"Ghost cities" versus boom towns: Do China's high-speed rail new towns thrive?

Lei Dong[a,b] Rui Du[c] Matthew Kahn[d,g] Carlo Ratti[b,e] Siqi Zheng[a,e,f]

[a]MIT Sustainable Urbanization Lab, Massachusetts Institute of Technology, USA

[b]MIT Senseable City Lab, Massachusetts Institute of Technology, USA

[c]Department of Economics, Oklahoma State University, USA

[d]National Bureau of Economic Research, USA

[e]Department of Urban Studies and Planning, Massachusetts Institute of Technology, USA

[f]Center for Real Estate, Massachusetts Institute of Technology, USA

[g]Department of Economics, University of South Califonia, USA

By Yi Wang, June 24, 2022

Journals and Authors

- **Regional Science and Urban Economics**, 89, 103682.
- Received 15 April 2020; Received in revised form 11 April 2021; Accepted 15 April 2021.
- https://doi.org/10.1016/j.regsciurbeco.2021.103682

- **Authors:**

Lei Dong

Contact Information

Email: arch.dongl(at)gmail.com | Homepage: donglei.org | Weibo: @LEI_DONG

Research Interests

City, complexy systems, computational social science, data science

Academic Positions

- 2017-2021, Post-Doctoral Fellow, School of Earth and Space Sciences, Peking University
- 2018-2020, Visiting Scholar and Research Affiliate, Senseable City Lab and Sustainable Urbanization Lab, MIT

Industry Experiences

- 2021- , *Founder*, Maptable
  - A new database, mapping tool, and data marketplace created for everyone
- 2017-2021, *Co-founder*, QuantUrban
  - Data science for urban development

RUI DU

Department of Economics
Assistant Professor

347 BUSINESS BUILDING
STILLWATER, OK 74078-4011
Phone: 405-744-4487

rui.du@okstate.edu

Education

- Ph D, Clark University, Economics, 2017
- MA, Clark University, Economics, 2014
- BA, Central University of Finance and Economics, Mathematical Economics and Finance, 2011

Publications

- Yingcheng Li and Rui Du. (2022). "Polycentric Urban Structure and Innovation: Evidence from a Panel of Chinese Cities". *Regional Studies*. (56), 1, 113-217.
- Rui Du and Junfu Zhang. (2021). "Super Bowl Participation and the Local Economy: Evidence from the Stock Market". *Growth and Change*.
- Lei Dong, Rui Du, Matthew E. Kahn, Carlo Ratti, and Siqi Zheng. (2021). "'Ghost Cities' versus Boom Towns: Do China's Hig speed Rail New Towns Thrive?". *Regional Science and Urban Economics*. (89), 103682.
- Siqi Zheng and Rui Du. (2020). "How Does Urban Agglomeration Integration Promote Entrepreneurship in China? Evidence from Regional Human Capital Spillovers and Market Integration". Amsterdam: *Cities*. (97), 102529.
- Rui Du and Siqi Zheng. (2020). "Agglomeration, Housing Affordability, and New Firm Formation: The Role of Subway Network". Amsterdam: *Journal of Housing Economics*. (48), 102529.
- Rui Du and Junfu Zhang. (2019). "Walled Cities and Urban Density in China". Hoboken, NJ: *Papers in Regional Science*. (98), 3, 1517–1539.

ADDRESS

Department of Economics
University of Southern California
3620 South Vermont Ave.
Kaprielian (KAP) Hall, 300
Los Angeles, CA 90089-0253

COMMUNITY + ECONOMIC DEVELOPMENT

Profile

Publications
Subjects

contact 9-323
(617) 253-3988
sqzheng@mit.edu

links Personal Website
Google Scholar
Research Projects

CITY DESIGN + DEVELOPMENT

Profile

Publications
Subjects

contact 9-216
(617) 253-7926
ratti@mit.edu

Overview

Introduction

1. Introduction

- Given **the unique land public finance scheme** in China, city leaders have strong incentives to **build "new towns" near HSR stations**.
- Building a new town around a new HSR station is regarded as an effective industrial policy to **attract new industries and population to boost the local economy**.
- While some HSR new towns have **enjoyed economic growth**, others have remained vacant for many years and **become "ghost cities"**.
- This study explores the determinants of **the new towns' economic vibrancy heterogeneity**.
- Using satellite imagery and online archives of government documents, we identify **180 HSR new towns**. We use several datasets to measure local economic growth **at a fine spatial scale**.

1. Introduction

- **The endogeneity issue**: transportation infrastructure investments are not randomly assigned to places.
- A shrewd decision-maker will consider the **benefits** and **costs** of creating a new transit hub in one location versus others.
- **Two identification strategies**: instrumental variable strategy and difference-in-differences approach.
- Each strategy estimates the treatment effect with **counterfactual locations** to study how the creation of a new HSR station stimulates local economic growth.

Background and data

2. Background and data

- **We conduct our analysis at the new-town level for three reasons:**
  i) Many prefectural-level cities have **multiple HSR stations**.Data at the prefectural level cannot fully reflect the very local economic impact.
  ii) Requires the analysis to be conducted **at a finer spatial resolution**.
  iii) Official economic and demographic statistics are almost exclusively aggregated at various administrative levels. **No official data source is available for new towns**.

2.1. High-speed rail and new towns in China

- "HSR new town": new towns around HSR stations
- Digitize HSR stations and lines based on **high-resolution transportation maps published between 2017 and 2018**.
- Collect the opening year and level (national, regional, or intercity level) of each line, and the opening time of each station from **Wikipedia and BaiduBaike**.
- Sample: 90 lines and 839 stations.

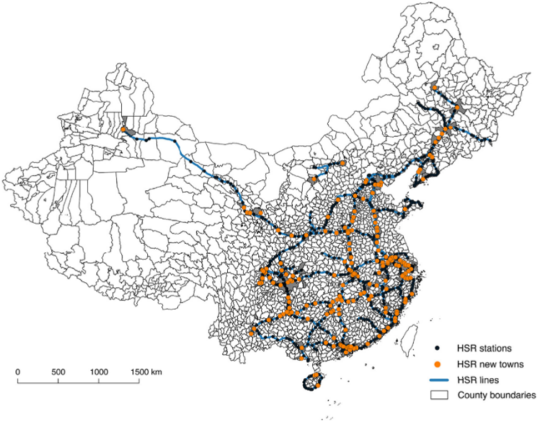


Fig. 1. Mainland China's HSR new towns, stations, and lines.

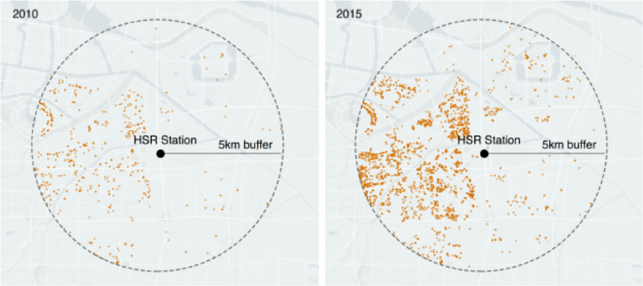
2.1. High-speed rail and new towns in China

- **Three-stage decision-making process of building up an HSR new town:**
  i) The central government (the State Council of the PRC and the Ministry of Railway) decides **which city receives the HSR connection**.
  ii) The city leaders then decide **whether to convert** an existing railway station (often in the existing city center) to an HSR stop or build a new station.
  iii) The city considers **whether to build a new town** near the HSR stop.**(Focus)**
- Identified as an HSR new town based on three criteria **(180 HSR new towns)**:
  i) newly built
  ii) active road networks and residential and industrial land development **observed** around the station (Google Earth satellite imagery)
  iii) **further** confirm new town development (collecting online archives of the government documents)

2.2. Firm data

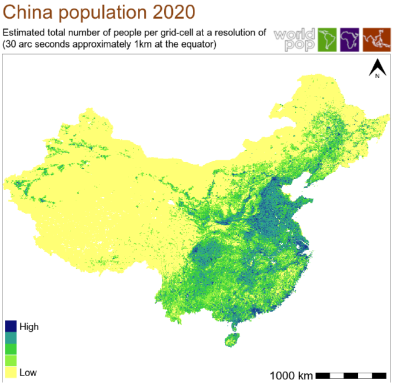
- From: **the National Enterprise Credit Information Publicity System**
- Include: firm's name, founding year, address, and capitalization
- Geocode firm addresses into longitudes and latitudes using **AMap API** and compute the total counts of the geo-located firms in **two buffer areas: 3 km and 5 km-radius circles around each HSR station**.

Figure 1: Spatial distribution of firms (orange points) around the Zhengdong HSR new town (Zhengzhou city, Henan province).



2.3. Population data

- From: **WorldPop** (https://www.worldpop.org).
- Include: 'unconstrained individual countries 2000–2020 (1 km resolution)' version,
- Aggregate the 1 km grid-cell-based population into **two different buffers: 3 km- and 5 km-radius circles around HSR stations**.



2.4. County-level data

- From: **County Economic Statistical Yearbooks from 2000 to 2018**.
- Include: such as GDP and population, **distance between an HSR station and the center of its host city(*)**
- **host city**: the counties(county-level city or county) where the HSR stations are located
- Define the city center as the location of the city government. Using the Haversine formula calculates the great-circle distance.

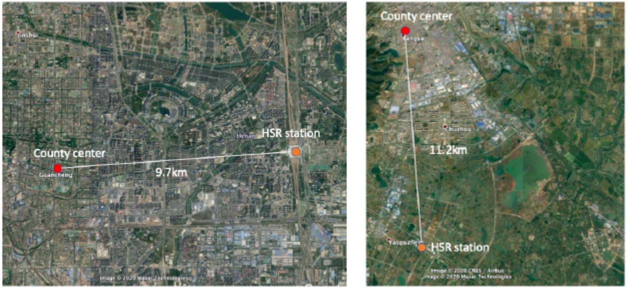


Figure A.1: The satellite imagery of two typical new towns: Zhengdong station of Zhengzhou city, Henan province (left); Chuzhou station of Chuzhou city, Anhui province (right). (Image Copyright: Google Earth)

2.5. Summary statistics

Table 1
Descriptive statistics.

	Obs.	Mean	Std. Dev.	Min.	Max.
New town establishments (5 km)	3667	489.90	1439.10	0	31,587
New town establishments (3 km)	3667	172.40	587.06	0	13,374
New town pop. (5 km, 10^4 people)	3667	15.22	14.10	0.61	92.45
New town pop. (3 km, 10^4 people)	3667	5.10	4.94	0.13	31.28
Dist. to city center (km)	3667	13.57	12.96	2.17	71.46
Least-cost distance (km)	3344	20.43	13.77	1.42	73.34
Gov't revenue (10^8 RMB)	3667	151.8	455.46	0.17	8630
City pop. (10^4 people)	3667	202.41	299.69	8.04	2483
Agri. employment share (%)	3667	4.02	5.74	0.01	38.12
City status	3667	0.64	0.48	0	1

Note: A city has a city status if there is a character "district" (*qu*) or "city" (*shi*) in its Chinese name.

The empirical strategy

3.1. Conceptual framework

- Assume that local government officials consider **maximize the net economic gains** by **selecting the location to build an HSR stop and develop a new town**.
  - **Benefits** depend on: agglomeration spillovers from the existing city center and the city's market access.
  - **Costs** depend on: reduce upfront construction and demolition costs as well as travel costs between the new town and nearby subcenters.
- **Empirically test three sets of key research questions**:
  i) What is the local impact of new town development?(**treatment effect**) What factors lead to the successes or failures of HSR new towns (treatment effect heterogeneity)? (expect **market access** and **proximity to the existing city center** are conducive to the new town economic boom)
  ii) Why do some cities build new towns in locations with high development costs?
  iii) Do some "ghost cities" eventually prosper?

3.2. IV estimation

- 3.2.1. Empirical model
  - **distance IV based on a least-engineering-cost path(LCP-based distance IV)**: address the concern about non-random HSR route placement
  - **characteristics of the ancient city wall**: address the concern about endogenous market access
- First stage regression:

$$D_{icsp} = \alpha_1 + \beta_{11} d_{icsp} + \beta_{12} m_{csp,t} + \boldsymbol{\gamma_1 X_{csp,t}} + \delta_{1p} + \lambda_{1t} + \epsilon_{1icsp,t} \quad (1)$$

$$M_{csp,t} = \alpha'_1 + \beta'_{11} d_{icsp} + \beta'_{12} m_{csp,t} + \boldsymbol{\gamma'_1 X_{csp,t}} + \delta'_{1p} + \lambda'_{1t} + \epsilon'_{1icsp,t} \quad (2)$$

  - station i, city center c, prefectural-level city s, province p, year t.
  - D_{icsp} actual HSR station location, $M_{csp,t}$ market access, $X_{csp,t}$ local characteristics, δ_{1p} province fixed effects, λ_{1t} year fixed effects, $\epsilon_{1icsp,t}$ error terms.
  - $\beta_{11} > 0, \beta_{12} > 0$

3.2. IV estimation

- 3.2.1. Empirical model
  - Key outcome equation:

$$y_{icsp,t} = \alpha_2 + \beta_{21}\widehat{D}_{icsp} + \beta_{22}\widehat{M}_{csp,t} + \gamma_2 X_{csp,t} + \delta_{2p+}\lambda_{2t} + \epsilon_{2icsp,t} \quad (3)$$

    - $y_{icsp,t}$ economic outcome.
    - $\beta_{21} < 0$, $\beta_{22} > 0$
  - long-difference equation:

$$y_{icsp,t1} - y_{icsp,t0} = \alpha_3 + \beta_{31}\widehat{D}_{icsp} + \beta_{32}\widehat{M}_{csp,t0} + \gamma_2 X_{csp,t0} + \delta_{3p} + \epsilon_{3icsp,t0} \quad (4)$$

    - $y_{icsp,t1} - y_{icsp,t0}$ change in the new town economic outcome from base year t0(2000) to a future year period t1(2018).
    - $\beta_{31} < 0$, $\beta_{32} > 0$

3.2. IV estimation

- 3.2.2. The LCP-based distance IV
- Our goal is to **construct a counterfactual location for Station B** so we can **calculate an exogenous distance** that is a strong predictor of the actual distance **between Station B and the center of City B**.

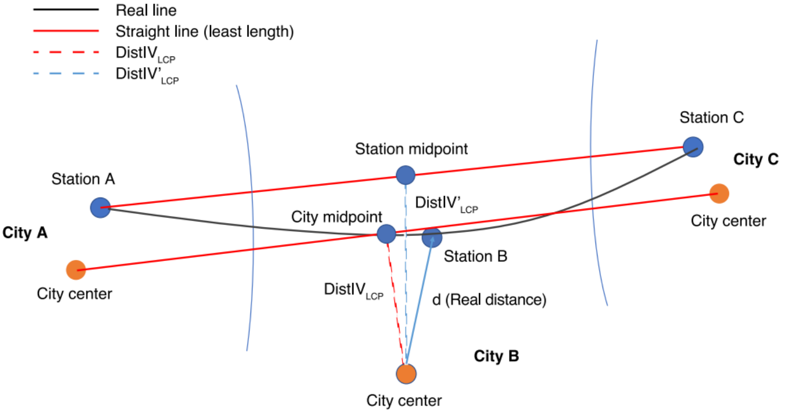


Fig. 2. An illustration of the construction of the distance IV.

3.2. IV estimation

- 3.2.3. Market access
- **Measure the market access** of a new town using the **inverse-distance weighted** sum of the **urban markets around the host city** of the HSR station.

$$MA_{is,t} = \sum_{j=1}^{N} Income_{js,t} \cdot e^{-\gamma d_{ijt}} \quad (5)$$

城市i和j之间车程

城市i的市场准入值

位于同一地级市s的城市j的市场规模，用市政府收入作为该市地方经济规模的代表

- **construct the wall-based market access IV**:

$$M_{csp,t} = \alpha + \beta \cdot \mathbf{wall}_c \times GDP_t + u_{csp,t}$$

- $M_{csp,t}$ market access, wall $_c$ a vector of ancient wall characteristics, GDP_t the national GDP level

3.3. Treatment effect estimation using a DID approach

- baseline DID regression equation:

$$y_{icspt} = \alpha_4 + \beta_{40} \cdot treated_{icspt} + \boldsymbol{\gamma_4 X_{cspt}} + \delta_{4i} + \lambda_{4t} + \mu_{pt} + \epsilon_{4icspt} \tag{6}$$

- Treated locations become increasingly vibrant for reasons other than the new town development(ex-ante main city fundamentals, relative location, and travel cost): **interaction**
- long-difference specification:

$$y_{icspt_1} - y_{icspt_0} = \alpha_5 + \beta_{50} \cdot treated_{icspt_0} + \boldsymbol{\gamma_5 X_{cspt_0}} + \theta_{5s} + \epsilon_{5icspt_0} \tag{7}$$

- Parallel trend assumption:the actual new town and the counterfactual new town location tend to **have similar trends in local economic activity before the HSR station opening**.

Results

4. Results

4.1. IV estimation results

Table 2

First-stage results: Determinants of new town economic growth.

	Short-run		Medium-run	
	Ln(distance) (1)	Ln(market access) (2)	Ln(distance) (3)	Ln(market access) (4)
Dist_IV$_{LCP}$	$\boxed{0.509^{***}}$	−0.009	$\boxed{0.504^{***}}$	−0.072
	$\boxed{(0.089)}$	(0.086)	$\boxed{(0.088)}$	(0.134)
Wall_IV	0.231***	$\boxed{0.575^{***}}$	0.259**	$\boxed{0.792^{***}}$
	(0.085)	$\boxed{(0.080)}$	(0.101)	$\boxed{(0.117)}$
Ln(city population)	0.069	0.808***	0.075	0.748***
	(0.053)	(0.101)	(0.063)	(0.152)
Agr. share	−0.007	−0.066***	0.002	−0.035*
	(0.006)	(0.009)	(0.011)	(0.019)
City status	0.033	0.089	0.014	0.086
	(0.080)	(0.139)	(0.115)	(0.263)
Underidentification test				
Kleibergen-Paap rk *LM* statistic	18.210	18.210	13.330	13.330
Weak identification test				
F statistic	19.706	19.706	14.699	14.699
10% maximal IV size	7.03	7.03	7.030	7.030
Province FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
City-level controls	Y	Y	Y	Y
4th order polynomial in geography	Y	Y	Y	Y
Obs.	3344	3344	176	176

4. Results

4.1. IV estimation results

Table 3
The short-run determinants of new town economic growth.

	OLS		IV	
	Ln(firm)	Ln(pop)	Ln(firm)	Ln(pop)
	(1)	(2)	(3)	(4)
Ln(distance)	−0.589***	−0.479***	−0.733***	−0.449***
	(0.123)	(0.099)	(0.182)	(0.172)
Ln(market access)	0.326***	0.109**	0.604***	0.135
	(0.067)	(0.044)	(0.168)	(0.122)
Ln(city population)	0.163	0.220**	−0.096	0.205
	(0.112)	(0.099)	(0.164)	(0.137)
Agr. share	−0.006	−0.002	0.012	−0.001
	(0.009)	(0.005)	(0.015)	(0.010)
City status	0.569***	0.207*	0.561***	0.222**
	(0.132)	(0.105)	(0.140)	(0.105)
Underidentification test				
Kleibergen-Paap rk *LM* statistic			18.210	18.210
Weak identification test				
Kleibergen-Paap Wald rk *F* statistic			19.706	19.706
10% maximal IV size			7.03	7.03
Province FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
City-level controls	Y	Y	Y	Y
4th order polynomial in geography	Y	Y	Y	Y
Obs.	3667	3667	3344	3344
Adj. R^2	0.714	0.533	0.716	0.519

4. Results

▶ 4.1. IV estimation results

Table 4

The medium-run determinants of new town economic growth.

	OLS		IV	
	△Ln(firm) (1)	△Ln(pop) (2)	△Ln(firm) (3)	△Ln(pop) (4)
Ln(distance)	−0.504***	−0.462***	−0.575***	−0.429**
	(0.141)	(0.112)	(0.218)	(0.171)
Ln(market access)	0.338***	0.109*	0.527***	0.126
	(0.085)	(0.057)	(0.122)	(0.099)
Ln(city population)	0.085	0.167	−0.078	0.162
	(0.135)	(0.130)	(0.142)	(0.137)
Agr. share	−0.009	−0.009	0.003	−0.008
	(0.013)	(0.010)	(0.014)	(0.011)
City status	0.603***	0.265*	0.674***	0.273**
	(0.208)	(0.155)	(0.189)	(0.136)
Underidentification test				
Kleibergen-Paap rk *LM* statistic			13.330	13.330
Weak identification test				
Kleibergen-Paap Wald rk *F* statistic			14.699	14.699
10% maximal IV size			7.030	7.030
Province FE	Y	Y	Y	Y
City-level controls	Y	Y	Y	Y
4th order polynomial in geography	Y	Y	Y	Y
Obs.	193	193	176	176
Adj. R^2	0.562	0.532	0.550	0.519

4. Results

- 4.2. DID estimation results

表5：高铁新城开发的处理效果

Table 5
The treatment effect of HSR new town development.

	Difference-in-differences		Long-difference	
	Ln(firm)	Ln(pop)	△Ln(firm)	△Ln(pop)
	(1)	(2)	(3)	(4)
Treated	$\boxed{0.113^{***}}$	−0.000	$\boxed{1.229^{***}}$	$\boxed{0.645^{***}}$
	(0.029)	(0.005)	(0.129)	(0.112)
Location FE	Y	Y	N	N
Prefectural-level city FE	N	N	Y	Y
Year FE	Y	Y	–	–
Province-year FE	Y	Y	–	–
Controls	Y	Y	Y	Y
Obs.	5206	5206	361	361
Adj. R^2	0.972	0.999	0.759	0.577

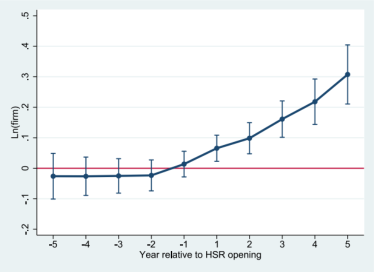
The results reported in this table are based on estimating Eqs. (6) and (7) in the text. Standard errors in parentheses are clustered at the location level. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

$e^{0.113}-1=12.0\%$
高铁新城促成了12%的企业发展

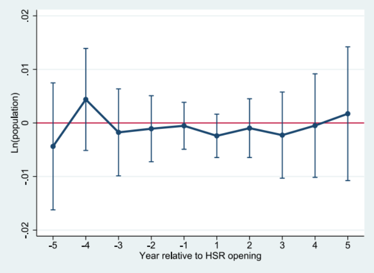
短期内可能存在“鬼城”

4. Results

- 4.2. DID estimation results



(a) Parallel trend test of the new town firm formation.



(b) Parallel trend test of the new town population.

4. Results

- 4.3. Explaining the ghost town phenomenon
- Ex-ante main city fundamentals

Table 6
Heterogeneous treatment effects by main city fundamentals.

	Ln(firm) (1)	Ln(pop) (2)	Ln(firm) (3)	Ln(pop) (4)	Ln(firm) (5)	Ln(pop) (6)
Treated	−1.115*	−1.096**	−0.915**	−0.889***	0.664***	−0.076
	(0.577)	(0.468)	(0.410)	(0.288)	(0.219)	(0.233)
New town (=1)	0.993***	0.573***	0.992***	0.572***	1.009***	0.578***
	(0.122)	(0.115)	(0.122)	(0.115)	(0.122)	(0.115)
Treated × Ln(market access)	0.103**	0.092**				
	(0.049)	(0.040)				
Treated × Ln(city population)			0.217**	0.187***		
			(0.087)	(0.061)		
Treated × Ln(dist. to prefectural center)					−0.194**	0.025
					(0.076)	(0.077)
Prefectural-level city FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Province-year FE	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y
Obs.	5206	5206	5206	5206	5206	5206
Adj. R^2	0.827	0.670	0.828	0.671	0.827	0.669

The results reported in this table are based on estimating Eq. (6) in the text. Standard errors in parentheses are clustered at the location level. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

4. Results

- 4.3. Explaining the ghost town phenomenon
- Ex-ante main city fundamentals

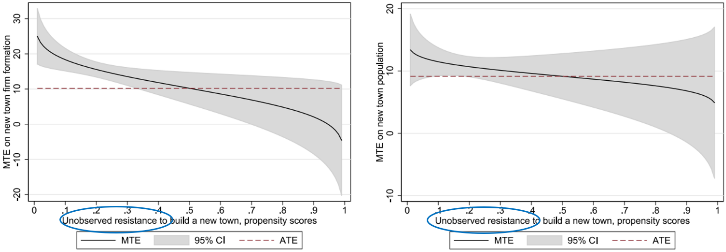
Table 7
Urban shadow and relative location size.

	DV: Ln(firm)		DV: Ln(pop)	
	Small neighbor (1)	Large neighbor (2)	Small neighbor (3)	Large neighbor (4)
Treated	0.071	0.193*	−0.261**	0.148*
	(0.098)	(0.102)	(0.110)	(0.087)
New town (=1)	1.129***	0.842***	0.751***	0.420***
	(0.143)	(0.186)	(0.149)	(0.146)
Treated × Q_1*_Dist to county center*	0.005	−1.196***	0.223*	−0.300
	(0.174)	(0.393)	(0.122)	(0.556)
Prefectural-level city FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Province-year FE	Y	Y	Y	Y
Controls	Y	Y	Y	Y
Obs.	3696	1440	3696	1440
Adj. R^2	0.823	0.744	0.704	0.610

The results reported in this table are based on estimating the variants of Eq. (6) in the text. Standard errors in parentheses are clustered at the location level. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

4. Results

- 4.3. Explaining the ghost town phenomenon
- Essential heterogeneity: estimate marginal treatment effects (MTE) via the method of local IVs.



(a) Parallel trend test of the new town firm formation.

(b) Parallel trend test of the new town population.

4. Results

- 4.3. Explaining the ghost town phenomenon
- additional evidence

Table 8
Testing for essential heterogeneity in treatment effects.

	Ln(firm) (1)	Ln(pop) (2)	Ln(firm) (3)	Ln(pop) (4)
Treated	−0.752**	−0.557**	−0.729**	−0.611***
	(0.320)	(0.272)	(0.323)	(0.233)
New town (=1)	1.642***	1.011***	1.642***	1.011***
	(0.112)	(0.104)	(0.112)	(0.104)
Treated × $Dist\_IV_{LCP}$	$\boxed{0.312^{***}}$	$\boxed{0.198^{**}}$		
	$\boxed{(0.117)}$	$\boxed{(0.097)}$		
Treated × $Dist\_IV'_{LCP}$			$\boxed{0.310^{***}}$	$\boxed{0.222^{**}}$
			$\boxed{(0.119)}$	$\boxed{(0.087)}$
Prefectural-level city FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Province-year FE	Y	Y	Y	Y
Controls	Y	Y	Y	Y
Obs.	5076	5076	5076	5076
Adj. R^2	0.756	0.585	0.756	0.585

The results reported in this table are based on estimating the variants of Eq. (6) in the text. Standard errors in parentheses are clustered at the location level. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

4. Results

- 4.4. Do some "ghost cities" eventually prosper?
- convert panel data to survival data
- focus on firm formation in our survival analysis

Table 9
Cox and parametric survival models for the treatment effect.

DV: Hazards ratio	Cox		Parametric	
	(1)	(2)	(3)	(4)
ATE (treated 1 vs. 0)	2.525***	2.057***	2.764***	2.207***
	(0.072)	(0.062)	(0.085)	(0.073)
Controls	N	Y	N	Y
Obs.	7011	7011	7011	7011

Exponentiated coefficients are reported in this table. Robust standard errors are presented in parentheses. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

与不开发新城相比，建设新城会增加经济繁荣的风险，约为2.1-2.8倍

Conclusion and Discussion

5. Conclusion and Discussion

- Study how the creation of a new HSR station stimulates local economic growth.
- Find that the location and local market access are key determinants of the success of new towns.
- "Ghost cities" are more likely to emerge if the new stations are located too far from the existing city center or the city itself has weak market access.