Place-Based Policies, Creation and Agglomeration Economies: Evidence from China's Economic Zone Program

Yi Lu, Jin Wang and Lianming Zhu

Online Appendix

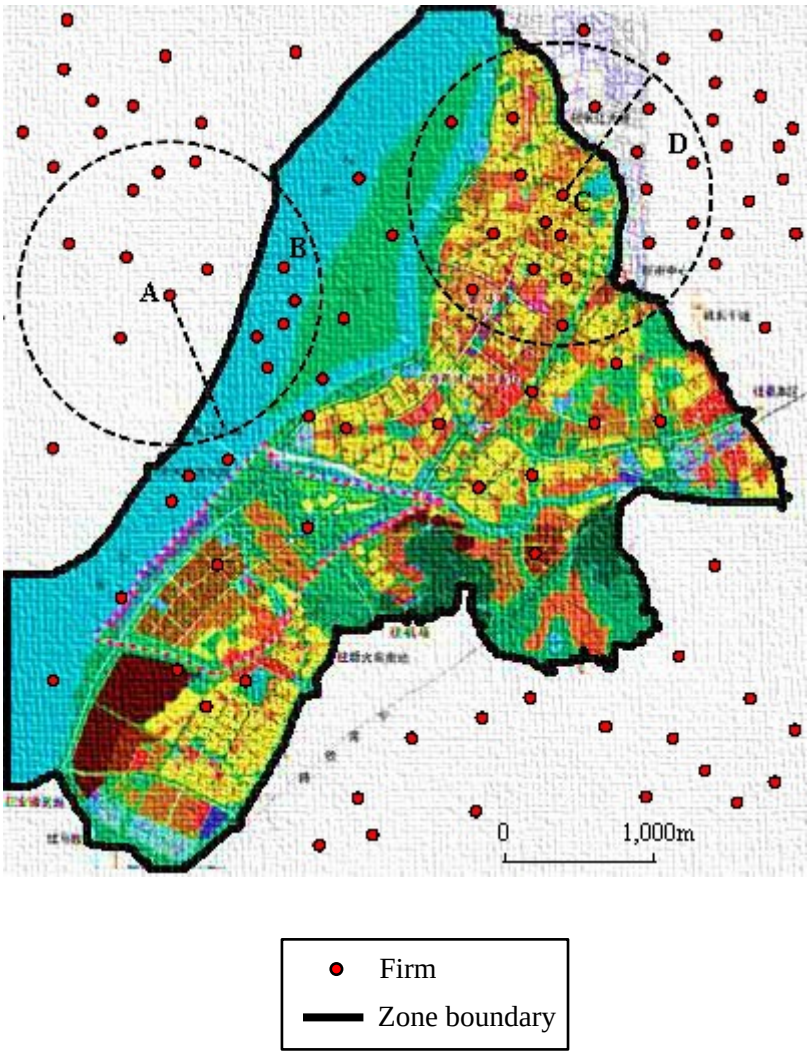
Figure A1 Searching for a Detailed Address with Google Maps

Notes: To obtain coordinates for firms that report detailed Chinese addresses, the address (for example, "157 Nandan Road, Xuhui District, Shanghai, China") is searched in Google Maps to get a red marker showing the specific location of the address. Once that location is confirmed, the latitude and longitude of the address are read from the map.

Figure A2 Searching for Villages and Communities with Google Maps

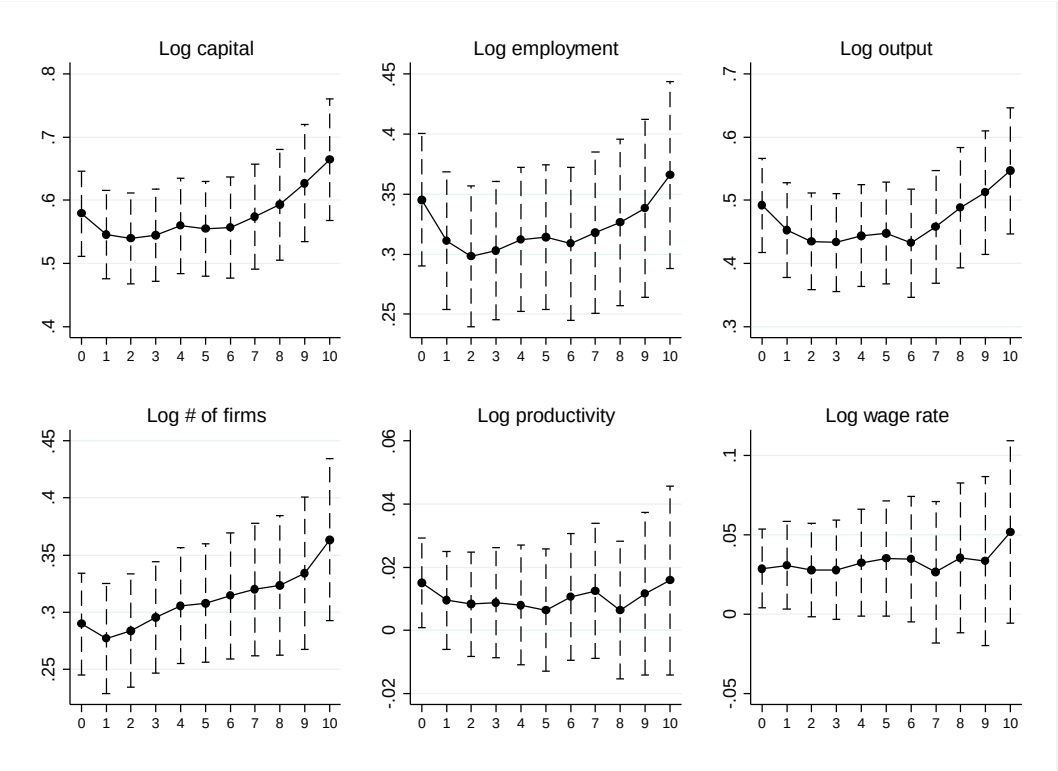
Notes: To obtain the coordinates for villages and communities, the name of the village or community is entered followed by the name of the town, city, and province to which the village or community belongs (for example, “Liunan Village, Liuhe Town, Taicang City, Suzhou, Jiangsu Province, China”). The specific location is then denoted by a red marker. Once that location has been confirmed, the latitude and longitude of the village or community are read from the map.

Figure A3 Firms Near a Zone Boundary



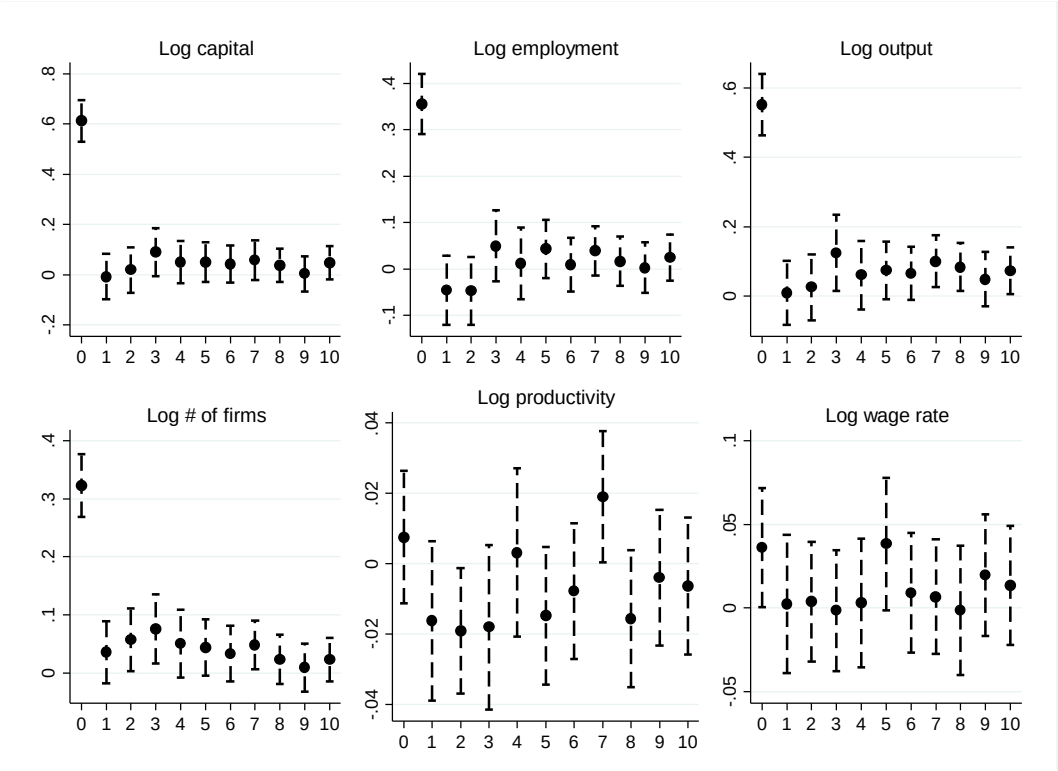
Notes: If a firm (firm A) is located outside the zone and within 1,000 meters, and there is another firm (firm B) located inside the zone, firm A is designated as located within 1,000 meters of the zone boundary. If a firm (firm C) is located inside the zone and within 1,000 meters, and there is another firm (firm D) located outside the zone, firm C is designated as located within 1,000 meters of the zone boundary.

Figure A4 Ring Analysis



Notes: In this analysis, we step-wisely exclude from the control group, the non-SEZ villages within the 2km of the SEZ villages, those within the 4km of the SEZ villages, continuing to the exclusion of those within the 20km of SEZ villages. The benchmark estimate is plotted in the horizontal line labeled as 0 without exclusion of any control SEZ villages. The estimates labeled from 1, 2, 3,…, 10 indicate exclusion of SEZ villages within 2, 4, 6,…, 20 kilometer distance of the SEZ villages.

Figure A5 Ring Analysis: An Alternative Model Specification



Notes: In this analysis, we provide an alternative spillover estimates using equation (4). The treatment effect on the SEZ village is plotted in the horizontal line labeled as 0. The estimates labeled from 1, 2, 3,…, 10 indicate the spillover externality effect on the nearby 1st, 2nd, 3rd,…10th ring between 0 and 2, 2 and 4, 4 and 6, …, 18 and 20 kilometers from its nearest SEZ village.

Table A1 Comparison between the Census and ASIF Data

	Economic Census				Annual Survey of Industrial Firms			
	Obs.	Mean	p10	p90	Obs.	Mean	p10	p90
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Year 2004								
Capital	1,321,026	14091	120	13800	256,999	62,093	2,731	89,541
Employment	1,321,026	64	4	118	256,999	221	25	426
Output	1,321,026	14692	21	14607	256,999	68,451	5,424	95,593
Panel B. Year 2008								
Capital	1,788,227	20558	303	20265	382,842	80,805	3,575	107,615
Employment	1,822,419	58	4	100	382,838	194	25	350
Output	1,738,045	27578	360	30765	383,779	111,150	6,340	160,229

Note: p10 and p90 denote the 10th and 90th percentiles. Sources: Economic Census and Anual Survey of Industrial Firms for 2004 and 2008.

Table A2 Village-Level Analysis (ASIF Data from 2004-2008)

	(1)	(2)	(3)	(4)
Dependent variable	Log capital	Log employment	Log output	Log # of firms
SEZ*post2006	0.294	0.223	0.261	0.195
	(0.035)	(0.031)	(0.038)	(0.021)
Covariates*year dummies	Yes	Yes	Yes	Yes
Village FEs	Yes	Yes	Yes	Yes
County-year FEs	Yes	Yes	Yes	Yes
Number of clusters	406	406	406	406
Observations	56,772	56,772	56,772	56,772

Note: All observations are at the village-year level. ASIF data from 2004 to 2008 are used. Covariates include village-level characteristics listed in Panel A, Table 2. The standard errors are clustered at the county level. All regressions control for village fixed effects and county-year fixed effects.

Table A3 Testing for Pretrends: Village-Level Analysis

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	Log capital	Log employment	Log output	Log # of firms	Log productivity	Log wage rate
SEZ*year2005	0.011	0.055	0.057	0.025	−0.001	0.016
	(0.027)	(0.027)	(0.033)	(0.017)	(0.007)	(0.014)
SEZ*year2006	0.170	0.150	0.184	0.107	0.015	0.023
	(0.037)	(0.035)	(0.041)	(0.023)	(0.009)	(0.017)
SEZ*year2007	0.296	0.239	0.294	0.196	0.014	0.039
	(0.045)	(0.039)	(0.048)	(0.027)	(0.010)	(0.018)
SEZ*year2008	0.489	0.415	0.442	0.367		0.054
	(0.065)	(0.058)	(0.067)	(0.037)		(0.021)
Covariates*year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Village FEs	Yes	Yes	Yes	Yes	Yes	Yes
County-year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Number of clusters	406	406	406	406	405	406
Observations	56,772	56,772	56,772	56,772	43,830	56,600

Note: In columns 1-4 and 6, ASIF data from 2004 to 2008 are used for the analysis. In column 5, ASIF data from 2004 to 2007 are used. Covariates include village-level characteristics listed in Panel A of Table 2. The standard errors are reported in parentheses, clustered by county. All regressions control for village fixed effects and county-year fixed effects.

Table A4 The SEZ Effects on Housing Prices

	(1)	(2)
Dependent variable	Log house prices	Log house prices
SEZ*post2006	0.015	0.024
	(0.026)	(0.026)
City FEs	Yes	Yes
Year FEs	Yes	Yes
Covariates*year dummies	No	Yes
Number of clusters	284	284
Observations	1,362	1,362

Note: All observations are at the city-year level. The dependent variable is the natural log of the housing prices. If a city has any county under its administration granted with SEZs, the treatment indicator equals one since 2006. Covariates include average county-level characteristics (as listed in Panel B of Table 2) for those under a city. The standard errors are reported in parentheses, clustered by city. All regressions control for city and year fixed effects.

Table A5 Heterogeneous Effects by Industrial Capital-Labor Ratio

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	Log capital	Log employment	Log output	Log # of firms	Log productivity	Log wage rate
Panel A. Capital-Intensive Industries						
SEZ*post2006	0.604	0.375	0.547	0.296	0.013	0.027
	(0.044)	(0.036)	(0.053)	(0.027)	(0.009)	(0.016)
Number of clusters	549	549	549	549	398	400
Observations	58,784	58,784	58,784	58,784	23,421	31,274
Panel B. Labor-Intensive Industries						
SEZ*post2006	0.498	0.308	0.438	0.272	0.003	0.025
	(0.035)	(0.027)	(0.035)	(0.022)	(0.009)	(0.016)
Number of clusters	560	560	560	560	392	395
Observations	98,540	98,540	98,540	98,540	30,615	39,696

Note: All observations are at the village-year level. In columns 1-4, Census data 2004 and 2008 are used for analysis. In column 5, ASIF data from 2004 to 2007 are used. In column 6, ASIF data from 2004 to 2008 are used. Capital and labor intensity are defined at the 4-digit level based on a capital-labor ratio above or below the median value in 2004. Standard errors are in parentheses. The standard errors are clustered at the county level.

Table A6 Heterogeneous Effects by Infrastructure

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	Log capital	Log employment	Log output	Log # of firms	Log productivity	Log wage rate
Panel A. SEZ Counties with Good Infrastructure						
SEZ*post2006	0.569	0.350	0.491	0.271	0.009	0.037
	(0.044)	(0.037)	(0.049)	(0.029)	(0.009)	(0.017)
Number of clusters	292	292	292	292	165	165
Observations	78,698	78,698	78,698	78,698	25,206	32,382
Panel B. SEZ Counties with Poor Infrastructure						
SEZ*post2006	0.614	0.347	0.502	0.339	0.025	0.016
	(0.053)	(0.041)	(0.056)	(0.034)	(0.012)	(0.019)
Number of clusters	288	288	288	288	240	241
Observations	41,200	41,200	41,200	41,200	18,624	24,218

Note: All observations are at the village-year level. In columns 1-4, Census data 2004 and 2008 are used for analysis. In column 5, ASIF data from 2004 to 2007 are used. In column 6, ASIF data from 2004 to 2008 are used. SEZ counties with good (poor) infrastructure index are those with infrastructure indices above (below) the median in 2004: a larger index indicates better infrastructure. Standard errors are in parentheses. The standard errors are clustered at the county level.

Table A7 Heterogeneous Effects by Firm Size

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable:	Log capital	Log employment	Log output	Log # of firms	Log productivity	Log wage rate
Panel A. Firms with Large Size						
SEZ*post2006	0.362	0.212	0.342	0.153	−0.006	0.011
	(0.030)	(0.024)	(0.035)	(0.014)	(0.008)	(0.017)
Number of clusters	560	560	560	560	385	394
Observations	71,022	71,022	71,022	71,022	21,300	27,953
Panel B. Firms with Small Size						
SEZ*post2006	0.297	0.157	0.313	0.103	−0.010	0.017
	(0.036)	(0.028)	(0.043)	(0.017)	(0.008)	(0.016)
Number of clusters	525	525	525	525	394	398
Observations	66,154	66,154	66,154	66,154	24,357	30,613

Note: All observations are at the village-year level. In columns 1-4, Census data 2004 and 2008 are used for analysis. In column 5, ASIF data from 2004 to 2007 are used. In column 6, ASIF data from 2004 to 2008 are used. Firms with large (small) size are those with sales above (below) the median in 2004. Standard errors are in parentheses. In panels A and B, the standard errors are clustered at the county level.