DISCUSSION PAPER SERIES

DP13632

HOW THE 1906 SAN FRANCISCO EARTHQUAKE SHAPED ECONOMIC ACTIVITY IN THE AMERICAN WEST

Philipp Ager, Katherine Eriksson, Casper Worm Hansen and Lars Lønstrup

ECONOMIC HISTORY

ISSN 0265-8003

HOW THE 1906 SAN FRANCISCO EARTHQUAKE SHAPED ECONOMIC ACTIVITY IN THE AMERICAN WEST

Philipp Ager, Katherine Eriksson, Casper Worm Hansen and Lars Lønstrup

Discussion Paper DP13632
Published 28 March 2019
Submitted 27 March 2019

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

HOW THE 1906 SAN FRANCISCO EARTHQUAKE SHAPED ECONOMIC ACTIVITY IN THE AMERICAN WEST

Abstract

This paper examines the long-run effects of the 1906 San Francisco Earthquake on the spatial distribution of economic activity in the American West. Using variation in the potential damage intensity of the earthquake, we show that more severely affected cities experienced lower population increases relative to less affected cities until the late 20th century. This long-lasting effect is largely a result of individuals' high geographical mobility at that time. Less affected areas became more attractive migration destinations in the immediate aftermath of the earthquake, which permanently changed the spatial distribution of economic activity in the American West.

JEL Classification: N9, O15, O40, R11, R12

Keywords: Economic Geography, Location of Economic Activity, migration, Natural Disasters, American West

Philipp Ager - phag@sam.sdu.dk
University of Southern Denmark and CEPR

Katherine Eriksson - kaeriksson@ucdavis.edu
UC Davis and NBER

Casper Worm Hansen - casper.worm.hansen@econ.ku.dk
University of Copenhagen

Lars Lønstrup - loe@sam.sdu.dk
University of Southern Denmark

Acknowledgements
The authors would like to thank Alberto Alesina, Sascha Becker, Carl-Johan Dalgaard, Francesco Drago, James Fenske,Walker Hanlon, Marc Klemp, Christopher Meissner, Stelios Michalopoulos, Nathan Nunn, Paul Sharp, Daniel Sturm, Holger Strulik, David Weil, Nico Voigtlaender, and seminar participants at the University of Copenhagen, the Fresh Meeting at the University of Southern Denmark, the Spanish Economic Association Meeting in Girona 2015, the 1st CREA workshop in Luxembourg, and at the University of Goettingen. The research of Philipp Ager and Casper Worm Hansen is generously funded by the Danish Research Council reference No. DFF – 4182-00043.

How the 1906 San Francisco Earthquake Shaped Economic Activity in the American West[*]

Philipp Ager   Katherine Eriksson   Casper Worm Hansen   Lars Lønstrup[†]

[This version: March 2019]

Abstract

This paper examines the long-run effects of the 1906 San Francisco Earthquake on the spatial distribution of economic activity in the American West. Using variation in the potential damage intensity of the earthquake, we show that more severely affected cities experienced lower population increases relative to less affected cities until the late 20th century. This long-lasting effect is largely a result of individuals' high geographical mobility at that time. Less affected areas became more attractive migration destinations in the immediate aftermath of the earthquake, which permanently changed the spatial distribution of economic activity in the American West.

Keywords: Economic Geography; Location of Economic Activity; Migration; Natural Disasters, American West

JEL codes: N9; O15; O40; R11; R12

---

[*] The authors would like to thank Alberto Alesina, Sascha Becker, Carl-Johan Dalgaard, Francesco Drago, James Fenske, Walker Hanlon, Marc Klemp, Christopher Meissner, Stelios Michalopoulos, Nathan Nunn, Paul Sharp, Daniel Sturm, Holger Strulik, David Weil, Nico Voigtlaender, and seminar participants at the University of Copenhagen, the Fresh Meeting at the University of Southern Denmark, the Spanish Economic Association Meeting in Girona 2015, the 1st CREA workshop in Luxembourg, and at the University of Goettingen. The research of Philipp Ager and Casper Worm Hansen is generously funded by the Danish Research Council reference No. DFF – 4182-00043.

[†] Contact: Philipp Ager, University of Southern Denmark and CEPR; phag@sam.sdu.dk. Katherine Eriksson, UC Davis and NBER; kaeriksson@ucdavis.edu; Casper Worm Hansen, University of Copenhagen; casper.worm.hansen@econ.ku.dk. Lars Lønstrup, University of Southern Denmark; loe@sam.sdu.dk.

1 Introduction

The location of economic activity is highly skewed. In 2017, the ten largest U.S. metropolitan areas accounted for roughly one-third of America's GDP. As economic activity is clustered within a few core metropolitan areas, a central question in economics is what determines the location of economic activity. One potential answer to why such clusters exist is economies of scale (e.g., Krugman, 1991), implying high degrees of persistence in the distribution of city sizes. Cities that were relatively large historically are predicted to be relatively large today if they were able to take advantage of scale economies. These theories also predict that large enough historical shocks would leave a permanent mark on the spatial distribution of city sizes. Thus, location theories based on scale economies can explain persistence in the placement of the population in the absence of shocks and, at the same time, explain why sufficiently large shocks may imply less persistence over long periods of time.[1]

This paper examines how the 1906 San Francisco Earthquake, a large but temporary shock, affected the long-run distribution of economic activity in the American West. Using newly-collected data on the population size of cities and towns in California, Oregon, and Nevada for every decade from 1890 to 1970, our estimation strategy compares city population sizes before and after the earthquake between more and less affected cities. Estimates from this differences-in-differences strategy reveal that more affected cities experienced a substantial decline in population size relative to less affected cities after the earthquake occurred. Our baseline estimate implies that a one-standard deviation increase in earthquake intensity is associated with a decrease in city population size of around 40 percent some 60 years after the disaster, corresponding to an average decline in annual population growth of 0.5 percentage points over the period 1906 to 1970. The conclusion that the earthquake caused a permanent shift in the relative size of cities is robust to controlling for possible time-dependent effects of a wide range of local geographical characteristics in addition to possible mean-reverting patterns unrelated to the shock. Importantly, the negative long-run effect of the earthquake also remains in place when using alternative measures of economic activity as outcome variables. By showing no sign of recovery more than half a century after the disaster, our empirical analysis supports the idea that the spatial distribution of economic activity observed today may partly be a result of sufficiently large historical shocks.

There are two main reasons for studying this particular event. The first reason is the magnitude of the disaster. With an estimated death toll of 3,000 people, more than 220,000 left homeless, and property damages that would exceed 10.5 billion U.S. dollars today, the 1906 San Francisco Earthquake is one of the worst disasters in the history of the United States (Hansen and Condon, 1989).

---

[1] Other competing explanations are that fundamentals or random growth instead play a crucial role for the spatial concentration of economic activity (e.g., Gabaix, 1999; Davis and Weinstein, 2002).

In light of the theories focusing on agglomeration forces from economies of scale, we consider the magnitude of this disaster as being large enough to alter the relative magnitudes of agglomeration forces from scale economies across cities. The second reason is that major settlements in the American West started only a few decades before. Since the 1906 earthquake happened during a period of mass migration to and within the United States (1850-1920), and that the American West was a popular destination for migrants (Hatton and Williamson, 1998), we argue that our baseline finding is likely driven by the high mobility of the population living in the United States during this time period.

To provide evidence in support of this argument, we study the settlement patterns of migrants arriving in the American West around the time of the earthquake, since these people were arguably less likely to have strong ties to specific places in the region. Our migration analysis is based on newly-digitized complete-count U.S. Census samples for the immediate decades before and after the earthquake (1900-1910). We further create a linked sample of adult males based on this data, which allows us to construct in- and out-migration rates for affected and non-affected areas. Our results show that the more earthquake-affected areas experienced a reduced inflow from internal migrants (born outside the current state of residence) and foreign-born residents relative to less earthquake-affected areas, which is in line with the idea that the earthquake diverted migrants to less affected areas of the American West.

This finding relates to empirical studies on the economic consequences of natural disasters in the United States during the first half of the 20th century, documenting that people moved away from the affected areas as a response to the shock (Boustan, Kahn and Rhode, 2012; Hornbeck, 2012; Hornbeck and Naidu, 2014; and Long and Siu, 2018).[2] Recent work by Boustan, Kahn, Rhode and Yanguas (2017) evaluates how natural disasters affected migration rates and other economic outcomes at the county level between 1920 and 2010. The authors show that counties after being hit by severe disasters experienced falling land prices, greater out-migration rates, and higher poverty rates. Our migration analysis, on the other hand, focuses on whether in-migration rates responded to a natural disaster that happened in an relative attractive area at that time, and it suggests that the 1906 San Francisco Earthquake left a long-lasting mark mainly because it diverted newly arriving migrants to less affected areas of the American West.

We also relate to recent studies that investigate the effect of city fires on U.S. city development. Hornbeck and Keniston (2017) focus on the Great Boston Fire of 1872 and Siodla (2015, 2017) on the 1906 San Francisco Fire. Both studies find beneficial long-run effects on city development from these shocks. While these papers investigate the development of a single city after reconstruction,

---

[2] Hornbeck and Naidu (2014) find for a sample of rural counties in the U.S. South that the Great Mississippi Flood of 1927 did not effect total population possibly due to the in-migration of white labor, yet the flood triggered black out-migration and agricultural modernization.

the present paper evaluates how a shock to the placement of people affected the relative size of cities and towns in a whole region over a relatively long time period.[3] In this respect, our finding speaks also to a broader literature that examines the effect of natural disasters on economic growth which finds, in general, mixed results (e.g., Strobl (2011), Loayza et al. (2012), Cavallo et al. (2013), and Imaizumi et al. (2016)).

Our paper also contributes to a growing literature that highlights the importance of historical events for understanding the present-day locations of economic activity. Bleakley and Lin (2012) and Henderson et al. (2018) show that fundamentals that were historically important at the time when cities were founded leave a permanent mark on the spatial distribution of economic activity due to scale economies. Kline and Moretti (2014), Jedwab and Moradi (2016), Hanlon (2017), and Michaels and Rauch (2018) all obtain similar conclusions by finding persistent effects of historical events on the geographic distribution of economic activity. Schumann (2014) even shows that a short-run policy can have long-lasting consequences for population levels. In contrast, Davis and Weinstein (2002), Brakman et al. (2004), Miguel and Roland (2011) find no long-run effects of war-related events on relative city sizes. Henderson et al. (2018) show that one possible explanation of these mixed findings is variation in the level of transport costs. The analysis presented in this paper suggests that the geographical mobility of affected people may play a similar role as transport costs in determining to what extent the location of economic activity we observe today is a result of historical events.

2 Historical background

On April 18, 1906, the earthquake struck at 5:12 a.m. local time without warning (Zoback, 2006). The total length of the rapture was 477 km (296 miles), and the main epicenter was located about 3 km off the shore of San Francisco. While the estimated moment magnitude places the 1906 San Francisco Earthquake just in the top 20 of the largest earthquakes in the history of the United States, the death toll, economic cost, and the associated fires classify it as one of the worst natural disasters on American soil. The 1906 earthquake is naturally associated with the city of San Francisco since it was almost entirely destroyed and the worst affected area. However, as also pointed out by Ellsworth (1990), the intensity of the earthquake was comparable to that felt in the city of San Francisco in many other places close to the San Andreas Fault line implying that the damage caused by the earthquake was widespread in the American West.

A contemporary U.S. Army relief report recorded 498 deaths in San Francisco, 64 deaths in Santa Rosa, and 102 deaths in and around San Jose (Greely, 1906). Hansen and Condon (1989)

---

[3] Since our empirical analysis is based on a relative comparison of city sizes, our findings are not incompatible with the positive long-run effects found in these studies.

revised these numbers and estimate an overall death toll of around 3,000 individuals. Besides the casualties, more than 225,000 people became homeless and 28,000 buildings were destroyed with estimated economic damage of 10.5 billion U.S. dollars today (Algermissen, 1972). In comparison, there were less than 200 earthquake casualties in California from 1812 to 1901 and the property damages of the next significant earthquake after 1906, the Long Beach Earthquake of 1933, was around 750 million U.S. dollars today (Toppozada and Branum, 2004; Coffman et al., 1982). From the 1930s until today, around 200 people have been killed by earthquakes in California (Toppozada and Branum, 2004). Together, these facts paint a picture of the 1906 earthquake as an unparalleled natural disaster in U.S. history.[4]

Furthermore, the U.S. Geological Survey states that the 1906 earthquake marked the onset of a scientific revolution in earthquake research, meaning that the timing and the location of the earthquake were unanticipated. This fact is expressed by Andrew Lawson, at that time a professor in Geology at the University of California, Berkeley, who wrote in the university newspaper in 1904 *"History and records show that earthquakes in this locality have never been of a violent nature, as so far as I can judge from the nature of recent disturbances and from accounts of past occurrences there is not occasion for alarm at present"* (cited in Zoback, 2006, p. 4).

3 Data and estimation strategy

3.1 Data

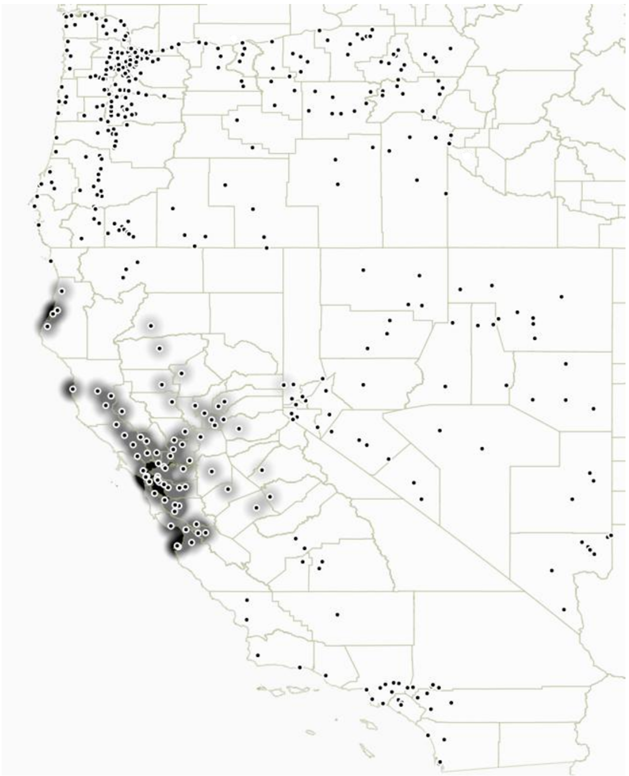
Our measure of the earthquake intensity is based on Boatwright and Bundock's (2005) ShakeMap. The ShakeMap is a smooth measure of the so-called Modified Mercalli Intensity (MMI) of the 1906 earthquake and is deduced from damage reports compiled by Lawson and Reid (1908) and augmented with intensities inferred from additional historical sources. This collection amounts to more than 600 sites with information on the *potential* damage intensity of the earthquake. Boatwright and Bundock deploy the ShakeMap methodology to produce a "heat map" from these sites, where the potential damage intensity ranges from none to very heavy (Wald et al., 1999).[5] We then assign every city an MMI value by overlaying the geographic coordinates (longitude and latitude) for every sampled city in California, Oregon, and Nevada with the ShakeMap using the software QGIS. Cities in the sample that are not depicted on the ShakeMap are assigned the value zero and, therefore, regarded as non-affected control cities. Figure 1 displays the location of the cities in the sample together with the potential damage intensity of the earthquake.

---

[4] The United States Geological Survey (USGS) provides a full list of the earthquakes in the region for the period 1800–1970; see, Coffman et al., 1982).

[5] Using the *potential* and not the actual earthquake damage should mitigate concerns that variation in the earthquake damage variable is endogenous to the size of cities.

Figure 1: Earthquake Intensity and Sampled Cities



NOTES: This map shows the cities in the American West included in the empirical analysis. Darker shaded areas indicate a higher earthquake intensity.

For the main outcome of interest, we compiled a newly digitized data set on the population history of cities and towns in the American West. Our source is Moffat (1986), who collected data on city sizes from publications of the U.S. Census Bureau and, in some instances, from local state and territorial censuses. Our database contains the population size of 746 cities covering the states of California, Oregon, and Nevada for every decade from 1890-1970. We include a city in the sample if it is at least observed in 1900 and 1910; the immediate decades before and after the earthquake occurred (similar results are obtained for a balanced sample of cities). We further use county-level manufacturing data from Haines (2010) as alternative measures of economic activity.[6] Controls for geography (i.e., latitude, longitude) are retrieved from Fishback et al. (2011). Distance to San Francisco, Los Angeles, or to the epicenter (N37.75, W122.55) are calculated using the "geodist" command in STATA.

The migration analysis is based on newly digitized complete-count U.S. Census samples for the immediate decades before and after the earthquake (1900-10). The data are retrieved from IPUMS (Ruggles et al., 2017) and consist of a repeated cross-section of individuals that reported

---

[6] No county-level manufacturing data are available for 1910.

their birthplace (state or country of birth) and place of residence (available at the county level). We use this information to calculate the fraction of the county population born outside the state.

The complete-count U.S. Census data further contain information about individuals' age, and their first and last names. With this information at hand, we can construct a linked sample of adult males (age 18-55), applying the fully automated matching approach pioneered by Ferrie (1996) and further developed by Abramitzky et al. (2012, 2014). We use a conservative matching strategy of this linking technique requiring individuals to be unique by name and state of birth within a five-year age band. This conservative approach reduces the likelihood that individuals are matched to the wrong person with similar attributes by around 50 percent (Abramitzky et al. 2019), which is crucial to do when studying migration decisions. We match all men aged 18 to 55 in 1900 to the Census of 1910, and then restrict to individuals living in one of our three states in 1900 or in 1910. The initial matched sample results in a match rate of about 15 percent, and a sample of over 137,000 men after the state of residence restriction. We then aggregate the data at the county level to construct in- and out-migration rates for 1900-10, the decade in which the earthquake occurred. Summary statistics of the main variables are shown in Online Appendix Table 1.

3.2 Empirical strategy

Our econometric model follows a differences-in-differences approach with a continuous measure of treatment using data at the city or county level for the decades 1890 to 1970. In the following section, we use event-study regressions to evaluate how population size varies by earthquake intensity across locations in California, Oregon, and Nevada before and after the disaster occurred. The baseline estimation equation takes the following form:

$$y_{ct} = \sum_{j=1890}^{1970} \beta_j Quake_c \times I_t^j + \sum_{j=1890}^{1970} \mathbf{X}_c' \mathbf{I}_t^j \Gamma_j + \lambda_c + \lambda_t + \varepsilon_{ct}, \tag{1}$$

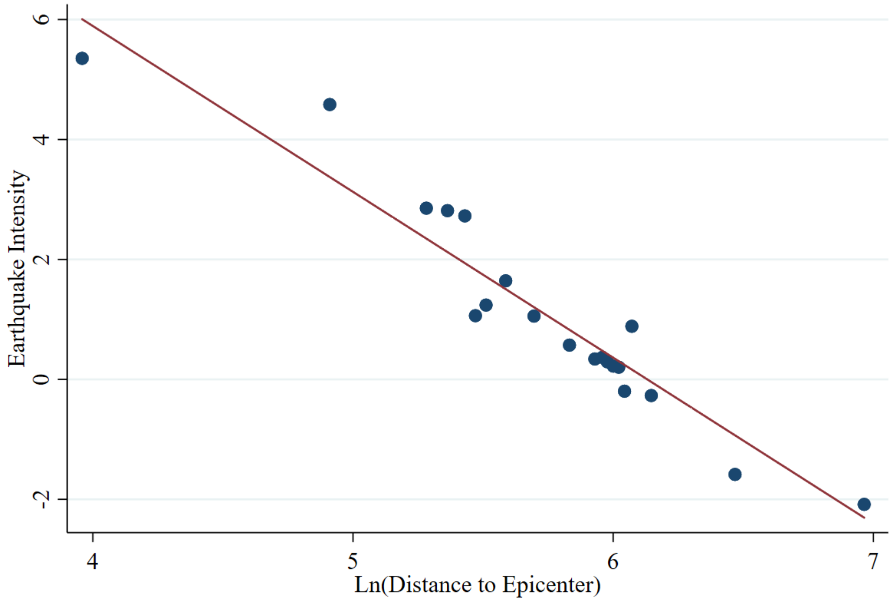
where y_{ct} is *ln* population size in city c in period t; $Quake_c$ is the potential damage intensity in city c interacted with a full set of time-period fixed effects, $\sum I_t^j$, where 1900 is the omitted year of comparison; $\mathbf{X}_c$ includes controls for latitude and longitude measured in logarithmic units interacted by time-period fixed effects, λ_c captures city fixed-effects, λ_t is a time-period fixed effect, and ε_{ct} is the error term. The standard errors are Huber robust and clustered at the city level.

The parameters of interest are the $\beta_j's$, which can be interpreted as the change in *ln* population size from 1900 to year j by earthquake intensity. For example, negative values of $\hat{\beta}_j$ for $j > 1900$ indicate that more affected cities experienced smaller population increases relative to less affected cities after the earthquake occurred.[7] The main identifying assumption is that cities with different

---

[7] By assumption, the Stable Unit Treatment Value Assumption (i.e., potential outcomes for any unit should be

Figure 2: First-stage Relationship between Earthquake Intensity and Distance to Epicenter



Notes: This figure (binscatter) displayes the first-stage relationship between earthquake intensity and logarithm of the distance to the epicenter controlling for longitude and latitude in logartihmic units.

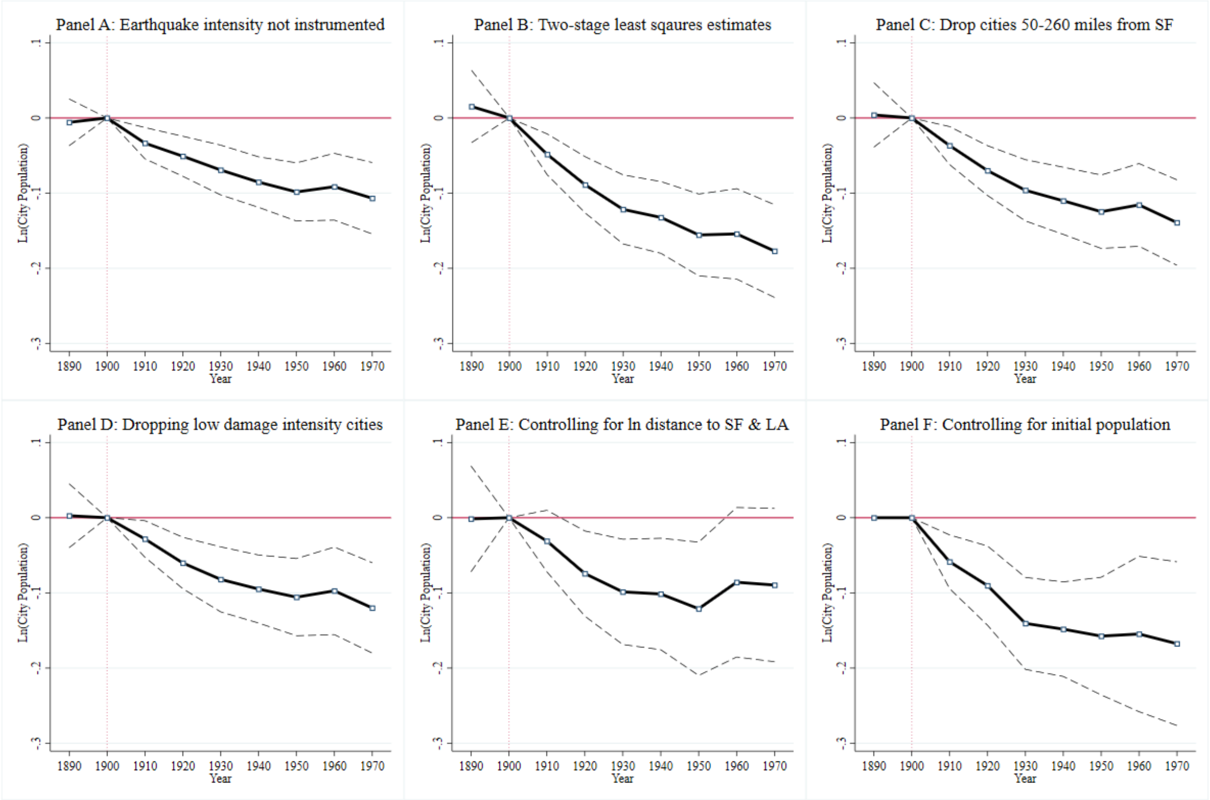
potential damage intensities would have evolved similar to non-affected cities had the earthquake not occurred. In the next section, we provide evidence that changes in the outcome variables across cities (or counties) in the decade before the earthquake are not systematically correlated with damage intensity, suggesting that pre-earthquake trends in observables are indeed parallel, supporting the main identifying assumption.

One concern with estimating equation (1) is that the potential damage intensity is likely to be measured with error and least-squares estimation would yield to attenuated effects (i.e., the parameters of interest, β_j's, would be biased towards zero). We address this issue by instrumenting the potential damage intensity with *ln* distance to the epicenter. It is plausible to assume that measurement error in the potential damage intensity is unrelated to measurement error in the distance to the epicenter. If this is the case, our approach would yield to consistent estimates of the parameters of interest, β_j's. Figure 2 presents a binscatter plot of the relationship between potential damage intensity and *ln* distance to the epicenter controlling for latitude and longitude. The relationship is negative and highly statistically significant.

---

unaffected by treatments assigned to other units) is not satisfied, since the null hypothesis under investigation is that the shock did not have any effect on *relative* population size in the American West. A similar approach is taken by Hanlon (2016).

Figure 3: Event-study of the effect of earthquake intensity on city population


Panel A: Earthquake intensity not instrumented

Panel B: Two-stage least sqaures estimates

Panel C: Drop cities 50-260 miles from SF

Panel D: Dropping low damage intensity cities

Panel E: Controlling for ln distance to SF & LA

Panel F: Controlling for initial population

Notes: Panels A-F show the dynamic effects of earthquake intensity on city population. The solid line depicts the effect on city population relative to the base year (1900). The different panels show modifications of the baseline estimates (Panel A). The method of estimation is least squares in Panel A and two-stages least squares in Panels B-F. Dashed lines indicate 95% confidence intervals.

4 Results

4.1 Effects on city population and manufacturing

Figure 3 presents the event-study estimates for *ln* city population as the outcome variable. All specifications (Panel A-F) include fixed effects for city and time and control for *ln* latitude and *ln* longitude interacted by time-period fixed effects. Panel A presents the baseline results of estimating equation (1) using least squares as the estimation method. This specification shows that more affected cities experienced smaller population increases relative to less affected cities after the 1906 earthquake. Interestingly, the gap in population sizes widens until the 1940s, after which it stabilizes. The event-study displayed in Panel A also reveals that there was no statistically significant relationship between earthquake intensity and population levels before the disaster took place, which is crucial for the identifying assumption described in Section 3.2.

Panel B displays our preferred specification. The potential damage intensity is now instrumented by *ln* distance to the epicenter. While the two-stage least squares coefficients have the same sign, they turn out to be larger (in absolute terms) for any post-earthquake decade, which possibly reflects potential classical measurement error in the earthquake intensity variable. The estimated magnitude of β_{1940} implies that a one standard-deviation increase in earthquake intensity is associated with a decrease in population size by around 40 percent relative to 1900. This effect is increasing slightly in magnitude until 1970, where our sample period ends. In addition to point estimates and standard errors displayed in Panel B, the Kleibergen-Paap F-statistic of instrument strength does not indicate any sign of instrumental weakness (not reported). As in Panel A, this event study does not reveal any signs of potential pre-trends in city population before the disaster occurred. Therefore, we find evidence of a very long-lasting effect on relative population sizes of the earthquake.

Panels C-F present modified specifications of the two-stages least squares approach used in Panel B. The event-study reported in Panel C drops any cities located within 50-260 miles from San Francisco and, in Panel D, we drop cities with low and intermediate damage intensities. Reassuringly, dropping these cities do not affect our results. We next address the concern of future correlated shocks. In particular, the specification in Panel E takes into account that cities in the proximity of San Francisco and Los Angeles might have evolved differently over time irrespective of the potential earthquake damage (e.g., the port of Los Angeles opened its gates in December 1907 and could have attracted population independently of the earthquake). Even when controlling for the *ln* distance to San Francisco and Los Angeles (both interacted by time-period fixed effects), the relative decline in population size of more affected cities after the earthquake remains in place, albeit the precision of the estimated coefficients is somewhat lower. In light of the close proximity of the epicenter to San Francisco (the raw correlation is -0.82), it is not so surprising that these coefficients are less precisely estimated.[8] Estimated coefficients reveal a similar pattern as shown in Panel B when taking potential mean reverting effects into account. Panel F displays this result when *ln* city population 1890 and 1900 (both interacted by time-period fixed effects) are added to estimating equation (1). Reassuringly, we also find similar results when collapsing the city analysis to the county level (available upon request).

Online Appendix Table 2 presents further robustness checks. We include additional geographic controls (access to water, climatic conditions, altitude, and cotton suitability) each interacted by time-period fixed effects, controlling for the diffusion of new knowledge about higher earthquake risk (i.e., a dummy if a city is located in a county on the San Andreas Fault line interacted by time-period fixed effects), or account for changes in the trading costs between any city in our sample and San Francisco. Importantly for the message of this paper, none of these additional control variables

---

[8] Results are very similar to the estimates reported in Panel B when only controlling for the *ln* distance to Los Angeles instead (not reported).

affect our main conclusion. The final specification reported in column (6) includes all the control variables introduced in columns (1)-(5). The negative effects of the earthquake intensity on relative city population growth in this specification are overall somewhat larger. However, we are cautious about interpreting these estimates as being causal effects of the earthquake intensity on city sizes due to the possible endogeneity of the control variables added to the baseline specification. The general conclusion from the robustness analysis is that the estimated effects remain stable when we control for additional factors that possibly influenced city sizes during the sample period.

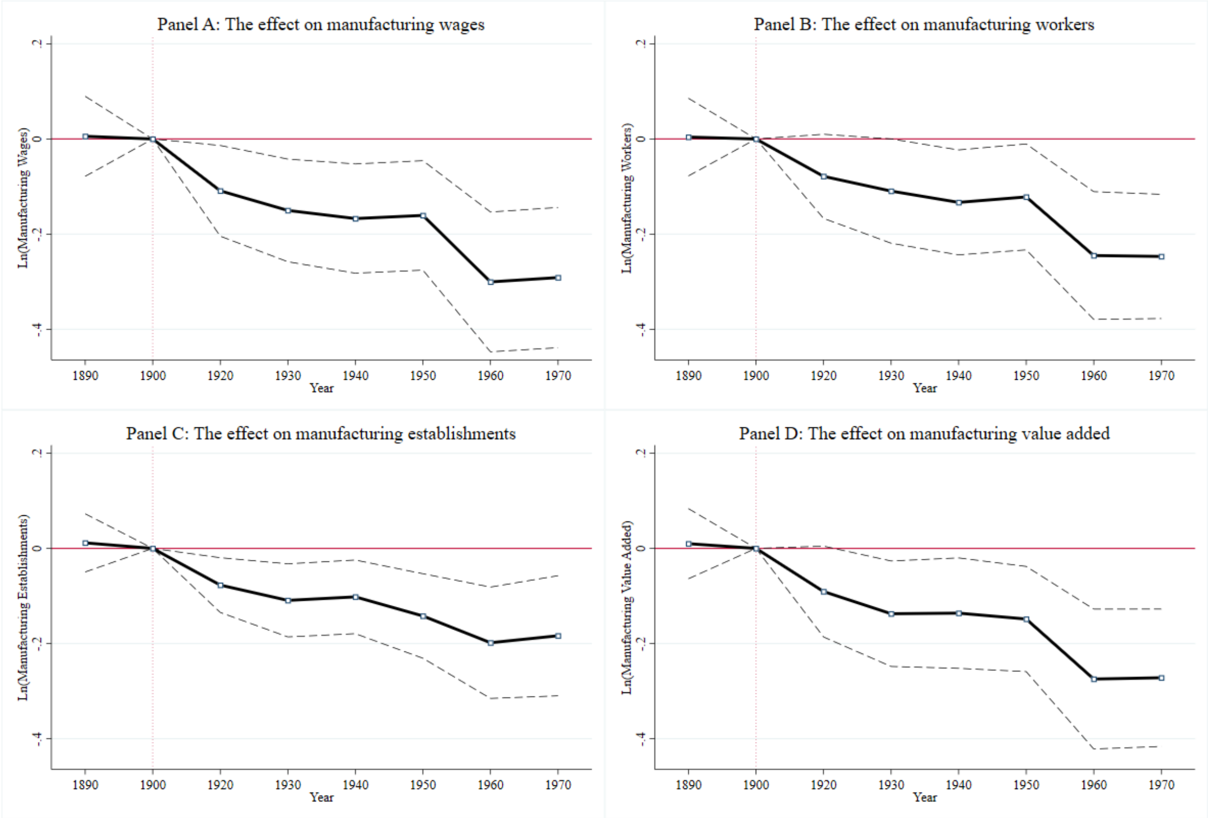
Figure 4 shows how various measures of manufacturing activity evolved in the period 1890 to 1970 by earthquake intensity. The analysis uses estimating equation (1) but is conducted at the county level. All specifications include fixed effects for county and time and control for *ln* latitude and *ln* longitude interacted by time-period fixed effects. The estimates reported in Panels A-D show no signs of differential pre-earthquake changes between more and less affected counties in wages (Panel A), workers (Panel B), establishments (Panel C), and value added (Panel D) in the decade before the disaster (the year 1900 is the omitted year of comparison). Beginning in 1920, which represents the first decade in this sample after the earthquake, the estimates reveal a clear negative association between the potential damage intensity and our measures of manufacturing activity. All four measures suggest that the earthquake left its mark in the manufacturing sector of the affected counties with no sign of recovery even at the end of the sample period. By comparing the effect on wages relative to workers, our estimates indicate small negative productivity effects, as measured by wages per worker. This is also the case for value added per worker. Overall, our empirical results reveal a substantial relative increase in economic activity (measured in terms of city population size and manufacturing) in less and non-affected counties after the earthquake, which is consistent with the migration pattern documented in the next section.

4.2 Why did the 1906 earthquake matter in the long-run?

One possible explanation for the persistent relative decline in economic activity in more earthquake-affected locations is the relative high geographic mobility of the population living in the United States at the beginning of the 20th century. In the following analysis, we evaluate whether migrants were more likely to settle in less affected counties in the immediate aftermath of the earthquake. While the American West was in general a popular destination for many international and internal migrants at that time, one would expect that new arrivals into this area found the affected counties relative less attractive in the aftermath of the disaster.

Table 1 presents the migration analysis, which is based on complete-count Census data and a corresponding linked sample of adult males. The analysis examines the evolution of the stock of people born outside the current state of residence and rates of out-migration and in-state migration

Figure 4: Event-study of the effect of earthquake intensity on the manufacturing sector



Notes: Panels A-D show the dynamic effects of earthquake intensity on manufacturing outcomes. The solid line depicts the effect on city population relative to the base year (1900). The different panels show the baseline estimates for different manufacturing outcomes. Dashed lines indicate 95% confidence intervals.

between 1900 and 1910 at the county level. As in the previous analysis, the potential earthquake damage intensity is instrumented by the *ln* distance to the epicenter. While county fixed effects difference-out in this specification, we still include controls for longitude and latitude to capture other time-varying changes of a county location over the sample period.

Column (1) of Table 1 shows, in line with our conjecture, there was a statistically significant decline in the share of people born outside the state of residence in more affected counties in the immediate aftermath of the earthquake. Column (2) yields similar results when only looking at the change in the share of foreign-born. Quantitatively, a one-standard deviation increase in potential earthquake damage led to a 2 percentage point decline in the share of people born outside the state.

Columns (3)-(4) present the results for the linked sample. The advantage of using a linked sample is that we can test whether parts of the result of column (1) is simply driven by composition bias or whether indeed fewer internal migrants moved into the disaster-struck counties. Column (3) shows that more affected counties experienced a lower rate of in-migration from other U.S. states. The effect is quantitatively sizable: a one-standard deviation increase in potential earthquake dam-

Table 1: Migration Analysis

	(1)	(2)	(3)	(4)
	Dependent Variable			
	Δ Migration Share	Δ Foreign-born Share	Out-of-state In-migration Rate	Out-migration Rate
$Quake\ Intensity_c$	-0.007***	-0.004***	-0.042***	0.004
	(0.002)	(0.001)	(0.016)	(0.004)
Observations	102	102	102	102
R-squared	0.145	0.014	0.155	0.070

NOTES: This table shows the impact of the 1906 earthquake on migration for the period 1900-10. The dependent variable in column 1 is the share of individuals born outside the current state of residence; the share of foreign-born (column 2); the out-of-state in-migration rate (column 3); and the out-migration rate (column 4). Columns 1-2 are based on complete-count Census data while columns 3-4 are based on a linked sample of adult males age 18-55 in 1900. The variable of interest, $Quake\ Intensity_c$ is the potential earthquake damage intensity in county c. All specifications control for *ln* longitude and *ln* latitude. Robust standard errors clustered at the county level in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

age reduced the in-migration rate of out-of-state migrants by 11 percentage points. Interestingly, we do not find any statistically significant impact of the disaster on out-migration rates; the estimated coefficient is nearly zero. Our results suggest that the earthquake diverted new arrivals to other, less affected, areas in the American West, but did not trigger large migratory responses of residents living already in the affected counties.

5 Concluding remarks

The 1906 San Francisco Earthquake fundamentally changed the spatial pattern of economic activity in the American West over the course of the 20th century. More affected areas experienced a persistent decline in population size and manufacturing activities relative to less and non-affected areas. The earthquake ignited a self-enforcing process of agglomeration in less and non-affected cities, which reinforces over time until the economies of scale might be outweighed by congestion costs. This finding is in line with a number of recent empirical studies arguing that historical events may have substantially impacted the location of economic activity due to the presence of economies of scale.

Our finding that a large but temporary shock had a long lasting negative effect on economic activity provides new insights to the literature on U.S. city development. Since individuals can move away or not into disaster-stuck areas as response to the shock, we argue that the effect of the earthquake was long-lasting because it diverted migrants in the immediate aftermath to less affected areas of the American West. As a consequence, the more affected areas experienced relative declines in population size and manufacturing activities. These results support the view that large historical shocks are more likely to have a persistent effect on the location of economic activity when geographical mobility is high as it was the case in the U.S. during the early 20th century. The high geographical mobility of the affected population played therefore an important role in determining to what extent history mattered for the spatial distribution of economic activity in the American West.

References

Abramitzky, Ran, Leah P. Boustan, and Katherine Eriksson, "Europe's tired, poor, huddled masses: Self-selection and economic outcomes in the age of mass migration," *American Economic Review*, 2012, *102* (5), 1832–56.

_ , _ , and _ , "A nation of immigrants: Assimilation and economic outcomes in the age of mass migration," *Journal of Political Economy*, 2014, *122* (3), 467–506.

_ , _ , _ , James Feigenbaum, and Pèrez Santiago, "Automated Linking of Historical Data," Unpublished Manuscript 2019.

Algermissen, Sylvester T., "A study of earthquake losses in the San Francisco Bay Area: Data and analysis," US Department of Commerce, National Oceanic & Atmospheric Administration 1972.

Bleakley, Hoyt and Jeffrey Lin, "Portage and path dependence," *The Quarterly Journal of Economics*, 2012, *127* (2), 587–644.

Boatwright, John and Howard Bundock, "Modified Mercalli intensity maps for the 1906 San Francisco earthquake plotted in ShakeMap format," 2005.

Boustan, Leah P., Matthew E. Kahn, and Paul W. Rhode, "Moving to higher ground: Migration response to natural disasters in the early twentieth century," *American Economic Review*, 2012, *102* (3), 238–44.

_ , _ , _ , and Maria L. Yanguas, "The effect of natural disasters on economic activity in us counties: A century of data," *NBER Working Paper 23410*, 2017.

Brakman, Steven, Harry Garretsen, and Marc Schramm, “The strategic bombing of German cities during World War II and its impact on city growth,” *Journal of Economic Geography*, 2004, *4* (2), 201–218.

Cavallo, Eduardo, Sebastian Galiani, Ilan Noy, and Juan Pantano, “Catastrophic natural disasters and economic growth,” *Review of Economics and Statistics*, 2013, *95* (5), 1549–1561.

Coffman, Jerry L., Carl A. Von Hake, and Carl W. Stover, “Earthquake history of the United States,” US Department of Commerce, National Oceanic and Atmospheric Administration 1982.

Davis, Donald R. and David E. Weinstein, “Bones, bombs, and break points: the geography of economic activity,” *American Economic Review*, 2002, *92* (5), 1269–1289.

Donaldson, Dave and Richard Hornbeck, “Railroads and American economic growth: A “market access” approach,” *The Quarterly Journal of Economics*, 2016, *131* (2), 799–858.

Ellsworth, William L., “The San Andreas Fault System, California,” *US Geol. Surv. Prof. Pap*, 1990, *1515*, 153–185.

Ferrie, Joseph P., “A new sample of males linked from the public use microdata sample of the 1850 US federal census of population to the 1860 US federal census manuscript schedules,” *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 1996, *29* (4), 141–156.

Fishback, Price V., Werner Troesken, Trevor Kollmann, Michael Haines, Paul W. Rhode, and Melissa Thomasson, “Information and the impact of climate and weather on mortality rates during the Great Depression,” in “The Economics of Climate Change: Adaptations Past and Present,” University of Chicago Press, 2011, pp. 131–168.

Gabaix, Xavier, “Zipf’s law for cities: an explanation,” *The Quarterly Journal of Economics*, 1999, *114* (3), 739–767.

Greely, Adolphus W., “Earthquake in California,” *Annual Reports of the War Department*, 1906, *1*, 189–205.

Haines, Michael R., “ICPSR Historical, Demographic, Economic, and Social Data: The United States, 1790-2002,” 2010.

Hanlon, W. Walker, “Temporary Shocks and Persistent Effects in Urban Economies: Evidence from British Cities after the US Civil War,” *Review of Economics and Statistics*, 2017, *99* (1), 67–79.

Hansen, Gladys C. and Emmet Condon, *Denial of disaster*, Cameron and Company San Francisco, 1989.

Hatton, Timothy J. and Jeffrey G. Williamson, *The age of mass migration: Causes and economic impact*, Oxford University Press on Demand, 1998.

Henderson, J. Vernon, Tim Squires, Adam Storeygard, and David Weil, "The global distribution of economic activity: nature, history, and the role of trade," *The Quarterly Journal of Economics*, 2017, *133* (1), 357–406.

Hornbeck, Richard, "The enduring impact of the American Dust Bowl: Short-and long-run adjustments to environmental catastrophe," *American Economic Review*, 2012, *102* (4), 1477–1507.

_ **and Daniel Keniston**, "Creative destruction: Barriers to urban growth and the Great Boston Fire of 1872," *American Economic Review*, 2017, *107* (6), 1365–98.

_ **and Suresh Naidu**, "When the levee breaks: black migration and economic development in the American South," *American Economic Review*, 2014, *104* (3), 963–90.

Imaizumi, Asuka, Kaori Ito, and Tetsuji Okazaki, "Impact of natural disasters on industrial agglomeration: The case of the Great Kantō Earthquake in 1923," *Explorations in Economic History*, 2016, *60*, 52–68.

Jedwab, Remi and Alexander Moradi, "The permanent effects of transportation revolutions in poor countries: evidence from Africa," *Review of economics and statistics*, 2016, *98* (2), 268–284.

Kline, Patrick and Enrico Moretti, "Local economic development, agglomeration economies, and the big push: 100 years of evidence from the Tennessee Valley Authority," *The Quarterly Journal of Economics*, 2014, *129* (1), 275–331.

Krugman, Paul, "Increasing returns and economic geography," *Journal of Political Economy*, 1991, *99* (3), 483–499.

Lawson, Andrew C. and Harry F. Reid, "The California Earthquake of April 18, 1906: Report of the State Earthquake Investigation Commission," Carnegie Institution of Washington 1908.

Loayza, Norman V., Eduardo Olaberria, Jamele Rigolini, and Luc Christiaensen, "Natural disasters and growth: Going beyond the averages," *World Development*, 2012, *40* (7), 1317–1336.

Long, Jason and Henry Siu, "Refugees from Dust and Shrinking Land: Tracking the Dust Bowl Migrants," *The Journal of Economic History*, 2018, *78* (4), 1001–1033.

Michaels, Guy and Ferdinand Rauch, "Resetting the urban network: 117–2012," *The Economic Journal*, 2018, *128* (608), 378–412.

Miguel, Edward and Gerard Roland, "The long-run impact of bombing Vietnam," *Journal of Development Economics*, 2011, *96* (1), 1–15.

Moffat, Riley Moore, *Population History of Western U.S. Cities and Towns, 1850-1990*, Scarecrow Press, 1996.

Ruggles, Steven, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek, "Integrated Public Use Microdata Series: Version 7.0 [dataset]," 2017.

Schumann, Abel, "Persistence of population shocks: Evidence from the occupation of West Germany after World War II," *American Economic Journal: Applied Economics*, 2014, *6* (3), 189–205.

Siodla, James, "Razing San Francisco: The 1906 disaster as a natural experiment in urban redevelopment," *Journal of Urban Economics*, 2015, *89*, 48–61.

_ , "Clean slate: Land-use changes in San Francisco after the 1906 disaster," *Explorations in Economic History*, 2017, *65*, 1–16.

Strobl, Eric, "The economic growth impact of hurricanes: evidence from US coastal counties," *Review of Economics and Statistics*, 2011, *93* (2), 575–589.

Toppozada, Tousson and David Branum, "California earthquake history," *Annals of Geophysics*, 2004, *47* (2-3).

Wald, David J., Vincent Quitoriano, Thomas H. Heaton, Hiroo Kanamori, Craig W. Scrivner, and Bruce C. Worden, "TriNet "ShakeMaps": Rapid generation of peak ground motion and intensity maps for earthquakes in southern California," *Earthquake Spectra*, 1999, *15* (3), 537–555.

Zoback, Mary Lou, "The 1906 earthquake and a century of progress in understanding earthquakes and their hazards," *GSA TODAY*, 2006, *16* (4/5), 4.

Appendix Tables

Appendix Table 1: Summary Statistics

	(1) N	(2) mean	(3) sd
Population	3,266	12,587	93,595
Quake	3,266	1.218	2.634
Latitude	3,266	41.79	3.790
Longitude	3,266	-121.0	2.558
Distance to Los Angeles	3,266	575.5	249.5
Distance to San Francisco	3,266	373.3	175.9
Distance to Epicenter	3,266	377.8	176.5
Δ Migration Share	102	0.052	0.079
Δ Foreign-born Share	102	0.013	0.034
Out-of-state In-migration Rate	102	0.518	0.722
Out-migration Rate	102	0.454	0.082

Appendix Table 2: Additional Robustness Checks Population Analysis

	(1)	(2)	(3)	(4)	(5)	(6)
			Dependent Variable			
			ln(City Population)			
Quake x Year 1890	0.0258	-0.00585	0.0302	0.00978	0.0147	0.00659
	(0.0231)	(0.0301)	(0.0247)	(0.0242)	(0.0258)	(0.0351)
Quake x Year 1910	-0.0508***	-0.0344*	-0.0522***	-0.0449***	-0.0444***	-0.0337
	(0.0137)	(0.0188)	(0.0147)	(0.0139)	(0.0148)	(0.0267)
Quake x Year 1920	-0.0947***	-0.0532**	-0.0955***	-0.0846***	-0.0925***	-0.0879***
	(0.0182)	(0.0229)	(0.0202)	(0.0189)	(0.0201)	(0.0297)
Quake x Year 1930	-0.129***	-0.0951***	-0.129***	-0.115***	-0.118***	-0.135***
	(0.0228)	(0.0288)	(0.0247)	(0.0233)	(0.0245)	(0.0346)
Quake x Year 1940	-0.137***	-0.121***	-0.138***	-0.128***	-0.126***	-0.151***
	(0.0234)	(0.0300)	(0.0256)	(0.0243)	(0.0256)	(0.0358)
Quake x Year 1950	-0.166***	-0.147***	-0.153***	-0.150***	-0.150***	-0.182***
	(0.0268)	(0.0340)	(0.0291)	(0.0277)	(0.0290)	(0.0425)
Quake x Year 1960	-0.165***	-0.137***	-0.159***	-0.147***	-0.149***	-0.190***
	(0.0291)	(0.0377)	(0.0323)	(0.0306)	(0.0320)	(0.0470)
Quake x Year 1970	-0.184***	-0.148***	-0.181***	-0.172***	-0.178***	-0.232***
	(0.0305)	(0.0393)	(0.0334)	(0.0316)	(0.0331)	(0.0528)
Observations	3,231	3,257	3,231	3,266	3,120	3,120
R-squared	0.931	0.930	0.930	0.927	0.928	0.935

NOTES: The observations are reported at the city level for the period 1890-1970 (every decade). The table reports differences in population size by damage intensity (i.e., quake) relative to the base year (1900). The variable of interest, *Quake* is the potential earthquake damage intensity interacted by time-period fixed effects. All regressions include city and year fixed effects, and *ln* latitude and *ln* longitude interacted with time-period fixed effects. The sample is restricted to cities that are, at the least, observed twice before the earthquake and once afterwards. The table adds cross-sectional controls interacted with year fixed effects. Column 1 adds proximity to rivers and a dummy if the county of the city has access to the Pacific Ocean. Column 2 adds the number of droughts and avg. temperature. Column 3 adds cotton suitability and altitude. Column 5 adds a dummy if the county of a city is placed on the San Andreas Fault line. Column 5 adds county trade costs in 1890 from Hornbeck and Donaldson (2016). Column 6 adds all controls from columns 1-6 altogether. Robust standard errors clustered at the county level in parentheses: *** p<0.01, ** p<0.05, * p<0.1.