

Online Appendix to “**The Spatial Scope of Human  
Capital Externalities and Firms’ Location in Brazil**”

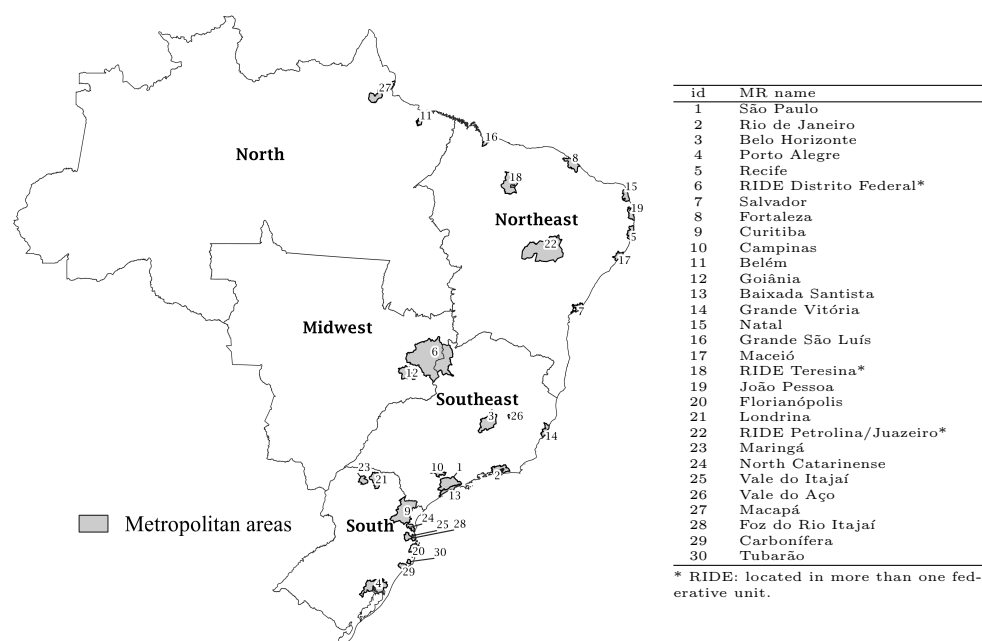
Edilberto Almeida, Raul Silveira Neto, Roberta Rocha, and Carlos  
Azzoni

## Appendix A Data - additional details

**Table A1.** Geocoded annual counts of new establishments and employment

Year	New Plants	New Employment	New Plants geocoded	New Empl. geocoded	Percent geo. of plants	Percent geo. of empl.
2007	11,637	84,890	11,578	84,789	99.49	99.88
2008	14,723	81,976	14,656	81,836	99.54	99.83
2009	15,553	77,947	15,478	77,392	99.52	99.29
2010	18,413	117,115	18,304	116,146	99.41	99.17
2011	16,870	99,932	16,796	99,760	99.56	99.83
2012	15,081	109,733	15,031	107,494	99.67	97.96
2013	15,955	94,346	15,894	93,573	99.62	99.18
2014	11,352	87,855	11,309	87,174	99.62	99.22
Total	119,584	753,794	119,046	748,164	99.55	99.25

Source: Prepared by the authors using information from RAIS.



**Figure A1.** Brazilian Metropolitan Areas in 2006

Source: Almeida et al. (2023a).

**Table A2.** Descriptive statistics

	2006		2012		All sample	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
<b># of new plants</b>						
New plants	0.838	2.131	0.709	1.801	0.870	2.1445
<b># of workers by ring</b>						
College-or-more, 0 to 1 km	32.969	130.723	43.820	233.960	38.859	185.554
College-or-more, 1 to 5 km	614.051	1,361.528	814.093	1,811.874	721.719	1615.707
College-or-more, 5 to 10 km	1,597.167	3,304.097	2,124.607	4,177.896	1,884.992	3,848.932
College-or-more, 10 to 20 km	4,672.737	8,243.357	6,229.426	10,281.9	5,520.539	9,532.884
All workers, 0 to 20 km	37,680.970	51,795.690	27,646.230	36,240.280	32,659.900	44,722.280
<b>Municipal-level controls</b>						
Population	1,576,742	2,936,272	1,508,957	2,917,027	1,544,217	2,923,261
Export rate	1,260	2,930	1,560	3,850	1,380	3,130
Import rate	1,010	3,420	2,330	5,330	1,740	4,220
Homicide rate	32.641	19.4048	33.5028	22.038	32.53956	21.114
Traffic fatality rate	21.095	10.341	20.161	10.226	20.657	9.937
Capital expenditures	124	142	260	195	194	172
Tax revenue	255	212	557	447	401	357
Housing expenses	146	171	301	221	221	191
<b>Cell-level controls (distances)</b>						
Airport	17.637	14.072	18.277	14.343	17.942	14.261
Public port	124.657	199.143	129.258	206.598	127.415	203.515
Railway	18.455	59.240	19.024	60.600	18.544	59.0342
Federal highway	9.539	15.817	9.569	15.662	9.546	15.747
State highway	2.623	2.866	2.660	2.965	2.636	2.905
River/lake	8.413	6.966	8.294	6.858	8.347	6.9116
# cells by district	19.934	36.852	22.500	40.364	20.983	38.370
# of cells	11,841	-	13,635	-	50,422	-
# of districts	594	-	606	-	617	-
# of municipalities	313	-	319	-	365	-

Notes: Municipal-level controls, except population, are per hundred thousand inhabitants. Export and Import rates are US\$/hundred thousand inhabitants. Cell-level controls correspond to Euclidean distance (km) from each cell's centroid to the nearest transport infrastructure/natural event.

**Table A3.** Technological classification of industry

CNAE 2.0 2-digt	Industry name	Tech level
10	Food manufacturing	low
11	Beverage production	low
12	Tobacco products manufacturing	low
13	Textile products manufacturing	low
14	Apparel manufacturing	low
15	Leather and leather products mfg	low
16	Wood products manufacturing	low
17	Paper manufacturing	low
18	Printing & Related Support activ.	low
19	Petroleum & biofuels mfg	m-low
20	Chemical manufacturing	m-high
21	Pharmaceuticals products mfg	high
22	Plastics & rubber products	m-low
23	Nonmetallic mineral products	m-low
24	Metalurgy	m-low
25	Metal products mfg	m-low
26	Computer & electronic products	high
27	Electrical machinery mfg	m-high
28	Machinery manufacturing	m-high
29	Motor vehicles manufacturing	m-high
30	Ttransport exc. motor vehicles	m-high
31	Furniture manufacturing	low
32	Miscellaneous manufacturing	low
33	Maintenance of machinery	m-low

Note: [Cavalcante \(2014\)](#)'s technological classification based on the compatibility of CNAE with the OECD technological classification.

## Appendix B Additional Results

**Table A4.** Control Function results -  
First stage residuals

# of workers with college-or-more	$\hat{\rho}_r$
0 to 1 km	-3.77e-04 (3.78e-04)
1 to 5 km	8.23e-05*** (1.68e-05)
5 to 10 km	4.87e-05*** (1.59e-05)
10 to 20 km	2.81e-05 (1.93e-05)

Notes: This table shows the estimated coefficients of all estimated residuals in equation 3 ( $r = 1, \dots, 4$ ). The model is estimated using 50,422 observations and includes all the control variables presented in Table 1. Significance levels: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ .