Place-Based Policy and Entrepreneurship: Evidence from Theme Park Openings

Chun-Yu Ho* Tingting Peng†

November 6, 2024

Click here to view most recent version.

Abstract

This paper examines the economic effects of large theme park openings in China from 2000 to 2020 using a newly compiled dataset on county-level theme parks and entrepreneurship activities. Leveraging the staggered openings of theme parks across various counties, we document three main findings. First, theme park openings lead to an approximately 14% increase in entrepreneurship activities, especially in tourism-related service sectors. This result is robust to the instrumental variable approach, the heterogeneous treatment effect, alternative specification, measurement, and selection methodology. Second, theme park openings generate spillover effects on neighboring counties within a 75-kilometer radius. Third, we identify tourism and agglomerations as the potential mechanisms driving these economic impacts. Back-of-the-envelop calculations suggest that theme park openings promote employment by 11% and overall economic activities by 2%-3%. This study sheds new light on the evaluation of the effectiveness of tourism-related place-based policies.

Keywords: Theme parks; Entrepreneurship; Tourism; Agglomeration; Regional development
JEL Classification: R10; R53; R58

*Department of Economics, University at Albany, State University of New York, cho@albany.edu
†Department of Economics, University at Albany, State University of New York, tpeng2@albany.edu

1 Introduction

Tourism was the largest service sector globally, valued at over $9 trillion, providing one in ten jobs and accounting for 10.4% of global GDP (World Bank 2022). Theme parks play a crucial role in the tourism industry, drawing visitors and investment and driving regional economic development. According to the International Association of Amusement Parks and Attractions (IAAPA), the theme park industry generated an economic impact of almost $13.9 billion in Europe in 2008, $122 billion in the United States in 2011, including $40 billion in total labor income and 1.3 million jobs, as well as $54.4 billion in regions such as the Middle East as well as part of Africa, Latin America, and the Caribbean in 2022. In an effort for local governments to attract theme park development to their jurisdictions, they offer incentives to theme park developers, i.e., a place-based policy. For example, the state governments of California and Florida offer tax breaks to Disney, and a state-owned enterprise in Shanghai provides equity investment to its local Disney park. Although it is imperative to evaluate the benefits brought by such place-based policy to local economies, little research has been done on this.

This paper examines the impacts of theme park openings on local economic development. Our analysis employs China as a case study, which provides two advantages for answering our research question. First, China has been experiencing substantial investment in the theme park industry since the first theme park was launched in 1988. By 2021, there were 156 theme parks in China. The staggered openings of theme parks across different parts of China provide a unique setting to understand the causal impacts of theme park openings on local economies.

Second, local governments in China have been actively providing subsidies and funding for theme park development as part of their place-based policies. For example, one of the leading theme park developers, Fantawild, has received over 5 billion RMB in government subsidies.[1] Moreover, many Fantawild theme parks are heavily financed by local governments, such as the Fantawild Park in Jingzhou City, where the Jingzhou Cultural Industry Investment Group holds a 95% stake, while Fantawild holds only 5%. Despite receiving sub-

[1]Source: An article titled "Theme Parks Continue to Operate at a Loss, Profits Mainly Reliant on Government Subsidies" from China Business Journal.

stantial government support, many theme parks still underperform. For instance, in 2018, major Fantawild parks reported significant losses, with Wuhu Fantawild losing 85.08 million RMB, Qingdao Fantawild 25.83 million RMB, and Shenyang Fantawild 8.95 million RMB. Additionally, the per-capita attendance in China is significantly lower than that in other developed economies, raising concerns about the benefits of developing theme parks for local economic development.[2] In response, the central government issued a guideline to curb theme park development in 2018, partly due to the rising local government debt attributed to these theme parks.[3] Given its contentiousness in policy-making, an evaluation of the economic benefits of theme park openings is particularly important.

Our empirical analysis employs a staggered difference-in-differences (DiD) design to analyze the causal impacts of theme park openings on entrepreneurship at the county level.[4] This approach allows us to compare changes in entrepreneurship over time between treated counties (those with large theme park openings) and control counties (those without openings), which addresses identification issues by accounting for unobserved time-invariant differences and common time trends in entrepreneurship. We focus on entrepreneurship as our main outcome variable because entrepreneurial activities drive economic growth (Stel et al. 2005). It is also relevant for China as small firms that typically developed from entrepreneurship are drivers of economic growth (Song et al. 2011). For constructing the treatment variable, we manually collect the openings of large and extra-large theme parks. We focus on this class of theme parks because they attract a greater number of visitors, generate higher revenue, and require more substantial infrastructure investments, making them significant contributors to regional economic development. These parks not only deliver direct economic benefits through ticket sales and on-site spending but also indirect benefits such as increased demand for local services and employment opportunities. Moreover, they are more likely to have spillover effects on the economies of neighboring regions.

Our empirical analysis first shows that theme park openings positively and significantly affect entrepreneurship. Specifically, after opening a large theme park, counties with theme park openings experience a 13.9% (= exp(0.130)-1) increase in the number of new business

[2]Source: China Theme Park Pipeline Report 2018 released by AECOM.

[3]See the press release of National Development Reform Commission (Link).

[4]We measure entrepreneurship activities using the number of new business registrations.

registrations. Although the DiD approach is able to address various identification issues, there still exist some identification concerns. There is a heterogeneous treatment timing which may bias our average treatment effect of theme park openings. We estimate our model with a stacked DiD approach as a robustness check. Also, treated counties could be different from control counties in many dimensions. We use a propensity score matching algorithm that matches treatment and control counties based on socioeconomic characteristics that could affect local entrepreneurship. Further, we apply a negative binominal method to account for our overdispersed outcome variable which has a larger standard deviation than its mean value. Encouragingly, our main results are robust to these checks.

A potential concern is that theme park openings may not be exogenous. To establish the causal effect of theme park openings on entrepreneurship, we construct a "Bartik-style" instrumental variable (Bartik 1991). Specifically, we exploit the aggregate annual variation in tourism growth, measured by the log of national travel expenditure, alongside the exogenous cross-sectional variation in a county's accessibility, which is measured by the county's distance to the nearest airports. The instrumental variable is constructed by interacting these two sources of variations. Using the control function method, we find an insignificant coefficient of the predicted residual obtained from the first stage, which suggests that endogeneity may not be a big problem. The positive and significant coefficient for theme park openings still suggests that they promote entrepreneurial activities. Additionally, we perform a placebo test to assess the possibility of a spurious time trend. In this test, we randomize the treatment timing and the treatment counties, without maintaining the original cohort structure.[5] If the placebo results align with the baseline results, it would indicate that the observed impact of theme park openings could be due to random fluctuations or pre-existing trends rather than the actual treatment. However, our results show a clear divergence between the baseline and placebo coefficients, suggesting that the baseline findings are not driven by spurious factors.

Second, we find that the positive impact of theme park openings on business creation varies across industries. Specifically, our analysis shows that tourism and travel-related

[5]The original cohort structure groups counties based on when they first experienced the treatment, i.e., theme park openings. Each group or "cohort" reflects the timing of treatment for these counties.

services, such as retail, restaurants, hotels, entertainment, and real estate, experience the most significant growth in business creation, while agricultural, manufacturing, construction, and utility industries are not significantly affected.[6]

Third, we find positive spillover effects on neighboring counties within a 75-kilometer radius. However, beyond 75 kilometers, these spillover effects diminish and become statistically insignificant. Additionally, we examine the spillover effects by industries. We find that tourism and travel-related services sectors and the real estate industry are mostly affected by the spillover effects while other industries show no spillover effects on neighboring counties regardless of the distance. These findings indicate that the spillover effects of theme parks are highly localized, impacting entrepreneurial activities primarily in nearby areas and certain industries.

Fourth, we explore tourism and agglomeration as potential mechanisms through which theme parks promote local business creation.[7] We find that three years after a theme park opens, the total number of tourists increases by 11%. Furthermore, we examine the land market and find an increase in land prices by 20% following the opening of a theme park, suggesting a significant appreciation in the property and land values in the host regions. These findings together highlight tourism as a potential channel that drives the positive effect of theme park openings on local economies. Turning to the agglomeration mechanism, we find that theme parks increase industrial specialization of cultural, sport, and entertainment in a host city by 16% four years after openings and this positive effect persists after four years, indicating that new businesses related to theme parks cluster together, potentially benefiting from proximity, shared knowledge, and infrastructure.

Finally, we put our empirical results into perspective. Specifically, we conduct a back-of-envelop calculation to show that the entrepreneurship effect of theme park openings can lead to 11% higher employment in the county hosting the theme park. Also, we use nighttime light as a proxy of overall economic activities, which not only include the new entrepreneurial activities but also the existing business activities. We find that theme park openings lead to

[6] According to WTO, tourism and travel-related services include services provided by hotels and restaurants (including catering), travel agencies and tour operator services, tourist guide services and other related services.

[7] For this part of the analysis, due to data limitation, we employ prefectural city-level data.

a 2%-3% increase in overall economic activities. These findings together assess the aggregate economic effects of theme park openings on local economies.

Related Literature. This paper contributes to the existing literature examining the economic impacts of place-based policies on developing new cultural, sport and recreational attractions, including sport stadium (Coates and Humphreys 1999), casino (Scavette 2023), national parks (Szabó and Ujhelyi 2024), and heritage sites (Bertacchini et al. 2024). Coates and Humphreys (1999) find no increase in local income growth from building a new professional sports stadium. Scavette (2023) shows that casino development in Atlantic City increases local employment, wages, and house prices. Szabó and Ujhelyi (2024) find that national park designation boosts local employment and income, primarily driven by tourism. Bertacchini et al. (2024) find that UNESCO World Heritage List inscriptions in Italy raise local income and property values through tourism and gentrification.[8]

Our study contributes to the literature on place-based policy in three aspects. First, we introduce a new dataset on an important tourism-related place-based policy in the world. As global tourism continues to evolve, the theme park market has been experiencing significant growth, with revenues rising from $51.67 billion in 2020 to $79.7 billion in 2024, representing an increase of 54%.[9] Second, our study is not only the first assessment of theme park development on local economies but also relevant for policy analysis as theme park development has been used as a place-based policy for regional development. Third, theme parks not only generate revenue through ticket sales but also generate revenue through catering, hotel, retail, and entertainment. In contrast to national and cultural parks, theme parks are specifically designed to facilitate a concentrated, high-density flow of visitors. This is because their business model relies on attracting large numbers of visitors and frequent consumer spending. Through clustered attractions, dining, shopping, and entertainment in one location, theme parks create an environment that naturally encourages visitors to spend within a compact area. This concentration, in turn, brings about agglomeration effects—the mechanism through which theme parks drive business creation and stimulate economic development. Our work exploits this unique setting to examine whether place-based tourism

[8]Addtionally, Fritsch et al. (2016) and Franco and Macdonald (2018) examine the price effects of a UNESCO World Heritage Site on housing prices.

[9]See Global Amusement Parks Market Report 2021-2024.

policies can promote local economic development through agglomeration, which remains understudied in this strand of literature.

Moreover, we provide new evidence on the efficacy of place-based policy in China. Previous studies have mainly investigated special economic zones (SEZs) and industrial parks. For example, Lu et al. (2019) find that the SEZs promote capital, employment, output, and firm entry within the zone boundary.[10] Zheng et al. (2017) find that industry parks promote local employment and wages, which in turn stimulate nearby local housing construction and retail store openings. Recently, Tian and Xu (2022) find that national high-tech zones foster local innovation output and entrepreneurial activities. Our study extends this strand of literature by examining a tourism-related place-based policy in China, i.e. theme park development. Our study shows theme park openings not only drive local business creation and economic development but also spillovers to neighboring regions. A novel finding of our study is that labor-intensive industries benefit more than capital-intensive industries from theme park openings, whereas previous work finds that industry parks benefit capital-intensive industries more than labor-intensive ones (Lu et al. 2019). These results suggest that different types of place-based policy should be considered for regional development according to the local relative factor endowment structure.

This paper also contributes to the literature on the impact of tourism on economic development. Faber and Gaubert (2019) examine tourism in Mexico and find that it generates local economic benefits through spillover effects to the manufacturing sector and generates national economic gains from tourism through market integration. Lanzara and Minerva (2019) explored the effects of tourism on Italian cities from 2001 to 2011, demonstrating that tourism not only increases firm entry and employment in the non-tradable sector. Nocito et al. (2023) assess how an international release of an Italian TV entertainment series affect the tourism and economic development of the regions where the series were shot. Their findings show that the series' release increases tourist numbers, tourist expenditures, and rental and property prices. We extend this literature by showing theme park openings promote tourism, which generates spillover of business creation in other industries and neighboring

[10] Earlier studies focus at the city level and also find positive economic effects of SEZs, see Wang (2013) and Alder et al. (2016).

regions. We also find tourism promotes agglomeration economies through specialization in related industries, such as hotels, catering, and retail.

The remainder of the paper is structured as follows. Section 2 provides an institutional background. Section 3 outlines data and the empirical strategy. Sections 4 and 5 discuss empirical results and potential mechanisms. Section 6 estimates the aggregate economic impacts and Section 7 concludes.

2 Background: China's Theme Park Development

A theme park is defined as a profit-oriented development that occupies a substantial land area and necessitates significant capital investment, operating under a closed management system.[11] Such parks are characterized by one or more distinct cultural or tourism themes and provide visitors with paid access to leisure experiences, cultural entertainment products, or services. Examples include amusement parks dominated by large-scale rides, extensive miniature landscape parks, and various film or animation cities designed to offer scenario simulations and immersive environmental experiences. On the contrary, publicly funded urban parks, botanical gardens, zoological parks, and other similar facilities constructed by the government are excluded from this definition of theme parks.

China's theme park industry has evolved significantly over the past 30 years. The first parks, Happy World, which opened in 1988, and Splendid China Folk Village, which opened in 1989, marked the beginning of this development. In the 1990s, the industry development was highlighted by the construction of Beijing World Park and the Window of the World in Shenzhen. Since then, the theme park industry in China has experienced rapid growth.

Figure 1 shows the number of large theme parks opened from 1988 to 2020, revealing a surge in openings since 2005.[12] From 2005 to 2020, an average of approximately

[11] See "Guiding Opinions on Regulating the Construction and Development of Theme Parks" (Document No. 400 [2018] of the National Development and Reform Commission), promulgated by five ministries on April 9, 2018.

[12] According to the "Guiding Opinions on Regulating the Construction and Development of Theme Parks," theme parks are categorized into three scales: super-large, large, and medium-small. A super-large theme park is defined by a total area of 2,000 acres (approximately 1.33 square kilometers) or more, or a total investment of more than 5 billion yuan. A large theme park has a total land area of 600 acres (approximately 0.4 square kilometers) or more but less than 2,000 acres, or a total investment of 1.5 billion yuan or more but less than 5 billion yuan. Medium-small theme parks are characterized by a total land area of 200 acres

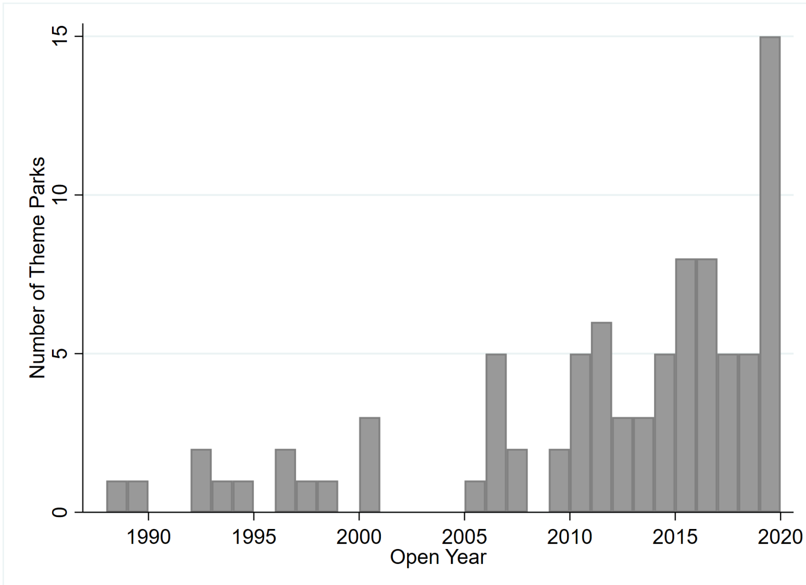


Figure 1: Number of Large Theme Parks from 1988 to 2020

Notes: The figure illustrates the number of theme parks opened from 1988 to 2020. The Y-axis represents the number of theme parks, while the X-axis indicates the years in which the parks were opened.

five new theme parks opened each year.[13] Table B.2 lists all the large theme parks that opened from 1988 to 2020. Notably, most of these parks are owned by three leading Chinese companies—OCT Group, Chimelong Group, and Fantawild Holdings—all of which rank among the world's top ten theme park operators, surpassing earlier developers from Japan, South Korea, and Singapore.

Local governments have supported the construction of theme parks by providing subsidies and tax incentives, as they consider those parks can create employment and stimulate local consumption, which in turn contributes to local economic development. Take Fantawild as an example: From 2013 to the first half of last year, it received more than RMB 5 billion in subsidies from the government. Meanwhile, preferential fiscal policy also indirectly helps these companies' profitability. Tax incentives accounted for 12.96%, 19.16%, and 19.91% of Fantawild's net profits for 2013, 2014, and 2015, respectively.

Moreover, local governments also directly finance theme park construction through debts. For instance, Locajoy, an amusement park located a two-hour drive from downtown Chongqing,

(approximately 0.13 square kilometers) or more but less than 600 acres, or a total investment of 200 million yuan or more but less than 1.5 billion yuan.

[13]No new theme parks were opened in 2008, which may be due to the financial crisis.

is owned by the Chongqing Tourism Investment Group, a state company wholly owned by the Chongqing municipal government.[14] In 2021, Chongqing municipal government invested over 10 billion RMB to support the construction of projects such as Locajoy Tourism Resort, which was included in the city's major cultural and tourism development projects for 2021-2023.[15]

3 Data and Empirical Strategy

We use a unique dataset of theme parks at the county level in China. We merge it with data on the number of new business registrations to examine the impact of theme park openings on entrepreneurship. We introduce the data sources for all variables used in the study in Section 3.1. In Section 3.2, we outline the empirical strategy, including our research design and approaches to addressing endogeneity concerns. Finally, we report the descriptive statistics for the main variables used for analyses, broken down by treated and control group in Section 3.3.

3.1 Data

Our empirical analysis draws on three main datasets at the county level covering 2000-2020 in China: the opening and location data of theme parks, business creation data, and DMSP/OLS Nighttime Light.[16] Supplementary data include city-level number of tourists, land prices, and city-industry level employment for additional analyses. We also include socioeconomic characteristics in initial years at the city- and county-levels as control variables.

3.1.1 Theme Parks

The opening year and location of theme parks are collected from various issues of the Evaluation of the Theme Park Competitiveness Reports, which are sourced from the Institute

[14] See Foreign Policy 2016, The Terrible Amusement Park That Explains Chongqing's Economic Miracle

[15] See Reply from the Chongqing Municipal Commission of Culture and Tourism regarding " Proposal on Supporting the Development of a Happy City in Western China."

[16] County-level administrative divisions are the second-level administrative regions in China. These divisions include districts under municipalities, county-level cities, counties, autonomous counties, banners, autonomous banners, special districts, and forestry districts.

for Theme Park Studies in China (http://www.our-themepark.com/index/baogao). These reports cover large and extra-large theme parks, defined as those with a total area of 600 acres or more or a total investment of RMB 1.5 billion.

Our sample contains 86 theme parks, with 76 openings over the period 2000-2020. We manually collect the opening years and addresses of these parks. Among the 76 theme parks, six are located in Beijing, Shanghai, Shenzhen, and Guangzhou—the country's largest and most economically advanced cities. We exclude the theme parks in four cities in our analysis because there are no comparable cities to serve as the control group for them. As a result, the final sample consists of 70 theme parks in 58 counties.

3.1.2 New Business Registrations and Nighttime Lights

We collect data on the number of new business registrations from the Tianyancha website, a platform that provides information on nearly 300 million entrepreneurship entities.[17] This dataset includes annual data on the number of new business registrations by industry across Chinese counties from 2000 to 2020.[18]

The nighttime light data used in this study are derived from the improved DMSP-OLS-like data for China (Wu et al. 2022). The data were processed using the "Pseudo-invariant Pixel" method for calibration. The data also considered temporal consistency between the DMSP-OLS data and SNPP-VIIRS data. The missing values in the original monthly SNPP-VIIRS data were addressed before compiling the annual SNPP-VIIRS dataset. The calibrated DMSP-OLS data (1992-2013) and the DMSP-OLS-like data (2013-2022), converted from SNPP-VIIRS, are then combined to produce the enhanced DMSP-OLS-like dataset spanning from 1992 to 2022.

[17]The data on the Tianyancha website comes from publicly available information from sources such as the National Enterprise Credit Information Publicity System, China Judgment Online, China Enforcement Information Public Website, the National Intellectual Property Administration, the Trademark Office, etc.

[18]The industry coverage includes agriculture, forestry, animal, and fisheries; mining; manufacturing; production and supply of electricity, heat, gas, and water; construction; wholesale and retail; transportation, storage, and postal services; accommodation and restaurants; information transmission, software, and information technology services; finance; real estate; leasing and business services; scientific research and technical services; water conservancy, environment, and public facilities management; resident services, repairs, and other services; education, health, and social work; culture, sports, and entertainment; public administration, social security, and social organizations; and international organizations.

3.1.3 Tourism, Employment, and Land Prices

Additionally, we collect the data on tourism and industrial employment from City Statistical Yearbooks, compiled by each city's Municipal Bureau of Statistics. This data on tourism is for the years 2002-2020, including information such as the year, the name of the prefecture-level city, the number of local tourists and tourists from other cities, the number of international tourists, the consumption expenditure of domestic tourists, the consumption expenditure of international tourists, and the number of hotels in the city.

Industrial employment data from 2003 to 2019 includes the year, the number of people employed in each industry, the industry name, and the prefecture-level city. We use employment in the cultural, sports, and entertainment industries to construct the specialization index, while employment in all other industries is used to calculate the diversity index.

Land Prices data at the prefecture city level are obtained from China National Land and Resources Statistical Yearbooks which are available from 2003 to 2017. They provide the city name, the area of land supplied in hectares, and the transaction price value of the land in ten thousand RMB.[19]

To explore the tourism mechanism, we merge theme park and tourism data at the prefecture-city-year level for the period from 2002 to 2020. To investigate the agglomeration mechanism, we merge theme park and employment data to construct a prefecture-city-year level dataset from 2003 to 2019.

3.1.4 Socioeconomic Characteristics

For covariates, the county-level socioeconomic characteristics are mainly collected from the China County Statistical Yearbook 2001 and supplemented by the Fiscal Statistical Compendium for all counties (*Quanguo Dishixian Caizheng Tongji Ziliao 2001*). City-level socioeconomic data for 2001 are drawn from the China Urban Statistical Yearbook 2001 and the same Fiscal Statistical Compendium. These characteristics include gross domestic product (GDP), secondary and tertiary sector GDP shares, and total population. We use these

[19]The land price data for Jiangxi Province in 2007 is unavailable; therefore, in the regression, the land prices for all cities in Jiangxi Province for that year are treated as missing.

data to measure local initial characteristics.[20,21] By including these covariates, our analysis can focus on the effect of theme park openings, rather than being confounded by existing differences in economic or population characteristics.

3.2 Empirical Strategy

We apply the difference-in-differences method to estimate the effects of theme park openings in a setting with multiple periods and varying treatment timings across different counties. Since the treatment (such as the opening of a theme park, which is staggered and permanent) occurs at different times for different counties and never reverts to 0, the assumptions of no anticipation and parallel trends are crucial. These assumptions require that, in the absence of the treatment, the treated and control groups would follow the same outcome trends, which can be tested by comparing pre-treatment trends between the two groups.

To cope with these concerns, we need to select a control group that is comparable to the treated group. We introduce how our research design helps select the control group in Section 3.2.1. We then present our estimation specification in Section 3.2.2 and discuss the identification issues in Section 3.2.3.

3.2.1 Research Design

The primary econometric challenge in this analysis is the non-random selection of theme park locations. Theme park companies choose sites based on their expectations of the present discounted value of future profits. This decision-making process is influenced by various observable and unobservable factors, including income level, population density, transportation infrastructure, and local amenities of the location. However, many of these factors may also be correlated with the outcome variables—entrepreneurship, which can lead to biased estimates when comparing counties with large theme parks to those without. To obtain unbiased estimates of the effect of theme park openings on local entrepreneurship, it is crucial

[20] The earliest available data on GDP shares for the secondary and tertiary sectors in *Quanguo Dishixian Caizheng Tongji Ziliao* is from 2001. Therefore, we use initial local characteristics data in 2001 instead of 2000.

[21] GDP serves as an indicator of the overall economic output of the region, reflecting its economic size and capacity. The shares of GDP from the secondary and tertiary sectors provide information on the development stage and economic focus of the region. Additionally, the population measures the region's demand.

to identify a control group that is similar to the treated group—counties hosting large theme parks—in terms of the key determinants of entrepreneurial activities.

One of the most important factors in choosing the location of a large theme park is accessibility. Improved accessibility reduces travel distances, thereby enhancing the attractiveness of the destination for visitors (Zhang et al. 2022). Many theme parks are built in the area with air or ground connectivity. Specifically, we geocode the airports and railway stations that were in operation before the theme park openings in China, selecting those with the shortest distances from the counties where large theme parks are located. Based on these selected airports and railway stations, we then draw a circle with a radius equal to the maximum distance between an airport and a county in the treated group.[22] All counties located within this circle were included in the control group. The identification assumption here is that counties within this designated proximity share similar economic characteristics, such as GDP, population density, and transportation infrastructure, which make them comparable to the treated counties.

Figure 2 shows the distribution of treated and control groups in our sample. The red dots denote the treated group while the blue dots denote the control group. In total, there are 58 counties in the treated group and 512 counties in the control group. Notably, the majority of our sample counties are concentrated in the eastern and central regions of China, including Jiangsu, Shandong, and Zhejiang provinces, with no sample counties coming from western areas such as Tibet, Xinjiang, or Ningxia. This geographic concentration underscores the importance of using the control group that shares similar regional development patterns in local economies.

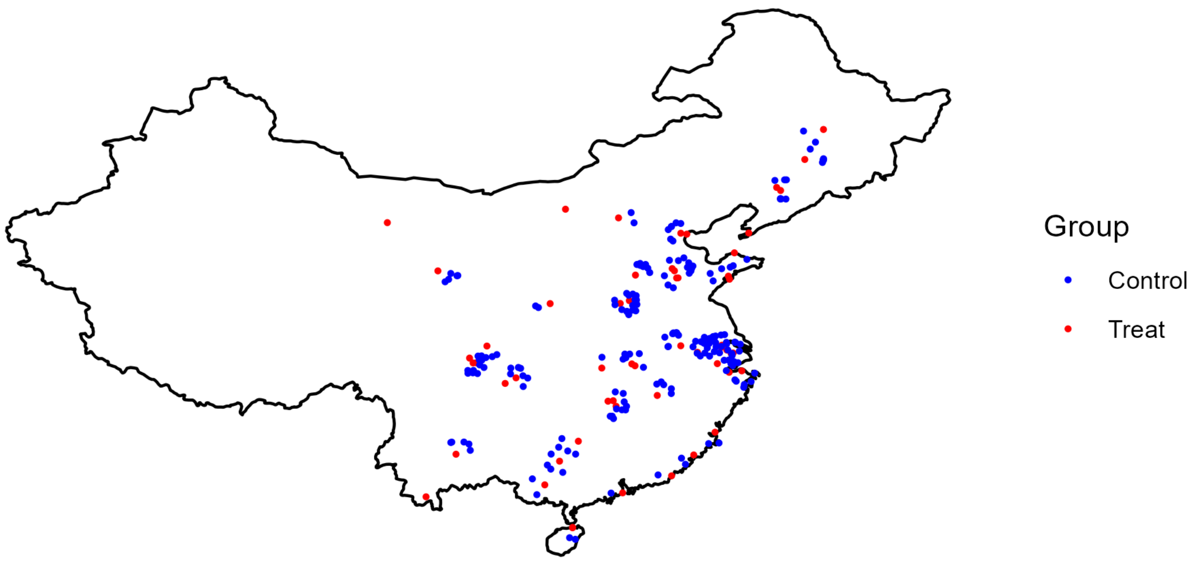
3.2.2 Difference-in-Differences (DiD)

To examine the effect of theme park openings on regional economic growth at the county level, we estimate the following DiD model:

$$Y_{it} = \alpha_i + \alpha_t + \beta_1 D_{it} + X_{it}\gamma + \epsilon_{it} \tag{1}$$

[22] In the sample, the largest distance between an airport and a treated county is 80.55 kilometers.

Figure 2: Distribution of Treated and Control Counties



Notes: The figure shows the map of Mainland China. Red dots indicate the counties that were exposed to large theme parks during the sample period from 2000 to 2020. The blue dots represent counties with no large theme parks during this period but are comparable to counties with theme parks. Counties in Beijing, Shanghai, Shenzhen, and Guangzhou are excluded.

where Y_{it} represents entrepreneurship, measured by the number of new business registrations in county i of in year t. Given that entrepreneurship is a count variable, we employ Poisson pseudo-maximum likelihood (PPML) which accounts for zero values in the estimation following Silva and Tenreyro (2006). $D_{it} = 1$ is a dummy variable, indicating a new theme park was opened in county i in year t and 0 otherwise. The coefficient β_1 is the coefficient of interest, which is expected to be positive and significant.

In Equation 1, we incorporate two-way fixed effects to account for potential confounding factors. First, we address concerns regarding invariant county-specific shocks, such as geographical features. For instance, counties located in developed regions or near coastlines may benefit from ports and trade, while mountainous regions face logistical challenges that could hinder development. These factors can influence entrepreneurial activities through urbanization or natural constraints. To account for this unobserved county heterogeneity, we include county fixed effects, α_i. Second, we include year-fixed effects, α_t, to control for aggregate trends, such as national business cycles and fiscal policies. The error term is denoted by ϵ_{it}, and standard errors are clustered at the county level to allow for serial correlation. Additionally, we include X_{it} which is the interaction term between time-invariant covariates W_{i0} with the time dummy variable.

To test for pre-trends and to understand the dynamic effects of theme park openings, we estimate the event study version of Equation 1 as:

$$Y_{it} = \alpha_i + \alpha_t + \sum_{\tau=-5}^{5} \beta_\tau D_{it}^{\tau} + X_{it}\gamma + e_{it} \tag{2}$$

We designate the event period as period 0, which is the year of theme park openings. The vector D_{it}^{τ} is composed of dummies for each period before and after the openings, ranging from 5 years before to 5 years after.[23] Additionally, we normalize the values for the preceding period leading up to the event to 0 (period -1). The other items are defined the same as above. If the coefficients are all insignificant in the pre-exit periods, then it can be concluded that there is no pre-trend.

3.2.3 Addressing Endogeneity Concerns

Reverse Causality. Building a large theme park typically takes a long time from planning and constructing to opening. For instance, Universal Studios in Beijing took 20 years from initial planning to its opening. The process began in 2001 when Universal Studios representatives met with Beijing officials to discuss the park and signed a joint venture agreement. After additional negotiations and adjustments, the project was approved by the State Council in 2014. The resort officially opened to the public in 2021.[24] Besides, Shanghai Disneyland also took 7 years starting from planning in 2009 to opening in 2016.

One potential concern is that theme parks are often established in regions that are already economically developed, raising the issue of selection bias. However, the fact that park locations are determined during the planning phase—long before the parks open—helps alleviate concerns about reverse causality by providing a clear temporal separation between the decision to locate and the actual economic outcomes observed post-opening.

[23]There are 20 years before and 32 years after the openings of theme parks. To simplify the analysis, years more than 5 years before and after the event are aggregated to -5 and 5, respectively.

[24]Construction of Beijing Universal Resort started in July 2018. By the end of 2019, the main construction was complete, and by December 2020, core systems and decorations were finished.

Omitted Variable Bias. Another potential concern is the presence of omitted variables that could bias the results. To address this issue, we incorporate county- and year-fixed effects to account for time-invariant omitted variables that could be correlated with the explanatory variable and aggregate time variables. In addition, we interact initial local characteristics which are time-invariant covariates, such as population and GDP, with year dummies to account for how these factors may vary over time, reducing the risk of omitted variable bias. For robustness, we include region-year fixed effects to capture differences in economic trends across regions that could influence the results, such as regional policy changes or economic shocks specific to certain areas.

Instrumental Variable (IV) Approach. To further alleviate the endogeneity concern and identify the causal effect of theme park openings on entrepreneurship, we construct an instrumental variable that is plausibly uncorrelated with local shocks to the creation of new businesses at the county level but is likely to affect the theme park openings. To this end, we employ a Bartik-style instrument following Bartik (1991), which exploits the interaction between national trends with a potentially exogenous time-invariant cross-sectional variable. The rationale behind this approach is that some plausibly exogenous aggregate time trends affect different spatial units systematically along some cross-sectional exposure variables.

Following this logic, we construct our instrument by starting with a plausibly exogenous aggregate trend—the national-level expenditure on travel. An increase in travel expenditure indicates a rise in demand for tourism-related services, such as theme parks. In this way, travel expenditure would affect theme park constructions and openings in each county. A potential concern about using national travel expenditure is that there could be other changes over time that can affect theme park openings, which could then confound the control function estimates. This concern can be potentially addressed by the year-fixed effects. However, since the national travel expenditure only varies by year, they will be collinear with year-fixed effects. Therefore, we interact national travel expenditure with the county's accessibility. Specifically, we apply the control function method using the constructed IV as below:

$$D_{it} = \alpha_i + \alpha_t + \delta_1 \ln travelexp_t \times \frac{1}{dist_i} + X_{it}\gamma + \eta_{it} \tag{3}$$

$$Y_{it} = \alpha_i + \alpha_t + \beta_1 D_{it} + \delta_2 \hat{\eta}_{it} + X_{it}\gamma + \epsilon_{it} \tag{4}$$

where $IV = \ln travelexp_t \times \frac{1}{dist_i}$, in which $\ln travelexp_t$ is the log of national travel expenditure in year t and $\frac{1}{dist_i}$ captures a county's accessiblity. If the county is more accessible, the county's distance to the nearest airport is lower, and $\frac{1}{dist_i}$ is larger. Therefore, we expect that δ_1 to be positive. Equation 3 shows the first-stage regression and Equation 4 shows the second-stage regression. We obtain the predicted value of residuals $\hat{\eta}_{it}$ in the first stage estimation and regress entrepreneurial activity on theme parking openings and the predicted residuals $\hat{\eta}_{it}$ in the second stage.

To ensure the validity of our instrument, the interaction of the aggregate time trends with the time-invariant exposure variable has to be independent of the error term. This independence could occur if either the time trend or the exposure variable is uncorrelated with the error term Goldsmith-Pinkham et al. (2020). In our context, the identifying assumption is that conditional on the controls, the interaction between national travel expenditure and county accessibility impacts entrepreneurship only through the openings of theme parks. More specifically, $\ln travelexp_t \times (dist_i)^{-1}$ must remain uncorrelated with county-specific time-variant shocks to entrepreneurship, ϵ_{it}. This would be true because, on the one hand, county accessibility is a time-invariant characteristic that reflects the geographic and infrastructural attributes—the proximity to the airports. Since these airports were constructed and operated before theme park openings, they are unlikely to respond directly to county-level shocks to entrepreneurship. On the other hand, national travel expenditure reflects broader macroeconomic conditions and consumer behaviors that are determined independently of any single county's entrepreneurial activity. These trends capture national or even international factors influencing tourism and leisure spending, which are external to individual counties and their economic fluctuations. This independence supports the argument that the time trend is unlikely to be influenced by county-specific entrepreneurial shocks.

3.3 Summary Statistics

By merging two main datasets—theme park and business creation—using county and year, the final dataset includes 263 counties, with 58 counties in the treated group and 205 counties in the control group from 2000 to 2020. Panel A in Table 1 presents the summary statistics of the county-level variables after selection. The mean values of the number of new business registrations for the whole industry and for the tertiary industry are approximately 5,256 and 4,530, respectively. In the sample counties, most of the newly registered firms are in the tertiary industry, accounting for 86%. In the treated group, the average is higher, with 7,743 newly registered firms per year, including 6,865 in the tertiary sector. In contrast, the control group has an average of 4,531 newly registered firms per year, with 3,850 in the tertiary industry.

Additionally, the mean value of the dummy variable for the theme park openings is around 0.081, which indicates that on average, 8.1% of the sample counties experienced a large theme park opening in China from 2000 to 2020, and in the treated group, there are 36% of the counties have large theme park openings over the sample period.

Panel B reports the descriptive statistics for county-level covariates. The mean values for the log of GDP and the log of population in 2001 for the final sample are 3.466 and 3.831, respectively. The mean shares of secondary and tertiary GDP in the final sample are 0.410 and 0.382, respectively. Notably, the mean values of these covariates in the control group are similar to those in the treated group, indicating that the selected control group is comparable to the treated group.

Panel C presents the descriptive statistics for city-level variables. The tourism variable, which is measured by the number of total tourists, spans the years from 2002 to 2020. The agglomeration variables, encompassing specialization and diversity, cover the period from 2003 to 2019, while land prices are reported from 2003 to 2017. On average, cities receive approximately 33.8 million tourists annually, with about 51 million tourists in the treated group and 19.7 million tourists in the control group.

For agglomeration variables, the mean value of the specialization index is 0.925, indicating a relatively high level of industry concentration in the cities studied. Notably, the treated group shows a higher mean specialization index of 1.096, suggesting that cities with large theme park openings tend to have a more concentrated industry structure compared to the control group, which has a mean of 0.785. This difference may imply that the presence of theme parks could enhance the focus of certain industries, potentially driving economic activity and attracting related businesses. The mean diversity index stands at 7.109, with the treated group averaging 8.317 and the control group averaging 6.114.

Regarding land prices, the overall mean is 13.75 billion RMB, but there is a stark contrast between the treated and control groups. The treated group averages 23.07 billion RMB, significantly higher than the control group's average of 6.13 billion RMB. This disparity suggests that cities with large theme parks not only experience increased tourist inflow but also see a substantial rise in land value, likely due to higher demand for real estate driven by increased economic activity and tourism-related investments. This highlights the potential economic benefits that large theme parks can bring to their local economies, influencing both agglomeration effects and the land market. Additionally, the mean value of land areas is 1,288 hectares, with 1,595 hectares in the treated group and 1,037 in the control group.

Panel D reports the descriptive statistics for city-level covariates in 2001. The mean values for the log of GDP and the log of population in 2001 for the final sample are 5.934 and 5.967, respectively. The mean shares of secondary and tertiary GDP in the final sample are 0.466 and 0.371, respectively. Same as county-level covariates, the mean values of city-level covariates in the control group are close to those in the treated group.

4 Empirical Results

This section first reports the baseline results in Section 4.1 and robustness checks in Sections 4.2 and 4.3. Second, it shows the heterogeneity effects by industry in Section 4.4. Third, it explores the spatial spillover effects of theme parks in Section 4.5.

4.1 Baseline Results

Table 2 reports results for newly registered firms using PPML estimation. Specifically, Column 1 reports the results with the inclusion of county FEs, Year FEs, and the interaction

Table 1: Summary Statistics

	Full Sample		Treated Group	Control Group
Variable	Mean	SD	Mean	Mean
Panel A: County-level Variables				
Year: 2000-2020				
Firms (Total)	5,256.079	6,802.429	7,743.337	4,531.150
Firms (Tertiary)	4,530.407	6,170.044	6,865.187	3,849.918
Light (Sum)	9,414.765	9,739.820	13,186.70	8,309.855
D_{it}	0.081	0.273	0.360	0
Number of Observations	5,565	5,565	1,218	4,347
Panel B: County-level Covariates				
Year: 2001				
Log(GDP)	3.466	1.004	3.626	3.420
GDP Share (Secondary Industry)	0.410	0.145	0.438	0.402
GDP Share (Tertiary Industry)	0.382	0.135	0.431	0.367
Log(Population)	3.831	0.603	3.737	3.858
Number of Observations	5,397	5,397	1,218	4,179
Panel C: City-level Variables				
Year: 2002-2020				
Tourist (Million)	33.767	51.764	51.042	19.692
$D_{j(i)t}$	0.204	0.403	0.455	0
Number of Observations	1,862	1,862	836	1,026
Year: 2003-2019				
Specialization	0.925	0.461	1.096	0.785
Diversity	7.109	1.802	8.317	6.114
$D_{j(i)t}$	0.199	0.400	0.441	0
Number of Observations	1,656	1,656	748	908
Year: 2003-2017				
Land Prices (10 Thousands RMB)	1,375,426	2,321,974	2,306,907	612,878.7
Land Areas (Hectare)	1,287.996	4,039.469	1,595.136	1,036.559
$D_{j(i)t}$	0.175	0.80	0.387	0
Number of Observations	1,444	1,444	650	794
Panel D: City-level Covariates				
Year: 2001				
Log(GDP)	5.934	0.858	6.331	5.611
GDP Share (Secondary Industry)	0.466	0.093	0.479	0.455
GDP Share (Tertiary Industry)	0.371	0.079	0.414	0.336
Log(Population)	5.967	0.703	6.023	5.921
Number of Observations	1,862	1,862	836	1,026

Note: For the county-level variables used in the regressions, panel A reports the mean and standard deviations for the final sample after selection from 2000 to 2020, as well as the mean values for the treated and control groups. Additionally, panel B reports the mean and standard deviations, the log of GDP, the share of GDP in the secondary industry, the share of GDP in the tertiary industry in 2001, and the log of the population in 2001, along with their respective mean values for both the treated and control groups at the county level. Panel C reports the mean and standard deviations for the variables at the city level, as well as the mean values for the treated and control groups. Specifically, the number of tourists is from 2002 to 2020, the specialization and diversity are from 2003 to 2019, and the land prices and areas data are from 2003 to 2017. The city-level covariates use the same sample as tourists.

terms between time-invariant covariates (W_{i0}) with year dummies. The findings indicate that theme park openings lead to a 13.9% (exp(0.130)-1) increase in the newly registered firms.

Some regions in the east of China might experience faster economic growth than others in western China, which could influence the outcome variable. To account for these heterogeneous regional trends and mitigate the risk of omitting important region-level variables that change over time but are not directly observed, we add region-year fixed effects to Equation 1 and rerun the regression.[25] The results, presented in column 2 in Table 2, show that the significance and magnitude of the coefficients remain consistent with those in column 1. This consistency confirms the robustness of the baseline results. By controlling for unobserved regional shocks—such as local policy changes or macroeconomic conditions—that could otherwise bias the estimates, region-year fixed effects allow for clearer isolation of the true effect of the variable of interest, reinforcing confidence in the validity of the baseline findings.

Panel (a) of Figure 3 further illustrates the dynamic estimates for the outcome variable across periods. County-fixed effects, year-fixed effects, and interaction terms between time-invariant covariates (W_{i0}) and year dummies are included in the models. Notably, there is no significant pre-trend before the opening of a large theme park, supporting the assumption of parallel trends. Specifically, panel (a) shows a significant 15% increase in newly registered firms three waves after the park opening. This positive effect becomes insignificant by the fourth wave.

In summary, these findings suggest that theme park openings have the potential to stimulate entrepreneurial activity, as reflected in the increases in the number of newly registered firms.

4.2 Heterogeneous Treatment Effects

Given that the timing of theme park openings varies across regions in this study, using traditional two-way fixed effects models to examine the impact of theme park openings as

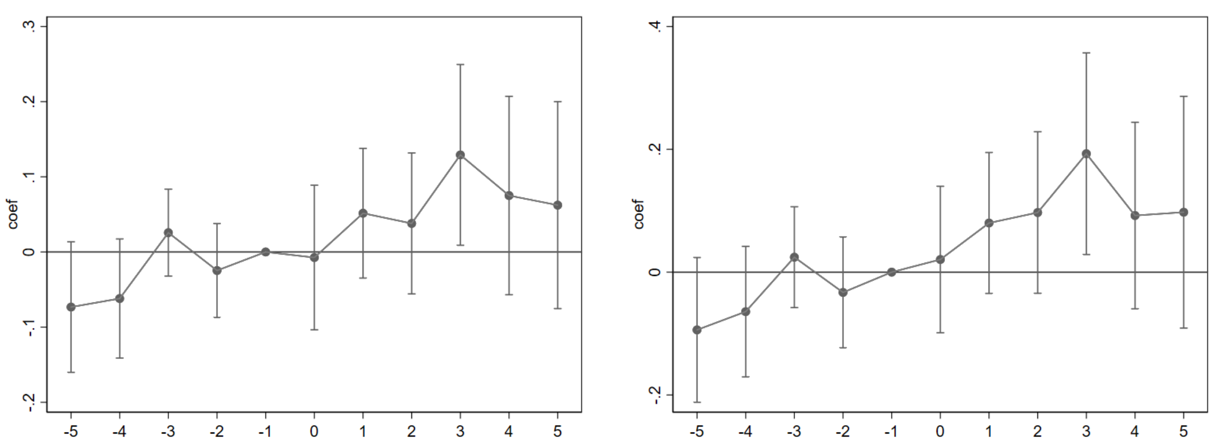
[25]Following Kahn et al. (2021), we divide China into three regions: the eastern region includes Beijing, Tian-jin, Shanghai, Liaoning, Hebei, Jiangsu, Zhejiang, Fujian, Shandong, and Guang-dong, Hainan, and Guangxi; the central region includes Inner Mongolia, Jilin, Heilongjiang, Shanxi, Anhui, Jiangxi, Henan, Hubei, and Hunan; the western region covers Shananxi, Gansu, Qinghai, Ningxia, Xinjiang, Chongqing, Sichuan, Guizhou, and Yunnan.

Table 2: Baseline Results

	DID		Stacked DID		DID	PSM-DID
	(1)	(2)	(3)	(4)	(5)	(6)
	PPML	PPML	PPML	PPML	Negative Binomial	PPML
D_{it}	0.130**	0.093*	0.176**	0.131*	0.125***	0.135**
	(0.058)	(0.056)	(0.069)	(0.069)	(0.025)	(0.061)
County FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	No	Yes	No	Yes	Yes
Region-Year FEs	No	Yes	No	Yes	No	No
$W_{i0}\times$Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes
pseudo R^2	0.882	0.888	0.865	0.872		0.931
N	5,397	5,397	68,082	68,082	5,397	2,331

Notes: Columns 1 and 2 display the results of business creation in all and tertiary industries by including county and year FEs. Columns 3 and 4 show the PPML results of business creation by including county and region-year FEs. I also include the interaction terms between time-invariant covariates and year dummy in all regressions. Standard errors, clustered at the county level, are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure 3: Baseline Results over Time



(a) Total newly registered firms

(b) Tertiary newly registered firms

Notes: Each point (and 95% CI) represents estimates from a regression for the dependent variable. Panel (a) plots estimates on the total number of newly registered firms in all industries while panel (b) focuses on only tertiary industries. In all panels, county- and year-FEs, as well as the interaction terms between time-invariant covariates W_{i0} and year dummies are included. We cluster the standard errors at the county level.

an exogenous shock on the economy may introduce estimation bias due to heterogeneous treatment effects. To address this issue, this study first decomposes the staggered DID estimates following Goodman-Bacon (2021).

The Bacon decomposition indicates that the estimated results primarily come from the treated group and the group never subjected to treatment, accounting for approximately 90% of the results. In this group, heterogeneity due to time-varying effects generally does not occur. However, there is still a 10% probability that the baseline results are affected by heterogeneous treatment effects. Therefore, we next conduct the staggered DID estimations using stacked DID following Cengiz et al. (2019).

The estimation results are reported in Table 2. Column 3 includes county and year-fixed effects, while column 4 adds county and region-year-fixed effects. Both columns incorporate interaction terms between time-invariant covariates and year dummies. Using stacked DID with PPML, the coefficients for newly registered firms are 0.176 and significant at the 5% level. After controlling for heterogeneous regional trends that change over time by adding the region-year fixed effects, the results are still consistent with the baseline results, demonstrating that estimation bias due to heterogeneous treatment effects does not pose a serious problem in this study. Thus, the baseline results are reliable.

4.3 Addtional Robustness Checks

This subsection performs several robustness checks to validate the baseline results. First, we employ a fixed effect negative binomial (NB) model to estimate the effect of theme park openings. Second, we implement propensity score matching (PSM) to refine the selection of the control group. Third, we conduct a placebo test to further evaluate the robustness of the baseline findings.

Alternative Specification. Following Hilbe (2011), which highlights the implicit restriction on the distribution of observed counts in the Poisson model - where the variance of the random variable is constrained to equal its mean - researchers often opt for more flexible specifications, such as the NB model. This model is commonly regarded as the standard for basic count data analysis. In our study, we estimate Equation 1 using the fixed effects NB model that incorporates county and year FEs, along with the interaction term between time-invariant covariates and year dummies. The results are presented in column 5 of Table 2. The coefficient for newly registered firms is 0.125, significant at the 1% level. These results are consistent with the baselines using PPML.

Alternative Selection Method. Another approach to selecting a comparable control group is to employ propensity score matching based on observable economic characteristics that are correlated to the business activity. Specifically, we estimate a logit model that incorporates several key variables: county GDP, the share of secondary GDP, the share of tertiary GDP, population size, and the average number of newly registered firms in both the overall industry and the tertiary sector in 2001.[26] After matching, we identify a treated group consisting of 54 counties and a control group of 51 counties, resulting in a sample with a total of 2,205 observations. This sample is notably smaller than the baseline sample but allows for a more robust comparison between treated and control counties. To estimate the effect of theme park openings on business creation using this matched sample, we include county and year-fixed effects in the regression model. Additionally, we control for interaction terms between time-invariant covariates and year dummies to account for any potential temporal effects.

The estimated results are presented in column 6 in Table 2. The coefficient is 0.158, indicating a positive impact of theme park openings on total newly registered firms. These findings remain consistent with the baseline results presented in column 1 of Table 2, suggesting the robustness of the baseline estimates.

IV Approach. We estimate Equation 3 using OLS with the county- and year-fixed effects and Equation 4 using PPML with the county- and year-fixed effects in the second stage. Additionally, the interaction terms between time-invariant covariates and year dummies are included in both stages. The results are presented in Table 3. Column 1 represents the results of the first stage and it shows that our IV has a positive and significant effect on the theme park openings, which is consistent with our expectation. More importantly, the F statistics in the first stage is 11.08, which suggests that our IV is a sufficiently strong predictor of the endogenous variable. Column 2 reports the results of the second stage and the

[26] Detailed information regarding the model specifications, variable definitions, and matching results can be found in Appendix A.1.

coefficient of the predicted residual is -0.186 and not statistically significant, which suggests that the endogeneity may not be a concern and our baseline results are reliable. As a result, the IV approach might not provide an advantage over the baseline estimates in this case, as the risk of bias from endogeneity appears minimal.

Table 3: IV Regressions

	(1) First Stage	(2) Second Stage
IV	1.180***	
	(0.355)	
D_{it}		0.303*
		(0.135)
Residuals		-0.186
		(0.164)
County FEs	Yes	Yes
Year FEs	Yes	Yes
$W_{i0}\times$Year Dummy	Yes	Yes
$F-stat$	11.08	
N	5,397	5,397

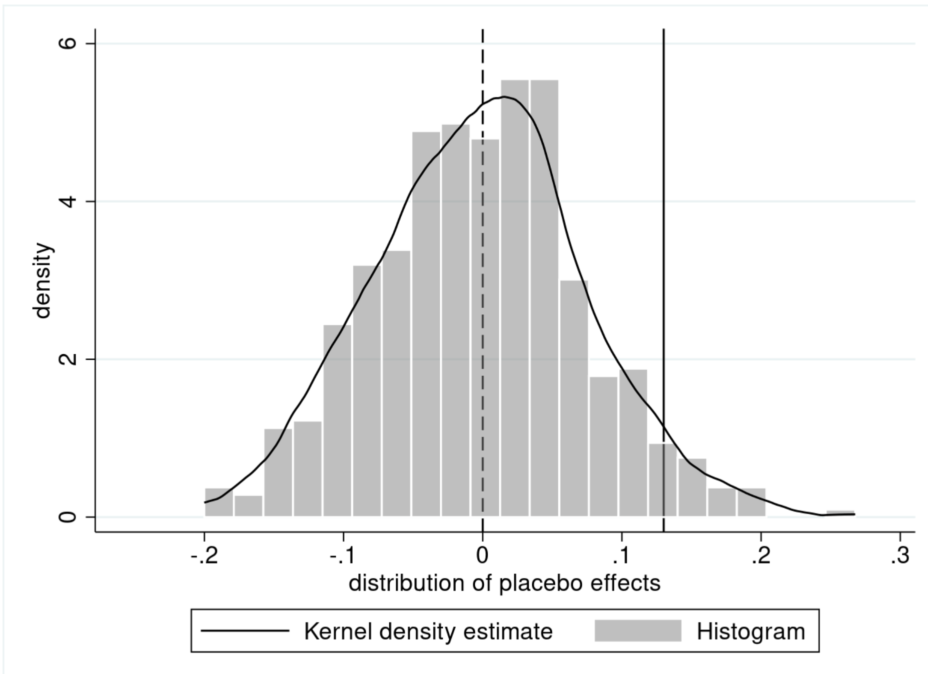
Notes: We run PPML regressions with IV and fixed effects using the control function method. Standard errors, clustered at the county level, are in parentheses. Significance levels are $*p < 0.1$, $**p < 0.05$, $***p < 0.01$. KP: Kleibergen-Paap, AR: Anderson-Rubin.

Placebo Tests. Although this study includes a set of fixed effects and control for county-level covariates, unobserved county-specific characteristics may still influence the baseline results. To address this, we conduct placebo tests following the approach of Ferrara et al. (2012). Using the range between the earliest and latest theme park opening years in the sample, we randomly assign a placebo treatment time to each county from a uniform distribution. We then estimate the two-way fixed effects (TWFE) model, repeating this process 500 times to generate the distribution of placebo effects, as illustrated in Figure 4.[27] The

[27] Based on Chen et al. (2023), we employ the unrestricted version of the mixed placebo test to account for the multi-period nature of theme park openings, using a staggered DID framework. In this approach, a fake treatment variable is generated in each iteration by randomly assigning both treatment units and treatment times. Unlike the original cohort structure, which maintains the number of units per cohort, the unrestricted version independently assigns a random treatment time to each unit.

placebo coefficients are mostly centered around zero, while the actual average treatment effect, represented by the vertical solid line, falls in the right tail of the distribution. In other words, it is an outlier relative to the placebo effects, indicating that the observed effect is likely due to the real impact of theme park openings rather than random variation. These findings confirm that the baseline results are robust.

Figure 4: Placebo Tests



Notes: The figure shows the density of the estimated coefficients from 500 simulations using false dates and false treatments of theme park openings. The dashed line represents zero effect. The solid line in panel (a) corresponds to the coefficient reported in column (1) of Table 2, while the solid line in panel (b) corresponds to the coefficient reported in column (1) of Table 4.

4.4 Industry Heterogeneity

This subsection investigates how the theme park openings' effects vary across industries. To explore the heterogeneous effects across industries, we first categorize industries into eight groups: agricultural resources (including agriculture, forestry, animal husbandry, and fishing); manufacturing; construction; utilities (comprising water, electricity, heating, and gas); hospitality & entertainment (encompassing restaurants, hotels, culture, sports, and entertainment); retail (covering wholesale and retail); real estate; and other services (including leasing and related business services, health, social work, resident services, repairs, transportation, warehousing, postal services, and others).[28]

Figure 5 illustrates the PPML estimates of business creation across these industries. The findings indicate that theme park openings primarily impact the tertiary sector. Specifically, the opening of a large theme park is associated with a 19% increase in hospitality & entertainment and a 13% increase in business creation in retail sectors. Additionally, business creation rises by 29% in real estate. The effects on other services are negligible and insignificant. At the same time, the non-tertiary sectors, such as construction and utility are not affected by theme park openings significantly. Agriculture and manufacturing sectors show negative and no significant changes in response to theme park openings. These results suggest that the positive effects of theme park openings on economic development are mainly driven by tourism-related industries, such as hospitality & entertainment, retail, and real estate.

Since several tertiary industries show significant impacts from theme park openings, we group those tertiary industries and estimate Equation 1 and 2 with the new firm registration in tertiary industries as the outcome variable. We report the results in Table 4. Column 1 shows that theme park openings lead to an increase in the tertiary industry, and the coefficient is larger than the one for the total newly registered firms (which is 0.130). Such results are robust to the use of alternative estimation methods (see columns 2-6) and a placebo test (see Panel (b) of Figure 4).[29]

Overall, the impact of theme park openings varies across industries, with a stronger positive impact on tourism-related tertiary industries.

4.5 Spatial Spillover Effects

In this subsection, we analyze whether exposure to large theme parks created positive spillovers in other counties. Or, rather, the positive effect on economic activity came at

[28] Financial services, education, and research are not included in other services category.

[29] We also plot the dynamic effects of theme park openings on business creation in tertiary industry. The results are presented in panel (b) of Figure 3, which shows a similar pattern as panel (a). Notably, there is no pre-trend.

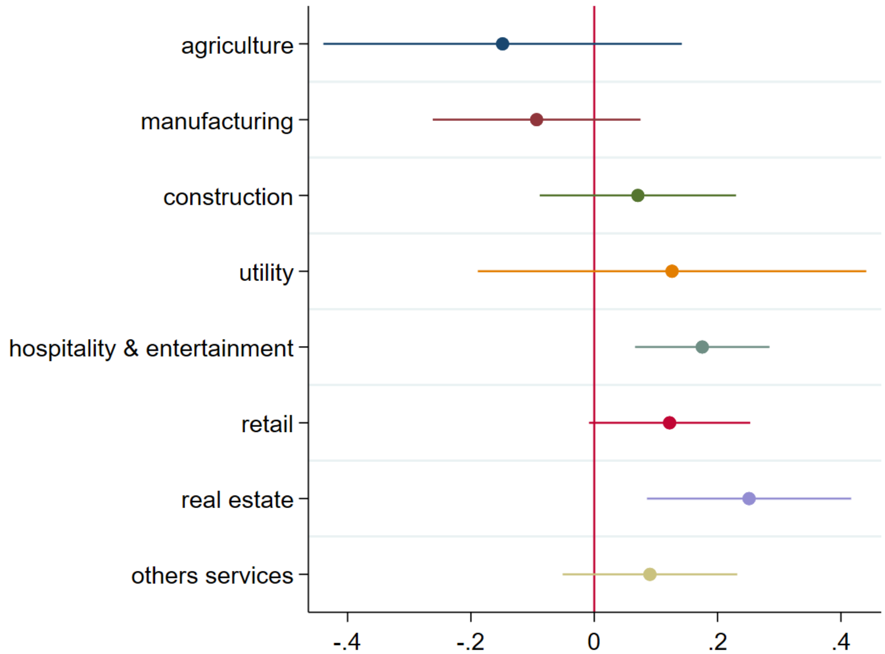


Figure 5: PPML estimates on various industries

Notes: The figure shows average treatment effects on business creation across 8 industries using PPML estimation. Each estimate includes year and county FEs. Covariates w_{i0} are included. Standard errors are clustered at the county level.

the expense of other counties (i.e., crowding out business from non-treated counties). To answer this question, we consider the neighboring counties as a treated group and compare the outcomes of these neighboring counties with the same control group used in the baseline regression. Specifically, we use the following specification:

$$Y_{i't}^{d} = \alpha_i + \alpha_t + \beta_1 D_{i't}^{d} + \beta_2 X_{i't}^{d} + \epsilon_{i't} \tag{5}$$

where $Y_{i't}^{d}$ represent the business creation in county i' located at a distance d from the treated counties i in the baseline regression, where $d = \{25, 50, 75, 100, 125, 150\}$. Specifically, when $d = 25$, it indicates that we consider the counties i' within the distance range (0, 25] from the treated counties i as the treated group for this specification, and $D_{i't}^{25}$ takes 1. When $d = 50$, we include counties within the distance range (25, 50] as the treated group, and this pattern continues for higher values of d. The control group remains consistent with the baseline sample, excluding the counties identified in the treated group. County- and

Table 4: Results on Tertiary Industries

	DID		Stacked DID		DID	PSM-DID
	(1)	(2)	(3)	(4)	(5)	(6)
	PPML	PPML	PPML	PPML	Negative Binomial	PPML
D_{it}	0.142**	0.105*	0.186***	0.140*	0.137***	0.157**
	(0.058)	(0.056)	(0.070)	(0.072)	(0.026)	(0.063)
County FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	No	Yes	No	Yes	Yes
Region-Year FEs	No	Yes	No	Yes	No	No
$W_{i0}\times$Year Dummy	Yes	Yes	Yes	Yes	Yes	Yes
pseudo R^2	0.882	0.888	0.865	0.872		0.930
N	5,397	5,397	68,082	68,082	5,397	2,331

Notes: Column 1 displays the results of business creation in tertiary industries by including county and year FEs while column 2 presents the results by including county and region-year FEs. Columns 3 and 4 show the PPML results of business creation using stacked DID. We also include the interaction terms between time-invariant covariates and year dummy in all regressions. Standard errors, clustered at the county level, are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Year-FEs, as well as the interaction terms between the time-invariant covariates and year dummies, are included. Standard errors are clustered at the county level.

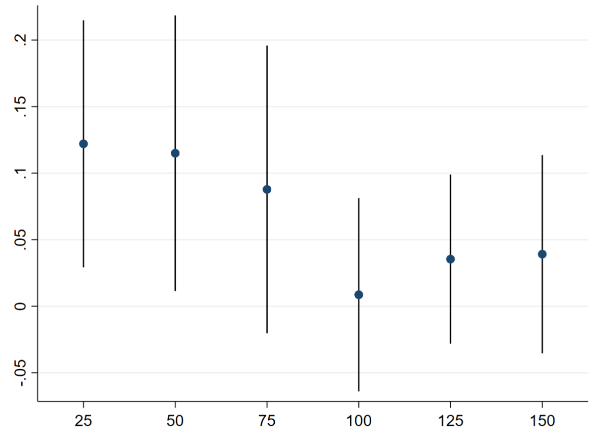
The results are presented in Figure 6. It shows that theme park openings have a significant effect on business creation in neighboring counties within a 50-kilometer radius. However, as the distance increases to 75 kilometers, the positive effects become insignificant. Beyond 75 kilometers, the effects diminish further and become negligible and statistically insignificant.

Overall, these findings imply that the substantial spillover effects of large theme parks on economic growth in neighboring counties are confined to areas within 50 kilometers of counties hosting theme parks. When the distance exceeds a certain threshold, the spillover effect diminishes and eventually becomes negligible.

4.5.1 Spatial Spillover by Industry

In subsection 4.5, we document a spillover effect of theme park openings on neighboring counties located 50 kilometers away from those with large theme parks. In this subsection,

Figure 6: Spatial Spillover Effects of Theme Parks



Notes: The x-axis represents the distance from the treated county in miles. Each point (and 95% CI) represents estimates from a regression for the dependent variable across distances. County and year FEs are included. All panels include the interaction terms covariates W_{i0} and year dummies. We cluster the standard errors at the county level.

we will explore how this spillover effect varies by industry, specifically identifying which industries are most and least affected. Following the industry categories discussed above, I re-estimate Equation 5 for each industry group and present the results in Figure 7. Panels (a) and (b) focus on the agriculture and manufacturing sectors, respectively. The results indicate no significant effect of theme park openings on business creation in these two industries, both in the host county and in the neighboring counties, regardless of the distances.

Panel (c) presents the results for the Hospitality and Entertainment industry, where we observe that, when the distance is within 25 kilometers, the impact on newly registered firms in the hotel, restaurant, culture, sports, and entertainment sectors increases significantly by 20%. As the distance extends to 75 kilometers, the positive effect decreases by one-half, around 10%. These findings suggest a spillover effect on business creation within this industry. Similarly, panel (d) displays the results for the retail industry, which also shows a spillover effect within a distance range of 50 kilometers. This is because these industries rely on visitors who prioritize convenience. Most tourists are unwilling to travel far away from the park for dining, shopping, or staying overnight. Therefore, the spillover effect in these industries diminishes at greater distances.

Panel (e) illustrates the results for the construction industry, which reveals no significant spillover effect in neighboring counties. Conversely, panel (f) presents the results for the real estate industry, where a significant spillover effect occurs in counties within the 50-75 kilometer range. Compared to the Leisure & Hospitality and retail industries, the spillover effect in the real estate sectors is at greater distances from theme parks. This is because people are more willing to commute longer distances for housing, especially if transportation infrastructure improves due to the theme park. When theme parks increase the land prices in the hosting counties, living 50-75 kilometers away from the park may be more affordable.

Furthermore, panel (g) depicts the utility industry, and panel (h) plots the results for other industries. We do not find significant spillover effects in these industries regardless of distance. Utilities are mainly driven by long-term investment and infrastructure planning, which are typically not influenced by localized economic activities such as theme park openings. Similarly, many industries in the "other" category, such as health services, logistics, warehouses, repairs, postal, and resident services, are less dependent on consumer-driven tourism. Therefore, these industries exhibit negligible spillover effects of theme park openings.

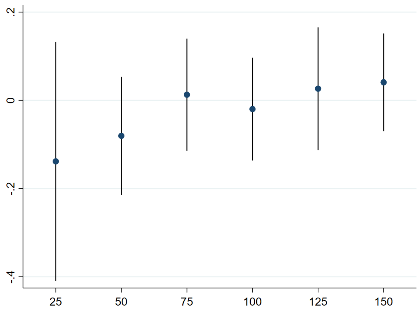
Based on these results, we can conclude that theme park openings have positive and significant spillover effects on neighboring counties within 75 kilometers, but these positive and significant spillover effects mainly occur in the tourism-related industries and real estate industry, which reflects the localized economic impacts of theme park openings. This also suggests that sectors closely tied to consumer-driven tourism and regional development are the primary beneficiaries of the economic stimulus generated by theme parks.

5 Mechanisms

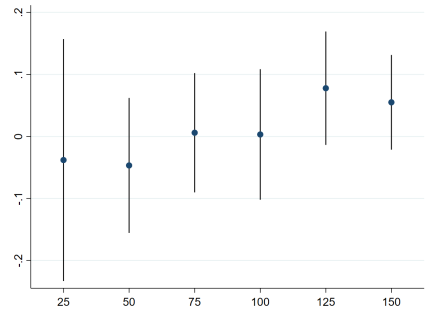
This section sheds light on the potential mechanisms that could drive the effects of theme park openings on local business creation we show in the previous section. Since the related data such as number of tourists, land prices, and industrial employment are at the city-level, this section provides evidence at the city-year level.[30]

[30] City-level administrative divisions are a primary level of local administration in China, positioned between provincial and county-level divisions. These include prefecture-level cities, prefectures, autonomous prefectures, and leagues.

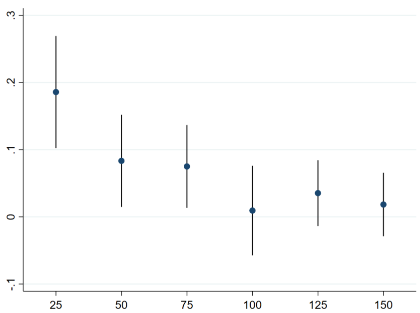
Figure 7: Spatial Spillover Effects of Theme Parks



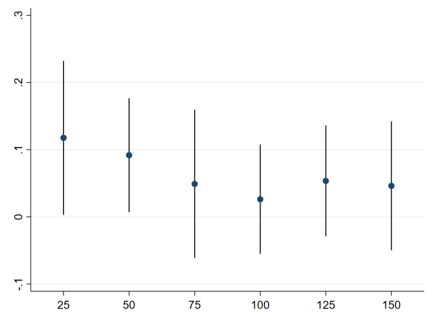
(a) Agriculture



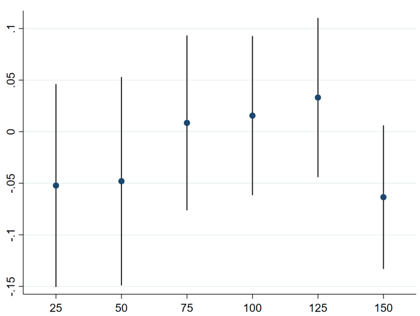
(b) Manufacturing



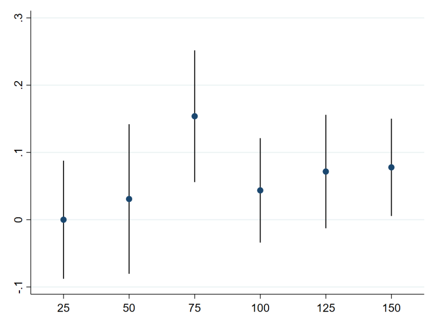
(c) Hospitality & Entertainment



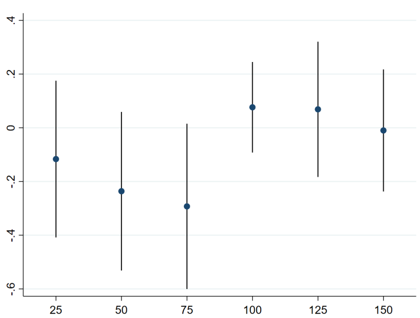
(d) Retail



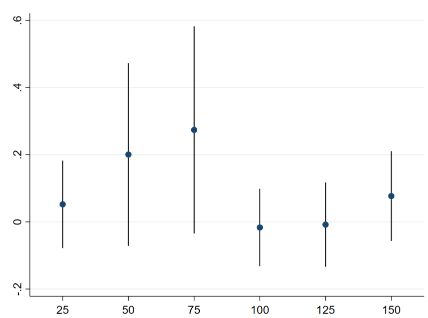
(e) Construction



(f) Real Estate



(g) Utility



(h) Others

Notes: The x-axis represents the distance from the treated county. Each point (and 95% CI) represents estimates from a regression for the dependent variable across distances. County and year FEs are included. All panels include the interaction terms covariates W_{i0} and year dummies. We cluster the standard errors at the county level.

5.1 Theme Parks and Tourism

One potential mechanism is that theme parks attract more tourists, both domestic and international, who contribute to increased consumption, thereby driving business creation and GDP growth. To investigate this, we examine the effect of theme park openings on tourism using the following specification:

$$tourism_{j(i)t} = \lambda_{j(i)} + \lambda_t + \gamma_1 D_{j(i)t} + X_{j(i)t}\gamma_2 + e_{j(i)t} \tag{6}$$

where $tourism_{j(i)t}$ denotes the total number of tourists, including both domestic and international tourists coming to the city j in year t. The indicator variable $D_{j(i)t} = 1$ if the city j, where the county i in our baseline sample is located, experiences a theme park opening in year t, and $D_{j(i)t} = 0$ if no theme park opening occurs in city j where county i is located during year t.

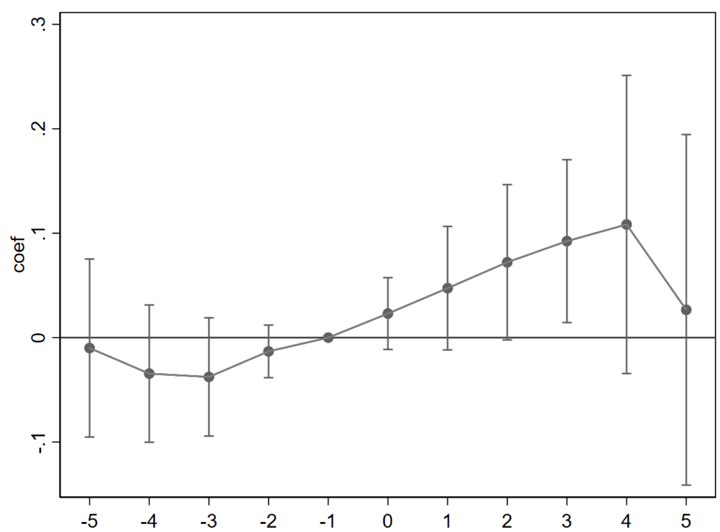
$\lambda_{j(i)}$ and λ_t are prefecture-city and year fixed effects, respectively. $\lambda_{j(i)}$ captures all the invariant unobserved variables across cities, such as cultural attributes, historical factors, geographic locations, and governance styles that may affect the tourism attractiveness of the locations. λ_t controls for any external factors that impact all cities in a given year, such as national economic growth, inflation, policy changes, or common events like financial crises or pandemics.

$X_{j(i)t}$ represents an interaction between key covariates from the initial years of the data and year dummies. Specifically, we interact the log of prefecture-city level GDP, the share of secondary industry GDP, the share of tertiary industry GDP, and the log of population in 2001 with year dummies. This allows for the capture of time-varying effects of initial economic conditions on the dependent variable in each year. The error term is denoted as $e_{j(i)t}$, and standard errors are clustered at the prefecture-city level to account for within-city correlations over time.

Figure 8 presents the results of tourism. Importantly, no pretrend is observed. The findings indicate that the theme park openings have no significant effect on tourist numbers in the first two years. However, in the third year after the opening, the total number of tourists increased by 11%, with domestic tourists increasing by 6% and international tourists by approximately 10%. This positive impact becomes insignificant four years after the event.

Overall, theme parks significantly boost tourism by attracting visitors and driving overall tourism growth in the host city. This surge in tourism fosters local business creation.

Figure 8: Effects on Tourists



Notes: Each point (and 90% CI) represents estimates from a regression for the dependent variable: the total number of tourists arriving in the city. City and year FEs are included. In all panels, the interaction terms between time-invariant city-level covariates W_{j0} and year dummies are included. We cluster the standard errors at the prefecture-city level.

To further understand the role theme park openings play in local development through tourism, we investigate their impacts on the land market. Previous research by Nocito et al. (2023) explores the relationship between tourism and the housing market, revealing that increased tourism expenditure can lead to capital gains through rising rental and selling prices of commercial properties. Drawing inspiration from this study, we posit that theme park openings similarly influence land markets, acting as a tourism multiplier. Given that theme parks typically require extensive acreage, their establishment is likely to increase demand for land in the surrounding area, resulting in higher land prices. Therefore, in this subsection, we estimate the effect of theme park openings on land prices as follows:

$$\ln price_{j(i)t} = \lambda_{j(i)} + \lambda_t + \gamma_5 D_{j(i)t} + \gamma_6 X_{j(i)t} + \zeta_{j(i)t} \tag{7}$$

where $price_{j(i)t}$ represents the selling prices of land in the unit of 10,000 yuan and is transformed in the log of transformation in the regression analysis. All other terms are the same as Equation 6. The parameter of interest is $gamma_5$, which is anticipated to be positive and statistically significant.

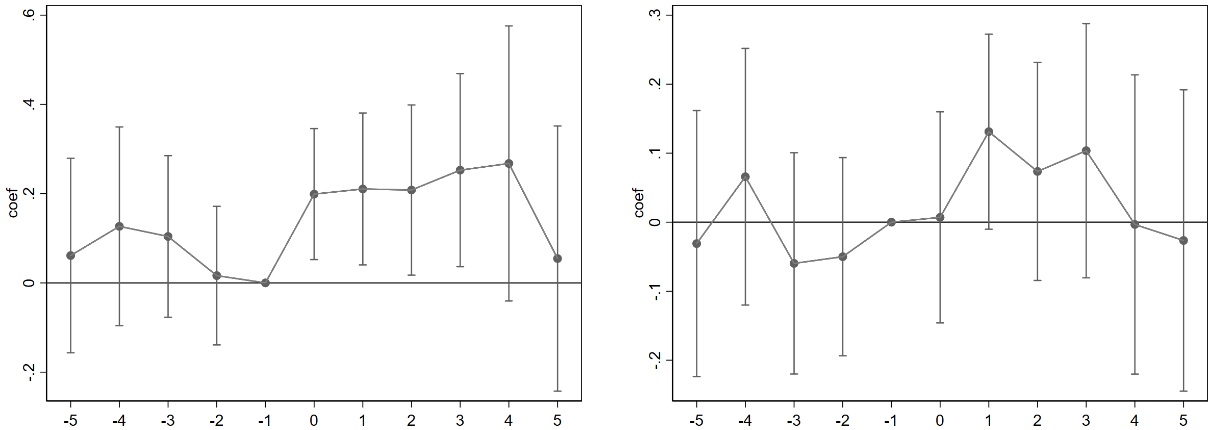
The results, illustrated in Figure 9, indicate that land prices increase by approximately 20% following the opening of a theme park. This positive effect not only persists but also intensifies over the three years after the theme park opens, reflecting a sustained positive influence on the local land market. The rise in land prices can stimulate business creation as higher property values may attract investors and entrepreneurs seeking to capitalize on the increased economic activity generated by the theme park. However, by the fourth year, the impact becomes statistically insignificant. This decline may suggest that the initial excitement surrounding the theme park wanes over time, or that the market reaches a new equilibrium as the supply of land adjusts to the increased prices. Overall, these findings align with expectations, highlighting that the opening of a theme park contributes significantly to rising land prices in the city, underscoring its potential as a catalyst for local economic development.

One potential concern is that the observed increase in land prices may be driven by a reduction in land supply. To address this, we analyze changes in land sale areas following the openings of theme parks at the prefecture-city level. Specifically, we estimate the impact of theme park openings on the log of total land sale area in each prefecture city using the regression model outlined in Equation 11. The specification includes prefecture-city fixed effects and year-fixed effects to control for unobserved, time-invariant heterogeneity across cities and macroeconomic shocks over time. Furthermore, we interact time-invariant city characteristics with year dummies. The results, shown in panel (b) of Figure 9, indicate that theme park openings do not have a statistically significant effect on land sale areas, alleviating concerns that the rise in land prices may be driven by a contraction in land supply.

5.2 Theme Parks and Agglomeration

Tao et al. (2019) documents the existence of agglomeration economies in the creative industry which includes the theme park industry. Therefore, theme park openings attract

Figure 9: Effects on Land Market



(a) Land Prices (b) Land Areas

Notes: Each point (and 95% CI) represents estimates from a regression for the log of land prices in panel (a) and for the log of land sale areas in panel (b). City and year FEs are included. The interaction terms between covariates W_{j0} and year dummies are included. I cluster the standard errors at the prefecture-city level.

business creation as firms can enjoy higher productivity through the agglomeration effects. In this section, we construct the industrial specialization and diversity indexes following Combes and Gobillon (2015) to examine the agglomeration effects. To do so, we collect the industrial employment data from the China City Statistical Yearbook. Since this data is obtained from 2003 to 2019, in which city-industry-level employment data is available.[31,32,33] If $specialization_{jt}$ is greater than 1, the city's Cultural, Sports, and Entertainment industries have a higher concentration than the national average. This suggests agglomeration or clustering of firms within the industry.

The specialization index measures the concentration of specific industries within a region compared to a broader benchmark. Therefore, it is constructed as follows:

$$specialization_{jt} = \frac{employment_{jkt}/employment_{jt}}{\sum_j employment_{jkt}/\sum_j employment_{jt}} \tag{8}$$

[31] According to http://www.tjnjw.com/hangyefb/c/zhongguo-chengshi-tongjinianjian-2021.html, China City Statistical Yearbook stops reporting the employment by industries in 2020.

[32] Before 2003, the China City Statistical Yearbook categorized industries differently, and employment in the Cultural, Sports, and Entertainment Industry was not included. Therefore, the analysis begins in 2003.

[33] The employment data on Chongqing in 2015, including total employment, and employment in each industry, are obtained from the National Bureau of Statistics of China.

where $specialization_{jt}$ measured by city j's employment in a particular industry k in year t. Here, k mainly denotes the Cultural, Sports, and Entertainment industry which includes the theme park industry. $employment_{jkt}$ is the number of people employed in city j in a particular industry k in year t while $employment_{jt}$ is the total employment in city j in year t.

The diversity index measures how evenly economic activities or employment are distributed across different industries in a region and it is constructed by the inverse of the Herfindahl index:

$$diveristy_{jt} = \frac{1}{[\sum_{k'} \frac{employment_{jk't}}{\sum_{k'} employment_{jk't}}]^2} \tag{9}$$

where $employment_{jk't}$ denotes the number of people employed in the city j in the industry k' in year t. k' contains all the industries in the Tertiary industry except for the Cultural, Sports, and Entertainment industries. $diveristy_{jt}$ is ranged from $[1, N_{k'}]$ where $N_{k'}$ is the number of industries in the Tertiary industry except for the Cultural, Sports, and Entertainment industries. Overall, specialization relates to the role of the industry's local share, and diversity relates to the role of the distribution of employment over all other industries. The two indices capture two different types of mechanisms. In particular, whereas specialization is a determinant of localization economies, the Herfindahl index is a determinant of urbanization economies.

By constructing the two indexes of agglomeration, I estimate the effect of theme park openings on agglomeration as follows:

$$\ln agglomeration_{j(i)t} = \lambda_{j(i)} + \lambda_t + \gamma_3 D_{j(i)t} + \gamma_4 X_{j(i)t} + \epsilon_{j(i)t} \tag{10}$$

where $agglomeration_{j(i)t} \in \{specialization_{j(i)t}, diversity_{j(i)t}\}$. Since $agglomeration_{j(i)t}$ is positive and does not include zero values, I take the log of $agglomeration_{j(i)t}$ in the regression analysis. All other terms are the same as Equation 6.

The results, displayed in panel (a) of Figure 10, indicate that theme park openings have no significant effect on industrial specialization in the first three years. However, by the fourth year, the openings increase industrial specialization in the host city by 13%. This positive and significant effect persists and increases to 20% five years after the openings. This suggests that theme parks introduce a new industry that gradually becomes more specialized over time. The increase in specialization reflects the agglomeration effects of theme parks, where related firms - such as those in tourism, entertainment, and hospitality—cluster together. This clustering allows businesses to benefit from shared infrastructure to achieve scale economies, proximity to knowledge spillovers, and a specialized labor market.

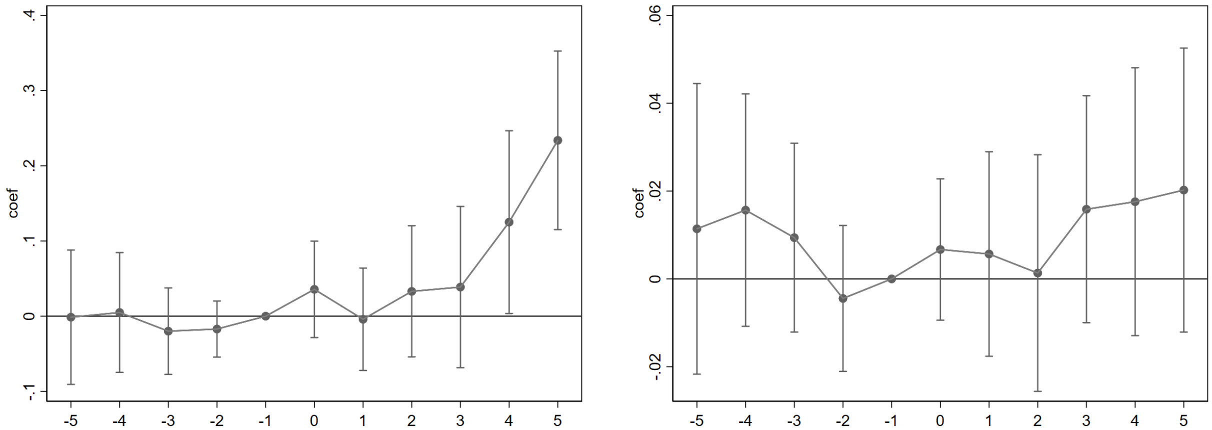
In addition, this delayed effect of specialization aligns with Duranton and Puga (2001)'s "nursery city" model, which argues that new industries often start in more diverse cities to experiment and innovate but gradually move toward specialized areas as they mature and require more concentrated resources. In this context, theme parks may initially drive general tourism, but after several years, a more specialized set of industries, such as specialized entertainment services, resort management, and advanced hospitality, emerges and begins to dominate the local economy. This shift toward specialization suggests that theme parks act as a catalyst for industrial transformation in their host cities. As Combes (2000) posits, this growing specialization can promote faster economic growth, particularly in smaller cities, as these regions become centers of focused economic activity.

Panel (b) of Figure 10 examines the effects of theme park openings on industrial diversity and finds no significant impact. This lack of change in diversity suggests that theme parks do not contribute to broadening the range of industries present in the city. Instead, their influence appears to be concentrated in specific sectors closely tied to their operations. It implies that the positive economic impact of theme parks is not driven by urbanization. Rather, theme parks foster localized growth in a narrower set of industries, contributing to regional specialization rather than diversification, which is consistent with Braun and Milman (1990) which demonstrates the localization nature of the theme park industry.

6 Aggregate Economic Impacts

This section provides evidence on the aggregate economic impacts of theme park openings on employment in Section 6.1 and on overall economic activities using nighttime lights in Section 6.2.

Figure 10: Effects on Agglomeration



(a) Specialization

(b) Diversity

Notes: Panels (a) and (b) present estimates on the log of the specialization index and the log of diversity index across time, respectively. Each point (and 95% CI) represents estimates from a regression for the dependent variable. City and year FEs are included. In all panels, the interaction terms between covariates W_{i0} and year dummies are included. I cluster the standard errors at the prefecture-city level.

6.1 Back-of-the-Envelope Calculation for Employment Impacts of New Businesses.

New businesses create new jobs, which is why policymakers place significant emphasis on business startups and have enacted various policies to encourage their establishment (Fritsch and Noseleit 2013; Faggio and Silva 2014). In this study, the baseline results suggest that counties experience a 14% increase in newly registered firms following the opening of a large theme park. However, little is known about the employment effect of this increase in business creation. This section aims to provide a back-of-the-envelope calculation to derive a rough estimate of the employment effect.

A 14% increase in the number of newly registered firms can be expressed in logarithmic terms as follows: $log(Newstartups) - log(Oldstartups) = log(1 + 0.14) = log(1.14) = 0.13$. According to Fritsch and Schindele (2011), a 1% increase in startup rate leads to an 82% increase in short-term employment contribution, which means that the change in the log of the startup rate per 1,000 employees is approximately equal to the change in the log the number of startups, $log(\frac{Newstartups}{1000}) - log(\frac{Oldstartups}{1000}) = log(1.14) = 0.13$. Since we know that a 1% increase in the startup rate (in logs) leads to an 82% increase in short-term employment contribution, we can scale this based on the actual increase in the startup rate. Thus, the effect on short-term employment contribution is 11% (= $0.13 \times 82\%$). Therefore, after the opening of a large theme park, the 14% increase in the number of newly registered firms corresponds to an estimated 11% increase in the short-term employment contribution.

6.2 Effects on Nighttime Lights

Having documented that theme park openings promote business creation and increase employment, in this subsection, we examine if the counties with theme park openings experienced faster economic growth. Many literatures use nighttime lights in empirical research and establish evidence that nighttime lights are strongly correlated with standard economic measures, including GDP (Chen and Nordhaus 2011; Henderson et al. 2012; Hodler and Raschky 2014; Storeygard 2016) and development indicators (Michalopoulos and Papaioannou 2013; Chor and Li 2024). Therefore, we gather data on nighttime nights to proxy local economic growth. Specifically, we regress the theme park openings on the log of nighttime lights using ordinary least squares (OLS) with two-way fixed effects. The regression specification is as follows:

$$\ln lights_{it} = \alpha_i + \alpha_t + \theta_1 D_{it} + \theta_2 X^l_{it} + u_{it} \tag{11}$$

where $\ln lights_{it}$ represents the nighttime lights in county i in year t; D_{it} is a dummy variable that equals one if a new theme park was opened in county i in year t and 0 otherwise. θ_1 is the coefficient of interest.

County- and year-fixed effects, α_i and α_t, are included, to account for unobserved county heterogeneity and to control for aggregate trends, such as national business cycles and fiscal policies, in nighttime lights. Additionally, following Tian et al. (2024), we not only include the county characteristics W_{i0}, such as GDP, secondary and tertiary sector GDP shares, and total population in 2001, but also include the county's intensity of nighttime lights in 2000 which is the beginning of our sample period (in logs). X^l_{it} represents the interaction terms between the mentioned county characteristics and year dummies to control for the differential time trends in nighttime lights between counties with different initial characteristics.

Table 5 presents the results. Column 1 shows that, following the opening of a large theme park, total nighttime light brightness in a county increases by approximately 7.4%. This result remains robust when applying a stacked DID approach in column 2, which shows a similar effect of 8.1%, consistent with the estimate in column 1. These findings suggest that the intensity of local nighttime lights grew more in counties with large theme park openings. Following Henderson et al. (2012) and Hodler and Raschky (2014), the back-of-the-envelope calculation suggests that a 7.4% increase in nighttime light brightness following the opening of a large theme park would lead to an increase in GDP by approximately 2.96% (=0.4×7.4%) at the sub-national level and 2.22% (=0.3×7.4%) at the national level.[34]

Table 5: Impacts on Nighttime Lights

	DID	Stacked DID
D_{it}	0.074*	0.081**
	(0.040)	(0.039)
County FEs	Yes	Yes
Year FEs	Yes	Yes
W_{i0}×Year Dummy	Yes	Yes
R^2	0.960	0.959
N	5,376	67,746

Notes: Columns 1 and 2 display the results of the log of nighttime lights using DID and stacked DID, respectively. County and year FEs are included. We also include the interaction terms between time-invariant covariates and year dummy in the regressions. Standard errors, clustered at the county level, are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

7 Conclusion

This paper examines the effects of theme park openings on regional economic development. We find that, after the opening of a large theme park, counties experience a 14% increase in

[34] Henderson et al. (2012) used national-level DMSP data and GDP data for 188 countries from 1992 to 2008 to estimate an elasticity of 0.3 for temporal changes in GDP with respect to changes in the NTL data. Hodler and Raschky (2014) used DMSP data for 1500 regions (mostly at the first sub-national level) from 82 countries from 1992 to 2009 and estimated that the elasticity for temporal changes in GDP with respect to temporal changes in DMSP data was 0.4.

business creation. This result is robust with an instrumental variable approach, alternative specifications, alternative measurements, and alternative selection methods. Furthermore, the positive impact of theme park openings on business creation varies across industries and distances. First, tourism and travel-related services, such as retail, restaurants, hotels, entertainment, and real estate experience the most significant growth in new businesses, while agricultural, manufacturing, and utility industries are not significantly affected. Second, theme park openings have a spillover effect on the business creation in the counties located within a 75-kilometer radius. Tourism and travel-related services sectors are mostly affected by the spillover effects while other industries in the neighboring counties are not affected regardless of the distance. These findings underscore the localized nature of theme parks' spillover effects on business creation.

Additionally, we examine two mechanisms—tourism and agglomeration—through which theme parks promote regional economic development. First, one year after a theme park opens, the number of tourists rises by 11%. These findings suggest that theme parks significantly boost tourism, which, in turn, stimulates tourism-related business creation and supports broader regional economic development. Second, theme parks increase industrial specialization in the host city by 16% four years after opening but do not affect industrial diversity significantly.

Finally, we gauge the aggregate economic impacts of theme park openings. A back-of-envelop calculation suggests that theme park openings increase employment by 11% and overall economic activities by 2%-3%. We further find an increase in land prices by 20% following the opening of a theme park. These results shed light on the policy implications of regional economic development. Theme park openings not only stimulate new businesses especially in tertiary industries but also generate positive spillover to neighboring counties. As such, our results provide an economic rationale for tourism-related place-based policies, such as public subsidies, aimed at boosting local economic development through theme park construction and related tourism sectors.

References

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

Bartik, T. J. (1991). Who benefits from state and local economic development policies?

Bertacchini, E., Revelli, F., and Zotti, R. (2024). The economic impact of unesco world heritage: Evidence from italy. *Regional Science and Urban Economics*, 105:103996.

Braun, B. M. and Milman, A. (1990). Localization economies in the theme park industry. *Review of Regional Studies*, 20(3):33–37.

Cengiz, D., Dube, A., Lindner, A., and Zipperer, B. (2019). The effect of minimum wages on low-wage jobs. *The Quarterly Journal of Economics*, 134(3):1405–1454.

Chen, Q., Qi, J., and Yan, G. (2023). Didplacebo: Stata module for in-time, in-space and mixed placebo tests for estimating difference-in-differences (did) models.

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

Chor, D. and Li, B. (2024). Illuminating the effects of the us-china tariff war on china's economy. *Journal of International Economics*, 150:103926.

Coates, D. and Humphreys, B. R. (1999). The growth effects of sport franchises, stadia, and arenas. *Journal of Policy Analysis and Management: The Journal of the Association for Public Policy Analysis and Management*, 18(4):601–624.

Combes, P.-P. (2000). Economic structure and local growth: France, 1984–1993. *Journal of urban economics*, 47(3):329–355.

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.

Duranton, G. and Puga, D. (2001). Nursery cities: urban diversity, process innovation, and the life cycle of products. *American Economic Review*, 91(5):1454–1477.

Faber, B. and Gaubert, C. (2019). Tourism and economic development: Evidence from mexico's coastline. *American Economic Review*, 109(6):2245–2293.

Faggio, G. and Silva, O. (2014). Self-employment and entrepreneurship in urban and rural labour markets. *Journal of Urban Economics*, 84:67–85.

Ferrara, E. L., Chong, A., and Duryea, S. (2012). Soap operas and fertility: Evidence from brazil. *American Economic Journal: Applied Economics*, 4(4):1–31.

Franco, S. F. and Macdonald, J. L. (2018). The effects of cultural heritage on residential property values: Evidence from lisbon, portugal. *Regional Science and Urban Economics*, 70:35–56.

Fritsch, M., Haupt, H., and Ng, P. T. (2016). Urban house price surfaces near a world heritage site: Modeling conditional price and spatial heterogeneity. *Regional Science and Urban Economics*, 60:260–275.

Fritsch, M. and Noseleit, F. (2013). Indirect employment effects of new business formation across regions: The role of local market conditions. *Papers in Regional Science*, 92(2):361–383.

Fritsch, M. and Schindele, Y. (2011). The contribution of new businesses to regional employment—an empirical analysis. *Economic Geography*, 87(2):153–180.

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review*, 110(8):2586–2624.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of econometrics*, 225(2):254–277.

Henderson, J. V., Storeygard, A., and Weil, D. N. (2012). Measuring economic growth from outer space. *American economic review*, 102(2):994–1028.

Hilbe, J. (2011). *Negative binomial regression*. Cambridge University Press.

Hodler, R. and Raschky, P. A. (2014). Regional favoritism. *The Quarterly Journal of Economics*, 129(2):995–1033.

Kahn, M. E., Sun, W., Wu, J., and Zheng, S. (2021). Do political connections help or hinder urban economic growth? evidence from 1,400 industrial parks in china. *Journal of Urban Economics*, 121:103289.

Lanzara, G. and Minerva, G. A. (2019). Tourism, amenities, and welfare in an urban setting. *Journal of Regional Science*, 59(3):452–479.

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):325–360.

Michalopoulos, S. and Papaioannou, E. (2013). Pre-colonial ethnic institutions and contemporary african development. *Econometrica*, 81(1):113–152.

Nocito, S., Sartarelli, M., and Sobbrio, F. (2023). A beam of light: Media, tourism and economic development. *Journal of Urban Economics*, 137:103575.

Scavette, A. (2023). The economic impact of a casino monopoly: Evidence from atlantic city. *Regional Science and Urban Economics*, 103:103952.

Silva, J. S. and Tenreyro, S. (2006). The log of gravity. *The Review of Economics and statistics*, pages 641–658.

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

Stel, A. v., Carree, M., and Thurik, R. (2005). The effect of entrepreneurial activity on national economic growth. *Small business economics*, 24:311–321.

Storeygard, A. (2016). Farther on down the road: transport costs, trade and urban growth in sub-saharan africa. *The Review of economic studies*, 83(3):1263–1295.

Szabó, A. and Ujhelyi, G. (2024). National parks and economic development. *Journal of Public Economics*, 232:105073.

Tao, J., Ho, C.-Y., Luo, S., and Sheng, Y. (2019). Agglomeration economies in creative industries. *Regional Science and Urban Economics*, 77:141–154.

Tian, W., Wang, Z., and Zhang, Q. (2024). Land allocation and industrial agglomeration: Evidence from the 2007 reform in china. *Journal of Development Economics*, 171:103351.

Tian, X. and Xu, J. (2022). Do place-based policies promote local innovation and entrepreneurship? *Review of Finance*, 26(3):595–635.

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

Wu, Y., Shi, K., Chen, Z., Liu, S., and Chang, Z. (2022). Developing improved time-series dmsp-ols-like data (1992–2019) in china by integrating dmsp-ols and snpp-viirs. *IEEE Transactions on Geoscience and Remote Sensing*, 60:3135333.

Zhang, Y., Li, X., Cárdenas, D. A., and Liu, Y. (2022). Calculating theme parks' tourism demand and attractiveness energy: a reverse gravity model and particle swarm optimization. *Journal of Travel Research*, 61(2):314–330.

Zheng, S., Sun, W., Wu, J., and Kahn, M. E. (2017). The birth of edge cities in china: Measuring the effects of industrial parks policy. *Journal of Urban Economics*, 100:80–103.

Appendix A Results Appendix

A.1 Propensity Score Matching (PSM)

We consider counties with large theme parks treated groups and those without large theme parks as control groups. However, to ensure that the control groups are comparable to the treated groups, we select the control groups by employing the PSM methodology. To estimate the propensity score, we adopt a Logit specification to model the likelihood of a large theme park opening in a county between 2000 and 2020:

$$\frac{P_i}{1-P_i} = W_{i0}\alpha + \eta_i, \tag{A.1}$$

where P_i represents the probability of county i being exposed to a large theme park. The set of explanatory variables W_{i0} includes the county-level invariant covariates that may affect the openings of large theme parks. For example, the county's population, GDP, share of GDP in the secondary industry, and share of GDP in the tertiary industry. Due to data availability, we used population GDP-related covariates in 2001.

Panel (a) in Figure A.1 shows the estimated propensity scores.[35] All counties in the treated group are on support while a fraction of the counties in the control group are off support. Overall, the assumption of common support is mostly verified. Further, to ensure the matched counties are useful and appropriate, I only keep the on-support groups without missing weights and use the weights as the inverse of the probability in the regression. Panel (b) shows the comparisons of the covariates before and after matching. The standardized bias is reduced and smaller than 20% for all the covariates. Panels (c) and (d) show the kernel density of the propensity scores before and after matching, respectively. The results indicate that, after matching, the mean and distribution of the control group are closer to those of the treated group compared to before matching, which suggests that the PSM method helps select a control group that is comparable to the treated group.

Table A.1 reports the descriptive statistics for the main variables in the analysis. On

[35] In the baseline sample, there are 58 counties in the treated group and 206 counties in the control group. On average, each treated county is paired with 4 untreated counties. Therefore, I use 4-nearest neighbors matching and set the caliper at 0.05 level.

average, there are 6,873 newly registered firms per year in a county within the final sample, with 6,069 of these in the tertiary industry. In the treated group, the average is higher, with 7,793 newly registered firms per year, including 6,918 in the tertiary sector. In contrast, the control group has an average of 5,937 newly registered firms per year, with 5,204 in the tertiary industry. The mean value of the dummy variable D_{it} is 0.174, indicating that, on average, 17.4% of the counties experienced large theme park openings during the sample period. In the treated group, this figure rises to 34.5%. For the covariates, the mean values for the log of GDP and population in 2001 for the final sample are 3.591 and 3.750, respectively. The mean shares of secondary and tertiary GDP in the final sample are 0.443 and 0.401, respectively, indicating a balanced distribution of economic activity between the two sectors in the final sample counties. Importantly, after matching, the mean values of the main variables and covariates in the control group are closer to those in the treated group, compared to their pre-matching mean values, which further indicates that the control group after selection using PSM is more comparable to the treated group.

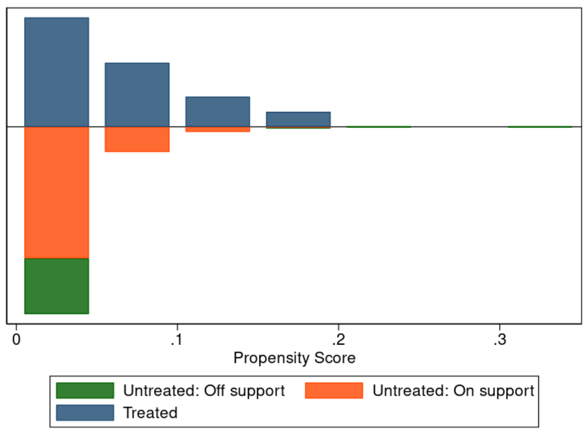
Figure A.2 plots the distribution of the treated and control counties after matching. 56 counties are matched in the treated group and 55 counties are matched in the control group. Most of the counties are located in the east of China, such as Jiangsu Province, Shandong Province, and Zhejiang Province. There is no county in the west of China, such as Tibet, Ningxia, or Xinjiang.

Table A.1: Summary Statistics

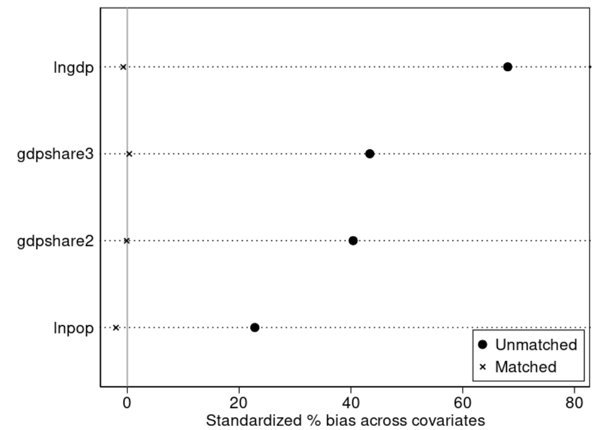
	Full Sample		Treated Group	Control Group	
				Before Matching	After Matching
Variable	Mean	SD	Mean	Mean	Mean
Main Variables					
Firms (Total)	6873.389	8940.347	7793.060	3522.399	5936.997
Firms (Tertiary)	6068.772	8259.100	6918.219	3031.153	5203.880
D_{it}	0.174	0.379	0.345	0	0
Covariates					
Log(GDP)	3.591	1.017	3.625	2.886	3.557
GDP Share (Secondary Industry)	0.443	0.159	0.447	0.371	0.440
GDP Share (Tertiary Industry)	0.401	0.143	0.418	0.352	0.424
Log(Population)	3.750	0.630	3.775	3.574	3.724
Number of Observations	2,331	2,331	1,176	47,502	1,155

Note: For variables used in regressions, we report the mean and the standard deviations for the full sample, as well as the mean value in the treated group and control group before and after matching. We also report the mean and standard deviations for the log of GDP, the share of GDP in the secondary industry, the share of GDP in the tertiary industry, and the log of population in 2001, as well as their mean values before and after matching in the treated group and control group.

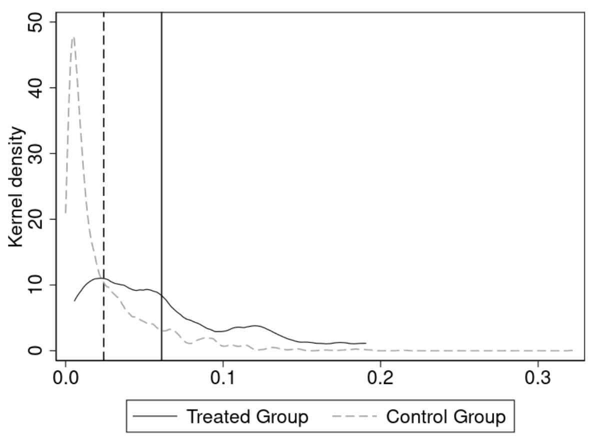
Figure A.1: Propensity Score Matching Results



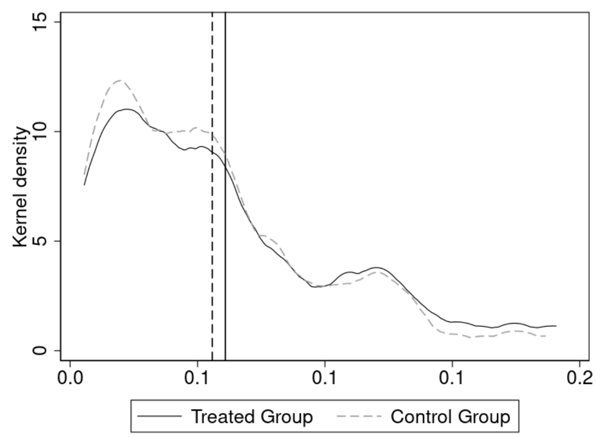
(a) PS Scores



(b) Balanced Test



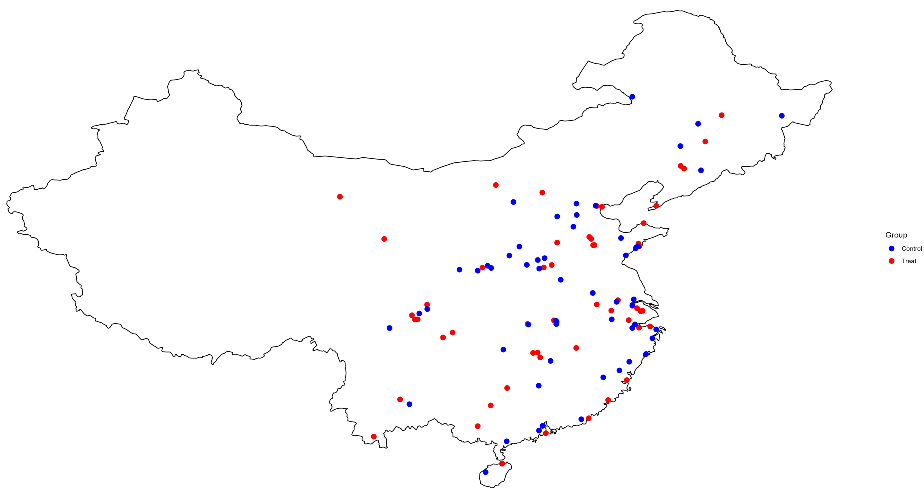
(c) Before Matching



(d) After Matching

Notes: Panel a depicts the propensity scores for treated and control observations. On-support means that I can find a matched market. Conversely, off-support means that there is no matched county can be found. We only keep the observations that are on support without missing weights in our estimation. Panel b plots the standard bias of covariates used in matching for unmatched samples and matched samples, respectively. Panel c plots the kernel density of the propensity scores for treated and control groups before matching while panel d plots the kernel density of the propensity scores for treated and control groups after matching.

Figure A.2: Treated and Control Counties After PSM



Notes: The figure shows the map of Mainland China in 2020. Red points indicate counties that were exposed to large theme parks during the sample period before 2020. The blue points represent counties that did not have any large theme parks before 2020. The counties in Beijing, Shanghai, Shenzhen, and Guangzhou are excluded from the list exposed to large theme parks.

Appendix B Data Appendix

B.1 Theme Park Data

Table B.2: Theme Park Data

Park	Province	City	County	Year
Guiyang Happy World	Guizhou	Guiyang	Baiyun District	1988
Splendid China Folk Village	Guangdong	Shenzhen	Nanshan District	1989
Shanghai Film Park	Shanghai	Shanghai	Songjiang District	1992
Dalian Laohutan Ocean Park	Liaoning	Dalian	Zhongshan District	1992
World Park	Beijing	Beijing	Fengtai District	1993
Window of the World, Shenzhen	Guangdong	Shenzhen	Nanshan District	1994
Hangzhou Songcheng Park	Zhejiang	Hangzhou	Xihu District	1996
Hengdian World Studios	Zhejiang	Jinhua	Dongyang City	1996
Changsha Window of the World	Hunan	Changsha	Kaifu District	1997
Happy Valley Shenzhen	Guangdong	Shenzhen	Nanshan District	1998
China Dinosaur Park	Jiangsu	Changzhou	Xinbei District	2000
Chongqing Ledu Theme Park	Chongqing	Chongqing	Yongchuan District	2000
Guilin Merryland Theme Park	Guangxi	Guilin	Xing'an County	2000
Changchun Movie Wonderland	Jilin	Changchun	Nanguan District	2005
Qingdao Polar Ocean World	Shandong	Qingdao	Laoshan District	2006
Beijing Happy Valley	Beijing	Beijing	Chaoyang District	2006
Fushun Royal Polar Ocean World	Liaoning	Fushun	Wanghua District	2006
Dalian Haichang Discoveryland	Liaoning	Dalian	Jinzhou District	2006
Guangzhou Chimelong Paradise	Guangdong	Guangzhou	Panyu District	2006
Wuhu Fantawild Adventure	Anhui	Wuhu	Jinghu District	2007
Baotou Amusement Park	Inner Mongolia	Baotou	Kundulun District	2007
Chengdu Happy Valley	Sichuan	Chengdu	Jinniu District	2009
Shanghai Happy Valley	Shanghai	Shanghai	Songjiang District	2009
Yancheng Spring and Autumn Theme Park	Jiangsu	Changzhou	Wujin District	2010
Wuhu Fantawild Oriental Heritage	Anhui	Wuhu	Jiujiang District	2010
Wuhu Fantawild Dreamland	Anhui	Wuhu	Jiujiang District	2010
Tai'an Fantawild Adventure	Shandong	Tai'an	Taishan District	2010
Shantou Fantawild Adventure	Guangdong	Shantou	Longhu District	2010
Zhuzhou Fantawild Adventure	Hunan	Zhuzhou	Yunlong Demonstration Zone	2011
Global Animation Joyland	Jiangsu	Changzhou	Wujin District	2011
Shenyang Fantawild Adventure	Liaoning	Shenyang	Shenbei New District	2011
Qingdao Fantawild Dreamland	Shandong	Qingdao	Chengyang District	2011
Wuhan Haichang Polar Ocean World	Hubei	Wuhan	Dongxihu District	2011
Feng Xiaogang Movie Town	Hainan	Haikou	Longhua District	2011
Wuhan Happy Valley	Hubei	Wuhan	Hongshan District	2012
Chengdu Guose Tianxiang Fairy Tale World	Sichuan	Chengdu	Wenjiang District	2012
Zhengzhou Fantawild Adventure	Henan	Zhengzhou	Zhongmu County	2012
Tianjin Happy Valley	Tianjin	Tianjin	Dongli District	2013
Xiamen Fantawild Dreamland	Fujian	Xiamen	Tong'an District	2013
Tai'an Sun Tribe Theme Park	Shandong	Tai'an	Daiyue District	2013
Wuhu Fantawild Water Park	Anhui	Wuhu	Jiujiang District	2010

Table B.2: Theme Park Data (Continued)

Park	Province	City	County	Year
Zhuhai Chimelong Ocean Kingdom	Guangdong	Zhuhai	Xiangzhou District	2014
Zhengzhou Fantawild Water Park	Henan	Zhengzhou	Zhongmu County	2014
Quancheng Europark Dreamworld	Shandong	Dezhou	Qihexian County	2014
Tianjin Fantawild Adventure	Tianjin	Tianjin	Binhai New Area	2014
Penglai Europark Dreamworld	Shandong	Yantai	Penglai City	2015
Jinan Fantawild Oriental Heritage	Shandong	Jinan	Huaiyin District	2015
Guian Happy World	Fujian	Fuzhou	Lianjiang County	2015
Zhengzhou Fantawild Dreamland	Henan	Zhengzhou	Zhongmu County	2015
Xishuangbanna Sunac Land	Yunnan	Xishuangbanna	Jinghong City	2015
Jiayuguan Fantawild Adventure	Gansu	Jiayuguan	Jiayuguan City	2015
Hangzhou Hello Kitty Park	Zhejiang	Huzhou	Anji County	2015
Lehua Happy World	Shaanxi	Xi'an	Xixian New District	2015
Shanghai Disneyland	Shanghai	Shanghai	Pudong New District	2016
Hefei Sunac Land	Anhui	Hefei	Binhu New District	2016
Ningbo Fantawild Oriental Heritage	Zhejiang	Ningbo	Cixi City	2016
Zhuzhou Fantawild Dreamland	Hunan	Zhuzhou	Yunlong Demonstration Zone	2016
Nanchang Sunac Land	Jiangxi	Nanchang	Xinjian District	2016
Oriental Land Cool Kingdom	Zhejiang	Shaoxing	Keqiao District	2016
Datong Fantawild Adventure	Shanxi	Datong	Pingcheng District	2016
Huayi Brothers Changsha Movie Town	Hunan	Changsha	Yuelu District	2016
Xiamen Fantawild Oriental Heritage	Fujian	Xiamen	Tong'an District	2017
Xiamen Fantawild Water Park	Fujian	Xiamen	Tong'an District	2017
Fushun Hotgo Jungle Park	Liaoning	Fushun	Wanghua District	2017
Chongqing Happy Valley	Chongqing	Chongqing	Yubei District	2017
Harbin Sunac Land	Heilongjiang	Harbin	Songbei District	2017
Colorful Yunnan Joy World	Yunnan	Kunming	Jinning District	2018
Suzhou Huayi Brothers Movie World	Jiangsu	Suzhou	Suzhou Industrial Park	2018
Liuzhou Cray Bay Water Park	Guangxi	Liuzhou	Liudong New District	2018
Shanghai Haichang Ocean Park	Shanghai	Shanghai	Pudong New District	2018
Changying Global 100 Fantasy Park	Hainan	Haikou	Xiuying District	2018
Handan Fantawild Colorful Spring and Autumn Theme Park	Hebei	Handan	Cixian County	2019
Bund Beyond Movie Park	Zhejiang	Jinhua	Dongyang City	2019
Zhengzhou Jianye Huayi Brothers Movie Town	Henan	Zhengzhou	Zhongmu County	2019
Xining Xinhua Union Dreamland	Qinghai	Xining	Huangzhong District	2019
Wuxi Sunac Land	Jiangsu	Wuxi	Binhu District	2019
Jinzhou Fantawild Oriental Heritage	Hubei	Jinzhou	Jingzhou District	2019
Jiayuguan Fantawild Silk Road Dreamland	Gansu	Jiayuguan	Jiayuguan City	2019
Changsha Fantawild Oriental Heritage	Hunan	Changsha	Ningxiang City	2019
Guangzhou Sunac Land	Guangdong	Guangzhou	Huadu District	2019
Chengdu Sunac Land	Sichuan	Chengdu	Dujiangyan City	2020
Suzhou Amusement Park Forest World	Jiangsu	Suzhou	Huqiu District	2020
Nanjing Happy Valley	Jiangsu	Nanjing	Qixia District	2020
Mianyang Fantawild Oriental Heritage	Sichuan	Mianyang	Jiangyou City	2020
Yinji Animal Kingdom	Henan	Zhengzhou	Xinmi City	2020
Nanning Fantawild ASEAN Legend	Guangxi	Nanning	Qingxiu District	2020