SDU

Shaking up the Equilibrium:
Natural Disasters, Economic Activity, and Immigration

by

Philipp Ager, Casper Worm Hansen and Lars Lønstrup

Discussion Papers on Business and Economics
No. 2/2018

FURTHER INFORMATION
Department of Business and Economics
Faculty of Business and Social Sciences
University of Southern Denmark
Campusvej 55, DK-5230 Odense M
Denmark

E-mail: lho@sam.sdu.dk / http://www.sdu.dk/ivoe

Shaking up the Equilibrium: Natural Disasters, Economic Activity, and Immigration*

Philipp Ager Casper Worm Hansen Lars Lønstrup†

[This version: January 2018]

Abstract

This paper examines the long-run effects on the spatial distribution of economic activity caused by historical shocks. Using variation in the potential damage intensity of the 1906 San Francisco Earthquake across cities in the American West, we show that more severely affected cities experienced lower population growth relative to less affected cities after the earthquake. This negative effect persisted until the late 20th century. The earthquake diverted migrants to less affected areas in the region, which, together with reinforcing dynamic agglomeration effects from scale economies, left a long-lasting mark on the location of economic activity in the American West.

Keywords: Economic Geography; Location of Economic Activity; Migration; Natural Disasters

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, Peter Sandholt Jensen, Marc Klemp, Christopher Meissner, Stelios Michalopoulos, Nathan Nunn, Pablo Selaya, Paul Sharp, Daniel Sturm, Holger Strulik, David Weil, Nico Voigtländer, 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; phag@sam.sdu.dk. Casper Worm Hansen, University of Copenhagen; casper.worm.hansen@econ.ku.dk. Lars Lønstrup, University of Southern Denmark; loe@sam.sdu.dk.

1 Introduction

What determines the location of economic activity? Two of the most influential theories of economic geography that seek to answer this question can be captured by the terms "first-nature" and "second-nature" effects (Cronon, 1991; Krugman, 1993a).[1] The importance of "first-nature" characteristics for the concentration of economic activity involves the intuitive idea that some locations have better natural endowments than others. The "second-nature" effects comprise the concept of economies of scale, which may arise from the ability to share indivisible facilities and gains from specialization, a thick labor market, and the creation and transmission of skills and ideas (Duranton and Puga, 2004). It seems well established that scale economies are the main reason why cities exist.[2]

Gains from economies of scale are also a central element in Krugman (1991) and the subsequent theoretical literature on New Economic Geography.[3] A central prediction of such theories is a high degree 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 the population. 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.[4]

This paper examines how the 1906 San Francisco earthquake affected the spatial distribution of economic activity in the American West to test the predictions of these theories. By digitizing data on the population size of cities (and towns) in California, Oregon, and Nevada from 1870 to 1970 we can compare the change in city sizes before and after the 1906 earthquake between more and less affected areas.[5] Estimates from this differences-in-differences design reveal that

[1]The unequal distribution of economic activity is best visualized by looking at light intensity from nighttime satellite images (see, e.g., Henderson et al., 2012).

[2]See e.g. Kim (1995) and Rosenthal and Strange (2004) and references therein for empirical studies on the role played by scale economies for agglomeration.

[3]See also Helpman (1998) who focuses on housing as the main mitigating force against spatial concentration.

[4]Krugman (1993b) provides simulations that illustrate how the initial distribution of the manufacturing sector may have persistent effects on city sizes due to increasing returns to scale at the firm level in the manufacturing sector.

[5]We show below that our main conclusions remain unchanged if we extend the sample by including all

cities that were more affected by the earthquake experienced a substantial decline in population growth relative to less affected cities.[6] Our baseline estimate implies that, on average, annual population growth between 1900 and 1970 declined by 0.75 percentage points in the most affected cities compared to non-affected cities.[7] 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.[8] By showing no sign of recovery more than half a century after the earthquake, our empirical analysis supports the idea that the geographic 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 main 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 10.4 billion U.S. dollars today, the 1906 San Francisco Earthquake is one of the worst disasters in the history of the United States (USGS, 2012).[9] 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 other main reason for studying this event is that major settlements in the American West started only a few decades before the shock hit. Based on this observation, we put forward the hypothesis that the effect of the earthquake on relative city sizes is driven by the high mobility of the population living in the American West at the time. To provide evidence in support of this hypothesis, we study the settlement patterns of migrants arriving to the American West around the time of the earthquake, since these people were arguably less likely to have strong ties to specific places in

counties of the American West (see Section 4.3).

[6]For other studies using this type of empirical strategy, see, e.g., Bleakley (2007), Nunn and Qian (2011), Hornbeck (2012), Hornbeck and Naidu (2014), Kline and Moretti (2014), and Ager et al. (2016).

[7]In addition, we provide estimates where we use the distance of a location(city/town or county) to the epicenter of the earthquake interacted with time as an instrumental variable for the earthquake intensity. Reassuringly, our main conclusions remain unaffected, although the instrumental variable estimates are quantitatively larger (in absolute terms) than the least squares estimates, which may reflect classical measurement error in the earthquake intensity variable.

[8]The key identifying assumption is that the population size of all the sampled cities with similar initial conditions would on average have evolved identically had the earthquake not occurred. To support this assumption, we show that there is no association between damage intensity of the earthquake and the population growth rate across cities before the earthquake took place; see Section 4.1.

[9]See http://www.measuringworth.com/ for the transformation to current U.S. dollars.

the region. Consistent with this hypothesis, we find that the more earthquake-affected areas experienced a persistent decline of migration inflows from U.S. natives born outside the current state of residence and foreign-born relative to less earthquake-affected areas.[10] Since the 1906 earthquake happened during a period of mass migration to the United States (1850–1920), and the American West was a popular destination for migrants at that time (Hatton and Williamson, 1998), we consider the change in the settlement patterns of migrants after the shock as a key explanation of the observed change in relative city caused by the earthquake.[11]

In light of the above-mentioned theories, at least two hypotheses holds the potential to explain our findings. The first hypothesis is that "first-nature" characteristics may have changed due to an update in the perception of earthquake risk associated with living in different locations. However, as shown in several tests below, the estimated negative effect of the earthquake on relative population growth rates remains almost unaffected after controlling for such possible effects. In contrast, three pieces of evidence support the hypothesis that forces of agglomeration due to economies of scale comprised the central mechanism by which the earthquake left a persistent mark on the spatial distribution of the population in the American West. Firstly, the estimated effect becomes larger over time until it settles at the end of the sample period, consistent with the idea that forces of agglomeration from scale economies were hampered in the more affected areas and ignited in the less affected areas. Secondly, we show that the earthquake only decreased relative city sizes for cities with more than 2,000 inhabitants. If the only reason for population movements was a change in fundamentals (due to changed perceptions of earthquake risk), the effect of the earthquake damage on population size should be independent of initial city size. On the other hand, the theories of economic geography that are based on scale economies state that small cities are indeed small because of the low degree of scale economies associated with the production processes in these cities. Consistent with our empirical findings, these theories predict that a shock which temporarily pushes people around between small cities/towns, would have insignificant (if any) long-run effects on the population

[10] This finding is in line with 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 et al., 2012; Hornbeck, 2012; Hornbeck and Naidu, 2014).

[11] Between 1910 and 1940, about one million migrants were processed through the port of San Francisco, and about half a million went through the immigration station at Angel Island (Lee and Yung, 2010, p.4). We also show in Section 4.4 that the earthquake had no effect on fertility which further lend credence to the proposed hypothesis of the vital importance of migration for the observed effects of the earthquake.

size of such cities/towns. Finally, we show that the number of manufacturing establishments, various measures of labor productivity in the manufacturing sector, and land prices all declined substantially in the more earthquake-affected areas. These findings are also consistent with the theoretical predictions of models in which the spatial distribution of economic activity is shaped by forces induced by economies of scale in the manufacturing sector (Krugman, 1991; Helpman, 1998).

Our paper contributes to the empirical literature that studies the importance of historical events for understanding the present-day locations of economic activity. Bleakley and Lin (2012) and Henderson et al. (2016) 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. Hanlon (2016), Jedwab and Moradi (2016), Jedwab et al. (2017), and Michaels and Rauch (2018) all obtain similar conclusions by finding persistent effects of historical events on the geographic distribution of economic activity.[12] In contrast, Davis and Weinstein (2002), Brakman et al. (2004), Miguel and Roland (2011), and Sanso-Navarro, Sanz and (2015) find no long-run effects of war-related events on relative city sizes. Henderson, Squires, Storeygard and Weil (2016) show that one possible explanation of these mixed findings is the level of transport costs. If transporting goods are sufficiently costly, first nature effects lock in the settlement of the population and shocks are unlikely to have persistent effects on the spatial distribution of population when trade costs are lower. The analysis presented in this paper suggests that the geographical mobility of affected people may play a similar role as transport cost of goods in determining to what extent the location of economic activity we observe today is a result of historical events.

Our paper also relates to a growing literature that investigates the effect of natural disasters on city development. Hornbeck and Keniston (2017) and Siodla (2015, 2017), focusing respectively on the Great Boston Fire of 1872 and the 1906 earthquake in San Fransisco, find beneficial long-run effects on city development from these shocks. In contrast to these studies which focus on the development of a single city, 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

[12] See also Holmes and Lee (2002) for the relative importance of first and second nature effects in the context of crop choice.

time. We also contribute to a broader literature that examines the effect of natural disasters on economic growth. Earlier studies report positive effects on income (e.g., Albala-Bertrand, 1993; Skidmore and Toya, 2002; Loayza et al., 2012; Noy, 2009), while Cavallo et al. (2013) find that a disaster affects economic growth only if it is followed by political turmoil. Hsiang and Jina (2014) show that tropical cyclones did not have any influence on subsequent rate of growth in GDP per capita. Using Italian regional data, Barone and Mocetti (2014) show that the short-run effects of two large-scale earthquakes in the 1970s and 1980s were negative, while the long-run effects were ambiguous.[13] Dupont and Noy (2015) find a lasting negative impact on population and GDP per capita of the 1995 earthquake in the Kobe region of Japan.

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. 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) 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

[13]Belloc et al. (2016) demonstrate that earthquakes retarded institutional transition from autocratic regimes to self-government in Italian cities between 1000 and 1300. Imaizumi, Ito, and Okazaki (2016) find that the Great Kanto Earthquake in 1923 had long-lasting effects on the spatial distribution of manufacturing activities within the Tokyo Prefecture.

destroyed with an estimated economic damage of 500 million 1906 U.S. dollars (NOAA, 1972; USGS, 2012). 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 40 million 1933 U.S. dollars. From the 1930s until today, around 200 people have been killed by earthquakes in California (USGS, 2014). Together, these facts paint a picture of the 1906 earthquake as an unparalleled disaster in U.S. history.[14]

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 was 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 that "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" (USGS, 2006).

3 Data and Estimation Strategy

3.1 Data

Figure 1 displays our measure of the earthquake intensity, which 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 (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). We then use Boatwright and Bundock's ShakeMap together with the county border files provided by the National Historical Geographic Information System (NHGIS) to obtain the county's average MMI using the software QGIS. The map spans all

[14] A full list of the earthquakes in the region for the period 1800–1950 is provided in the supplementary appendix.

counties of California, Oregon, and Nevada.[15] Each city is assigned the average potential damage intensity of the county it is located in. The location of the towns and cities considered in the analysis is shown in Figure 2.

The main outcome variable of interest is city size, which we obtained by digitizing data from Moffat (1996).[16] Our data set contains the population size of 746 cities measured every decade from 1870 to 1970. We include a city in the sample if it is observed at least one time before the earthquake and at least one time after the earthquake. Additional county-level data to measure manufacturing activities are the number of manufacturing establishments, manufacturing workers, manufacturing wages, and manufacturing value added. These data are retrieved from the ICPSR file 2896 (Haines, 2010). Additional outcome variables are from the Integrated Public Use Microdata Series (IPUMS) database (Ruggles et al., 2015) and include the stock of migrants (defined as people born outside the state of current residence) per capita,[17] and measures of fertility (defined as the child-woman ratio). In these specifications, 1890 is replaced by 1880 since the 1890 Census microdata are unavailable (these data were lost in a fire). The analyses based on these variables cover only the period until 1940 since county identifiers are not available in the IPUMS public use sample for the years 1950 and after. We also constructed a panel data set from 1892 to 1926 of the inflow of immigrants at the county level by using information on individuals' year of arrival to the United States for the census years 1910 to 1930 from IPUMS.[18]

The geographical controls (i.e., latitude, longitude, average temperature, maximum elevation, distance to rivers, and altitude) are retrieved either from Fishback et al. (2011) or are based on own calculations using QGIS. We further obtained cost-effective freight routes between counties and San Francisco in the year 1900 based on the calculations of Donaldson and Hornbeck (2016). Each city is assigned the cost-effective freight route between its county centriod and San Fransisco in 1900. Summary statistics are shown in Table 1.

[15] Counties that are not depicted on the ShakeMap are assigned the value zero and, therefore, regarded as non-affected control counties.

[16] Moffat (1996) draws his information on city sizes primarily from data published by the U.S. Bureau of Census and, in some instances, from local state and territorial censuses.

[17] In particular, this variable measures the stock of in-migrants and immigrants out of the total county population. However, it does not capture in-migration within a given state, since the IPUMS only provides information on the state (or country) of birth.

[18] To construct this variable, we need to assume that immigrants arrived in the county of residence.

Table 2 reports pre-earthquake differences in city population size by earthquake intensity based on estimating:

$$y_{c,t} = \alpha + \gamma_t Quake_c + X_c + \epsilon_{c,t}, \quad (1)$$

where $y_{c,t}$ is the natural logarithm of city population in city c measured in the years 1870, 1880, 1890, and 1900. $Quake_c$ is a continuous measure of the potential earthquake damage intensity caused by the 1906 earthquake, X_c includes geographic controls, and ϵ_c denotes the error term. The estimated coefficient, $\hat{\gamma}_t$, measures the cross-sectional association between the log of population size in year t and the intensity of the earthquake in 1906. The estimates of γ_t are reported in columns 1-3 of Table 2. Consistent with the historical accounts on the development of the American West prior to the 1906 earthquake, the estimates reveal that, on average, larger cities were more affected by the earthquake. Column 2 shows that this is still the case after controlling for longitude and latitude, which are the baseline geographical controls used in our analysis. Column 3 shows that the positive correlation between earthquake intensity and pre-earthquake city size is not solely driven by being close to the cities of San Francisco and Los Angeles. In addition, Table 2 shows that the difference in city size by earthquake intensity is stable over time in the period before the earthquake.

The fact that larger cities were, on average, more affected by the 1906 earthquake may raise two concerns. The first concern is that the variation in the earthquake damage variable is endogenous to the size of cities (e.g., if the structure of larger cities as well as the quality of houses make them more vulnerable to earthquakes). However, these concerns are arguably limited, since the Boatwright and Bundock's ShakeMap reports only the *potential* and not the actual earthquake damage. The second concern is that the earthquake intensity may capture potential effects that the initial city size as well as unobserved effects correlated with initial city size, have on subsequent population growth. To address this concern, we control for the initial value (the value in 1900) of the city size interacted with a full set of time-period fixed effects to capture possible mean-revering dynamics unrelated to the earthquake intensity.[19] Furthermore, we show that the potential damage intensity is uncorrelated with changes in outcome variables

[19] When other outcome variables are used, we also control for the initial values of these outcomes interacted with a full set of time-period fixed effects.

prior to the earthquake.

Figure 3, Panel (a) shows the evolution of (demeaned) population size over time split by averages of non-affected and affected cities in the sample, whereas Panel (b) only considers averages of non-affected and severely affected cities. As these figures make clear, the average population size in the three groups of cities evolve similarly up until 1900 (the last observation before the earthquake), while after 1906 the population growth rates in non-affected cities accelerates relative to the more affected cities.

[Figures 1-3 and Tables 1 & 2 about here]

3.2 Empirical Strategy

We use a differences-in-differences (DiD) estimation strategy with a continuous measure of treatment intensity to capture any differential development in the outcomes of interest between more and less earthquake-affected areas for the decades 1870 to 1970. The empirical analysis uses data at the city or county level. We exploit the differential intensity of the earthquake across places in California, Oregon, and Nevada using a so-called flexible model to evaluate how the outcome variables vary with the earthquake intensity before and after the disaster occurred. The baseline estimation equation takes the following form:

$$y_{ct} = \sum_{j=1890}^{1970} \beta_j \, Quake_c \cdot I_t^j + \sum_{j=1890}^{1970} \mathbf{X}_c' \mathbf{I}_t^j \Gamma_j + \sum_{j=1890}^{1970} \pi_j y_{c,1900} \cdot I_t^j + \delta_c + \phi_t + \varepsilon_{ct}, \tag{2}$$

where y_{ct} is the outcome in city or county c in period t; $Quake_c$ is the potential damage intensity in city or county c which is interacted with a full set of time-period fixed effects, I_t^j (the year 1900 is the omitted period of comparison); $\mathbf{X}_c$ is a set of geographical controls; $y_{c,1900}$ is the pre-shock outcome in 1900; δ_c captures city or county fixed-effects, ϕ_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 or county level.[20]

The parameters of interest are the $\beta_j' s$. They can be interpreted as the change in population growth in a given year associated with a marginal change in the earthquake intensity

[20] This type of clustering allows the residuals to be arbitrarily serially correlated within cities or counties.

among cities with similar geographic characteristics and similar initial population sizes. For example, negative values of $\hat{\beta}_j$ for $j > 1900$ imply that after the earthquake more affected cities experienced lower population growth rates relative to less affected cities. This implies that the estimated coefficients measure the effects of the earthquake in more affected cities relative to less and unaffected cities in the sample.[21]

To interpret the estimates of β_j as respresenting causal effects we need to assume that cities with different potential damage intensities would have evolved parallel with non-affected cities in the region had the earthquake not occurred. Below, we provide evidence that the growth rates of outcome variables across cities (or counties) in the decades before the earthquake are not systematically correlated with damage intensity suggesting that pre-earthquake trends in observables are indeed parallel.

4 Results

4.1 Effects on City Population

The flexible estimates of the effects that the earthquake had on city population growth are reported in Table 3. Column 1 presents the results for all cities in the sample, while column 2 only includes cities where population is recorded at least twice before 1906 and at least once after the event. This distinction makes no difference for the main conclusion of the analysis, which is that more affected cities experienced slower population growth relative to less affected cities after the 1906 earthquake. Crucially for the identifying assumption, there was no statistically significant relationship between earthquake intensity and city population growth before the disaster took place.

Interestingly, the gap in population growth across more and less affected cities widens until the 1940s, after which it stabilizes. This finding is consistent with the hypothesis that the earthquake ignited a self-enforcing process of agglomeration in less and non-affected cities, which reinforces over time until the economies of scale are outweighed by congestion costs. On

[21] Thus, by assumption, the Stable Unit Treatment Value Assumption (i.e., potential outcomes for any unit should be 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).

the other hand, this finding is more difficult to reconcile with the hypothesis that the earthquake solely affected the distribution of the population via a change in the perceived "first-nature" characteristics due to a change in people's preferences for living in cities that were more affected by the earthquake. Already by 1910, it seemed to be well understood that the San Andreas Fault line (which basically passes through the Western coastline) is an important marker for earthquake risk (Zoback, 2006), which implies that cities located along this line would be considered to be roughly subject to the same earthquake risk. However, compared to rational expectations, some individuals may attach too much weight on recent earthquakes when they form expectations about the risk of future earthquakes (Bordalo, Gennaioli, and Shleifer, 2012). This would imply that the effect would be largest directly after the earthquake followed by a decreasing trend over time until it eventually would die out. Since, we find the complete opposite pattern, this hypothesis is also not easily reconciled with the data.

Columns 3 controls for any time-dependent effects on city population growth of being close to the city of San Francisco and Los Angeles.[22] Reassuringly, these controls have minor effects on the estimated relationship between earthquake intensity and the subsequent population growth rate. Column 4 investigates the role played by cities adjacent to the most affected areas and column 5 shows that the effect is mainly driven by cities that were severely affected by the earthquake. Columns 6 and 7 reveal that the estimated effect is driven entirely by cities with a population above 2,000 in 1900. We interpret this as evidence in favor of the proposed hypothesis that effects arising from scale economies were crucial for the long-term consequences that the earthquake had on city population size in the American West. Such theories would predict no effects of shocks in small cities and towns, since these agglomerations are small because the production processes in these places have no or small benefits from scale. Finally, Figure 4 displays partial correlation plots between the earthquake shock and city population growth rates for different time spans. The main purpose of Figure 4 is to illustrate that the findings are not driven by a few outliers (e.g., the cities of San Francisco and Los Angeles).

[Table 3 and Figure 4 about here]

[22] We use 1/(1+distance) as a measure of proximity, to capture the diminishing marginal effects on population size of a city from being located farther away from San Francisco and Los Angeles.

4.2 Robustness Analysis

Table 4 presents robustness checks. The general conclusion from the robustness analysis is that the estimated effects remain stable when we control for additional factors that possibly influenced population growth during the sample period.

Columns 1-3 introduce additional geographic controls. All further controls are interacted with time fixed effects to capture any time-dependent effects of these geographical characteristics. In column 1, we study the role of access to water, both rivers and the Pacific Ocean, which may relate to trade-induced economic activities. In column 2, we control more directly for climatic conditions, which may have time-dependent effects on population growth. In column 3, we further include cities' altitude and cotton suitability to account for the posibility that specialization in cotton production, and its effect on the economy and thereby population growth, is correlated with earthquake intensity.

In column 4, we control for any time-dependent effects that the diffusion of new knowledge about higher earthquake risk associated with living closer to the San Andreas Fault line had on population growth. Since the main results remain unaffected, we find no empirical support for the hypothesis that a changed perception of "first-nature" characteristics, due to a rational update of earthquake risk associated with living in a given location, caused the long-run effects of the 1906 earthquake.

Next, we explore the importance of having economic ties to San Francisco. While we already captured spillover effects on other cities caused by the almost entire destruction of San Francisco by controlling for the proximity to San Francisco, this may be an imperfect measure of the economic connections that any city in our sample has with San Francisco. To address this issue, we control in column 5 for the lowest-cost route between any city in our sample and San Francisco in the year 1900. While this seems to lower the estimated effect of the earthquake on relative city sizes, the qualitative conclusions of the analysis remain unaffected. In column 6 we explore the possibility that measuring the initial conditions of a city only in terms of the size of its population in 1900 may be imperfect due to measurement errors and short-term fluctuations in city size. However, controlling for city size in 1890 and 1900 does not affect the results.

In column 7, we include all the control variables introduced in columns 1-6. The estimates reveal that the negative effects of the earthquake intensity on relative city population growth nearly double. However, we are cautious about interpreting these estimates as being causal effects of the earthquake on city size due to the possible endogeneity of the control variables added to the baseline specification. Importantly for the message of this paper, the qualitative results remain unaffected by these additional robustness checks.

In addition, we have conducted an equivalent analysis at the county level and performed a series of additional robustness checks by controlling for time-dependent effects of other geographical characteristics (number of bays, number of beaches, average number of droughts, a dummy for counties north of San Francisco), accounting for potential spatial dependence between the error terms (Conley standard errors), and including counties located in Colorado, Idaho, Montana, Utah, Washington, and Wyoming as additional control units. The main conclusion is robust to these additional checks (i.e., in all these specifications we find evidence in support of the conclusion that the 1906 earthquake diverted population away from more-affect areas to less-affected areas). These results are available from the authors upon request.

[Table 4 about here]

4.3 Migration

This subsection examines how the 1906 earthquake affected the patterns of migration in the American West. The analysis is based on county level data.[23] Table 5 summarizes the results for the sample period 1860-1940. All specifications control for county and time fixed effects and longitude and latitude interacted with time fixed effects.

Column 1, Panel (a), reports the natural logarithm of the number of immigrants per capita by earthquake intensity in a given year relative to 1900 (omitted). The estimates for the earliest decades indicate that, over the entire pre-earthquake period counties, more affected were on an upward trend in terms of their immigrants per capita relative to less affected counties. However–and crucially for our assumption that more and less affected counties with similar geographical characteristics (longitude and latitude) would have had similar changes in

[23] Migrants are measured as the stock of people born outside the current state of residence. These include both in-migrants (individuals born in another state in the U.S.) and immigrants (individuals born in another country).

immigrants per capita had the earthquake not occurred–we find that the change in immigrants per capita in the period from 1800 and 1900 is, conditional on longitude and latitude, unrelated to earthquake intensity. After the earthquake, the estimates reveal that more-affected counties experienced a decline in the number of immigrants per capita relative to less and non-affected counties. This effect is relatively stable in the decades that followed the earthquake suggesting that the earthquake diverted immigrants away from the more affected areas primarily in the immediate aftermath of the disaster. Column 2 shows that this result is not driven by the specific functional form of the outcome variable. Columns 3 and 4 in Panel (a) decompose the effect presented in column 2 into effects from internal migration and migrants from abroad. While losing statistical significance at the conventional levels, this analysis shows that both foreign born and internal immigrants contributed to the overall effect.

Panel (b) of Table 5 presents the result on the inflow of immigrants. To conduct this analysis, we use the fact that the U.S. Census collected information on the year of arrival of immigrants in the period 1900-1930. We use this information to construct a 5-year non-overlapping panel of the inflow of immigrants from 1891 to 1926.[24] While all the post-earthquake point estimates are negative, they do not become statistically significant until 1921.

[Table 5 about here]

4.4 Fertility

In Table 6, we investigate to what extent the 1906 earthquake affected fertility behavior. In particular, we study the change in the number of children below the ages of one and five per woman of childbearing age by earthquake intensity in a given year relative to 1900. In contrast to Finlay (2009) and Nobles et al. (2014), we do not find evidence of a fertility response to this particular natural disaster.

[Table 6 about here]

[24] We count the total number of immigrants for each of the 5-year periods by place of residence reported in the census.

4.5 Other Measures of Economic Activity

Table 7 presents how various measures of economic activity evolved in the period 1870-1970 by earthquake intensity across counties. As in the previous analyses, all specifications include county and year fixed effects in addition to latitude and longitude interacted by time fixed effects. The estimates for the manufacturing sector reported in columns 1-3 show no signs of differential pre-earthquake changes between more and less affected counties in the number of establishments, wages per worker, and value added per worker in the decade before the disaster (1900 is the omitted year of comparison). Beginning in 1920, which represents the first decade in this sample after the earthquake (there are no county-level data available for the manufacturing sector in 1910), the estimates reveal a clear negative association between the potential damage intensity and the manufacturing outcome variables. While the negative effect of the earthquake on the number of manufacturing establishments increases over time, the negative effect on average wages and value added per worker, both of which are proxies for the average labor productivity in manufacturing, is relatively stable over the entire period (the effect declines somewhat at the end of the period). In column 4, we study the changes in the natural logarithm of farm value by earthquake intensity. This analysis reveals that the change in farm land values was significantly smaller in more affected counties relative to less affected counties after the earthquake. Our empirical results reveal a substantial increase in economic activity and labor productivity after the earthquake in less and non-affected counties, which is consistent with the hypothesis that these counties experienced a relative boom from larger economies of scale in the manufacturing sector due to the effect that the earthquake had on city sizes in the immediate aftermath of the shock. Overall, economies of scale seemed to play a central role for the long-run effect that the 1906 earthquake had on the spatial distribution of economic activity in the American West.

[Table 7 about here]

5 Concluding Remarks

The 1906 San Francisco Earthquake 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 the number of manufacturing establishments and measures of labor productivity relative to less and non-affected areas. 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.

Since individuals might move away from disaster-stuck areas (Boustan et al., 2012; Hornbeck, 2012, Hornbeck and Naidu, 2014), one hypothesis why the 1906 San Francisco Earthquake had a long-lasting effect is that it diverted migrants to less affected areas of the American West. Our findings support this "moving-out hypothesis" and suggest that large historical shocks are more likely to have a persistent effect on the location of economic activity when geographical mobility is high. Indeed, our empirical evidence suggests that the geographical mobility of affected people may play a similar role to that of transport cost of goods (Henderson et al. 2016) in determining to what extent history matters for the spatial distribution of economic activity that we observe today.

References

Ager, P., Brueckner, M., and Herz B., 2016. The boll weevil plague and its effect on the southern agricultural sector, 1889–1929. Explorations in Economic History.

Ager, P., Hansen, C.W., and L. Lønstrup 2016. Church Membership and Social Insurance: Evidence from the Great Mississippi Flood of 1927. Discussion Papers of Business and Economics 7/2016, Department of Business and Economics, University of Southern Denmark.

Ahlfeldt, G., Redding, S., Sturm, D., and Wolf, N., 2015. The Economics of Density: Evidence from the Berlin Wall. Econometrica, 83(6), 2127-2189.

Albala-Bertrand, J.M., 1993. Natural disaster situations and growth: A macroeconomic model for sudden disaster impacts. World Development, 21(9), 1417–1434.

Barone, G., and Mocetti, S., 2014. Natural disasters, growth and institutions: A tale of two earthquakes. Journal of Urban Economics, 84(C), 52–66.

Beeson, P.E., and Troesken, W., 2006. When Bioterrorism Was No Big Deal, NBER Working Paper, 12636.

Belloc, M., Drago, F., and Galbiati, R., 2016. Earthquakes, Religion, and Transition to Self-Government in Italian Cities. The Quarterly Journal of Economics, 131(4), 1875-1926.

Bleakley, H., 2007. Disease and development: evidence from hookworm eradication in the American South. Quarterly Journal of Economics, 122 (1), 73–117.

Bleakley, H., and Lin, J., 2012. Portage and Path Dependence. Quarterly Journal of Economics, 127(2), 587–644.

Boatwright, J., and Bundock, H., 2005. Modified Mercalli Intensity Maps for the 1906 San Francisco Earthquake Plotted in ShakeMap Format. U.S. Geological Survey Open-File Report, 2005-1135.

Boustan L.P., Kahn, M.E., and Rhode, P.W., 2012. Moving to Higher Ground: Migration Response to Natural Disasters in the Early Twentieth Century. American Economic Review: Papers & Proceedings, 102(3), 238–244.

Brakman, S., Garretsen, H., and Schramm, M., 2004. The Strategic Bombing of German Cities During World War II and Its Impact on City Growth. Journal of Economic Geography, 4(2), 201–218.

Cavallo, E., Galiani, S., Noy, I., and Pantano, J., 2013. Catastrophic natural disasters and economic growth. Review of Economic and Statistics 95, 1549–1561.

Cronon, W., 1991. Nature's Metropolis: Chicago and the Great West. New York: W. W. Norton.

Daniels, R., 1990. Coming to America. A History of Immigration and Ethnicity in American Life, 1st ed. New York (NY): Harper Collins.

Daniels, R., 1995. Asian America, Chinese and Japanese in the United States since 1850. University of Washington Press: Seattle and London.

Davis, D., and Weinstein, D., 2002. Bones, Bombs, and Break Points: The Geography of Economic Activity. American Economic Review, 92(5), 1269–1289.

Donaldson, D., and Hornbeck, R., 2016. Railroads and American economic growth: A "market access" approach. The Quarterly Journal of Economics, 131(2), 799-858.

Dupont, W., and Noy, I., 2015. What Happened to Kobe? A Reassessment of the Impact of the 1995 Earthquake in Japan. Economic Development and Cultural Change, 63(4), 777–812.

Duranton, G., and Puga, D. (2004). Micro-foundations of urban agglomeration economies. In Handbook of regional and urban economics, 4, 2063-2117.

Ellsworth, W. L., 1990. The San Andreas Fault System, California. US Geol. Surv. Prof. Pap, 1515, 153-185.

Finlay, J.E., 2009. Fertility Response to Natural Disasters: the Case of Three High Mortality. Earthquakes. World Bank Policy Research Working Paper No. 4883.

Fujita, M., Krugman, P., and Venables, A., 1999. The Spatial Economy: Cities, Regions, and International Trade. Cambridge, MA: MIT Press.

Greely, A.W., 1906. Special Report of Maj. Gen. Adolphus W. Greely, U.S.A., Commanding the Pacific Division, on the Relief Operations Conducted by the Military Authorities of the United States at San Francisco and other Points.

Haines, M., 2010. Historical, Demographic, Economic, and Social Data: The United States, 1790-2002. ICPSR02896-v3. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2010-05-21. http://doi.org/10.3886/ICPSR02896.v3.

Haines, M., Fishback, and P., Rhode, P., 2014. United States Agriculture Data, 1840-2010. ICPSR35206-v1. Ann Arbor, MI: Bibliographic Citation: Inter-university Consortium for Political and Social Research [distributor], 2014-12-22. http://doi.org/10.3886/ICPSR35206.v1.

Hanlon, W.W., 2016. Temporary Shocks and Persistent Effects in Urban Economies: Evidence from British Cities after the U.S. Civil War, Review of Economics and Statistics, (forthcoming).

Hansen, G., and Condon, E., 1989. The untold story and photographs of the San Francisco earthquake and fire of 1906. Denial of Disaster, Cameron and Co., San Francisco.

Hatton, T., and Williamson, J., 1998. The Age of Mass Migration: Causes and Economic Impact. New York, NY: Oxford University Press.

Helpman, E., 1998. The Size of Regions, in: D. Pines, E. Sadka, and I. Zilcha, I. (eds.), Topics in Public Economics: Theoretical and Applied Analysis, Cambridge University Press: Cambridge.

Henderson, J., 1974. The Sizes and Types of Cities. The American Economic Review, 64(4), 640-656.

Henderson, J. V., Squires, T. L., Storeygard, A., and Weil, D. N. (2016). The Global Spatial Distribution of Economic Activity: Nature, History, and the Role of Trade. National Bureau of Economic Research (No. w22145).

Holmes, T. J., and Lee, S., 2012. Economies of density versus natural advantage: Crop choice on the back forty. Review of Economics and Statistics, 94(1), 1-19.

Hornbeck, R., 2012. The Enduring Impact of the American Dust Bowl: Short- and Long-Run Adjustments to Environmental Catastrophe, American Economic Review 2012, 102(4): 1477–1507.

Hornbeck, R., Keniston, D., 2017. Creative Destruction: Barriers to Urban Growth and the Great Boston Fire of 1872. American Economic Review, 107(6) 1365-98.

Hornbeck, R., and Naidu, S., 2014. When the Levee Breaks: Black Migration and Economic Development in the American South. American Economic Review, 104, 963–990.

Hsiang, S.M., and Jina A.S., 2014. The Causal Effect of Environmental Catastrophe on Long-Run Economic Growth: Evidence From 6,700 Cyclones. NBER Working Paper No. 20352.

Imaizumi, A., Ito, K., and Okazaki, T., 2016. Impact of natural disasters on industrial agglomeration: The case of the Great Kantō Earthquake in 1923. Explorations in Economic History, 60, 52-68.

Jedwab, R., and Moradi, A., 2016. The permanent effects of transportation revolutions in poor countries: evidence from Africa. Review of economics and statistics, 98(2), 268-284.

Jedwab, R., Kerby, E. and Moradi, A., 2017. History, Path Dependence and Development: Evidence from Colonial Railways, Settlers and Cities in Kenya. Econ J.

Kline, P., and Moretti, E., 2014. Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority. Quarterly Journal of Economics, 129(1), 275–331.

Kim, S., 1995. Expansion of markets and the geographic distribution of economic activities: the trends in US regional manufacturing structure, 1860–1987. The Quarterly Journal of Economics, 110(4), 881-908.

Krugman, P., 1991. Increasing returns and economic geography. Journal of Political Economy 99(3):484–499.

Krugman, P., 1993a. First nature, second nature, and metropolitan location. Journal of Regional Science 33(2):129–144.

Krugman, P., 1993b. On the number and location of cities. European economic review, 37(2-3), 293-298.

Lange, F., Olmstead, A. L., and Rhode, P. W., 2009. The impact of the boll weevil, 1892–1932. The Journal of Economic History, 69(3), 685-718.

Lawson, A.C., 1908. The California Earthquake of April 18, 1906. Reprinted 1969 by the Carnegie Institution of Washington, D.C.

Lee, E., and Yung, J., 2010. Angel Island: Immigrant Gateway to America. Oxford University Press.

Lin, J., 2012. Geography, History, Economies of Density, and the Location of Cities. Federal Reserve Bank of Philadelphia Business Review, Third Quarter.

Lin, J., 2015. The Puzzling Persistence of Place. Federal Reserve Bank of Philadelphia Business Review, Second Quarter.

Loayza, N., Olaberría, E., Rigolini, J., and Christiansen, L., 2012. Natural Disasters and Growth-going beyond the Averages. World Development, 40(7), 1317–1336.

McKinsey Global Institute, 2011. Urban world: Mapping the economic power of cities.

Michaels, G. and Rauch, F., 2018. Resetting the Urban Network: 117–2012. The Economic Journal, 128, 378–412.

Miguel, E., Roland, G., 2011. The long-run impact of bombing Vietnam. Journal of Development Economics, 96(1), 1–15.

Moffat, R. (1996). Population History of Western U.S. Cities and Towns, 1850-1990, Scarecrow Press; First Edition; First Printing edition (June 18, 1996).

NOAA, 1972. A study of earthquake losses in the San Francisco Bay Area - Data and Analysis, A report prepared for the Office of Emergency Preparedness: U.S. Department of Commerce.

Nobles, J., Frankenberg, E., Thomas, D., 2014. The Effects of Mortality on Fertility: Population Dynamics after a Natural Disaster. NBER Working Paper No. w20448.

Noy, I., 2009. The macroeconomic consequences of disasters. Journal of Development Economics 88, 221–231.

Nunn, N., Qian, N., 2011. The potato's contribution to population and urbanization: Evidence from a historical experiment. Quarterly Journal of Economics, 126, 593–650.

Redding, S.J., Sturm, D., 2008. The Costs of Remoteness: Evidence from German Division and Reunification. American Economic Review, 98(5), 1766–1797.

Redding, S.J., Sturm, D., Wolf, N., 2011. History and Industry Location: Evidence from German Airports. Review of Economics and Statistics, 93(3), 814-831.

Rosenthal, S. S., and Strange, W. C. (2004). Evidence on the nature and sources of agglomeration economies. Handbook of regional and urban economics, 4, 2119-2171.

Ruggles, S., Genadek, K., Goeken, R., Grover, and J. Sobek, M., 2015. Integrated public use microdata series, current population survey: Version 6.0.[machine-readable database]. Minneapolis: University of Minnesota.

Siodla, J., 2015. Razing San Francisco: The 1906 Disaster as a Natural Experiment in Urban Redevelopment. Journal of Urban Economics, 89, 48–61.

Siodla, J., 2017. Clean slate: Land-use changes in San Francisco after the 1906 disaster. Explorations in Economic History, 65, 1-16.

Skidmore, M., and Toya, H., 2002. Do natural disasters promote long-run growth? Economic Inquiry 40, 664–687.

Staiger, D., and Stock, J., 1997. Instrumental Variables Regression with Weak Instruments, Econometrica, 65(3), 557–586.

USGS, 2006. People, Land & Water - 100th Anniversary of the 1906. United States Geological Survey, page URL: http://www.usgs.gov/homepage/science_features/plw_1906.asp.

USGS, 2012. The Great 1906 San Francisco Earthquake. United States Geological Survey, page URL: http://earthquake.usgs.gov/regional/nca/1906/18april/revolution.php.

USGS, 2014. Deaths from U.S. Earthquakes: 1811-2014. United States Geological Survey, Page URL: http://earthquake.usgs.gov/earthquakes/states/us_deaths.php.

Wald, D., Quitoriano, V., Heaton, T., Kanamori, H., Scrivner, C, and Worden, C., 1999. TriNet "ShakeMaps": Rapid Generation of Peak Ground-motion and Intensity Maps for Earthquakes in Southern California. Earthquake Spectra, 15(3), 537–556.

Zoback, M., 2006. The 1906 earthquake and a century of progress in understanding earthquakes and their hazards. GSA Today, 16(4-5), 4–11.

Appendix

List of major earthquakes in the studied region, 1800–1950.

(Date - Place, Magnitude - Fatalities)

1812 12 08 - Southwest of San Bernadino County, California - M 6.9 Fatalities 40

1812 12 21 - West of Ventura, California - M 7.1 Fatalities 1

1836 06 10 - South San Francisco Bay region, California - M 6.5

1838 06 - San Francisco area, California - M 6.8

1857 01 09 - Fort Tejon, California - M 7.9 Fatalities 1

1865 10 08 - Santa Cruz Mountains, California - M 6.5

1868 10 21 - Hayward, California - M 6.8 Fatalities 30

1872 03 26 - Owens Valley, California - M 7.4 Fatalities 27

1873 11 23 - California - Oregon Coast - M 7.3

1890 02 24 - Corralitos, California - M 6.3

1892 02 24 - Imperial Valley, California - M 7.8

1892 04 19 - Vacaville, California - M 6.4 Fatalities 1

1892 04 21 - Winters, California - M 6.4

1897 06 20 - Calaveras fault, California - M 6.3

1898 03 31 - Mare Island, California - M 6.3

1898 04 15 - Mendocino County, California - M 6.8

1899 04 16 - Eureka, California - M 7.0

1899 12 25 - San Jacinto, California - M 6.7 Fatalities 6

1901 03 03 - Parkfield, California - M 6.4

1906 04 18 - San Francisco, California - M 7.8 Fatalities 3000

1910 08 05 - Oregon - M 6.8

1911 07 01 - Calaveras fault, California - M 6.5

1915 10 03 - Pleasant Valley, Nevada - M 7.1

1915 06 23 - Imperial Valley, California - M 6.3 Fatalities 6

1918 04 21 - San Jacinto, California - M 6.8 Fatalities 1

1922 01 31 - Eureka, California - M 7.3

1922 03 10 - Parkfield, California - M 6.1

1923 01 22 - Humbolt County, California - M 7.2

1925 06 29 - Santa Barbara, California - M 6.8 Fatalities 13

1926 06 29 - Santa Barbara, California - M 5.5 Fatalities 1

1926 10 22 - Monterey Bay, California - M 6.1

1927 11 04 - Lompoc, California - M 7.1

1932 06 06 - Eureka, California - M 6.4 Fatalities 1

1932 12 21 - Cedar Mountain, Nevada - M 7.2

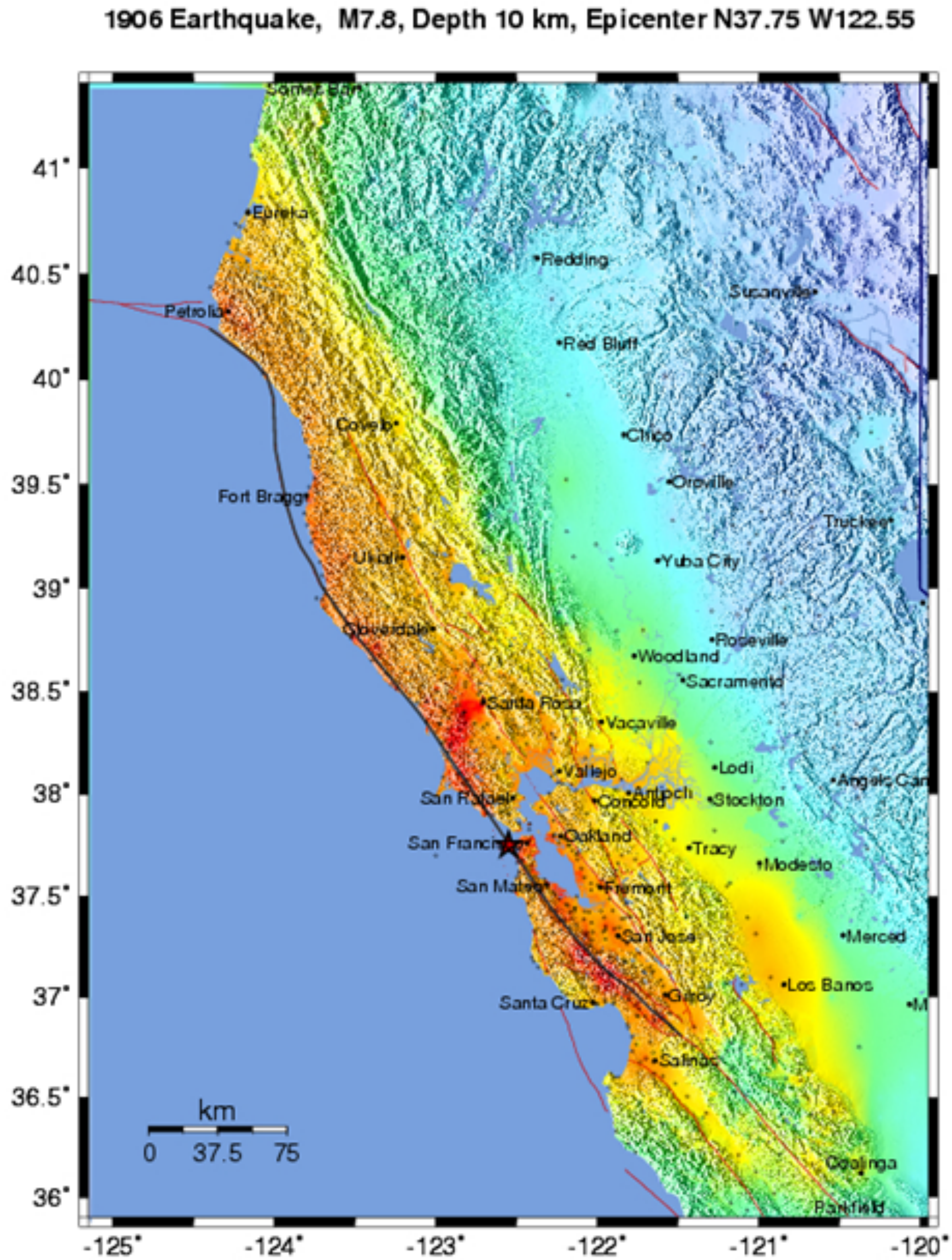
1933 03 11 - Long Beach, California - M 6.4 Fatalities 115

1934 01 30 - Excelsior Mountains, Nevada - M 6.5

1934 06 08 - Parkfield, California - M 6.1

1940 05 19 - Imperial Valley, California - M 7.1 Fatalities 9

Source: http://www.co.monterey.ca.us/Home/ShowDocument?id=44288.



PERCEIVED SHAKING	Not felt	Weak	Light	Moderate	Strong	Very strong	Severe	Violent	Extreme
POTENTIAL DAMAGE	none	none	none	Very light	Light	Moderate	Moderate/Heavy	Heavy	Very Heavy
PEAK ACC.(%g)	<.17	.17-1.4	1.4-3.9	3.9-9.2	9.2-18	18-34	34-65	65-124	>124
PEAK VEL.(cm/s)	<0.1	0.1-1.1	1.1-3.4	3.4-8.1	8.1-16	16-31	31-60	60-116	>116
INSTRUMENTAL INTENSITY	I	II-III	IV	V	VI	VII	VIII	IX	X+

Figure 1: Earthquake Intensity of the 1906 Earthquake

Source: Boatwright and Bundock (2005)

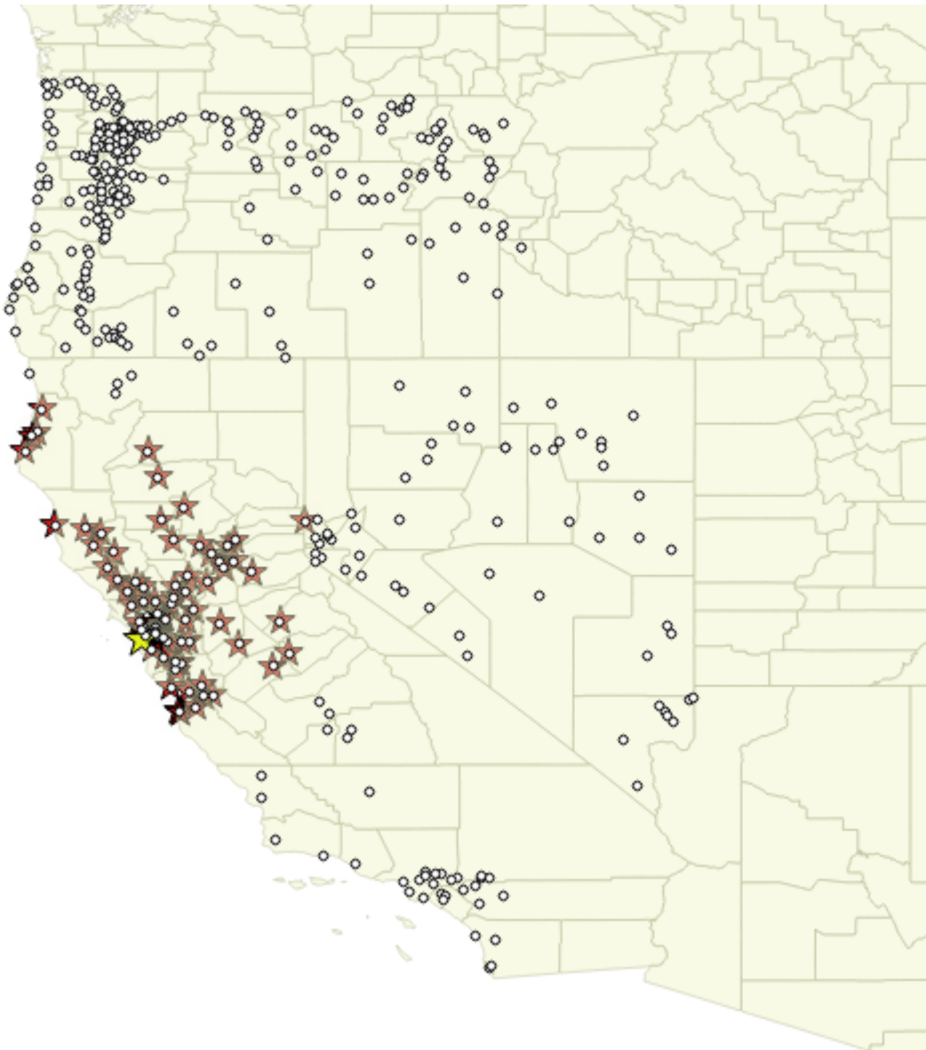
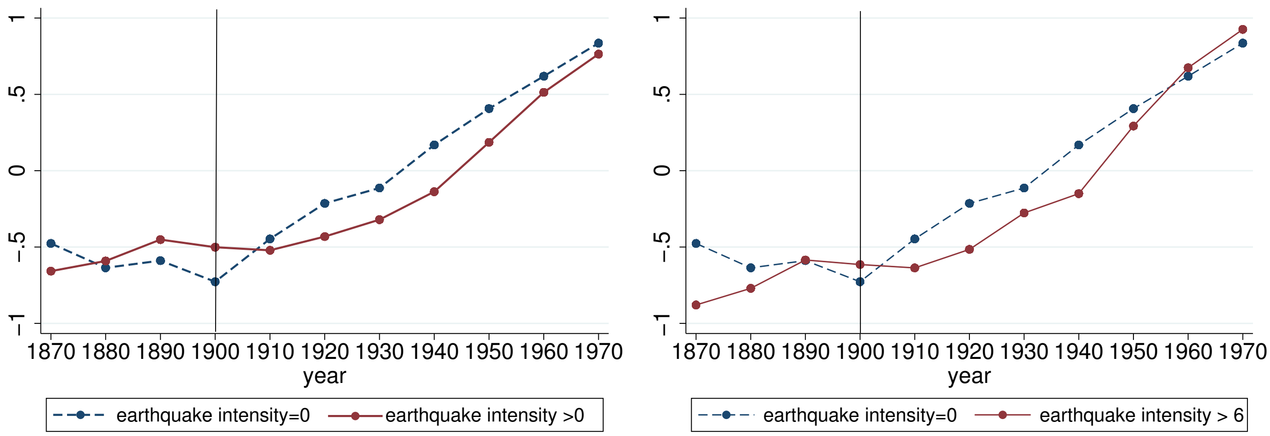


Figure 2: Location of studied towns and cities

Notes: This figure shows a map of the cities and towns considered in the analysis. Affected cities are marked with a star.

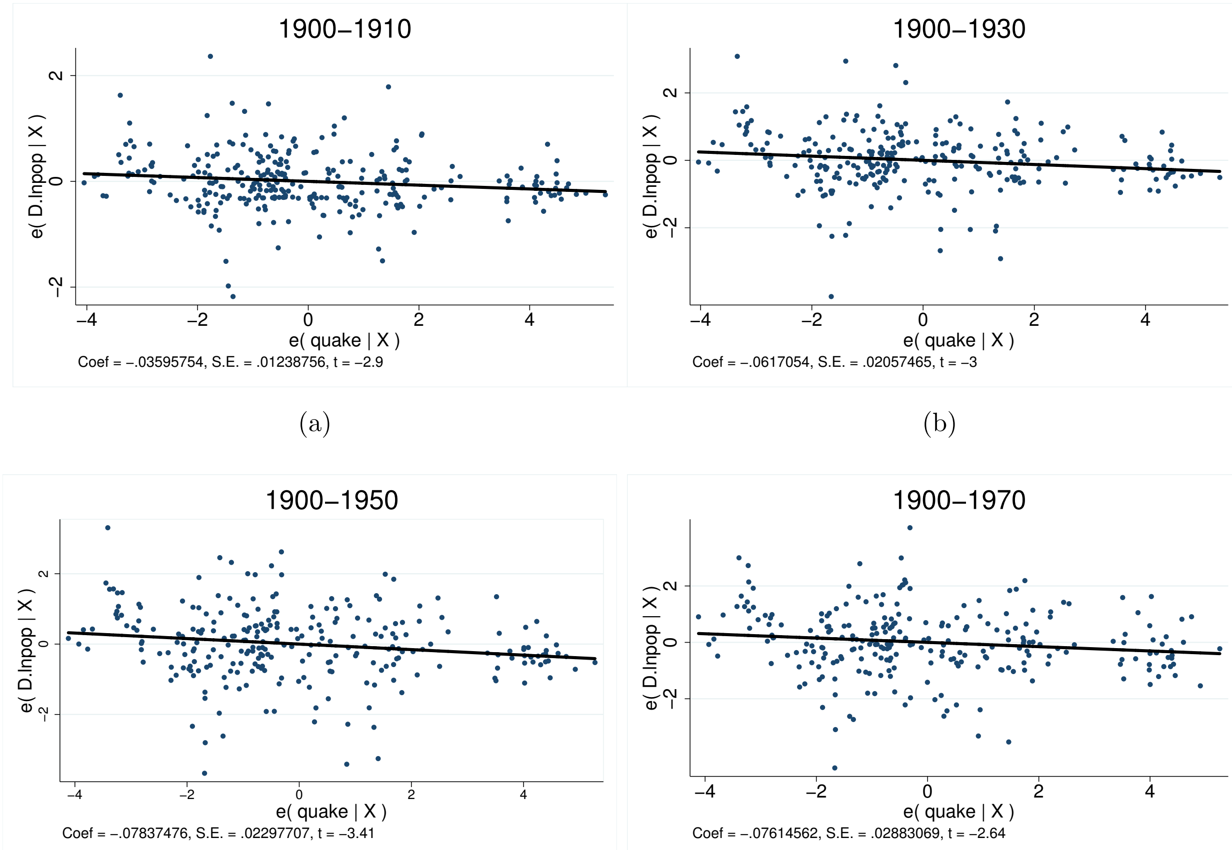


(a)

(b)

Figure 3: Evolution over Time in Affected and Non-affected Cities

Notes: Panel (a) plots for each year the logarithm of the average population size for unaffected and affected cities relative to their respective group averages for the total period. Panel (b) considers only non-affected and severely affected cities (measured by an earthquake intensity above 6).



(a) (b)

(c) (d)

Figure 4: Partial Correlations between Earthquake Intensity and Population Growth Rate across Cities

Notes: The figure shows the partial associations between population growth rates and earthquake intensity across cities (which are observed at least two times before the earthquake and one time afterwards) for different time spans after 1900 when controlling for longitude and latitude. The standard errors (S.E.) of the coefficients are robust and clustered at the city level.

Table 1: Summary Statistics

Variables	(1) N	(2) mean	(3) sd	(4) min	(5) max
Earthquake intensity	746	17	2.9	0	9.4
Latitude	738	39.5	4.3	32.6	46.2
Longitude	738	-120.4	2.5	-124.5	-114.1
City on fault line (no=0, yes=1)	746	0.3	5	0	1
Average population of cities, 1870-1970	746	10,164	44,815	33.8	1,006,000
City population in 1870	120	2,464	13,745	9	149,473
City population in 1880	247	2,112	15,153	28	233,959
City population in 1890	304	2,688	17,915	25	298,997
City population in 1900	426	2,573	18,249	25	342,782
City population in 1910	509	4,050	26,069	25	416,912
City population in 1920	518	5,834	37,220	23	576,673
City population in 1930	554	8,837	62,499	16	1,238,000
City population in 1940	526	10,851	74,714	46	1,504,000
City population in 1950	571	14,234	93,542	19	1,970,000
City population in 1960	626	19,663	110,072	30	2,479,000
City population in 1970	665	24,736	121,433	12	2,812,000

Table 2: Pre-earthquake Differences in Population

Variables	(1) ln(population)	(2) ln(population)	(3) ln(population)
Quake*1870	**0.319***** (0.0508)	**0.185***** (0.0585)	**0.141***** -0.0455
Quake*1880	**0.305***** (0.0338)	**0.201***** (0.0419)	**0.176***** (0.0340)
Quake*1890	**0.292***** (0.0297)	**0.159***** (0.0368)	**0.141***** (0.0312)
Quake*1900	**0.281***** (0.0313)	**0.157***** (0.0390)	**0.142***** (0.0332)
Observations	**912**	**912**	**912**
R-squared	**0,265**	**0,373**	**0,435**

Notes: The analysis is conducted at the city/town level with decadal observations during the pre-earthquake period 1870-1900 in California, Nevada, and Oregon. This table reports the cross-sectional difference in each pre-quake year from 1870 to 1900 by earthquake damage intensity. All regressions include year fixed effects. Column 2 includes the baseline geographical characteristics - longitude and latitude - interacted with year fixed effects, while column 3 also controls for proximity to San Fransisco and Los Angeles interacted with year fixed effects. Robust standard errors of the point estimates are clustered at the city level and reported in parentheses. $***\,p < 0.01$, $**\ p < 0.05$, $*\,p < 0.1$

Table 3: City Size by Earthquake Intensity over Time

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variables	**ln(population)**	**ln(population)**	**ln(population)**	**ln(population)**	**ln(population)**	**ln(population)**	**ln(population)**
Quake*1870	**-0.0924**	**-0.0762**	**-0.119**	**-0.0780**	**-0.119**	**-0.0790**	**-0.326*** **
	(0.0715)	(0.0726)	(0.0741)	(0.0761)	(0.0730)	(0.0738)	(0.0896)
Quake*1880	**-0.0285**	**-0.0101**	**-0.0271**	**-0.0011**	**-0.0309**	**0.0017**	**-0.0888** **
	(0.0386)	(0.0382)	(0.0383)	(0.0402)	(0.0366)	(0.0435)	(0.0363)
Quake*1890	**-0.0064**	**0.0017**	**-0.0009**	**-0.0071**	**-0.0061**	**-0.0465**	**0.0134**
	(0.0161)	(0.0115)	(0.0119)	(0.0136)	(0.0116)	(0.0330)	(0.0188)
Quake*1910	**-0.0270* **	**-0.0360*** **	**-0.0323** **	**-0.0357** **	**-0.0296** **	**-0.0646** **	**-0.0145**
	(0.0144)	(0.0124)	(0.0127)	(0.0143)	(0.0125)	(0.0253)	(0.0223)
Quake*1920	**-0.0530*** **	**-0.0411** **	**-0.0350** **	**-0.0456** **	**-0.0354** **	**-0.0749* **	**0.0088**
	(0.0165)	(0.0164)	(0.0166)	(0.0192)	(0.0160)	(0.0397)	(0.0256)
Quake*1930	**-0.0808*** **	**-0.0643*** **	**-0.0560*** **	**-0.0835*** **	**-0.0546*** **	**-0.0954* **	**-0.0019**
	(0.0181)	(0.0206)	(0.0210)	(0.0222)	(0.0202)	(0.0485)	(0.0386)
Quake*1940	**-0.0887*** **	**-0.0817*** **	**-0.0730*** **	**-0.0920*** **	**-0.0742*** **	**-0.0908** **	**-0.0372**
	(0.0178)	(0.0197)	(0.0201)	(0.0239)	(0.0195)	(0.0433)	(0.0326)
Quake*1950	**-0.0978*** **	**-0.0813*** **	**-0.0708*** **	**-0.0847*** **	**-0.0699*** **	**-0.109** **	**-0.0152**
	(0.0192)	(0.0229)	(0.0235)	(0.0253)	(0.0227)	(0.0513)	(0.0436)
Quake*1960	**-0.0950*** **	**-0.0754*** **	**-0.0615** **	**-0.0576* **	**-0.0587** **	**-0.0948**	**-0.0320**
	(0.0210)	(0.0267)	(0.0277)	(0.0308)	(0.0268)	(0.0591)	(0.0533)
Quake*1970	**-0.112*** **	**-0.0817*** **	**-0.0657** **	**-0.0479**	**-0.0638** **	**-0.0921**	**-0.0402**
	(0.0219)	(0.0287)	(0.0298)	(0.0330)	(0.0295)	(0.0618)	(0.0532)
Observations	**4.972**	**2.753**	**2.753**	**2.192**	**2.544**	**598**	**2.155**
R-squared	**0,908**	**0,898**	**0,902**	**0,911**	**0,899**	**0,903**	**0,834**

Notes: The analysis is conducted at the city/town level with decadal observations during the period 1870-1970 in California, Nevada, and Oregon. In addition to city and year fixed effects, all regressions include the value in 1900 of the outcome variable, latitude, and longitude interacted with year fixed effects. The table reports differences in the logarithm of the population size across cities by the damage intensity of the 1906 San Fransisco Earthquake (quake) in a year relative to 1900. Column 1 include all cities, while columns 2-7 restrict the sample to cities that are observed at least two times before the earthquake and at least one time after the earthquake. In columns 3-7, the additional control variables are i) proximity to the city of San Fransisco interacted with year fixed effects and proximity to the city of Los Angeles interacted with year fixed effects. Column 4 excludes cities that are located between 50 and 260 miles from the city of San Fransisco. Column 5 excludes cities with low damage intensity, specifically with damage intensities in the interval [0,6). Column 6 (7) only considers the sample where the population size of the cities are larger (smaller) than 2000 in the year 1900. Robust standard errors of the point estimates are clustered at the city level and reported in parentheses. $***p < 0.01$, $**\ p < 0.05$, $*p < 0.1$

Table 4: City Size by Earthquake Intensity over Time - Robustness Analysis

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variables	**ln(population)**	**ln(population)**	**ln(population)**	**ln(population)**	**ln(population)**	**ln(population)**	**ln(population)**
Quake*1870	**-0.0898**	**-0.119***	**-0.121**	**-0.0660**	**-0.123**	**0.0135**	**0.0343**
	(0.0632)	(0.0659)	(0.0801)	(0.115)	(0.0847)	(0.0531)	(0.0788)
Quake*1880	**-0.0086**	**-0.0216**	**-0.0469**	**0.0067**	**-0.0446**	**0.0436**	**0.0402**
	(0.0356)	(0.0343)	(0.0439)	(0.0481)	(0.0436)	(0.0312)	(0.0407)
Quake*1890	**0.0027**	**-0.0110**	**0.0041**	**0.0095**	**-0.0029**	**-**	**-**
	(0.0116)	(0.0151)	(0.0136)	(0.0121)	(0.0127)		
Quake*1910	**-0.0387*****	**-0.0331****	**-0.0400*****	**-0.0426*****	**-0.0304****	**-0.0374*****	**-0.0636*****
	(0.0128)	(0.0151)	(0.0138)	(0.0133)	(0.0136)	(0.0118)	(0.0228)
Quake*1920	**-0.0412****	**-0.0283**	**-0.0475****	**-0.0477****	**-0.0359***	**-0.0334****	**-0.0748****
	(0.0168)	(0.0190)	(0.0188)	(0.0188)	(0.0183)	(0.0166)	(0.0295)
Quake*1930	**-0.0636*****	**-0.0543****	**-0.0642****	**-0.0740*****	**-0.0499****	**-0.0596*****	**-0.105*****
	(0.0211)	(0.0232)	(0.0260)	(0.0231)	(0.0236)	(0.0217)	(0.0375)
Quake*1940	**-0.0776*****	**-0.0788*****	**-0.0835*****	**-0.0877*****	**-0.0647*****	**-0.0709*****	**-0.123*****
	(0.0205)	(0.0229)	(0.0235)	(0.0228)	(0.0220)	(0.0206)	(0.0360)
Quake*1950	**-0.0780*****	**-0.0837*****	**-0.0740****	**-0.0916*****	**-0.0631****	**-0.0741*****	**-0.147*****
	(0.0232)	(0.0248)	(0.0289)	(0.0259)	(0.0265)	(0.0245)	(0.0412)
Quake*1960	**-0.0723*****	**-0.0703****	**-0.0692****	**-0.0839*****	**-0.0495**	**-0.0643****	**-0.150*****
	(0.0266)	(0.0280)	(0.0339)	(0.0297)	(0.0307)	(0.0291)	(0.0458)
Quake*1970	**-0.0736****	**-0.0679****	**-0.0711****	**-0.0873*****	**-0.0562***	**-0.0669****	**-0.145*****
	(0.0288)	(0.0302)	(0.0359)	(0.0313)	(0.0329)	(0.0313)	(0.0479)
Observations	**2.723**	**2.743**	**2.723**	**2.753**	**2.646**	**2.753**	**2.646**
R-squared	**0,911**	**0,912**	**0,904**	**0,904**	**0,902**	**0,920**	**0,933**

Notes: The analysis is conducted at the city/town level with decadal observations during the period 1870-1970 in California, Nevada, and Oregon. In addition to city and year fixed effects, all regressions include the value in 1900 of the outcome variable, latitude, and longitude interacted with year fixed effects. The sample is restricted to cities that are observed at least two times before the earthquake and at least one time after the earthquake. The table reports differences in the logarithm of the population size across cities by the damage intensity of the 1906 San Fransisco Earthquake (quake) in each year relative to 1900. Each column adds cross-sectional control variables, which are all interacted with year fixed effects. These control variables are as follows. Column 1: proximity to rivers and a dummy if the county of the city has access to the Pacific Ocean. Column 2: number of draughts and avg. temperature. Column 3: cotton suitability and altitude. Column 4: a dummy if the county of city is placed on the San Andreas Fault line. Column 5: trade cost in 1890 from each city to the city of San Fransisco provided by Donaldson and Hornbeck (2016). Column 6: population size in 1890 and population size in 1900. Column 7: all controls from columns 1-6. Robust standard errors of the point estimates are clustered at the city level and reported in parentheses.
$***p < 0.01$, $**\ p < 0.05$, $*p < 0.1$

Table 5: Immigration Across Counties by Earthquake Intensity

(a) Stock of Immigrants

Variables	(1) ln(Stock of immigrants per capita)	(2) Stock of immigrants per capita	(3) Stock of internal immigrants per capita	(4) Stock of external immigrants per capita
Quake*1860	0.0665**	0.0388**	0.0512***	-0.0124
	(0.0285)	(0.0159)	(0.0108)	(0.0161)
Quake*1870	0.0429***	0.0262***	0.0169	0.0096
	(0.0140)	(0.00747)	(0.0104)	(0.00891)
Quake*1880	0.0063	0.0040	0.0092	-0.0052
	(0.0160)	(0.00865)	(0.00814)	(0.00530)
Quake*1910	-0.0274**	-0.0148***	-0.0082	-0.0066*
	(0.0113)	(0.00553)	(0.00519)	(0.00383)
Quake*1920	-0.0302**	-0.0139**	-0.0090*	-0.0049
	(0.0121)	(0.00595)	(0.00536)	(0.00483)
Quake*1930	-0.0226*	-0.0107*	-0.0046	-0.0061
	(0.0117)	(0.00560)	(0.00549)	(0.00474)
Quake*1940	-0.0238**	-0.0122**	-0.0025	-0.0097**
	(0.00953)	(0.00482)	(0.00570)	(0.00405)
Observations	649	649	649	649
R-squared	0,428	0,512	0,564	0,525

(b) Flow of Immigrants

Variables	(1) ln(Inflow of immigrants)
Quake*1891-1896	-0.0603*
	(0.0334)
Quake*1897-1901	0.0363
	(0.0349)
Quake*1907-1911	-0.0345
	(0.0306)
Quake*1912-1916	-0.0500
	(0.0338)
Quake*1917-1921	-0.0918**
	(0.0361)
Quake*1922-1926	-0.0804**
	(0.0371)
Observations	686
R-squared	0,873

Notes: The analyses are conducted at the county level. In addition to city and year fixed effects, all regressions include the value in 1900 of the outcome variable, latitude, and longitude interacted with year fixed effects. Panel (a) reports differences in the stock of immigrants across counties by the damage intensity of the 1906 San Fransisco Earthquake (quake) in a year relative to year 1900, whereas panel (b) considers the inflow of immigrants during a 5-year period relative to the period 1902-1906. Robust standard errors of the point estimates are clustered at the city level and reported in parentheses.$***p < 0.01$, $**\ p < 0.05$, $*p < 0.1$

Table 6: Fertility Across Counties by Earthquake Intensity

Variables	(1) Children below age 1 per woman	(2) ln(Children below age 1 per woman)	(3) Children below age 5 per woman	(4) ln(Children below age 5 per woman)
Quake*1860	-0.0146	-0.0775	-0.181**	-0.122**
	(0.0306)	(0.0480)	(0.0817)	(0.0521)
Quake*1870	-0.0114	-0.0590	-0.0974*	-0.0816*
	(0.0156)	(0.0485)	(0.0544)	(0.0435)
Quake*1880	-0.0068	-0.0108	-0.0351**	-0.0415*
	(0.00924)	(0.0409)	(0.0142)	(0.0210)
Quake*1910	-0.0039	-0.0315	-0.0139	-0.0319
	(0.00843)	(0.0314)	(0.0190)	(0.0230)
Quake*1920	-0.0063	-0.0211	-0.0115	-0.0245
	(0.00668)	(0.0320)	(0.0140)	(0.0209)
Quake*1930	-0.0019	-0.0120	-0.0080	-0.0138
	(0.00661)	(0.0317)	(0.0134)	(0.0247)
Quake*1940	-0.0010	-0.0082	-0.0002	-0.0155
	(0.00517)	(0.0294)	(0.0106)	(0.0181)
Observations	539	510	539	532
R-squared	0,406	0,586	0,488	0,558

Notes: The analysis is conducted at the county level with decadal observations during the period 1860-1940 in California, Nevada, and Oregon. In addition to city and year fixed effects, all regressions include the value in 1900 of the outcome variable, latitude, and longitude interacted with year fixed effects. The table reports differences in measures of fertility across counties by the damage intensity of the 1906 San Fransisco Earthquake (quake) in a year relative to year 1900. Robust standard errors of the point estimates are clustered at the county level and reported in parentheses. $***p < 0.01$, $**\ p < 0.05$, $*p < 0.1$

Table 7: Economic Activity Across Counties by Earthquake Intensity

	(1)	(2)	(3)	(4)
Variables	**ln(stock of establishments in manufacture)**	**ln(wages per worker in manufacturing)**	**ln(value added per worker in manufacturing)**	**ln(farm value)**
Quake*1870				**0.0979*** (0.0347)**
Quake*1880	**0.0653** (0.0400)	**-0.0366*** (0.0194)	**-0.0562***** (0.0188)	**0.0414** (0.0268)
Quake*1890	**0.0269** (0.0240)	**-0.0016** (0.0108)	**0.0042** (0.0139)	**0.0102** (0.0124)
Quake*1910				**-0.0630***** (0.0203)
Quake*1920	**-0.0639**** (0.0264)	**-0.0264*** (0.0145)	**-0.0241** (0.0208)	**-0.0602***** (0.0212)
Quake*1930	**-0.104***** (0.0323)	**-0.0363***** (0.0133)	**-0.0439*** (0.0224)	**-0.0793***** (0.0291)
Quake*1940	**-0.0889**** (0.0348)	**-0.0361***** (0.0122)	**-0.0233** (0.0192)	**-0.0531** (0.0331)
Quake*1950	**-0.112***** (0.0414)	**-0.0226**** (0.0103)	**-0.0273*** (0.0150)	**-0.0661*** (0.0377)
Quake*1960	**-0.151***** (0.0490)	**-0.0277*** (0.0140)	**-0.0141** (0.0155)	**-0.0288** (0.0342)
Quake*1970	**-0.145***** (0.0499)	**-0.0264**** (0.0130)	**-0.0206** (0.0154)	
Observations	**769**	**721**	**721**	**952**
R-squared	**0.875**	**0.974**	**0.948**	**0.889**

Notes: The analysis is conducted at the county level with decadal observations during the period 1870-1970 in California, Nevada, and Oregon. In addition to city and year fixed effects, all regressions include the value in 1900 of the outcome variable, latitude, and longitude interacted with year fixed effects. The table reports differences in the logarithm of various measures of measures economic by the damage intensity of the 1906 San Fransisco Earthquake (quake) in a year relative to 1900. Robust standard errors of the point estimates are clustered at the county level and reported in parentheses. $***p < 0.01$, $**\ p < 0.05$, $*p < 0.1$