# Do Place-based Policies Work? Micro-level Evidence from China's Economic Zone Program

Yi Lu[a], Jin Wang[b], Lianming Zhu[c]

[a]National University of Singapore;
[b]Hong Kong University of Science and Technology;
[c]Kyoto University

July 2015

# Introduction

## Place-Based Programs

- Place-based development programs have grown popular and been pursued by many governments around the world
  - referred to as economic development policies aimed at fostering economic growth of an area within its jurisdiction
  - e.g., enterprise zones in the US, and the regional development aid within the European Union.
- Increasingly rigorous theory and empirical analyses on the program incidence and effectiveness mostly in the context of US and Europe (Neumark and Simpson, 2014)
  - Agglomeration and redistribution (Kline and Moretti, 2014; Glaeser, Rosenthal and Strange, 2010; Gaubert, 2014; Brinkman, Coen-Pirani, and Sieg, 2015)
  - Heterogeneity in the treatment effect (Becker et al., 2013; Briant et al., 2014)

- Little attention has been paid to place-based interventions in developing countries.
  - Who benefits and who loses from place-based interventions?
  - Which factors determine the effectiveness of such interventions?
  - Would the assumptions and conceptual approaches from the U.S. and Europe place-based policy literature still hold?
- There has been little progress in addressing these issues
  - due to the lack of longitudinal studies in developing countries.
  - in particular, research that traces a place-based program's effect on micro-level units such as firms and workers.

- We exploit a natural experiment in the establishment of China's Special Economic Zones(SEZs).
  - As a world-wide development strategy: 3,000 SEZs in 135 countries in 2008.
  - SEZs have not been uniformly successful: successes in East Asia and Latin America, failures in Africa (World Bank, 2008).
  - It is largely unknown about the micro-foundations of SEZs and factors that account for the heterogeneity in the zone effectiveness.
- We intend to be the first of a series of micro-level studies to examine the place-based policies in developing countries.
  - extend Alder et al. (2013); Wang (2013); Cheng (2014).

# Introduction

## Research Question

- Using firms as a unit of analysis, our paper investigates the impact of the prominent SEZ program:
  - assess whether SEZs generate local benefits including productivity, investment and employment.
  - elucidate the mechanisms: incumbent firms, relocated firms, and entries/exiters.
  - examine heterogeneous zone effects depending on program features and characteristics of the areas being targeted.
- Methologically:
  - more comprehensive firm-level data: 2004 and 2008 economic censuses
  - disaggregated analsysis: geocoding of firm location, SEZ areas measured at the village and community level
  - difference-in-differences (DD), boundary discontinuity (BD), and a combined BD and DD (BD-DD).

# Roadmap

## Outline

- SEZ Program Background
- Identification Strategy and Data
- Empirical Results
- Conclusion

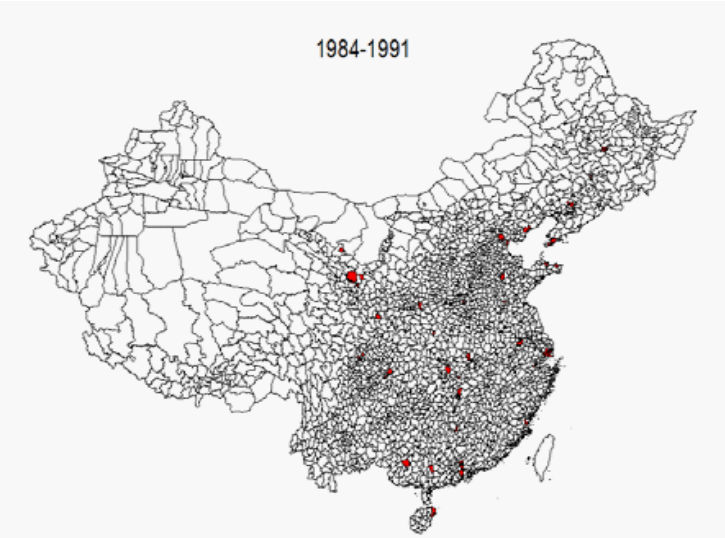
# Background

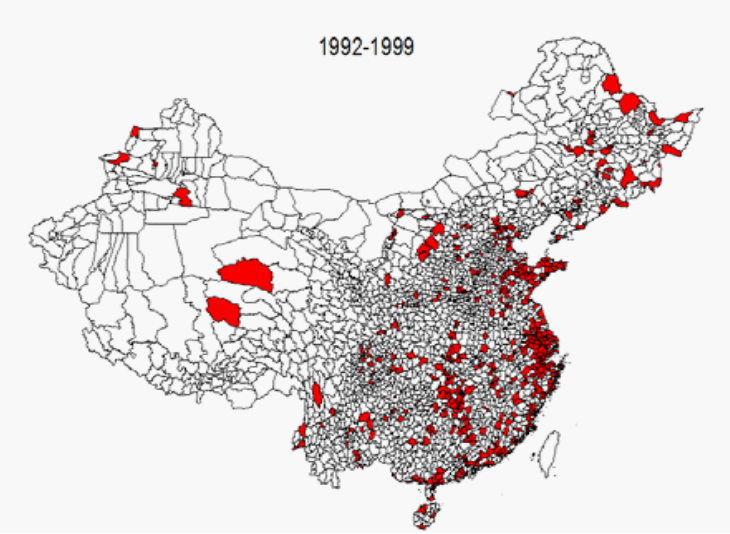
## SEZ Background

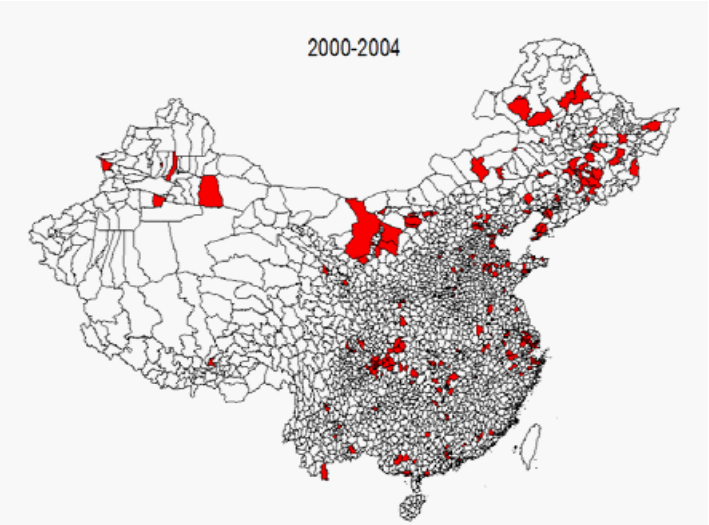
- Economic zones have been widely adopted by the Chinese government
  - special policies in an area within a jurisdiction
  - to increase foreign direct investment (FDI), domestic investment, international trade, technological cooperation and innovation, and employment.
  - zones have certain degrees of own authority to define preferential policies
- Two categories of zones
  - state-level and province-level development zones
- Preferential policies
  - tax and customs duty deduction
  - discounted land-use price
  - special treatment on bank loans

# Background

## SEZ Establishment Waves

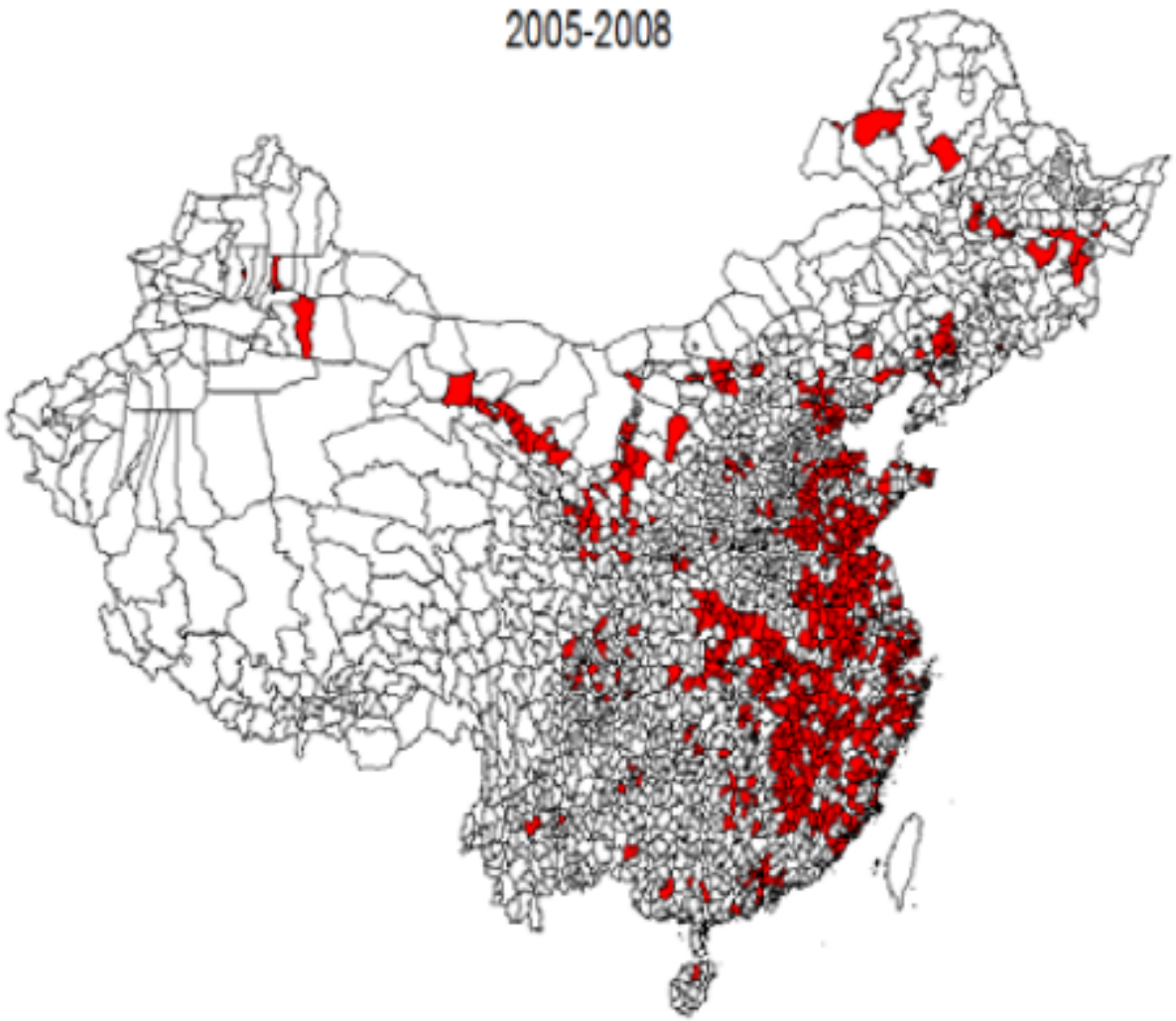



# Background

## SEZ Establishment



# Background

## SEZ Establishment Waves

| Granting Period | 1979-1983 | 1984-1991 | 1992-1999 | 2000-2004 | 2005-2008 |
|---|---|---|---|---|---|
| Number of Zones Newly established | 4 | 66 | 559 | 261 | 682 |
| # Comprehensive SEZs | 4 | | | | |
| # State-level Economic Zones in which: | | 46 | 93 | 64 | 19 |
| By Types | | | | | |
| 1. Economic and Technological Development Zone | | 12 | 23 | 17 | |
| 2. High-tech and Industrial Development Zone | | 26 | 27 | | |
| 3. Export Processing Zone | | | 1 | 39 | 18 |
| 4. Bonded Zones | | 4 | 11 | 6 | 1 |
| 5. Border Economic Cooperation Zones | | | 15 | 1 | |
| 6. Other | | 4 | 16 | 1 | |
| By Regions | | | | | |
| 1. Coastal Region | | 36 | 60 | 39 | 15 |
| 2. Central Region | | 6 | 18 | 12 | 2 |
| 3. Western Region | | 4 | 15 | 13 | 2 |
| # Province-level Economic Zones in which: | | 20 | 466 | 197 | 663 |
| By Types | | | | | |
| 1. Economic and Technological Development Zone | | 16 | 401 | 112 | 279 |
| 2. Industrial Development Zone | | 4 | 65 | 85 | 384 |
| in which: High-tech and Industrial Development Zone | | 3 | 29 | 14 | 19 |
| By Regions | | | | | |
| 1. Coastal Region | | 7 | 277 | 76 | 323 |
| 2. Central Region | | 7 | 138 | 71 | 267 |
| 3. Western Region | | 6 | 51 | 50 | 73 |

# Estimation Strategy

## Identification

- Three alternative strategies
- DD

$$y_{vct} = \gamma SEZ_{vct} + \lambda_v + \lambda_{ct} + \varepsilon_{vct}$$

  - villages from the same county
  - villages from the same town

- Identification issue
  - nonrandom selection $\longrightarrow$ non-SEZ villages are not good counterfactual
  - two periods $\nrightarrow$ no checks on pre-trends
  - placebo test: randomize villages' SEZ status

# Estimation Strategy

## Identification

- BD

$$y_{az} = \gamma SEZ_{az} + \lambda_z + \varepsilon_{az}$$

  - poineered by Black (1999) and widely applied in the literature (e.g., Bayer, Ferreira, and McMillan, 2007; Dell, 2010; Duranton, Gobillon and Overman, 2011)
  - we restrict our analysis to a sample of firms located in a narrow buffer on the two sides of the zone boundary
  - the benchmark analysis uses a bandwidth of 1,000 meters
  - two different windows (i.e., 500 meters and 2,000 meters) in the robustness checks

- Identification issue
  - zone boundary was not randomly selected $\longrightarrow$ unbalancing between areas

# Estimation Strategy

## Identification

- BD-DD: a combined BD and DD analysis
  - include a year of data when zones were not established, as a control group

$$y_{azt} = \gamma SEZ_{azt} + \lambda_a + \lambda_{zt} + \varepsilon_{azt}$$

  - $\hat{\gamma}_{BD,Control} = \eta_z$ from the control group
  - $\hat{\gamma}_{BD,Treatment} = \gamma + \eta_z$ from the treatment group
  - $\hat{\gamma}_{BD-DD} = \hat{\gamma}_{BD,Treatment} - \hat{\gamma}_{BD,Control} = \gamma$
- Identifying assumption
  - the underlying location characteristics are fixed over time except for the zone policies
  - placebo tests: both outside and both inside

# Estimation Strategy

## Data–Economic Census

- Economic census
  - first and second waves by the NBS of China in 2004 and 2008
  - covering all manufacturing firms in China
  - containing firms' full basic information, such as address, location code, industry affiliation, and ownership
  - three financial and operational variables: employment, output and capital.
- Matching 2004 and 2008, for 794,386 continuing firms
  - with unique IDs, match them by firm ID (92.7%)
  - with duplicate IDs, use firm name to link observations (4.7%)
  - with a new ID due to restructuring, mergers, or privatization, use firm name, location code, name of legal person representative, phone number (2.7%)

# Estimation Strategy

## Data–Coordinates Data

- Coordinates data
  1. Firms with precise location geographical information (i.e., province, city, road name, and numbering): coordinates of firm location (50.5%).
  2. Other firms (49.5%) with incomplete information, road updates, and reporting errors
     - village or community, which corresponds to a 12-digit location code reported by the firm

# Estimation Strategy

## Data–Coordinates Data

- *Case 1*: firms with precise location geographical information (i.e., province, city, road name, and numbering)
  - we obtain firm coordinates by searching their addresses using Google's Geocoding API services
    - e.g., "238 East Nandan Road, Xuhui District, Shanghai, China" in the Google Map
    - a map with a red marker showing the specific location of the address
    - visually confirm the search: district (i.e., Xuhui District, Shanghai, China) and road (i.e., East Nandan Road)
    - then collect latitude and longitude of the address from the Google Map

# Estimation Strategy

## Data–Coordinates Data

# Estimation Strategy

## Data–Coordinates Data

- *Case 2*: other firms with incomplete addresses, road updates, and reporting errors
  - we search the name of the village or community, which corresponds to a 12-digit location code reported by the firm
    - e.g., "Zhanan Village, Liuhe Town, Taicang City, Suzhou, Jiangsu Province" in the data
    - as precise address information is not available, we rely on 12-digit location code reported by firms (e.g., “320585102202” in the data)
    - from the 12-digit location code, we can obtain which neighborhood committee that a firm belongs to (e.g., “320585102202” corresponds to “Liunan Neighborhood, Liuhe Town, Taicang City, Suzhou, Jiangsu Province”).
    - then collect latitude and longitude of the neighborhood committee from the Google Map

# Estimation Strategy

## Data–Coordinates Data

# Estimation Strategy

## Data–Firm SEZ Mapping

- Firm SEZ status
  - firms do not directly report information about their SEZ status
- Sources to identify the villages and communities within the adminstrative boundary
  - a comprehensive SEZ boundary dataset from the Ministry of Land and Resources of China.
  - various economic zones' official websites
  - NBS's administrative division codes of the People's Republic of China at the village and community level 2008
  - the Ministry of Civil Affairs website

# Estimation Strategy

## Data–Firm SEZ Mapping

- Three cases
  1. with an independent zone admin code
     - For example, Anhui Nanling Industrial Zone has an independent 12-digit administrative location code: 340223100400 (Anhui Nanling Industrial Zone Community)
  2. equivalent to a town/street, i.e., all villages/communities under the town/street will be within the zone's boundaries
     - For example, Shandong Fei County Industrial Zone administrates Tanxin Town (administrative location code: 371325105)
  3. resides in several villages and communities
     - For example, Hubei Yunmeng Economic Development Zone administrates the following eight villages and one community: Xinli Village (administrative location code: 420923100201), Heping Village (420923100202), Qianhu Village (420923100203), Hebian Village (420923100204), Zhanqiao Village (420923100205), Quhu Village (420923103220), Zhaoxu Village (420923103223), Sihe Village (420923104209), and Qunli Community (420923100007)

# Estimation Strategy

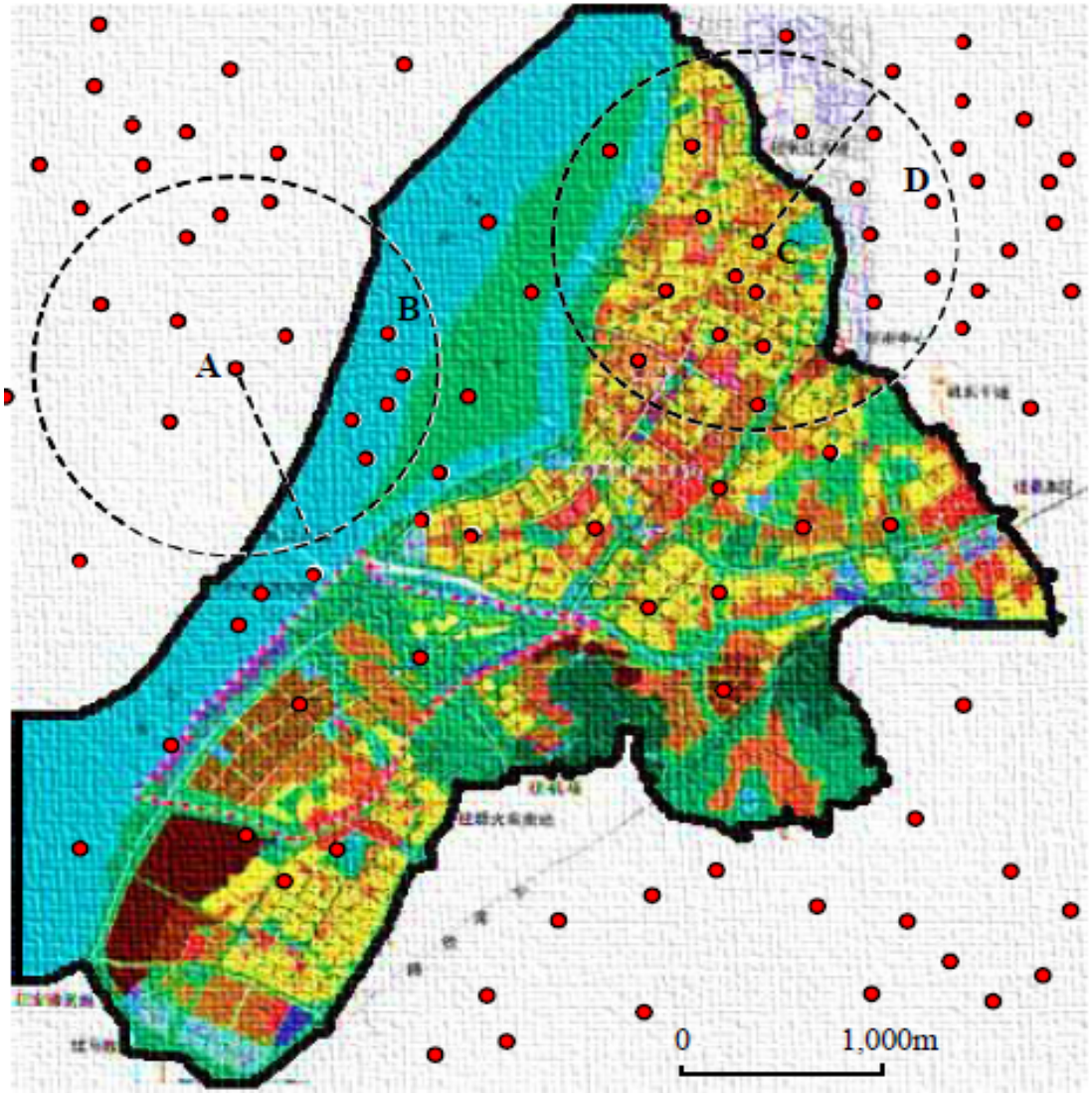
## Data–Firm SEZ Data

- The 2005-2008 wave: 663 province-level SEZs
  - a year of data before and a year of data after the zone establishment for the BD and BD-DD analyses
  - 323 SEZs in the coastal area, 267 in the central area and 73 in the western area
  - 615 Economic and Technological Development Zones (ETDZs), 48 High-technology and Special Industrial Development Zones (HSDZs)

- Regression data
  - the distance of a firm from its neighboring SEZ boundary
  - no geocodes of each SEZ boundary due to data limitation
  - instead follow the approach used by Duranton, Gobillon and Overman (2011)
    - e.g., within a 1,000 meters range from the zone boundary
    - search for a radar of 1,000 meters from the concerned firm
    - if we find there are neighboring firms located both inside and outside the zone within that radar, we designate the concerned firm as located within the 1,000 meters range from the zone boundary
    - manually doing this for all firms in the census data, we are left with a final regression sample of 587 areas with 163,069 firms located within the 1,000 meters range from the zone boundaries

# Estimation Strategy

## Data–Regression Data



# Empirical Results

## Baseline—DD Estimates

| | Within the same county | Within the same town |
|---|---|---|
| | (1) | (2) |
| Panel (a). Dependent variable: (log) employment | | |
| *InsideZone* | 0.273*** | 0.263*** |
| | (0.020) | (0.025) |
| Observations | 121,564 | 30,028 |
| Panel (b). Dependent variable: (log) output | | |
| *InsideZone* | 0.330*** | 0.331*** |
| | (0.026) | (0.032) |
| Observations | 121,564 | 30,028 |
| Panel (c). Dependent variable: (log) capital | | |
| *InsideZone* | 0.336*** | 0.340*** |
| | (0.025) | (0.031) |
| Observations | 121,564 | 30,028 |
| Panel (d). Dependent variable: (log) number of firms | | |
| *InsideZone* | 0.195*** | 0.141*** |
| | (0.015) | (0.018) |
| Observations | 121,564 | 30,028 |

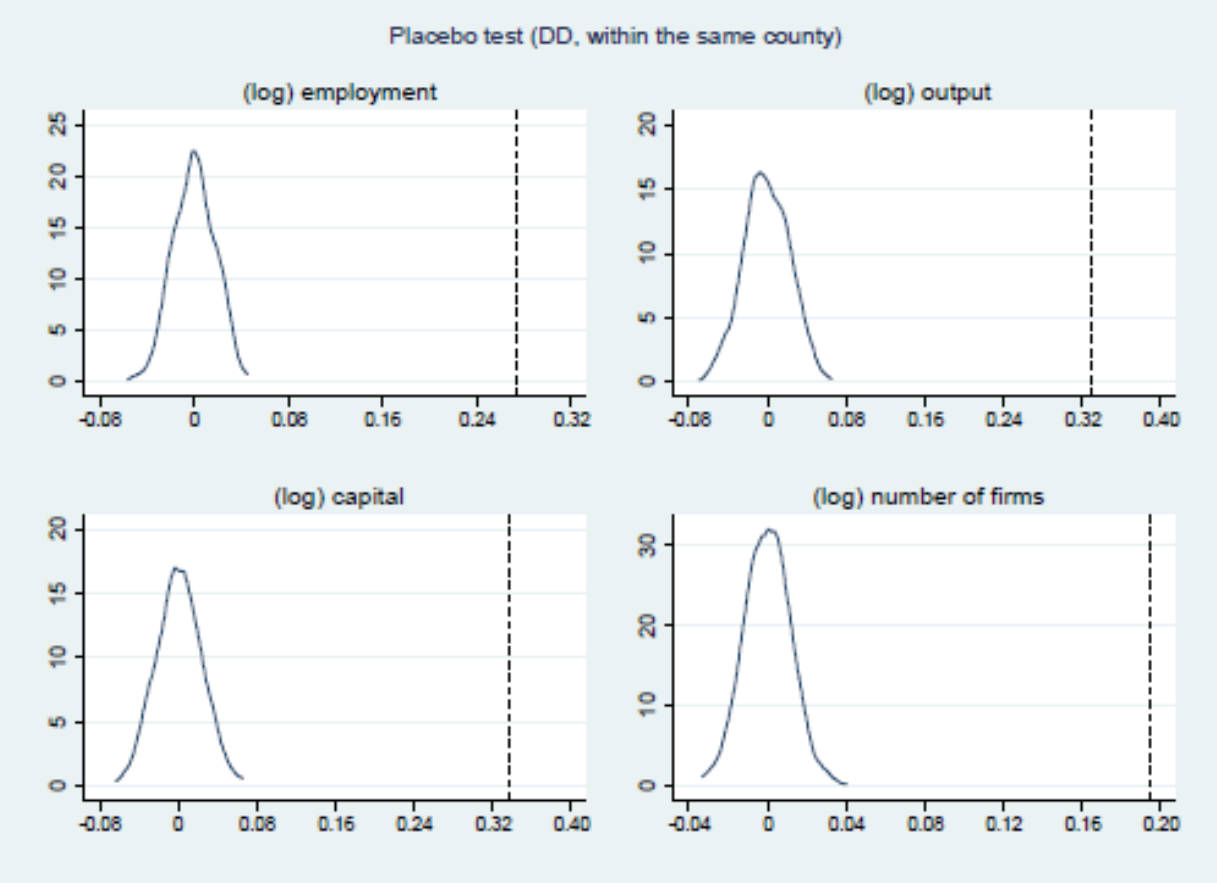
# Empirical Results

## Baseline—BD and BD-DD Estimates

| Within 1,000 meters | BD | BD-DD |
|---|---|---|
| | (1) | (2) |
| Panel (a). Dependent variable: (log) employment | | |
| *InsideZone* | 0.084 | 0.471*** |
| | (0.098) | (0.040) |
| Observations | 1,174 | 2,348 |
| Panel (b). Dependent variable: (log) output | | |
| *InsideZone* | 0.261** | 0.553*** |
| | (0.123) | (0.056) |
| Observations | 1,174 | 2,348 |
| Panel (c). Dependent variable: (log) capital | | |
| *InsideZone* | 0.307** | 0.547*** |
| | (0.122) | (0.054) |
| Observations | 1,174 | 2,348 |
| Panel (d). Dependent variable: (log) number of firms | | |
| *InsideZone* | −0.350*** | 0.233*** |
| | (0.078) | (0.031) |
| Observations | 1,174 | 2,348 |

# Empirical Results

## Robustness Checks—Randomly Assigned SEZ Status



# Empirical Results

## Robustness Checks

| | I. Alternative distance ranges | | II. 0-1000 meters vs. 1000-2000 meters | |
|---|---|---|---|---|
| | Within 2000 meters | Within 500 meters | Outside the zone | Inside the zone |
| | (1) | (2) | (3) | (4) |
| Panel (a). Dependent variable: (log) employment | | | | |
| *Dummy* | 0.494*** | 0.412*** | −0.023 | −0.080 |
| | (0.036) | (0.044) | (0.038) | (0.056) |
| Observations | 2,428 | 2,256 | 2,188 | 1,384 |
| Panel (b). Dependent variable: (log) output | | | | |
| *Dummy* | 0.567*** | 0.447*** | −0.073 | −0.082 |
| | (0.051) | (0.062) | (0.049) | (0.074) |
| Observations | 2,428 | 2,256 | 2,188 | 1,384 |
| Panel (c). Dependent variable: (log) capital | | | | |
| *Dummy* | 0.584*** | 0.423*** | −0.025 | −0.159** |
| | (0.050) | (0.060) | (0.051) | (0.079) |
| Observations | 2,428 | 2,256 | 2,188 | 1,384 |
| Panel (d). Dependent variable: (log) number of firms | | | | |
| *Dummy* | 0.267*** | 0.203*** | −0.010 | −0.060 |
| | (0.028) | (0.035) | (0.028) | (0.044) |
| Observations | 2,428 | 2,256 | 2,188 | 1,384 |

# Empirical Results

## Mechanism

- Decomposition the SEZ effects into
  - the effects due to new entry and exit, or extensive margin effects from entry and exit
  - the effects due to firm relocation, or extensive margin effects from relocation
  - the effects due to incumbents, or intensive margin effects
- Our BD-DD estimator $\hat{\gamma}_{BD-DD}$ is

$$\begin{aligned}
\hat{\gamma}_{BD-DD} &= \frac{\partial \ln Y}{\partial SEZ} = \frac{\partial \ln \left[ Y^{entry/exit} + Y^{inc} + Y^{re} \right]}{\partial SEZ} \\
&= \frac{Y^{entry/exit}}{Y} \frac{\partial \ln Y^{entry/exit}}{\partial SEZ} + \frac{Y^{inc}}{Y} \frac{\partial \ln Y^{inc}}{\partial SEZ} + \frac{Y^{re}}{Y} \frac{\partial \ln Y^{re}}{\partial SEZ} \\
&= \omega^{entry/exit} \hat{\gamma}^{entry/exit} + \omega^{inc} \hat{\gamma}^{inc} + \omega^{re} \hat{\gamma}^{re}
\end{aligned}$$

# Empirical Results

## Mechanism

| Dependent variable: | (log) employment | (log) output | (log) capital | (log) number of firms |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Entrants and Exiters | 0.281 | 0.283 | 0.305 | 0.204 |
| Incumbents | 0.094 | 0.134 | 0.114 | 0.000 |
| Relocaters | 0.078 | 0.114 | 0.111 | 0.028 |

# Empirical Results

## Heterogenous Effects—Industrial Capital-Labor Ratios

| Dependent variable: | (log) employment | (log) output | (log) capital | (log) number of firms |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Panel (a). Capital-Intensive Industries | | | | |
| *InsideZone* | 0.524*** | 0.667*** | 0.638*** | 0.291*** |
| | (0.051) | (0.073) | (0.071) | (0.033) |
| Observations | 2,124 | 2,124 | 2,124 | 2,124 |
| Number of industries | 242 | 242 | 242 | 242 |
| Panel (b). Labor-Intensive Industries | | | | |
| *InsideZone* | 0.424*** | 0.458*** | 0.483*** | 0.216*** |
| | (0.041) | (0.053) | (0.051) | (0.032) |
| Observations | 2,260 | 2,260 | 2,260 | 2,260 |
| Number of industries | 243 | 243 | 243 | 243 |
| Panel (c). Difference | | | | |
| *InsideZone* | 0.100 | 0.209** | 0.155* | 0.075 |
| | (0.065) | (0.090) | (0.088) | (0.046) |

# Empirical Results

## Heterogenous Effect

- The market potential $MP_z$ of zone $z$ is defined as

$$MP_z = \frac{\sum_{c \in PROV} GDP_c / dist_{zc}}{\sum_{c \in PROV} GDP_c},$$

  - $PROV$ denotes province, $c$ denotes prefecture-level city, $GDP_c$ stands for city $c$'s GDP, and $dist_{zc}$ is the distance between the zone's administrative committee $z$ and city $c$

- Access to Transportation Infrastructure:
  - We compute the distance of each zone's administrative committee to its nearest airport and rank from largest to smallest the distance ($rank\_airport$). Similarly, we construct ($rank\_highway$).
  - The zone's infrastructure accessibility is constructed as $rank = (rank\_airport + rank\_highway)/2$, with a low index value indicating the zone is further away from transportation infrastructure.

# Empirical Results

## Heterogenous Effects—Market Potential

| Dependent variable: | (log) employment | (log) output | (log) capital | (log) number of firms |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Panel (a). Zones with High Market Potential | | | | |
| *InsideZone* | 0.490*** | 0.599*** | 0.577*** | 0.230*** |
| | (0.053) | (0.074) | (0.073) | (0.042) |
| Observations | 1,192 | 1,192 | 1,192 | 1,192 |
| Number of zones | 309 | 309 | 309 | 309 |
| Panel (b). Zones with Low Market Potential | | | | |
| *InsideZone* | 0.451*** | 0.506*** | 0.517*** | 0.237*** |
| | (0.059) | (0.084) | (0.079) | (0.046) |
| Observations | 1,156 | 1,156 | 1,156 | 1,156 |
| Number of zones | 310 | 310 | 310 | 310 |
| Panel (c). Difference | | | | |
| *InsideZone* | 0.039 | 0.093 | 0.059 | −0.008 |
| | (0.080) | (0.112) | (0.107) | (0.062) |

# Empirical Results

## Heterogenous Effects—Access to Transportation Infrastructure

| Dependent variable: | (log) employment | (log) output | (log) capital | (log) number of firms |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | Panel (a). Spatially-integrated Zones | | | |
| *InsideZone* | 0.490*** | 0.545*** | 0.517*** | 0.252*** |
| | (0.050) | (0.069) | (0.068) | (0.040) |
| Observations | 1,196 | 1,196 | 1,196 | 1,196 |
| Number of zones | 311 | 311 | 311 | 311 |
| | Panel (b). Spatially-isolated Zones | | | |
| *InsideZone* | 0.451*** | 0.561*** | 0.579*** | 0.214*** |
| | (0.062) | (0.089) | (0.084) | (0.047) |
| Observations | 1,152 | 1,152 | 1,152 | 1,152 |
| Number of zones | 308 | 308 | 308 | 308 |
| | Panel (c). Difference | | | |
| *InsideZone* | 0.039 | −0.015 | −0.063 | 0.038 |
| | (0.080) | (0.113) | (0.108) | (0.062) |

# Conclusion

- This paper investigates the place-based development strategy in developing country
- We identify the effect of a prominent progrom (i.e., SEZ) in China from DD, BD and BD-DD estimations
- We find that after two years of establishment, SEZs have increased
  - employment by 47.1%
  - output by 55.3%
  - capital by 54.7%
  - the number of firms by 23.3%

# Conclusion

- By decomposing the firms into three sub-samples, we find that
  - a sizable effect on major outcomes such as employment, capital, and output associated with firm births and deaths.
  - incumbent firms in the zone show a significant improvement in performance.
  - relocation also plays a role in the total SEZ effects.
- The SEZ effects are found to be heterogeneous
  - large effects in capital-intensive industries
  - no significantly larger effects for zones with higher market potential and infrastructure accessibility