Online Appendix

Decomposition exercise

This subsection makes two points: 1) how we can account for the population results using the aggregate net-migration rate; and 2) how we can transform the migration rate by birthplace n in county c into an aggregated county-level migration rate.

We start our decomposition exercise by noting that the population of county c must develop according to the basic demographic equation:

$$P_{ct+1} = P_{ct} - D_{ct} + B_{ct} + I_{ct}. \tag{1}$$

Next, assuming that the number of deaths (D_{ct}) and births (B_{ct}) offset each other (or that they are uncorrelated with the earthquake) and that I_t is defined as net-migration, this equation can be rewritten as:

$$\frac{P_{ct+1} - P_{ct}}{P_{ct}} = \frac{I_{ct}}{P_{ct}}, \tag{2}$$

such that population growth of county c is equal to the net-migration rate (I_{ct}/P_{ct}), which we can use to compare the migration-rate estimate with the population-growth estimate (i.e., how much can migration account for the reduced-form population finding). Let the aggregated migration rate be equal to the sum of the migration rates of each birthplace n:

$$\frac{I_{ct}}{P_{ct}} = \frac{I^1_{ct}}{P_{ct}} + \frac{I^2_{ct}}{P_{ct}} + \cdots + \frac{I^N_{ct}}{P_{ct}} = \sum_{n=1}^{N} \frac{I^n_{ct}}{P_{ct}}. \tag{3}$$

Since nobody is born as a migrant in the US, the stock of migrants from birthplace n in a county, F^n_{ct}, develops as (ignoring deaths):

$$\begin{aligned} F^n_{ct+1} &= F^n_{ct} + I^n_{ct} \\ \frac{F^n_{ct+1} - F^n_{ct}}{P_{ct}} &= \frac{I^n_{ct}}{P_{ct}}. \end{aligned} \tag{4}$$

First, this shows that we can use the change in the number of migrants from t to $t+1$ scaled with county population size in t as a measure of the net-migration rate. Second, if we want to study how the earthquake influenced the aggregate immigration rate (I_{ct}/P_{ct}), using our birthplace-county estimation equation, we need to multiply the estimated coefficient, which reflects the average effects of each birthplace group, with the number of birthplace groups according to equation (3).

Appendix Tables and Figures

Appendix Table 1: Summary Statistics

	(1) N	(2) mean	(3) sd
Population	3,616	11,594	89,124
Quake	3,616	1.211	2.629
Latitude	3,616	41.82	3.749
Longitude	3,616	-121.0	2.564
Distance to Los Angeles	3,616	577.5	247.0
Distance to San Francisco	3,616	373.3	175.6
Distance to Epicenter	3,616	377.5	176.1
Δ Migration Share	7,150	0.005	0.019
In-migration Rate	102	0.518	0.722
Out-migration Rate	102	0.454	0.082

Appendix Table 2: Corresponding event-study estimates of Figure 3

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	ln(Population)					
$Quake_{1870}$	-0.196	0.0501	0.0272	0.0529	-0.0242	
	(0.188)	(0.247)	(0.208)	(0.244)	(0.333)	
$Quake_{1880}$	-0.0556	0.0827	0.0900	0.0835	0.0139	
	(0.0990)	(0.131)	(0.110)	(0.125)	(0.166)	
$Quake_{1890}$	-0.0187	0.0182	-0.00495	0.0130	-0.0233	
	(0.0398)	(0.0554)	(0.0541)	(0.0553)	(0.0825)	
$Quake_{1910}$	-0.0854***	-0.123***	-0.0980***	-0.110***	-0.0790	-0.147***
	(0.0270)	(0.0341)	(0.0339)	(0.0338)	(0.0513)	(0.0557)
$Quake_{1920}$	-0.135***	-0.225***	-0.186***	-0.208***	-0.189***	-0.183**
	(0.0347)	(0.0454)	(0.0436)	(0.0444)	(0.0678)	(0.0733)
$Quake_{1930}$	-0.180***	-0.309***	-0.256***	-0.284***	-0.251***	-0.296***
	(0.0431)	(0.0562)	(0.0542)	(0.0556)	(0.0854)	(0.0946)
$Quake_{1940}$	-0.227***	-0.332***	-0.292***	-0.309***	-0.256***	-0.335***
	(0.0433)	(0.0578)	(0.0594)	(0.0578)	(0.0894)	(0.0902)
$Quake_{1950}$	-0.254***	-0.394***	-0.331***	-0.363***	-0.309***	-0.344***
	(0.0503)	(0.0669)	(0.0660)	(0.0665)	(0.108)	(0.115)
$Quake_{1960}$	-0.235***	-0.390***	-0.306***	-0.350***	-0.217*	-0.309**
	(0.0576)	(0.0763)	(0.0765)	(0.0760)	(0.125)	(0.136)
$Quake_{1970}$	-0.275***	-0.449***	-0.369***	-0.414***	-0.228*	-0.318**
	(0.0616)	(0.0797)	(0.0812)	(0.0802)	(0.129)	(0.141)
Panel	A	B	C	D	E	F
Observations	3,616	3,616	2,962	3,383	3,616	1,208
R^2	0.016	-0.000	0.012	0.000	-0.017	-0.036

NOTES: This table shows the corresponding estimates of Figure 3. Standard errors in parentheses are clustered at the city level: *** p<0.01, ** p<0.05, * p<0.1.

Appendix Table 3: County-level estimates and long-run results

VARIABLES	(1) ln(population)	(2) Δln(population)
$Quake_{1880}$	0.293***	
	(0.0964)	
$Quake_{1890}$	0.0145	
	(0.0365)	
$Quake_{1910}$	-0.0804*	
	(0.0423)	
$Quake_{1920}$	-0.149**	
	(0.0654)	
$Quake_{1930}$	-0.235***	
	(0.0871)	
$Quake_{1940}$	-0.248***	
	(0.0901)	
$Quake_{1950}$	-0.279**	
	(0.115)	
$Quake_{1960}$	-0.346**	
	(0.132)	
$Quake_{1970}$	-0.361**	
	(0.141)	
Quake		-0.357**
		(0.156)
Sample	County	City
Observations	938	338
R^2	0.054	0.318

NOTES: This table shows county level estimates in column (1) and long-run estimates for the period 1900-1970 at the city level in column (2). The dependent variable is ln(population) in column (1) and the change in ln(population) between 1900 and 1970 in column (2). The method of estimation is two-stage least squares. The county-level specification controls for longitude and latitude interacted by time fixed effects. The city level specification includes the same control variables as Panel F of Appendix Figure 2. Standard errors in parentheses are clustered at the county level in column (1) and at the city level in column (2): *** p<0.01, ** p<0.05, * p<0.1.

Appendix Table 4: Corresponding event-study estimates of Figure 5

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				ln(Population)			
$Quake_{1870}$	0.545	0.0618	0.0642	0	0.169	5.344	-0.0897
	(0.565)	(0.332)	(0.191)	(0)	(0.218)	(5.359)	(0.388)
$Quake_{1880}$	0.359	-0.165	0.189	1.270	0.260**	0.882	-0.0157
	(0.238)	(0.202)	(0.120)	(1.290)	(0.113)	(0.639)	(0.211)
$Quake_{1890}$	0.0337	0.0552	0.00356	0.671	0.0504	-0.141	0.0808
	(0.0917)	(0.0996)	(0.0602)	(0.613)	(0.0555)	(0.154)	(0.0685)
$Quake_{1910}$	-0.212***	-0.139***	-0.156**	-0.154	-0.151***	-0.212	-0.148***
	(0.0665)	(0.0476)	(0.0701)	(0.131)	(0.0410)	(0.142)	(0.0421)
$Quake_{1920}$	-0.426***	-0.270***	-0.176*	-0.413**	-0.186***	-0.466***	-0.290***
	(0.0959)	(0.0633)	(0.0891)	(0.180)	(0.0554)	(0.157)	(0.0578)
$Quake_{1930}$	-0.592***	-0.351***	-0.273**	-0.608***	-0.269***	-0.536***	-0.362***
	(0.139)	(0.0803)	(0.119)	(0.231)	(0.0696)	(0.195)	(0.0728)
$Quake_{1940}$	-0.566***	-0.368***	-0.303**	-0.643***	-0.265***	-0.562***	-0.404***
	(0.134)	(0.0827)	(0.123)	(0.195)	(0.0737)	(0.184)	(0.0735)
$Quake_{1950}$	-0.783***	-0.448***	-0.311**	-0.833***	-0.291***	-0.852***	-0.444***
	(0.167)	(0.0989)	(0.125)	(0.249)	(0.0836)	(0.259)	(0.0847)
$Quake_{1960}$	-0.969***	-0.417***	-0.345**	-0.845***	-0.267***	-0.966***	-0.436***
	(0.197)	(0.114)	(0.135)	(0.279)	(0.0968)	(0.301)	(0.0945)
$Quake_{1970}$	-1.151***	-0.486***	-0.359**	-0.961***	-0.295***	-1.025***	-0.553***
	(0.213)	(0.116)	(0.158)	(0.294)	(0.102)	(0.305)	(0.0985)
$Quake_{1870} \times Bay$	-0.654						
	(0.539)						
$Quake_{1880} \times Bay$	-0.331*						
	(0.195)						
$Quake_{1890} \times Bay$	-0.0408						
	(0.0787)						
$Quake_{1910} \times Bay$	0.122*						
	(0.0622)						
$Quake_{1920} \times Bay$	0.275***						
	(0.0987)						
$Quake_{1930} \times Bay$	0.386**						
	(0.152)						
$Quake_{1940} \times Bay$	0.318**						
	(0.146)						
$Quake_{1950} \times Bay$	0.527***						
	(0.185)						
$Quake_{1960} \times Bay$	0.785***						
	(0.217)						
$Quake_{1970} \times Bay$	0.949***						
	(0.234)						
Sample	All	Below 2,500	Above 2,500	No Banks	Banks	Below MA	Above MA
Observations	3,616	3,071	545	2,020	1,596	1,620	1,841
R^2	-0.068	-0.006	-0.008	-0.050	0.025	-0.125	0.004

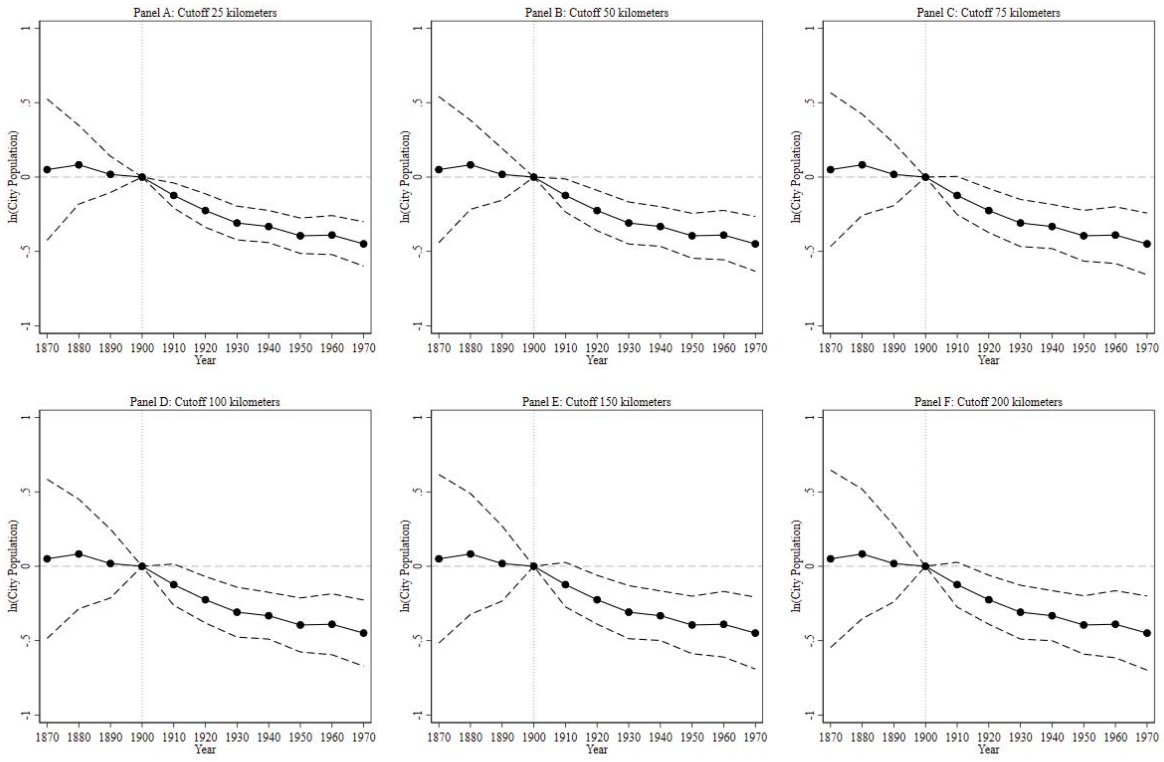
NOTES: This table shows the corresponding estimates of Figure 5. Standard errors in parentheses are clustered at the city level: *** p<0.01, ** p<0.05, * p<0.1.

Appendix Table 5: Migration Analysis – Pooled Results 1900-1940

VARIABLES	(1) Δ Migration-Rate	(2) Δ Migration-Rate
$Quake_{1900-10}$	-0.00111***	-0.000593***
	(0.000195)	(0.000213)
$Quake_{1910-20}$	-0.000921***	-0.000664***
	(0.000118)	(0.000124)
$Quake_{1920-30}$	-0.00113***	-0.000792***
	(0.000108)	(0.000113)
$Quake_{1930-40}$	-4.65e-05	2.33e-05
	(7.12e-05)	(7.80e-05)
$Nationality\ Share_{1900} \times Quake_{1900-10}$		-0.0250***
		(0.00541)
$Nationality\ Share_{1910} \times Quake_{1910-20}$		-0.00823***
		(0.00172)
$Nationality\ Share_{1920} \times Quake_{1920-30}$		-0.00841***
		(0.00166)
$Nationality\ Share_{1930} \times Quake_{1930-40}$		-0.00197***
		(0.000572)
$Nationality\ Share_{1900}$		0.0773***
		(0.0120)
$Nationality\ Share_{1910}$		0.0261***
		(0.00386)
$Nationality\ Share_{1920}$		0.0255***
		(0.00361)
$Nationality\ Share_{1930}$		-0.00131*
		(0.000695)
Observations	32,335	32,335
R-squared	0.003	0.016

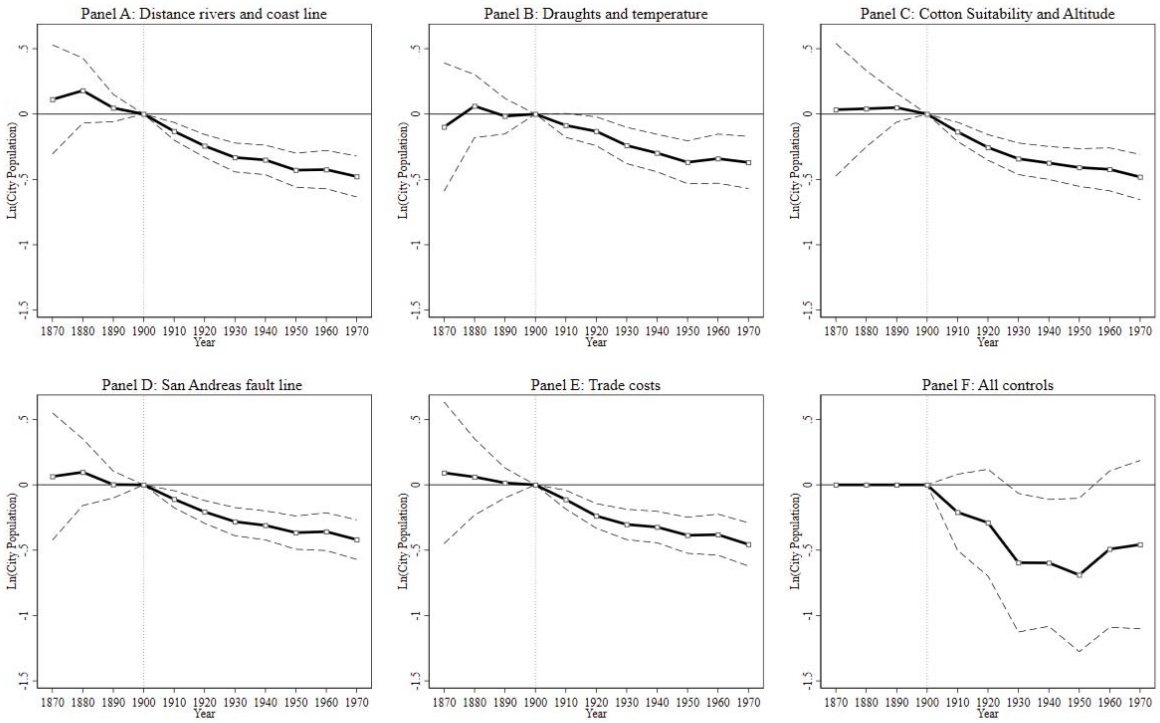
NOTES: This table extends the analysis presented in Table 1 columns (1)-(2). The sample includes the decades 1900-1910, 1910-1920, 1920-1930, and 1930-1940. The estimation equation is (2) and the estimation method two-stage least squares. Standard errors in parentheses are clustered at the birthplace by county level: *** p<0.01, ** p<0.05, * p<0.1.

Appendix Figure 1: Conley standard errors with different spatial cutoffs



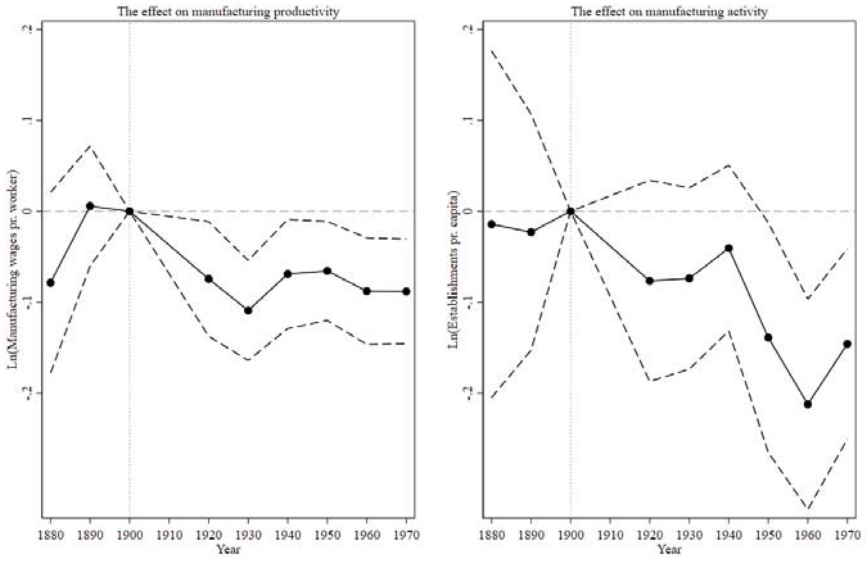
Notes: Panels A-F display the baseline regression of Figure 3 Panel B with Conley standard errors instead of clustering standard errors at the city level.
Each panel refers to different spatial cutoffs (25, 50, 75, 100, 150 and 200 kilometers). The solid line depicts the dynamic effects of earthquake intensity on city popul relative to the base year (1900). Dashed lines indicate the corresponding 95% confidence intervals based on Conley standard errors.

Appendix Figure 2: Robustness checks city population analysis



Notes: The different panels show modifications of the baseline estimates (Figure 3, Panel B). The method of estimation is two-stages least squares. Panels A-F show the dynamic effects of earthquake intensity on city population including the following controls (each interacted by time fixed effects): distance to rivers and coast line (Panel A), draughts and temperature (Panel B), cotton suitability and altitude (Panel C), a dummy == 1 if a city is located in a county on the San Andreas fault line (Panel D), trade costs (Panel E), and including the controls together (Panel F). The solid line depicts the effect on city population relative to the base year (1900). Dashed lines indicate 95% confidence intervals.

Appendix Figure 3: Manufacturing results using Conley standard errors



Notes: Panels A and B display the baseline regression of Figure 4 using conley standard errors with a spatial cutoff of 50 kilometers. The method of estimation is two-stages least squares. The solid line depicts the effect on various manufacturing outcomes relative to the base year (1900). Dashed lines indicate the corresponding 95% confidence intervals based on Conley standard errors.