Online Appendix for "Internal Geography, International Trade, and Regional Specialization," Coşar and Fajgelbaum (2014)

Robustness

To check the robustness of the baseline results, we first replicate the intensive margin analysis with a specification that allows us to include observations with zero employment in the estimation. Santos Silva and Tenreyro (2006) have shown that the Poisson pseudo-maximum likelihood (PPML) estimator is well behaved in datasets with many zeros. Their implementation is motivated by the need to estimate gravity equations without excluding bilateral country pairs with zero trade flows, but the methodology is applicable to a wide range of estimations with nonnegative and continuous dependent variables. Table A1 reports the results. In these specifications, the dependent variable is the absolute level of employment and capital stock at the industry-prefecture level, including observations with a zero. While it is not possible to directly compare the coefficients from the non-linear PPML estimators with our baseline linear estimates in Table 2, we still find a highly significant negative effect of the distance-exports interaction on regional specialization patterns.

In Table A2, we replicate the specification presented in columns 2 and 4 of Table 2 by excluding from the sample 34 prefectures in three large, autonomous interior provinces of China (Inner Mongolia, Tibet and Xinjiang, marked with red triangles in Figure A1), or 53 coastal provinces with zero distance. The three interior regions that we exclude (columns 1 and 3) are among the least industrialized parts of China. Their autonomy may also subject them to a different policy environment. Since many large Chinese cities are historically on the eastern seaboard, excluding coastal prefectures (columns 2 and 4) is a robustness check on the effect of cities. Another concern addressed by this robustness check is that the first special export zones (SEZs) established in

the early 1980s as part of China's gradual reforms were all coastal.[1]

Excluding autonomous interior provinces significantly increases the absolute value of the interaction coefficient. For example, when employment is the dependent variable and US labor intensity is the trade measure, the distance gradient for export-oriented industries increases from 2.9 to 5.25. Similarly, when the dependent variable is capital, the gradient increases from 2.55 to 4.15. These changes suggest that while activity in export-oriented industries shrinks sharply from the coast to the immediate interior of the country, the decline is less pronounced as one reaches the remote interior.

Excluding coastal prefectures from the sample has the opposite effect: the interaction coefficient shrinks in absolute value, indicating the high levels of export orientation in coastal prefectures. In the case of capital, it drops from 0.216 to 0.057 and is only significant at the 10 percent level when Chinese industry-wide export-revenue ratio is the trade measure (4th columns in the top panel of Tables 2 and A2). Still, the coefficient is of the expected sign and highly significant in other cases. Thus, while coastal prefectures may display disproportionately high levels of activity in export-orientated industries, the baseline intensive margin results do not exclusively rely on their presence in the sample.

In the extensive margin, however, the message is different. Table A3 replicates the results of the fourth column in Table 3 (the linear probability model). The first two columns in Table A3 impose the same sample restrictions by excluding the remote interior and the coast. As in the intensive margin, the gradient of industry presence gets steeper when interior provinces are dropped, but the coefficient reverses sign and loses significance when coastal prefectures are excluded.

The last two columns of Table A3 check the robustness of the extensive margin results to excluding state-owned enterprises (column 3) and to clus-

[1]The four SEZs established in 1979 were coastal. An additional 14 port cities in eight coastal provinces were declared SEZs in 1984. The policy experiment, however, quickly spread to the rest of the country. By 1992, more than half of Chinese municipalities had a SEZ with the majority being in the interior (Wang 2013).

tering the standard errors at the more conservative level of provinces.[2] The loss of significance in the case of province-level clustering is consistent with the result obtained when coastal prefectures are excluded: most of the variation in the baseline extensive margin seems to be coming from the steep drop in the probability of observing an industry as one moves away from coastal prefectures.[3] Thus, our robustness checks suggest that immediate proximity to the seaboard and ports seems to be a big advantage in attracting export-oriented industries in China. Conditional on regional presence, however, there is still a significant decline in industry activity beyond the coast.

In Table A4, we assess the robustness of the panel data outcomes to the four alternative specifications tried above. The first two and bottom panels show that the results are robust to excluding state-owned enterprises and autonomous interior provinces from the sample, as well as clustering at the province-level. As in the extensive margin analysis, however, the panel data results are not robust to excluding coastal prefectures. All 3 columns in Panel C indicate the lack of a statistically significant relationship between export-orientation and industry location over time.

Finally, as an alternative measure of foreign market access, we use the natural logarithm of travel time (in hours) to the nearest international port obtained from Google Maps. Table A5 confirms that the interaction term remains negative and significant.

References

Cullinane, K., and Wang, T. F. (2006). Port Governance in China. Research in Transportation Economics, 17, 331-356.

[2]The intensive margin results are robust to both of these changes. In order to keep the analysis focused, we do not report them in Table A2, but all additional results are available from the authors upon request.

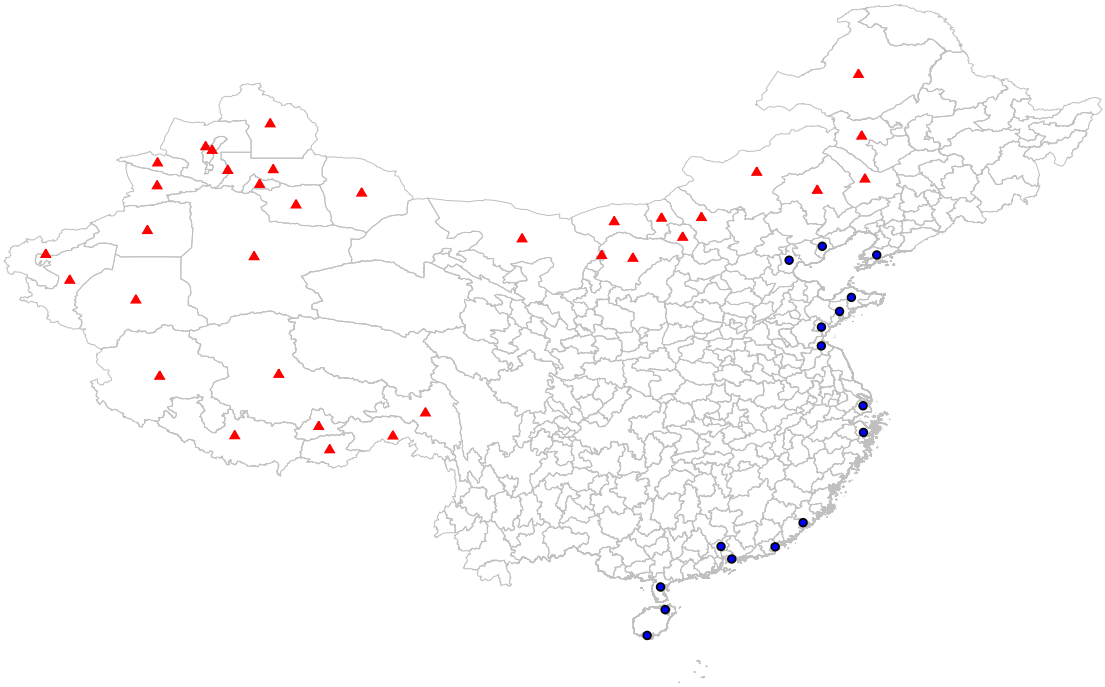
[3]When estimated with the probit model, the extensive margin coefficient remains significantly negative when we cluster at the province-level. To be conservative, we only report the outcomes that are not robust to alternative specifications.

Silva, J. S., and Tenreyro, S. (2006). The Log of Gravity. The Review of Economics and Statistics, 88(4), 641-658.

Wang, J. (2013). The Economic Impact of Special Economic Zones: Evidence from Chinese Municipalities. Journal of Development Economics. 101, pp. 133-147.

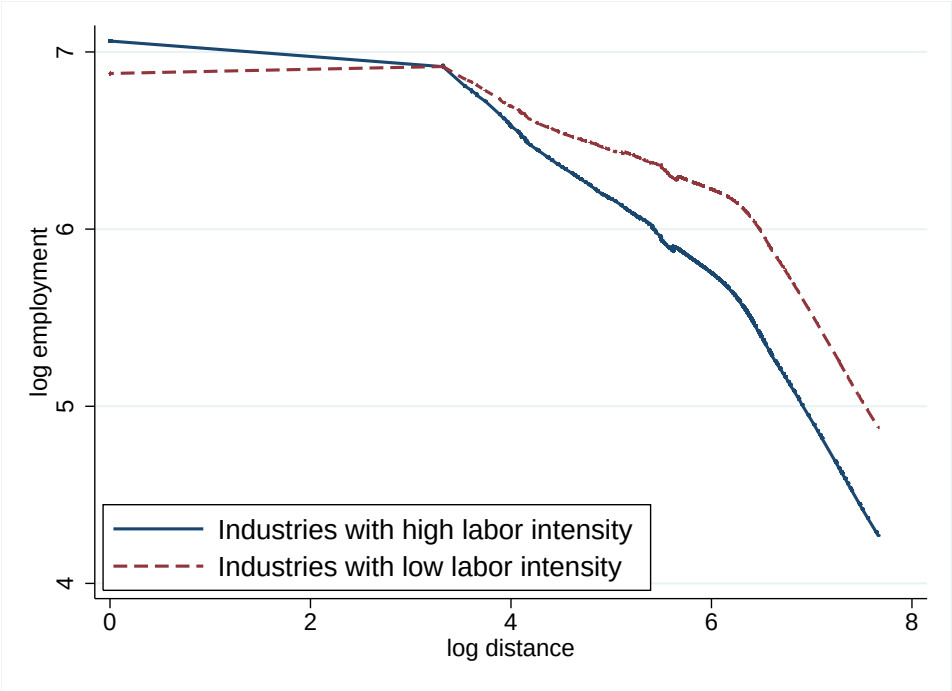
Tables and Figures

Figure A1: **Map of Chinese Prefectures and Ports**



Notes: Prefecture boundaries. Blue dots are top seaports by cargo throughput in 2005 (Cullinane and Wang 2007). Red triangles are the prefectures of Inner Mongolia, Tibet, and Xinjiang provinces which are excluded from the sample in robustness checks.

Figure A2: **Non-parametric Lowess Fit of Employment over Distance**



Notes: Non-parametric fit of log employment over log distance for industries with above or below median labor intensity.

Table A1: **Robustness: Poisson Pseudo-Maximum Likelihood**

	Employment, L		Capital, K	
	(1)	(2)	(3)	(4)
Panel A: Chinese Export-Revenue Ratio				
Distance	-0.058***		-0.079***	
	(0.01)		(0.0264)	
$CNXR$	1.746***		-0.287	
	(0.287)		(0.361)	
$Distance \times CNXR$	-1.421***	-1.111***	-1.615***	-0.925***
	(0.169)	(0.117)	(0.237)	(0.128)
Observations	147,030	147,030	147,030	147,030
R^2	0.022	0.411	0.008	0.353
Panel B: US Labor-Capital Ratio				
Distance	-0.101***		-0.119***	
	(0.025)		(0.039)	
$USLK$	7.157**		-46.17***	
	(2.849)		(4.566)	
$Distance \times USLK$	-14.15***	-11.31***	-14.25***	-8.256***
	(1.516)	(1.174)	(2.177)	(1.236)
Observations	39,884	39,884	39,884	39,884
R^2	0.037	0.442	0.024	0.450
Industry fixed effects		Yes		Yes
Prefecture fixed effects		Yes		Yes

Notes: All regressions are estimated with Poisson Pseudo-Maximum Likelihood. Dependent variables are employment L or capital K at industry-prefecture cells. R^2 is computed as the square of the correlation between the dependent variable and its fitted values. Prefecture-clustered standard errors in parentheses. Significance: * 10 percent, ** 5 percent, *** 1 percent.

Table A2: **Robustness to Sample Restrictions: Intensive Margin**

	Employment, $\ln(L)$		Capital, $\ln(K)$	
	(1)	(2)	(3)	(4)
Panel A: Chinese Export-Revenue Ratio				
$Distance \times CNXR$	-0.410***	-0.123***	-0.310***	-0.057*
	(0.045)	(0.026)	(0.045)	(0.029)
Observations	58,157	45,762	57,780	45,387
R^2	0.347	0.328	0.343	0.319
Panel B: US Labor-Capital Ratio				
$Distance \times USLK$	-5.253***	-1.703***	-4.145***	-1.708***
	(0.531)	(0.410)	(0.609)	(0.607)
Observations	24,779	20,922	24,661	20,789
R^2	0.482	0.471	0.477	0.455
Robustness	No Remote Interior	Non-coastal	No Remote Interior	Non-coastal
Industry fixed effects	Yes	Yes	Yes	Yes
Prefecture fixed effects	Yes	Yes	Yes	Yes

Notes: All regressions are estimated with OLS. Dependent variable is $\ln(L)$ in Columns 1-2, and $\ln(K)$ in Columns 3-4. Columns 1-3 exclude the 34 prefectures in 3 large interior provinces (Inner Mongolia, Tibet, Xinjiang). Columns 2-4 exclude 53 coastal prefectures with zero distance. Significance: * 10 percent, ** 5 percent, *** 1 percent.

Table A3: **Robustness to Sample Restrictions: Extensive Margin**

	(1)	(2)	(3)	(4)
Panel A: Chinese Export-Revenue Ratio				
$Distance \times CNXR$	-0.033***	0.005	-0.01**	-0.01
	(0.007)	(0.003)	(0.004)	(0.01)
Observations	132,240	123,975	146,258	147,030
R^2	0.428	0.419	0.433	0.434
Panel B: US Labor-Capital Ratio				
$Distance \times USLK$	-0.551***	0.0341	-0.273***	-0.137
	(0.127)	(0.069)	(0.074)	(0.150)
Observations	35,872	33,630	39,766	39,884
R^2	0.445	0.465	0.480	0.461
Robustness	No Remote Interior	Non-coastal	Privately Owned	Province Clustered
Industry fixed effects	Yes	Yes	Yes	Yes
Prefecture fixed effects	Yes	Yes	Yes	Yes

Notes: All regressions are estimated with LPM. Dependent variable is $\ln(L)$ in Columns 1-2, and $\ln(K)$ in Columns 3-4. Column 1 excludes the 34 prefectures in three autonomous interior provinces (Inner Mongolia, Tibet, Xinjiang). Column 2 excludes 53 coastal prefectures with zero distance. Column 3 excludes state-owned enterprises in the construction of the regional data. Column 4 clusters standard errors at the province level. Significance: * 10 percent, ** 5 percent, *** 1 percent.

Table A4: **Robustness to Sample Restrictions: Time Variation**

	(1) Employment, ln(L)	(2) Capital, ln(K)	(3) Extensive Margin
Panel A: Private Enterprises Only			
$CNXR$	0.233**	-0.059	-0.026***
	(0.114)	(0.119)	(0.007)
$Distance \times CNXR$	-0.151***	-0.120***	-0.0033***
	(0.032)	(0.038)	(0.001)
Observations	313,718	314,829	1,347,304
R^2	0.812	0.788	0.719
Panel B: Exclude Autonomous Interior Provinces			
$CNXR$	0.435***	0.215	0.008
	(0.145)	(0.132)	(0.011)
$Distance \times CNXR$	-0.259***	-0.217***	-0.012***
	(0.0502)	(0.051)	(0.003)
Observations	350,977	351,814	1,225,857
R^2	0.817	0.800	0.728
Panel C: Exclude Coastal Prefectures			
$CNXR$	-0.306**	-0.523***	-0.0249***
	(0.151)	(0.152)	(0.008)
$Distance \times CNXR$	-0.0473	-0.0202	-0.002
	(0.0360)	(0.0424)	(0.001)
Observations	258,216	258,937	1,137,868
R^2	0.808	0.790	0.718
Panel D: Clustering at Province Level			
$CNXR$	0.369	0.155	-0.005
	(0.274)	(0.226)	(0.012)
$Distance \times CNXR$	-0.206**	-0.176**	-0.005*
	(0.078)	(0.064)	(0.003)
Observations	361,580	362,652	1,363,315
R^2	0.818	0.800	0.732
Year fixed effects	Yes	Yes	Yes
Industry-prefecture fixed effects	Yes	Yes	Yes

Notes: All panel regressions are estimated with year and industry-prefecture fixed effects. Dependent variables are the logarithm of employment L or capital K at industry-prefecture-year cells with positive observations in the first two columns. Dependent variable in the third column is a binary variable that takes the value one if $L > 0$ and zero otherwise. Panel A excludes state-owned enterprises in the construction of the regional data. Panel B excludes the prefectures in three autonomous interior provinces (Inner Mongolia, Tibet, Xinjiang). Panel C excludes 53 coastal prefectures with zero distance. Panel D clusters standard errors at the province level. Significance: * 10 percent, ** 5 percent, *** 1 percent.

Table A5: **Robustness to Alternative Market Access Measure**

	Employment, ln(L)		Capital, ln(K)	
	(1)	(2)	(3)	(4)
Panel A: Chinese Export-Revenue Ratio				
$TTime$	-0.013***		-0.017***	
	(0.003)		(0.004)	
$CNXR$	0.966***		-0.327*	
	(0.173)		(0.175)	
$TTime \times CNXR$	-0.099***	-0.090***	-0.103***	-0.072***
	(0.015)	(0.014)	(0.016)	(0.013)
Observations	60,505	60,505	60,100	60,100
R^2	0.035	0.344	0.034	0.341
Panel B: US Labor-Capital Ratio				
$TTime$	-0.033***		-0.037***	
	(0.005)		(0.006)	
$USLK$	-10.96***		-37.65***	
	(2.047)		(2.245)	
$TTime \times USLK$	-0.800***	-0.873***	-0.763***	-0.793***
	(0.184)	(0.193)	(0.228)	(0.227)
Observations	26,229	26,229	26,091	26,091
R^2	0.080	0.476	0.100	0.471
Industry fixed effects		Yes		Yes
Prefecture fixed effects		Yes		Yes

Notes: All regressions are estimated with OLS. $TTime$ is travel time (in hours) to the nearest international port. Dependent variable is ln(L) in Columns 1-2, and ln(K) in Columns 3-4. Prefecture-clustered standard errors in parentheses. Significance: * 10 percent, ** 5 percent, *** 1 percent.