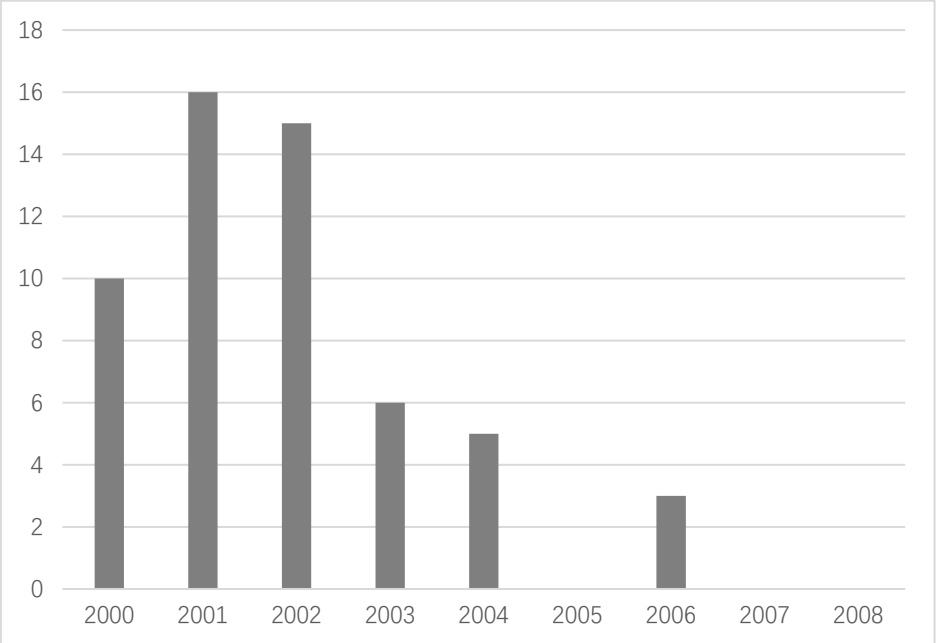
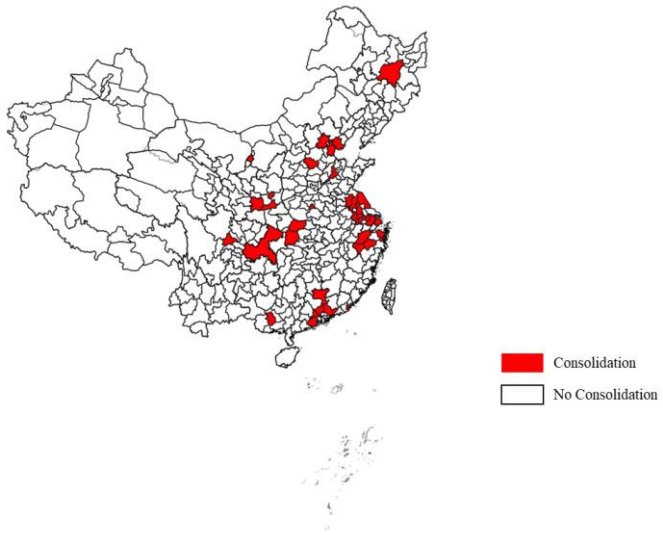
Appendix A

Figure A1. The number of consolidation cases in China (2000-2008)



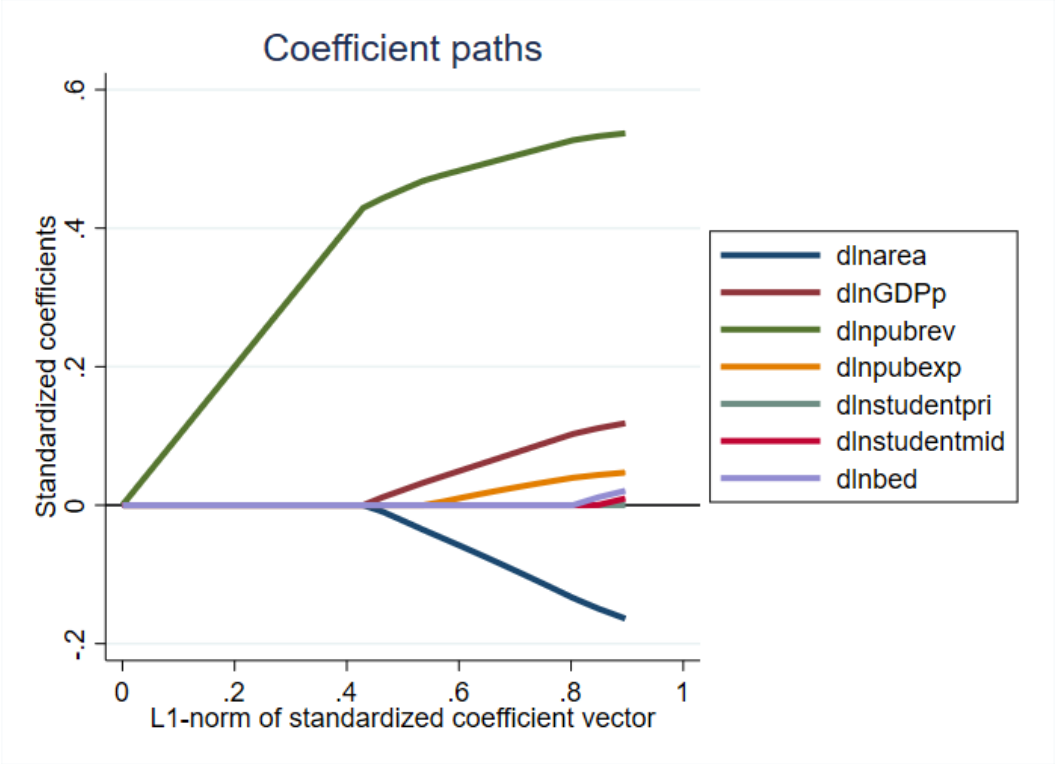
Note: The figure displays the annual number of consolidations from 2000 to 2008.

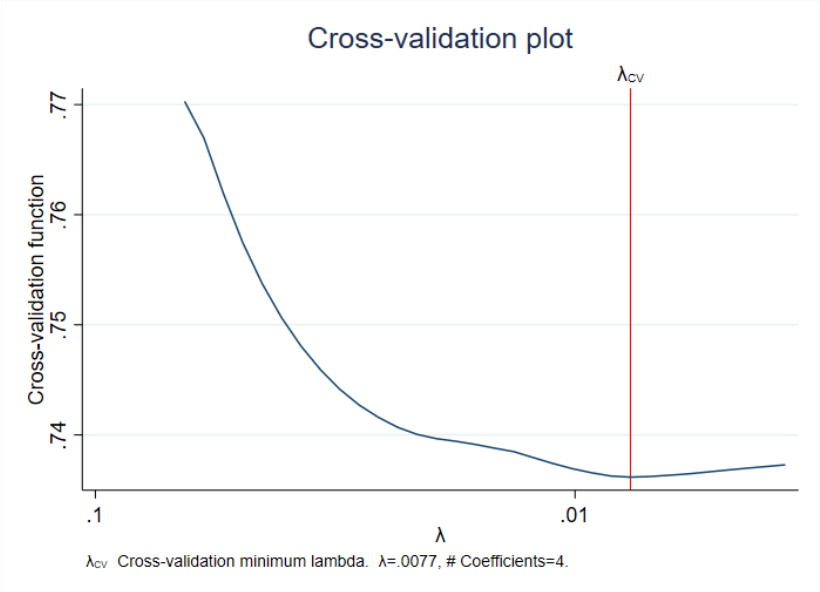
Figure A2. The distribution of “city-county consolidation” across China (2000-2008)



Note: The figure displays the prefectures where “city-county consolidation” occurred from 2000 to 2008.

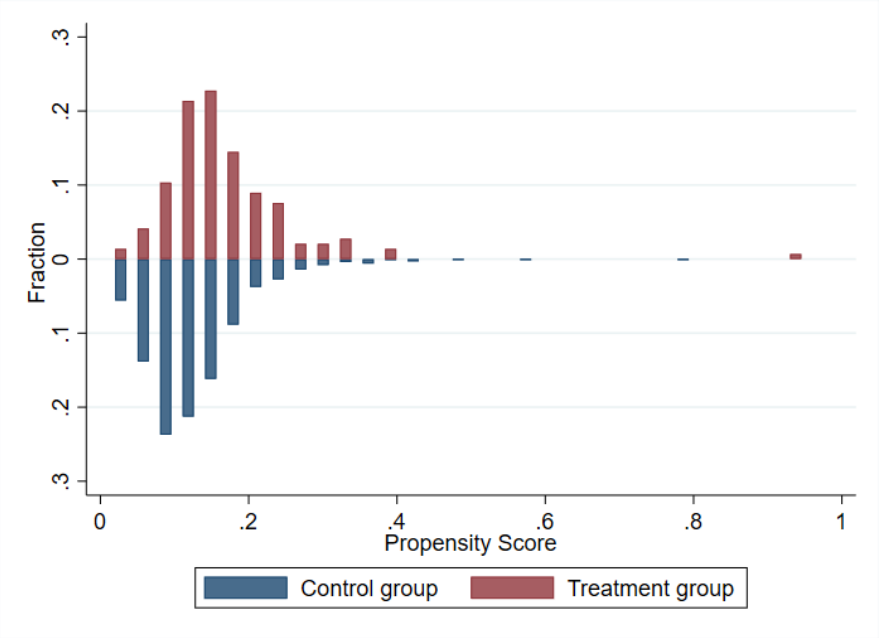
Figure A3. Coefficient paths and cross-validation of LASSO





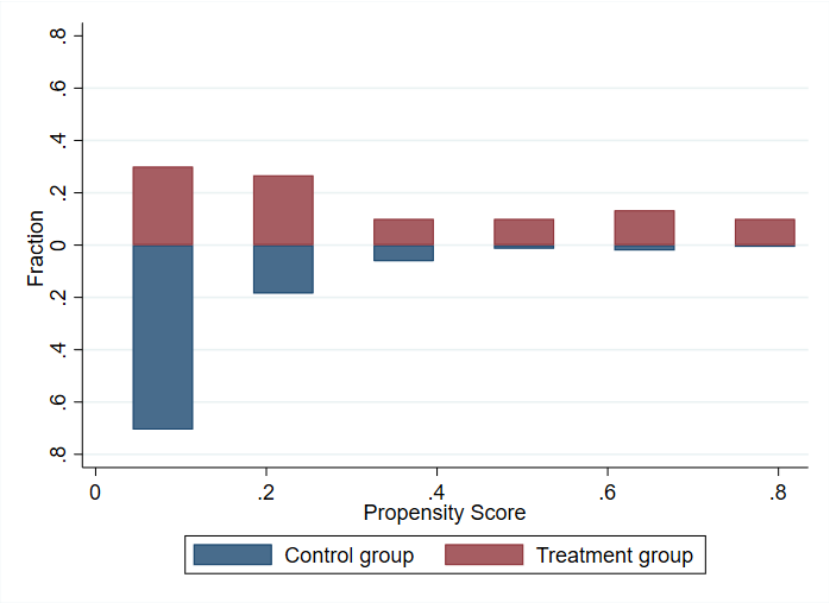
Note: The figures provide the coefficient paths for each regressor and the cross-validation plot. The results suggest that 4 regressors (average change of: area, GDP per capita, government revenue, government expenditure) are left to estimate propensity score according to LASSO.

Figure A4. The distribution of propensity score by groups



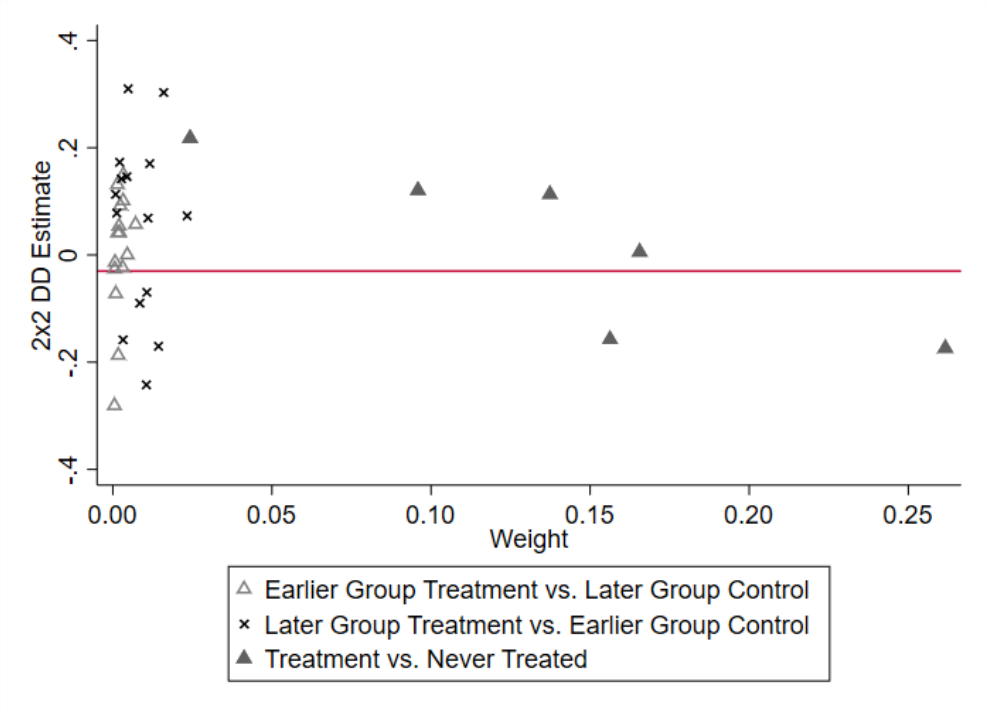
Note: The figure provides the distribution of propensity scores for the treatment and control groups.

Figure A5. The distribution of propensity score by groups (prefecture)



Note: The figure provides the distribution of propensity score for the treatment group and the control group.

Figure A6. Goodman-Bacon Decomposition without controls



Note: According to the difference-in-difference decomposition theorem proposed by Goodman-bacon (2018), when the treated units receive treatment at different times, the general two-way fixed effects DiD estimator is a weighted average of all possible 2*2 DiD estimators. The figure plots each 2*2 difference-in-difference estimator and their weight in the data. The closed triangles are terms in which one treated unit acts as the treatment group and the never treated unit acts as the control group. The open triangles are terms that an earlier treated unit act as the treatment group and a later treated unit acts as the control group. The x symbols are terms that a later treated unit act as the treatment group and an earlier treated unit acts as the control group. The average difference-in-difference estimate is -0.030, which is indicated by the red line. The overall treatment effect is the average of the y-axis values weighted by their x-axis values. According to Goodman-bacon, the overall estimates is biased by the coefficients of x symbols.

Table A1 Summary statistics (counties)

Definition	Total		Control Group		Treatment group	
	(N=22740)		(N=19820)		(N=2920)	
GDP per capita	17059.59	21185.66	15317.62	17492.3	28892.22	35481.45
Area	2128.084	2000.386	2176.649	2079.869	1797.516	1292.068
Government expenditure (RMB)	1.13E+09	1.48E+09	1.05E+09	1.25E+09	1.66E+09	2.48E+09
Government revenue (RMB)	4.94E+08	1.12E+09	4.08E+08	7.45E+08	1.07E+09	2.36E+09
# of primary students	45804.98	37257.84	46333.6	38497.42	42256.37	27268.88
# of middle students	31225.79	22090.13	31195.06	22595.97	31431.66	18348.01
# of hospital beds	1040.453	765.5778	999.9442	724.1624	1309.211	955.3404

Source: China County Statistics, China Data Online (https://www.china-data-online.com)

Table A2. Summary statistics (prefecture)

Definition	Total		Control Group		Treatment group	
	(N=4140)		(N=3540)		(N=600)	
GDP per capita	24340.57	23150.95	21777.27	19930.17	39424.77	32984.96
Area (10,000 km^2)	1.462157	1.475647	1.50847	1.547153	1.197402	0.92807
Government expenditure (100 million RMB)	142.4964	184.2281	121.0294	133.2814	267.3232	331.5311
Government revenue (100 million RMB)	80.42939	146.9153	58.74075	88.42587	200.1783	284.1939
# of primary students (10,000)	35.0321	24.14262	33.99861	24.38919	41.04355	21.719
# of middle students (10,000)	24.44679	15.6405	23.58032	15.57729	29.4868	15.0569
# of hospital beds (10,000)	1.328802	1.009864	1.201639	0.799158	2.05405	1.606

Source: China City Statistics Yearbook

Table A3. Prefecture level analysis

VARIABLES	(1) Ln(GDP per capita)	(2) Ln(GDP per capita)	(3) Ln(GDP per capita)	(4) Ln(GDP per capita)
Treatment	0.00338			
	(0.0393)			
Treat(-3)		0.00308		
		(0.0256)		
Treat(-2)		-0.00793		
		(0.0205)		
Treat(-1)		-0.00175		
		(0.0108)		
Treat(+1)		-0.000405		
		(0.0127)		
Treat(+2)		-0.00694		
		(0.0225)		
Treat(+3)		-0.0187		
		(0.0305)		
Treat(+4)		-0.0192		
		(0.0368)		
Treat(+5)		-0.00505		
		(0.0427)		
Treat(+6)		0.000184		
		(0.0469)		
Treat(+7)		0.0108		
		(0.0489)		
Treat(+8)		-0.0181		
		(0.0509)		
Treat(+9)		-0.0249		
		(0.0523)		
Treat(+10)		0.0339		
		(0.0531)		
Treat intensity (pop share of urban core)			-0.0496	
			(0.0730)	
Treat intensity (pop share of prefecture)				-0.0819
				(0.104)
County FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Controls	Y	Y	Y	Y
Observations	1,119	1,119	1,119	1,119
R-squared	0.982	0.982	0.982	0.982

Notes: The dependent variable is ln(GDP per-capita) of each rural county. Column (1) reports the base-DiD results. Column (2) includes the pre- and post-treatment-year indicators to assess pre-trends, treat(-t) means the indicator for t years before treatment, while treat(+t) means the indicator for t years after treatment. The years beyond -3 and +10 are grouped into -3 and +10 respectively. Column (3) shows the results of adding the treatment-intensity variable to the base model: i.e., ratio of the consolidated-county's(ies') population to the urban-core population before consolidation; Column (4) is the result of adding to the base model the treatment-intensity variable: i.e., ratio of the consolidated-county's population to the entire prefecture's population. Robust standard errors clustered at the county level are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A4. Results of excluding neighbors of the treated prefectures

VARIABLES	(1) Ln(GDP per capita)	(2) Ln(GDP per capita)	(3) Ln(GDP per capita)	(4) Ln(GDP per capita)
Treatment	-0.0871***			
	(0.0306)			
Treat(-3)		0.0309		
		(0.0274)		
Treat(-2)		0.0270		
		(0.0187)		
Treat(-1)		0.00754		
		(0.0107)		
Treat(+1)		-0.0386***		
		(0.0134)		
Treat(+2)		-0.0246		
		(0.0159)		
Treat(+3)		-0.0574***		
		(0.0210)		
Treat(+4)		-0.0973***		
		(0.0250)		
Treat(+5)		-0.103***		
		(0.0289)		
Treat(+6)		-0.104***		
		(0.0324)		
Treat(+7)		-0.119***		
		(0.0358)		
Treat(+8)		-0.117***		
		(0.0371)		
Treat(+9)		-0.110***		
		(0.0389)		
Treat(+10)		-0.0920**		
		(0.0403)		
Treat intensity (pop share of urban core)			-0.289***	
			(0.0606)	
Treat intensity (pop share of prefecture)				-0.560***
				(0.145)
County FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Controls	Y	Y	Y	Y
Observations	4,684	4,684	4,684	4,684
R-squared	0.968	0.968	0.968	0.968

Notes: The dependent variable is ln(GDP per-capita) of each rural county. Column (1) reports the base-DiD results. Column (2) includes the pre- and post-treatment-year indicators to assess pre-trends, treat(-t) means the indicator for t years before treatment, while treat(+t) means the indicator for t years after treatment. The years beyond -3 and +10 are grouped into -3 and +10 respectively. Column (3) shows the results of adding the treatment-intensity variable to the base model: i.e., ratio of the consolidated-county's(ies') population to the urban-core population before consolidation; Column (4) is the result of adding to the base model the treatment-intensity variable: i.e., ratio of the consolidated-county's population to the entire prefecture's population. Robust standard errors clustered at the county level are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A5. Results of using caliper matching approach

VARIABLES	(1) Ln(GDP per capita)	(2) Ln(GDP per capita)	(3) Ln(GDP per capita)	(4) Ln(GDP per capita)
Treatment	-0.0616***			
	(0.0220)			
Treat(-3)		0.0194		
		(0.0221)		
Treat(-2)		0.0149		
		(0.0156)		
Treat(-1)		0.000187		
		(0.00919)		
Treat(+1)		-0.0314***		
		(0.0120)		
Treat(+2)		-0.0139		
		(0.0126)		
Treat(+3)		-0.0431***		
		(0.0157)		
Treat(+4)		-0.0821***		
		(0.0178)		
Treat(+5)		-0.0848***		
		(0.0199)		
Treat(+6)		-0.0807***		
		(0.0218)		
Treat(+7)		-0.0906***		
		(0.0247)		
Treat(+8)		-0.0889***		
		(0.0252)		
Treat(+9)		-0.0809***		
		(0.0265)		
Treat(+10)		-0.0563**		
		(0.0276)		
Treat intensity (pop share of urban core)			-0.230***	
			(0.0437)	
Treat intensity (pop share of prefecture)				-0.470***
				(0.121)
County FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Controls	Y	Y	Y	Y
Observations	22,408	22,408	22,408	22,408

R-squared	0.966	0.966	0.966	0.966

Notes: The dependent variable is ln(GDP per-capita) of each rural county. Column (1) reports the base-DiD results. Column (2) includes the pre- and post-treatment-year indicators to assess pre-trends, treat(-t) means the indicator for t years before treatment, while treat(+t) means the indicator for t years after treatment. The years beyond -3 and +10 are grouped into -3 and +10 respectively. Column (3) shows the results of adding the treatment-intensity variable to the base model: i.e., ratio of the consolidated-county's(ies') population to the urban-core population before consolidation; Column (4) is the result of adding to the base model the treatment-intensity variable: i.e., ratio of the consolidated-county's population to the entire prefecture's population. Robust standard errors clustered at the county level are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A6. Results of using kernel matching approach

VARIABLES	(1) Ln(GDP per capita)	(2) Ln(GDP per capita)	(3) Ln(GDP per capita)	(4) Ln(GDP per capita)
Treatment	-0.0675***			
	(0.0224)			
Treat(-3)		0.0248		
		(0.0222)		
Treat(-2)		0.0171		
		(0.0158)		
Treat(-1)		0.00139		
		(0.00926)		
Treat(+1)		-0.0326***		
		(0.0120)		
Treat(+2)		-0.0160		
		(0.0127)		
Treat(+3)		-0.0455***		
		(0.0158)		
Treat(+4)		-0.0844***		
		(0.0181)		
Treat(+5)		-0.0875***		
		(0.0203)		
Treat(+6)		-0.0839***		
		(0.0224)		
Treat(+7)		-0.0943***		
		(0.0254)		
Treat(+8)		-0.0925***		
		(0.0259)		
Treat(+9)		-0.0848***		
		(0.0273)		
Treat(+10)		-0.0591**		
		(0.0282)		
Treat intensity (pop share of urban core)			-0.238***	
			(0.0443)	
Treat intensity (pop share of prefecture)				-0.488***
				(0.121)
County FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Controls	Y	Y	Y	Y
Observations	22,388	22,388	22,388	22,388
R-squared	0.966	0.966	0.966	0.966

Notes: The dependent variable is ln(GDP per-capita) of each rural county. Column (1) reports the base-DiD results. Column (2) includes the pre- and post-treatment-year indicators to assess pre-trends, treat(-t) means the indicator for t years before treatment, while treat(+t) means the indicator for t years after treatment. The years beyond -3 and +10 are grouped into -3 and +10 respectively. Column (3) shows the results of adding the treatment-intensity variable to the base model: i.e., ratio of the consolidated-county's(ies') population to the urban-core population before consolidation; Column (4) is the result of adding to the base model the treatment-intensity variable: i.e., ratio of the consolidated-county's population to the entire prefecture's population. Robust standard errors clustered at the county level are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.