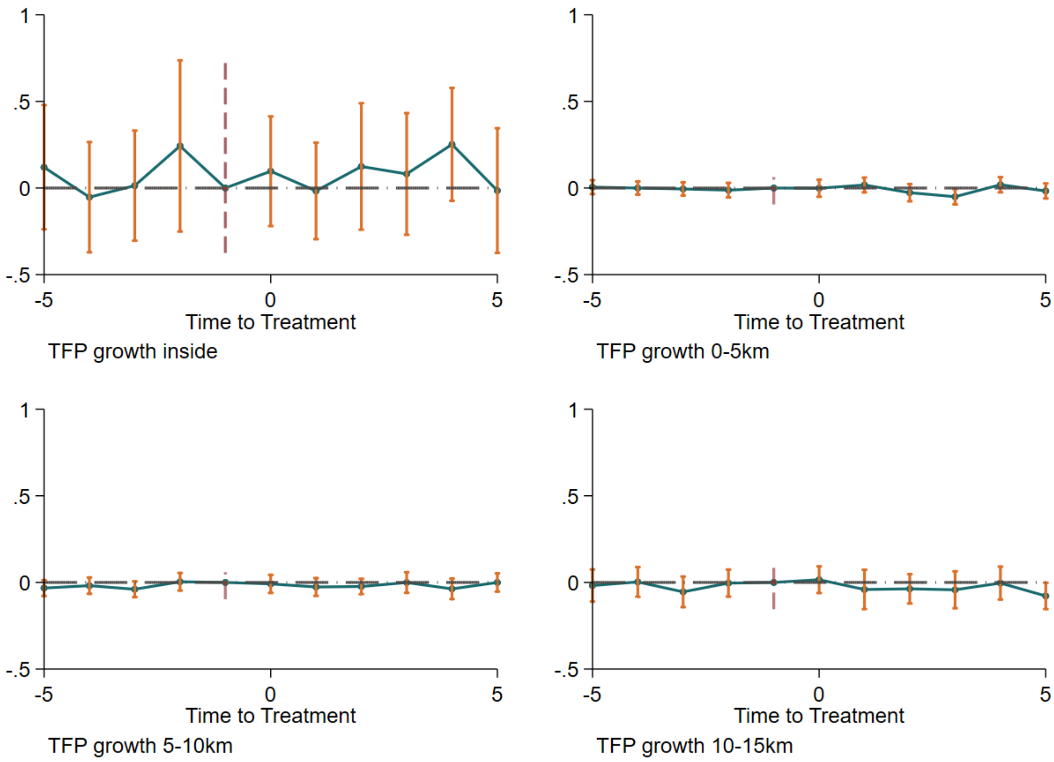

Figure A.4: Event study graph for TFP growth. 95% confidence interval is reported. The sample of treated firms is balanced for the event window [-5, 5].

B Appendix: Tables

Table B.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 B.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 B.3: Summary statistics of firms by industry.

	Frequency	Percent	Cum. percent
Crop & animal production	395	2.343	2.343
Forestry & logging	408	2.420	4.763
Fishing & aquaculture	3	0.0178	4.781
Mining of coal & lignite	9	0.0534	4.834
Extraction of crude petroleum & natural gas	16	0.0949	4.929
Mining of metal ores	23	0.136	5.066
Other Mining & quarrying	96	0.569	5.635
Food	650	3.856	9.491
Beverages	149	0.884	10.37
Tobacco	16	0.0949	10.47
Textiles	568	3.369	13.84
Wearing apparel	35	0.208	14.05
Leather	61	0.362	14.41
Wood	40	0.237	14.65
Paper	185	1.097	15.74
Printing & reproduction of recorded media	14	0.0830	15.83
Coke & refined petroleum products	68	0.403	16.23
Chemicals	661	3.921	20.15
Pharmaceuticals	290	1.720	21.87
Rubber & plastics products	363	2.153	24.02
Other non-metallic mineral	243	1.441	25.47
Basic metals	587	3.482	28.95
Fabricated metal products	240	1.424	30.37
Computer, electronic & optical products	244	1.447	31.82
Electrical equipment	354	2.100	33.92
Machinery & equipment	401	2.379	36.30
Motor vehicles	230	1.364	37.66
Other transport equipment	41	0.243	37.90
Furniture	11	0.0653	37.97
Other manufacturing	105	0.623	38.59
Electricity, gas etc. supply	163	0.967	39.56
Water collection	1	0.00593	39.57
Construction of buildings	408	2.420	41.99
Civil engineering	304	1.803	43.79
Specialized construction activities	18	0.107	43.90
Wholesale & retail trade & repair of motorvehicles and motorcycles	58	0.344	44.24
Other wholesale & retail trade	2470	14.65	58.89
Retail trade	81	0.480	59.37
Land transport & transport via pipelines	39	0.231	59.60
Air transport	18	0.107	59.71
Warehousing & support activities for transportation	142	0.842	60.55
Postal and courier activities	10	0.0593	60.61
Accommodation	249	1.477	62.09
Food and beverage service activities	3	0.0178	62.11
Publishing activities	55	0.326	62.43
Music publishing activities	60	0.356	62.79
Programming and broadcasting activities	2	0.0119	62.80
Telecommunications	82	0.486	63.29
Computer programming	453	2.687	65.97
Information service activities	63	0.374	66.35
Financial service activities	4620	27.41	93.75
Insurance, reinsurance and pension funding	3	0.0178	93.77
Other financial activities	297	1.762	95.53
Real estate activities	3	0.0178	95.55
Activities of head offices	216	1.281	96.83
Architecture & engineering activities	102	0.605	97.44
Scientific research & development	2	0.0119	97.45
Advertising & market research	60	0.356	97.81
Other scientific activities	6	0.0356	97.84
Employment activities	20	0.119	97.96
Travel agency etc. activities	29	0.172	98.13
Office administrative etc. activities	81	0.480	98.61
Public administration & defence	7	0.0415	98.65
Education	28	0.166	98.82
Residential care activities	109	0.647	99.47
Sports activities	26	0.154	99.62
Activities of membership organizations	42	0.249	99.87
Repair of computers	22	0.131	100
Total	16858	100	

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

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

Table B.5: Number of treated firms for a balanced sample of 10 years before and after the treatment for each distance band.

	inside	0-5km	5-10km	10-15km
Number of firms	25	564	307	160

Table B.6: Number of treated firms for a balanced sample of 5 years before and after the treatment for each distance band.

	inside	0-5km	5-10km	10-15km
Number of firms	70	1332	770	522

Table B.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 B.8: Time-varying treatment effect of SEZs on export probability.

	(inside) Exporter dummy	(0-5km) Exporter dummy	(5-10km) Exporter dummy	(10-15km) Exporter dummy
SEZs	-0.204***	-0.127	-0.177***	-0.118
	(0.0766)	(0.0843)	(0.0523)	(0.0806)
Age 2005	0.00475***	-0.0127**	-0.00228**	0.00349
	(0.00135)	(0.00555)	(0.00104)	(0.00244)
Service 2005	0.202***	0.197	-0.0182	0.0352
	(0.0773)	(0.169)	(0.0760)	(0.0756)
Manufacturing 2005	0.258***	0.330	-0.0476	0.122**
	(0.0798)	(0.215)	(0.0651)	(0.0540)
Foreign ownership 2005	0.459***	0.406***	0.359***	-0.196
	(0.0804)	(0.0871)	(0.0436)	(0.103)
Constant	-0.106**	0.504**	0.407***	-0.0349
	(0.0438)	(0.251)	(0.0967)	(0.118)
N	4761	15596	10637	7147
Year 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 the sum of earnings from exports of goods and services. 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 code. The weights are derived from the logistic regression using Equation 3. Standard errors are clustered at the firm level.

Table B.9: Robustness check for a winsorized sample at the 1st and the 99th percentile.

	(inside) TFP growth	(0-5km) TFP growth	(5-10km) TFP growth	(10-15km) TFP growth
SEZs	0.0910	-0.103	-0.0969	-0.0569***
	(0.0772)	(0.0545)	(0.0522)	(0.0220)
Age 2005	0.00222	-0.00333**	-0.00433	-0.000937
	(0.00262)	(0.00159)	(0.00262)	(0.000883)
Service 2005	0.377***	-0.232***	0.0324	0.105
	(0.0829)	(0.0851)	(0.0406)	(0.0543)
Manufacturing 2005	0.442***	-0.206**	0.0445	-0.0276
	(0.0894)	(0.0935)	(0.0407)	(0.0420)
Foreign ownership 2005	-0.0239	0.0620**	0.0364	-0.00342
	(0.0304)	(0.0292)	(0.0304)	(0.0150)
Constant	-0.164	0.461***	0.263**	0.0974**
	(0.0841)	(0.130)	(0.128)	(0.0484)
N	4395	14808	10090	6727
Year Fixed Effects	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

Note: Variables are winsorized at the 1st and the 99th percentile. 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 code. The weights are derived from the logistic regression using Equation 3. Standard errors are clustered at the firm level.

Table B.10: Robustness check using alternative TFP measure: Wooldridge (2009) approach.

	(inside) TFP growth WRDG	(0-5km) TFP growth WRDG	(5-10km) TFP growth WRDG	(10-15km) TFP growth WRDG
SEZs	-0.285***	-0.0739***	-0.0635	0.0106
	(0.0764)	(0.0232)	(0.0389)	(0.0480)
Age 2005	0.0142***	-0.00246***	-0.00320**	-0.00213
	(0.00189)	(0.000908)	(0.00125)	(0.00221)
Service 2005	0.927***	-0.373***	-0.00906	-0.0118
	(0.0903)	(0.0538)	(0.0370)	(0.0915)
Manufacturing 2005	1.026***	-0.307***	-0.00470	-0.0992
	(0.0805)	(0.0557)	(0.0387)	(0.0859)
Foreign ownership 2005	-0.279***	0.0317	0.00962	-0.00264
	(0.0393)	(0.0166)	(0.0252)	(0.0329)
Constant	0.0780	0.405***	0.139**	0.137
	(0.0701)	(0.0526)	(0.0682)	(0.139)
N	4395	14808	10090	6727
Year 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 Wooldridge (2009) approach. Time-invariant covariates include age, a foreign ownership dummy, dummies for manufacturing and service industries in 2005 and state code. The weights are derived from the logistic regression using Equation 3. Standard errors are clustered at the firm level.

Table B.11: Robustness check using an alternative control group - firms located further than 30 kilometers away.

	(inside) TFP growth	(0-5km) TFP growth	(5-10km) TFP growth	(10-15km) TFP growth
SEZs	0.0688	-0.123**	-0.0845	-0.0945***
	(0.186)	(0.0513)	(0.0633)	(0.0257)
Age 2005	0.00639	-0.00327***	-0.00250	-0.00140
	(0.00798)	(0.000643)	(0.00297)	(0.00104)
Service 2005	0.509***	-0.219**	0.239	0.111**
	(0.164)	(0.107)	(0.135)	(0.0516)
Manufacturing 2005	0.592***	-0.178	0.249	-0.0450
	(0.160)	(0.127)	(0.141)	(0.0324)
Foreign ownership 2005	-0.0845	0.0366	0.0124	-0.0263
	(0.0696)	(0.0244)	(0.0302)	(0.0148)
Constant	0.165	0.478***	-0.0128	0.207***
	(0.224)	(0.152)	(0.206)	(0.0538)
N	5185	15633	10915	7552
Year 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 code. The weights are derived from the logistic regression using Equation 3. Standard errors are clustered at the firm level. Control group is restricted to firms located further than 30 kilometers away from the zones.

Table B.12: Robustness check using alternative distance bands.

	(inside) TFP growth	(0-7km) TFP growth	(7-14km) TFP growth	(14-21km) TFP growth
SEZs	-0.0698	-0.123***	-0.0631***	-0.0489
	(0.124)	(0.0355)	(0.0225)	(0.0621)
Age 2005	0.00570	-0.00494***	-0.00294***	0.00343***
	(0.00334)	(0.000685)	(0.000483)	(0.000896)
Service 2005	0.683***	-0.251***	-0.0606	-0.128***
	(0.107)	(0.0560)	(0.0950)	(0.0445)
Manufacturing 2005	0.792***	-0.244***	-0.0270	-0.0752**
	(0.117)	(0.0598)	(0.101)	(0.0365)
Foreign ownership 2005	-0.135**	0.0736***	0.0357**	-0.0178
	(0.0662)	(0.0237)	(0.0178)	(0.0395)
Constant	0.0468	0.662***	0.341***	-0.0430
	(0.0865)	(0.0686)	(0.109)	(0.0558)
N	4395	17957	9509	4882
Year 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 code. The weights are derived from the logistic regression using Equation 3. Standard errors are clustered at the firm level.

Table B.13: 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	-0.0698	-0.134***	-0.144***	-0.0712***
	(0.124)	(0.0410)	(0.0426)	(0.0177)
Age 2005	0.00570	-0.00514***	-0.00867***	-0.000782
	(0.00334)	(0.000764)	(0.00182)	(0.00108)
Service 2005	0.683***	-0.233***	0.115	0.0756
	(0.107)	(0.0715)	(0.0830)	(0.0728)
Manufacturing 2005	0.792***	-0.238***	0.143	-0.0646
	(0.117)	(0.0838)	(0.0894)	(0.0515)
Foreign ownership 2005	-0.135**	0.0697***	0.0572	-0.000911
	(0.0662)	(0.0245)	(0.0344)	(0.0163)
Constant	0.0468	0.665***	0.453***	0.198***
	(0.0865)	(0.0913)	(0.123)	(0.0716)
N	4395	14701	10074	6727
Year 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 code. The weights are derived from the logistic regression using Equation 3. Standard errors are clustered at the firm level.

Table B.14: Robustness check using 50% threshold to define foreign ownership.

	(inside) TFP growth	(0-5km) TFP growth	(5-10km) TFP growth	(10-15km) TFP growth
SEZs	-0.187	-0.147***	-0.113**	-0.0735***
	(0.176)	(0.0410)	(0.0519)	(0.0176)
Age 2005	0.000627	-0.00474***	-0.00658**	-0.000799
	(0.00807)	(0.000832)	(0.00263)	(0.00102)
Service 2005	0.760***	-0.343***	0.113	0.0720
	(0.133)	(0.0739)	(0.0998)	(0.0747)
Manufacturing 2005	0.898***	-0.274***	0.148	-0.0642
	(0.163)	(0.0588)	(0.107)	(0.0504)
Foreign ownership 2005	-0.146	0.0664	0.0500	0
	(0.0902)	(0.0350)	(0.0657)	(.)
Constant	0.224	0.673***	0.332**	0.208***
	(0.218)	(0.0631)	(0.163)	(0.0703)
N	4395	14808	10090	6656
Year 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 code. The weights are derived from the logistic regression using Equation 3. Standard errors are clustered at the firm level.

C Appendix: TFP estimation

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

$$q_{it} \;=\; \beta_l \, l_{it} \,+\, \beta_k \, k_{it} \,+\, \omega_{it} \,+\, \epsilon_{it}, \tag{5}$$

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:

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

By inverting (6), productivity is expressed as:

$$\omega_{it} \;=\; h_t(l_{it}, k_{it}, m_{it}) \tag{7}$$

The estimation then proceeds in two steps. In the first stage, Equation (5) is estimated, where ω_{it} is substituted with its proxy from Equation (7). Thus, the estimation equation is as follows:

$$q_{it} \;=\; \Phi_t(l_{it}, k_{it}, m_{it}) + \epsilon_{it}, \tag{8}$$

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:

$$\omega_{it}(\beta) \;=\; \hat{\Phi}_{it} - \beta_l \, l_{it} - \beta_k \, k_{it}. \tag{9}$$

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}:

$$\omega_{it} \;=\; g_t(\omega_{it-1}) + \xi_{it}. \tag{10}$$

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:

$$E\left(\xi_{it}(\beta)\begin{pmatrix} l_{it-1} \\ k_{it} \end{pmatrix}\right) = 0 \tag{11}$$

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

$$\hat{\omega}_{it} = \hat{\Phi}_{it} - \hat{\beta}_l \, l_{it} - \hat{\beta}_k \, k_{it}. \tag{12}$$

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).

Combined industries:

- Crop & Animal production and Fishing & Aquaculture
- Mining of coal & lignite and Extraction of crude petroleum & natural gas
- Manufacturing of Food, Beverages and Tobacco
- Manufacturing of Paper and Printing & reproduction of recorded media
- Land transport & transport via pipelines and Air transport
- Music publishing activities and Telecommunications
- Computer programming and Information service activities
- Manufacturing of furniture and Other manufacturing
- Electricity, gas supply and Water collection
- Warehousing & support activities for transportation and Postal & courier activities
- Accommodation and Food service activities
- Real estate activities, Legal & accounting activities and Activities of head offices
- Scientific research & development, Advertising & market research and Other professional, scientific & technical activities
- Financial service activities, except insurance & pension funding and Other financial activities

[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 $asinh(0) = 0$.

- Employment activities, Office administrative & other business support activities, Public administration & defence, Education, Residential care activities, Sports activities & amusement and Repair of computers & personal/household goods

Additionally, alternative measures of TFP, namely the approaches of Wooldridge (2009) and Levinsohn and Petrin (2003) are calculated and presented in the correlation Table C.15.

Table C.15: 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.

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

Allcott, H. (2015). Site Selection Bias in Program Evaluation. *The Quarterly Journal of Economics*, 130(3):1117–1165.

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.

Blalock, G. and Gertler, P. J. (2009). How Firm Capabilities Affect Who Benefits from Foreign Technology. *Journal of Development Economics*, 90(2):192–199.

Bos, M. J. and Vannoorenberghe, G. (2018). Total Factor Productivity Spillovers from Trade Reforms in India. *Canadian Journal of Economics/Revue canadienne d'économique*, 51(2):549–606.

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.

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

Chaurey, R. (2017). Location-Based Tax Incentives: Evidence from India. *Journal of Public Economics*, 156:101–120.

Combes, P.-P. and Gobillon, L. (2015). The Empirics of Agglomeration Economies. In *Handbook of Regional and Urban Economics*, volume 5, pages 247–348. Elsevier.

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 Silva, D. G. and McComb, R. P. (2012). Geographic Concentration and High-Tech Firm Survival. *Regional Science and Urban Economics*, 42(4):691–701.

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

Girma, S. and Görg, H. (2021). Productivity Effects of Processing and Ordinary Export Market Entry: A Time-varying Treatments Approach. Technical report, KCG Working Paper.

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.

Griffith, R., Redding, S. J., and Simpson, H. (2002). Productivity Convergence and Foreign Ownership at the Establishment Level. *Available at SSRN 388802*.

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.

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.

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.

Mukherjee, A. and Bhardwaj, B. (2016). Imposition of MAT on SEZs: Concerns and the Way Forward.

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.

Nazarczuk, J. and Umiński, S. (2018). The Impact of Special Economic Zones on Export Behaviour: Evidence from Polish Firm-level Data. *E+ M Ekonomie a Management*, 21(3):4–22.

Neumark, D. and Simpson, H. (2015). Place-based Policies. In *Handbook of Regional and Urban Economics*, volume 5, pages 1197–1287. Elsevier.

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.

Rosenthal, S. S. and Strange, W. C. (2004). Evidence on the Nature and Sources of Agglomeration Economies. In *Handbook of Regional and Urban Economics*, volume 4, pages 2119–2171. Elsevier.

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

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, S., Sun, W., Wu, J., and Kahn, M. (2015). The Birth of Edge Cities in China: Measuring the Spillover Effects of Industrial Parks (No. w21378). *National Bureau of Economic Research.*