Discussion Papers

Department of Economics

University of Copenhagen

No. 15-17

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

Philipp Ager, Casper Worm Hansen & Lars Lønstrup

Øster Farimagsgade 5, Building 26, DK-1353 Copenhagen K., Denmark

Tel.: +45 35 32 30 01 – Fax: +45 35 32 30 00

http://www.econ.ku.dk

ISSN: 1601-2461 (E)

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

Philipp Ager Casper Worm Hansen Lars Lønstrup†

[This version: October 2015]

Abstract

This paper investigates the effects of a large temporary shock on the agglomeration of economic activity. Using variation in the potential damage intensity of the 1906 San Francisco earthquake across counties in the American West, we find that the earthquake persistently decreased various measures of economic activity, such as population size and total wage expenditures. The main reason for this long-lasting effect is that the earthquake changed the location choice of migrants, who decided to settle in less affected areas of the American West. Our findings suggest that a large temporary shock can have a persistent effect on the location of economic activity.

Keywords: Natural Disasters; Economic Development; Location of Economic Activity, Immigration

JEL codes: N9; O15; O40 R11; R12

*The authors would like to thank Carl-Johan Dalgaard, Peter Sandholt Jensen, Marc Klemp, Stelios Michalopoulos, Pablo Selaya, Nico Voigtländer, and seminar participants at the University of Copenhagen. The research of Philipp Ager and Casper Worm Hansen is 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

The location of economic activity is highly skewed. In 2007, 600 urban centers generated 60 percent of the world GDP; 190 Northern American cities contributed around 20 percent alone (McKinsey Global Institute, 2011, p.1).[1] As economic activity is clustered within a few core locations, a central question in the field of economic geography is whether fundamental differences between locations determine the destiny of their economies, or whether chance can play a role under some circumstances.

The objective of this paper is to empirically investigate how a large negative temporary shock can change the distribution of economic activity within a region and relate the findings to the main theories of economic geography (Krugman, 1991; Fujita et al., 1999; Davis and Weinstein, 2002). In particular, we study the implications of the 1906 San Francisco Earthquake for the distribution of people and production in the American West during the first half of the 20th century. The main motivation for analyzing the implications of this particular shock is that the magnitude of this disaster was so considerable that it had the potential to cause a persistent effect on the geographic distribution of economic development in the region. With an estimated death toll of 3,000 people, more than 220,000 left homeless, and property damages of more than \$500 million in 1906 dollars, the 1906 San Francisco Earthquake was one of the worst natural disasters in the history of the United States (USGS, 2012). The main advantage for our empirical strategy is that this shock was largely unexpected. Before 1906 there was almost no understanding of how and where earthquakes occurred, nor of the risk they implied for settling the American West (USGS, 2006, 2012).

We use the average potential damage intensity of the 1906 earthquake as a continuous measure of "treatment" intensity (Boatwright and Bundock, 2005), which generates variation across affected counties located in the states of California, Oregon, and Nevada. The timing of the shock (1906) is the second source of variation. What is crucial for identification is that the timing of the earthquake makes it possible to rule out that the affected areas were already on different growth paths relative to non-affected areas before the shock occurred. The empirical strategy exploits these two sources of variation to provide differences-in-differences (DiD) es-

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

timates of the shock by comparing economic outcomes before and after the 1906 earthquake between affected and less (or non-) affected areas in the American West.[2]

We find that the earthquake had a negative effect on population size and aggregate economic outcomes of manufacturing and agriculture from 1900 to 1950. This conclusion remains true after we have controlled for potential convergence patterns before the shock, and time-varying effects of geographical characteristics. Our estimates are consistent with earthquake damage being more severe in urban areas, as we find larger effects for the urban population (relative to total population) and the manufacturing sector (relative to agriculture). With no sign of recovery of affected counties almost half a century after the earthquake, our findings show that a temporary shock can have a persistent effect on the spatial distribution of economic activity.

What sets this research apart from the existing literature is the explanation we provide for the reported effects of the earthquake on the spatial distribution of economic development in the region. The earthquake happened during a period of mass migration to the United States (1850–1920), and the American West was a popular destination for immigrants at that time (Hatton and Williamson, 1998).[3] We find that the earthquake affected the settlement decision of immigrants to the American West. Locations in the less affected areas gained a critical mass of people and economic activity due to the effect that the earthquake had on the settlement of immigrants coming to the American West. In particular, more affected counties experienced a significantly lower inflow of migrants after the earthquake. The change in settlement patterns may partly be driven by the hostile treatment of immigrants in the affected areas in the aftermath of the earthquake. As Daniels (1995, p.120) writes about the situation in San Fransisco, "most contemporary observers felt that the earthquake somehow contributed to an atmosphere in which a large number of extralegal attacks on Japanese and other Asians occurred". As an illustrative example of this explanation, our empirical analysis shows a striking boost in the population of Los Angeles that takes off in the years after the earthquake (Figure I).

Our approach relates to the empirical literature that studies how historical events affected

[2] For other studies using this type of DiD strategy, see, e.g., Bleakley (2007), Nunn and Qian (2011), Hornbeck and Naidu (2014), Kline and Moretti (2014), and Ager et al. (2015).

[3] Between 1910 and 1940, about 1 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).

the geographical distribution of economic activity. One example of such an event is Germany's division and reunification after World War II, which is shown to have a persistent effect on the location of economic activity by Redding and Sturm (2008), Redding et al. (2011), and Ahlfeldt et al. (2015). Bleakley and Lin (2012) find persistent effects on population density from historical advantages of certain locations, even though these advantages have vanished today due to technological change. Their approach thus makes a strong case for theories predicting that historical events may leave a long-lasting legacy on the way economic activity is distributed within an area. In contrast, the studies by Davis and Weinstein (2002), Brakman et al. (2004), and Miguel and Roland (2011) find no persistent effect of bombings on relative city size.

In terms of studying effects from the 1906 earthquake on economic activities, the present paper relates to Siodla (2015), who examines how the fires associated with the 1906 earthquake affected the city structure of San Francisco. Along the same lines, Hornbeck and Keniston (2014) demonstrate that the land values of burned plots within Boston increased more relative to unburned plots after the Great Boston Fire of 1872. The results in both papers suggest that a natural disaster could provide an opportunity to reconstruct cities in more favorable ways. Our paper has a different focus, as we analyze how the 1906 earthquake affected the spatial distribution of economic activity in a whole region rather than within a particular city. We also use a different identification strategy, as we exploit time variation due to the earthquake together with cross-sectional variation in the potential damage intensity across counties of the American West.

Finally, our paper relates to the empirical literature on natural disasters and economic growth. Early studies found 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. 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. Our findings suggest that the earthquake in 1906 only affected population movements and aggregate production but not productivity.[4]

[4] Bentzen (2015) exploits different earthquakes at different points in time to show that religiosity increases significantly (across subnational world districts) after such a natural disaster.

2 The Setting

On April 18, 1906, the earthquake struck California at 5:12 a.m. local time without warning (Zoback, 2006). The total length of the rapture was 477 km, and the main shock epicenter was located about 3 km off the shore of the city 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 and economic cost of the earthquake and resulting fires rank it as one of the worst natural disasters in US history. In particular, a contemporary US Army relief report recorded 498 deaths in San Francisco, 64 deaths in Santa Rosa, and 102 deaths in and arround San Jose (Greely, 1906), whereas a later study by Hansen and Condon (1989) revises these numbers and estimates a death toll of around 3,000. More than 225,000 people became homeless and 28,000 buildings were destroyed with an estimated economic damage of $500,000,000 in 1906 dollars (NOAA, 1972; USGS, 2012). In comparison, the total number of earthquake casualties in California from 1812 to 1901 are less than 200 deaths (USGS, 2014). With property damages of an estimated $40 million in 1933 dollars and 115 people killed, the next significant earthquake after 1906 was the Long Beach Earthquake of 1933, whereas from the 1930s to the present about 200 people have been killed by earthquakes in California (USGS, 2014). These numbers paint a picture of the 1906 earthquake as a once-in-a-century disaster in terms of magnitude and destructions.

The US Geological Survey argues that the 1906 earthquake marked the onset of a scientific revolution in earthquake research. For seismologists, earthquake risk seemed to have played a relatively minor role before the disaster in California. 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). In terms of our empirical framework, this statement is particularly attractive since it provides a strong indication that the settlement of the American West in 1906 was not related to earthquake risk.

Finally, anecdotal evidence suggests that the 1906 earthquake in fact had a long-lasting effect on the development of the American West. For example, it has been argued that "though San Francisco would rebuild quickly, the disaster would divert trade, industry and population growth south to Los Angeles, which during the 20th century would become the largest and most important urban area in the West".[5] This argument is supported by comparing the population development in the two cities. Figure I shows that the population growth path of Los Angeles jumped upwards after 1900, and by 1920 Los Angeles was more populous than San Francisco. Since Los Angeles was not affected by the earthquake, whereas 80 percent of San Francisco was destroyed, the two areas are representative of a highly treated area and a non-treated area in our study. Figure I illustrates that the earthquake might have changed the spatial pattern of relative population densities in the American West. Moreover, Figure II displays the development of the manufacturing sector of the two cities between1904 and 1929 with data for every 5th year from the Census of Manufactures. As measured by the number of establishments, workers, wages, and value added, the manufacturing sector in San Francisco was stagnating over this time period compared to the manufacturing sector in Los Angeles.

The same conclusion emerges from the dataset we use to conduct the empirical analysis. In particular, in Figures III and IV we group counties in California, Oregon, and Nevada into affected (treatment) and non-affected (control) areas and show that urban population, total population, and total number of manufacturing establishments decreased in the treatment group relative to the control group after the earthquake. The remainder of this paper is devoted to empirically testing the hypothesis more formally.

[Figures I–IV about here]

3 Data and Identification

3.1 Data Description and Differences by Damage Intensity

This section describes the main variables used in the empirical analysis. Our continuous measure of treatment intensity is based on Boatwright and Bundock's (2005) ShakeMap, which is

[5] Source: http://www.sf-info.org/history/d7/1906-earthquake-and-fire.

displayed in Figure V. Boatwright and Bundock produce a smooth measure of the so-called Modified Mercalli Intensity (MMI) of the San Francisco Earthquake of 1906. The particular ShakeMap is deduced from damage reports compiled by Lawson (1908), augmented with intensities inferred from additional historical sources. This amounts to more than 600 sites with information on observed (potential) damage intensity of the earthquake. Boatwright and Bundock then deploy recent ShakeMap methodology (Wald et al., 1999) to produce the map of intensities from these sites, where the (potential) damage intensity ranges from none to very heavy (see Figure V). Using this ShakeMap and county borders from IPUMS, we obtain the county average MMI with the zonal-statistics function in QGIS. Including all counties in California, Oregon, and Nevada, Figure VI shows the data from this procedure, which we use in the empirical analysis.[6]

The potential damage intensity of the earthquake may be correlated with pre-earthquake population size, as the potential damage is inferred from observed effects. As also argued by Boatwright and Bundock, in some distant sites, such as Sacramento and San Joaquin Valleys, where the earthquake was only "slightly felt", there are not as many damage descriptions compared to sites in the bay area, for example. To address this limitation, we use distance from the centroid of each county to the epicenter of the earthquake (N37.70; W122.50) as an instrumental variable for the potential damage intensity. This instrumental variable is constructed using QGIS. Our empirical analysis also takes into account potential differences in initial (economic) conditions of counties. For example, in the specifications where population is the outcome variable, we control for the county population size immediately before the earthquake interacted with a full set of period fixed effects. In addition, we report various empirical tests demonstrating that the damage intensity is in fact uncorrelated with changes in observable characteristics before the 1906 earthquake.

[Figures V-VII about here]

County-level data on population, manufacturing, and agriculture are drawn from the US census of population, manufacturing, and agriculture, respectively (Haines, 2010; Haines et al.,

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

2014). From these sources, we use the following outcome variables: urban population (measured as the population size for places with a population of 2500 or more), total population, dwellings, manufacturing establishments, manufacturing workers, manufacturing wages, manufacturing value added, farms, average value of farmland per acre, and agricultural value added.[7] Additional outcome variables are constructed using census data from the Integrated Public Use Microdata Series (IPUMS) database (Ruggles et al., 2010). These variables are the number of men/women in the age group 20-40 per capita, the number of immigrants (measured as individuals born outside the state of current residence) per capita, fertility (measured as the child-women ratio), and the number of married individuals per capita.[8] Geographical controls (i.e., latitude, longitude, average temperature, number of severe droughts, number of beaches and bays, and ocean dummy) are retrieved from Fishback et al. (2011).

The samples are balanced panels of 35 to 98 counties (depending on the outcome variable) from 1890 to 1950.[9] When studying the effect on urban population, the sample is restricted to counties classified as urban in 1890. Figure VII displays the spatial distribution of the 35 urban counties by their log urban population size in 1900. Summary statistics of the main variables used in the empirical analysis are shown in Table I.

[Table I about here]

In order to describe baseline differences in initial (pre-earthquake) county characteristics by damage intensity, we estimate the following equation:

$$y_c^{1900} = \alpha + \gamma \ln(1+quake_c) + X_c + \epsilon_c, \tag{1}$$

[7] For the 1910 Census, no manufacturing data are available at the county level.

[8] In these specifications 1890 is replaced by 1880 as the 1890 individual-level census is unavailable (the data were lost in a fire).

[9] While Figures I, III, and IV use data from 1870 to 1950, the empirical analysis starts in 1890 such that we have a balanced panel with a cross-section of reasonable size. Our empirical analysis ends in 1950, as several measures of interest are only available at the county level until 1950. For example, the IPUMS does not provide county identifier in the public use sample from 1950 onwards. The Census also changed the definition of what is considered to be an urban place after 1950. In our sample 14 counties changed their border between 1890 and 1950. These counties are Colusa, Mariposa, San Bernardino, San Diego, and Tulare in California; Esmeralda, Humboldt, and Lincoln in Nevada; and Benton, Crook, Gilliam, Grant, Polk, and Wasco in Oregon. These results are available upon request from the authors.

where y_c^{1900} is the (log) outcome measured six years before the earthquake in 1900, $quake_c$ is the potential damage intensity, X_c includes geographical controls, and ϵ_c is the error term. The estimates of γ reflect pre-earthquake level differences between affected and non-affected counties and are reported in columns 1 and 2 of Table II. Consistent with the historical accounts on the development of the American West prior to the 1906 earthquake, the evidence reveals that the potential damage intensity of the earthquake was higher in areas with more economic mass as measured by population size and other indicators of economic development, such as manufacturing wages or the value of farmland per acres. This is also the case once we add the geographical controls in column 2.

As already mentioned, the identifying assumption of our strategy is that affected and non-affected counties share common pre-trends in the outcome variables. Before reporting more rigorous tests of this, we explore whether there are differences in pre-earthquake changes in the outcomes by damage intensity. The estimating equation is:

$$\Delta y_c^{1890-1900} = \alpha + \gamma \ln(1+quake_c) + X_c + \epsilon_c, \tag{2}$$

where $\Delta y_c^{1890-1900}$ is the change in the (log) outcome prior to the earthquake between 1890 and 1900, and the remaining variables are defined as in equation (1). The estimates, reported in columns 3 and 4 of Table II, show that more affected and less (or non-) affected counties had similar changes in all outcomes but one prior to the shock once we control for the geographical characteristics. In particular, only for agricultural value added do we find a negative estimate which is statistically significant at the 5 percent level.

[Table II about here]

3.2 Identification

We use a differences-in-differences (DiD) setup with a continuous measure of treatment intensity (which is the potential damage intensity) to capture any possible differential development in the outcomes of interest between affected and unaffected counties. This estimation technique exploits the differential intensity of the earthquake across counties in California, Oregon, and Nevada and compares the outcomes before and after the shock. The baseline estimation equation takes the following form:

$$y_{ct} = \sum_{j=1890}^{1950} \beta_j \ln(1 + quake_c) \times I_t^j + \sum_{j=1890}^{1950} \mathbf{X}_c' \mathbf{I}_t^j \Gamma_j + \sum_{j=1890}^{1950} y_{c,1900} I_t^j \pi_j + \delta_c + \phi_t + \varepsilon_{ct}, \tag{3}$$

where y_{ct} is the outcome in county c at period t, $quake_c$ is the potential damage intensity, $\sum I_t^j$ is the full set of time-period fixed effects, where 1900 is the omitted period of comparison, $\mathbf{X}_c$ is a set of geographical controls, $y_{c,1900}$ is the outcome of interest measured in the period prior to the earthquake, i.e. in 1900 (we refer to this variable as the initial outcome), δ_c and ϕ_t are county and time-period fixed effects, and ε_{ct} is the error term. We compute standard errors that are Huber robust and clustered at the county level.

While the specification of the flexible model in equation (3) allows us to investigate the identifying assumption of common pre-trends to "treatment" (i.e., the earthquake) along with the dynamic behavior of the outcome due to the shock, we also estimate a "fixed model" in which the time-period fixed effects, in the first term on the right-hand side of equation (3), are replaced with an indicator variable I_t^{post} which equals one for the periods after 1900 (i.e., $t \geq 1910$). The estimates from this model specification can be used to quantify the average effect of the earthquake over a 50-year period. In order to take into account the possibility that $quake_c$ is measured with error, we also estimate the effect of the earthquake by two-stage least squares (2SLS) using the (log) physical distance from a county's centroid to the epicenter as the excluded instrument. The 2SLS results are reported for the fixed-model specification in the bottom of the regression tables.

4 Results

4.1 Main Estimates

The estimates of the effect that the 1906 San Francisco Earthquake had on population development and housing are shown in Table III. We present results for urban population in columns

1 and 2, total population in columns 3 and 4, and total number of dwellings in columns 5 and 6.[10] The first specification, reported in the odd-numbered columns, includes time-period fixed effects, county fixed effects, and geographical controls (i.e., latitude, longitude, average temperature, number of severe droughts, number of beaches and bays, and an ocean dummy) interacted with time-period fixed effects. In the second specification, reported in the even-numbered columns, we also include the initial value of the outcome variable interacted with time-period fixed effects. Focusing on panel A, we find no differences in how population and housing changed between affected and non-affected counties prior to the 1900s, demonstrating that the non-affected counties constitute a relevant control group since they, at least on observed outcomes, were not on a significantly different growth path of development before the shock. The estimates reveal that from 1900 to 1910 the affected counties experienced a decline in their population and number of dwellings relative to unaffected counties. The remaining estimates in panel A reveal that the short-run decline in population and dwellings persisted (or even increased moderately) at least until 1950. Based on this analysis, we conclude that the 1906 earthquake had a persistent effect on the spatial pattern of population density in the American West.

The estimates from the fixed-model specification are reported in panel B of Table III. They can be interpreted as DiD estimates of the shock using a continuous measure of the intensity of treatment. In line with the flexible estimates, the fixed-model estimates also show that, on average, the disaster influenced both population and dwellings negatively. The estimated coefficients are all statistically significant at least at the 10 percent level. According to the estimate reported in column 2, increasing the potential damage intensity (i.e.,1+ *quake*) by one percent decreases urban population by 0.33 percent. In column 4, a one percent increase in potential damage intensity is associated with a decrease in total population of 0.21 percent. Thus, perhaps unsurprisingly, urban counties responded significantly stronger to the earthquake because the consequential damages are likely to be more severe in urban areas. For dwellings, a one percent increase in potential damage intensity reduces the housing stock by 0.18 percent (column 6). Panel C reports the corresponding 2SLS estimates using log physical distance to

[10] When considering urban population, the sample is restricted to 35 counties as we require that a county was urban before the earthquake to be included in our sample.

the epicenter of the earthquake as an instrumental variable for the potential damage intensity. In terms of instrument quality, our instrumental variable strategy yields a strong first-stage fit. The Kleibergen Paap F-statistic, reported in the bottom of Table III, is 24 or above, and hence in excess of the rule-of-thumb criterion provided in Staiger and Stock (1997) for instruments to be declared weak. In general, we obtain the same conclusions as from the least squares estimates in panel B, although the numerical values of the 2SLS estimates are generally larger in absolute terms. Attenuation bias or potential underreporting in rural areas where the damages from the earthquake were less severe could yield to downward-biased least squares estimates. Finally, Figure VIII displays the partial correlation plot between the earthquake shock and urban population by estimating a long-difference model using only data from the initial period, 1900, and the end year, 1950. This plot illustrates that our finding is not driven only by differences between San Francisco and Los Angeles, nor any other potential outliers.

[Table III and Figure VIII about here]

Tables IV and V provide estimates of the effects that the 1906 earthquake had on outcomes in the manufacturing and agricultural sectors. The presented results follow the same outline as Table III. In line with the previous findings, the estimates for the manufacturing sector reveal no differential pre-trend across affected and non-affected counties. This is the case for the number of establishments, workers, total wage expenditures, and total value added (panel A of Table IV).[11] By 1920, however, there is a clear negative association between the potential damage intensity and the outcome variables, and by 1930 the effects become even stronger in magnitude. The remaining estimates in panel A reveal that the declines from 1900 to 1930 generally persist through 1950. The fixed-model setup of panel B, shows a significant negative effect of the shock for all the considered outcome variables. The estimated impact on manufacturing establishments is -0.38 (column 2), for manufacturing workers it is -0.4 (column 4), for manufacturing wages it is -0.71 (column 6), and for manufacturing value added it is -0.55 (column 8). Following the same instrumental variable strategy as in panel C of Table III, we find that the estimated coefficients are all negative and statistically significant and again larger

[11] The year 1920 is the first time period we observe after the shock, as county-level data from manufacturing in 1910 are not available.

in absolute terms. Moreover, these findings suggest that productivity, measured as wage per manufacturing worker, is largely unaffected, which is confirmed by regressing the log of wages per worker on our shock variable (not reported).

If the effect of the 1906 earthquake was larger in urban areas, as suggested by the estimates presented in Table III, it is likely that the relative effect in the manufacturing sector, which is typically located in urban areas, would be significantly larger than in the agricultural sector. Table V considers number of farms, farmland prices, and total agricultural value added as outcome variables. Apart from agricultural value added, reported in the last two columns, there is no sign of significant pre-trends between affected and non-affected counties. Thus, only after 1900 do we find statistically significant negative effects on land prices and value added from the earthquake. Comparing the DiD estimates of panel B to the manufacturing results reveals that the agricultural sector was less affected by the earthquake. For example, column 6 reports that a one percent increase in damage intensity is associated with a 0.42 percent decrease in value added, while the same estimate for the manufacturing sector is -0.67 (column 8 of Table IV).[12] Finally, the instrumental variable analysis, reported in panel C, raises doubt whether the earthquake had any effect on the total number of farms, as the point estimate drops from 0.11 (standard error = 0.06) to 0.03 (standard error = 0.08).

[Tables IV and V about here]

4.2 Demographic effects

Table VI presents results on how the 1906 earthquake affected the demographics of the American West. The data for this analysis are obtained from IPUMS, implying that the years 1890 and 1950 are not included. Thus, in order to check the assumptions of common pre-trends of outcome variables in affected and non-affected counties, we extend the dataset to include 1880. Table VI reports the results for men (women) in the age group 20-40 per capita, immigrants per capita,[13] fertility measured by the log child-women ratio, and marriages per capita. The

[12] Given the positive and statistically significant coefficient in the pre-treatment period for agricultural value added, the computed DiD estimate, reported in panels B and C, over-estimates the effect of the shock.

[13] We measure immigrants by counting the number of people not born in the current state of residence.

reported estimates are conditional on county and time-period fixed effects, a set of geographical controls, and initial-outcome measures both interacted by time. The flexible estimates in panel A of Table VI largely confirm the validity of the identifying assumption. Panels B and C show that the number of immigrants per capita decreased as a consequence of the earthquake (column 3), suggesting that the total number of immigrants decreased disproportionately more than the total population according to the significant negative effects documented in Section 3 on total population size. The estimates also show *no* robust effects on the remaining demographic variables (the number of men (or women) aged 20 to 40, fertility, or marriages per capita.).[14] Thus, unlike, e.g., Finlay (2009) and Nobles et al. (2014), we find no evidence suggesting a fertility response to a natural disaster. In sum, we conclude that effects on population is driven by the earthquake diverting migrants to the less and non-affected areas of the American West.

[Table VI about here]

5 Concluding remarks

This paper establishes that the 1906 San Francisco Earthquake changed the spatial pattern of relative economic densities in the American West during the first half of the 20th century. We show evidence which supports the hypothesis that the shock had a long-lasting effect by diverting the settlement of migrants to less affected areas. Our results suggest that people living in areas affected by a natural disaster are more likely to stay there and rebuild what was lost compared to migrants, who have fewer, if any, ties to specific locations. These findings are consistent with multiple steady-state distributions of economic activity.

[14] We use these particular age groups since the individuals of these ages have the highest degree of mobility and are thus ex-ante more likely to migrate due to the shock.

References

Ager, P., Hansen, C., Lønstrup, L., 2015. Church Membership and Social Insurance: Evidence from the Great Mississippi Flood of 1927. Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2530868.

Ahlfeldt, G., Redding, S., Sturm, D., Wolf, N., 2015. The Economics of Density: Evidence from the Berlin Wall. Econometrica, (forthcoming).

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

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

Bentzen, J., 2015. Acts of God: Religiosity and Natural Disasters Across Subnational World Districts. Working paper.

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

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

Boatwright, J. 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.

Brakman, S., Garretsen, H., 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., Pantano, J., 2013. Catastrophic natural disasters and economic growth. Review of Economic and Statistics 95, 1549–1561.

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

Davis, D., Weinstein, D., 2002. Bones, Bombs, and Break Points: The Geography of Economic Activity. American Economic Review, 92(5), 1269–89.

Donaldson, D., Hornbeck, R., 2015. Railroads and American Economic Growth: A 'Market Access' Approach. Quarterly Journal of Economics, (forthcoming).

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., 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, 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.

Hansen, G., 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.

Henderson, V., Storeygard A., Weil, D., 2012. Measuring Economic Growth From Outer Space. American Economic Review, 102(2): 994–1028.

Hornbeck, R., Keniston, D., 2014. Create Destruction: Barriers to Urban Growth and the Great Boston Fire of 1872. Working paper.

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

Kline, P., 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.

Krugman, P., 1991. History Versus Expectations. Quarterly Journal of Economics, 106(2), 651–667.

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

Lee, E., 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., 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., Rauch F., 2013. Resetting the Urban Network: 117–2012. Centre for Economic Policy Research Discussion Paper, 9760.

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

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, 220.

Nobles, J., Frankenberg, E., and 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.

Ruggles, S., Alexander, T., Flood, S., Genadek, G., Schroeder, M. B., Trampe, B., Vick, R., King, M., 2010. Integrated public use microdata series, current population survey: Version 3.0.[machine-readable database].

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

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

Staiger, D., Stock, J., 1997. Instrumental Variables Regression with Weak Instruments, Econometrica, 65(3), pp. 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.

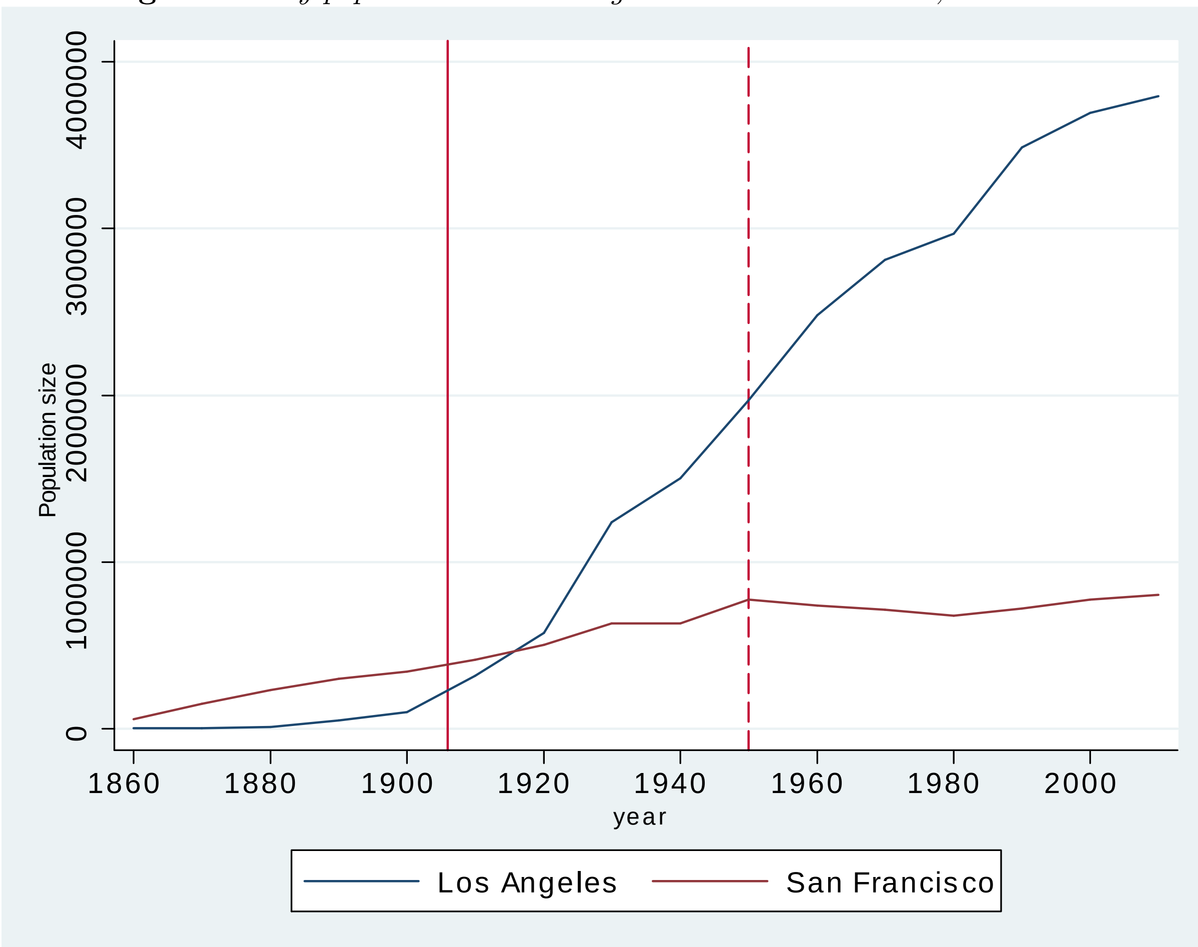
USGS, 2012. The Great 1906 San Francisco Earthquake. United States Geological Survey, page URL: http://earthquake.usgs.gov/regional/nca/1906/18april/revolution.php.

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, 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.

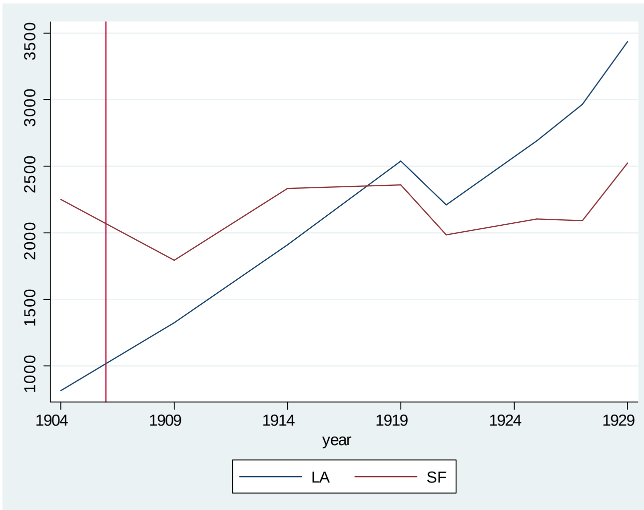
Zoback, M. The 1906 earthquake and a century of progress in understanding earthquakes and their hazards. GSA Today, 16(4-5), 4–11.

Figure I: *City population in Los Angeles and San Francisco, 1860–2010*

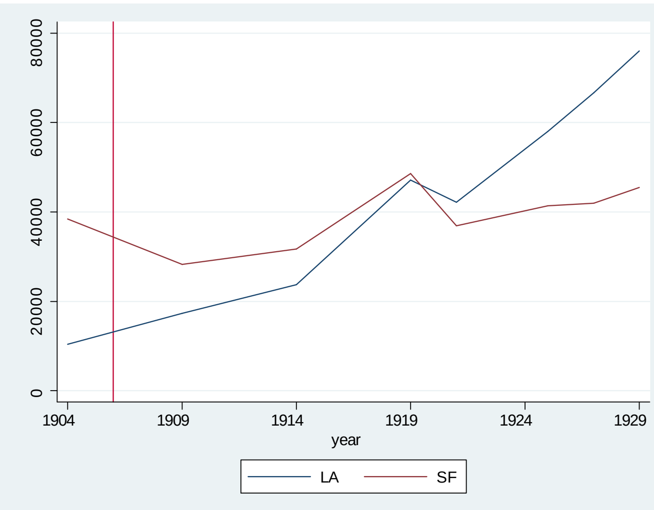


Notes. This figure depicts the city population sizes in Los Angeles and San Francisco every 10th year from 1860 to 2000. The solid vertical line indicates the year of the earthquake (1906), whereas the dashed vertical line indicated the final year of our sample in the empirical analysis.

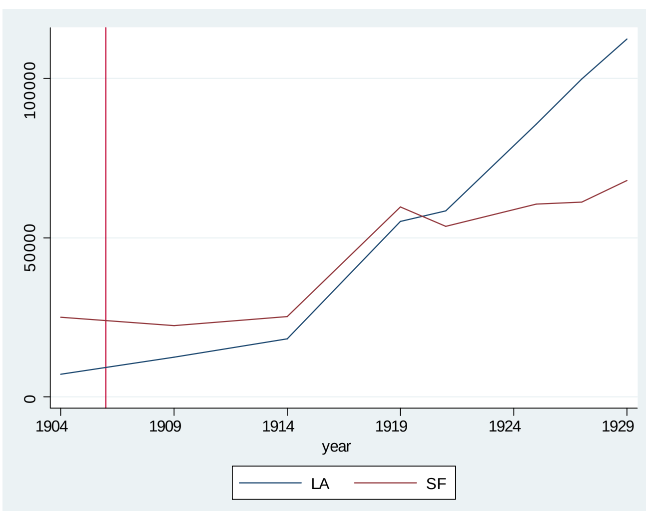
Figure II: *Manufacturing sectors in Los Angeles and San Francisco, 1904–1929*



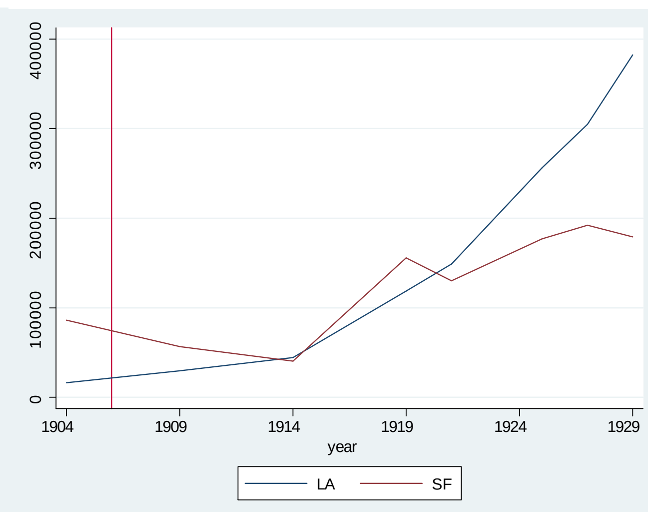
Panel A: Number of mfg. establishments.



Panel B: Number of mfg. workers



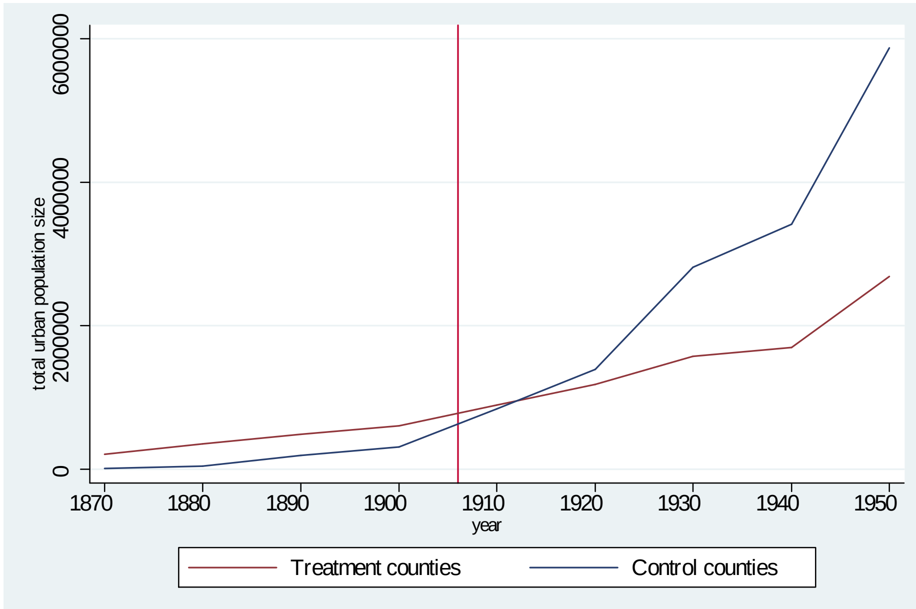
Panel C: Mfg. wages



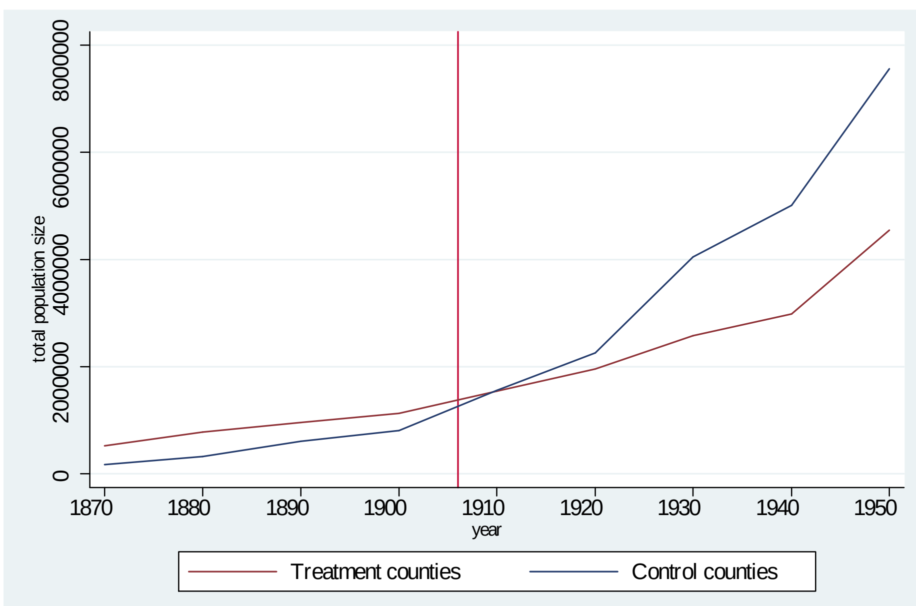
Panel D: Mfg. value added

Notes: These figures show the development in the total number of manufacturing establishments (panel A), total number of manufacturing workers (panel B), total wage expenditures in manufacturing (panel C), and total value added in manufacturing for the cities of Los Angeles and San Francisco. The years are: 1904, 1909, 1914, 1919, 1921, 1925, 1927 and 1929. The vertical indicates the year of the earthquake (1906).
Source: US Census of Manufactures (various years).

Figure III: *Population development in the affected and non-affected counties, 1870–1950*



Panel A: Urban population

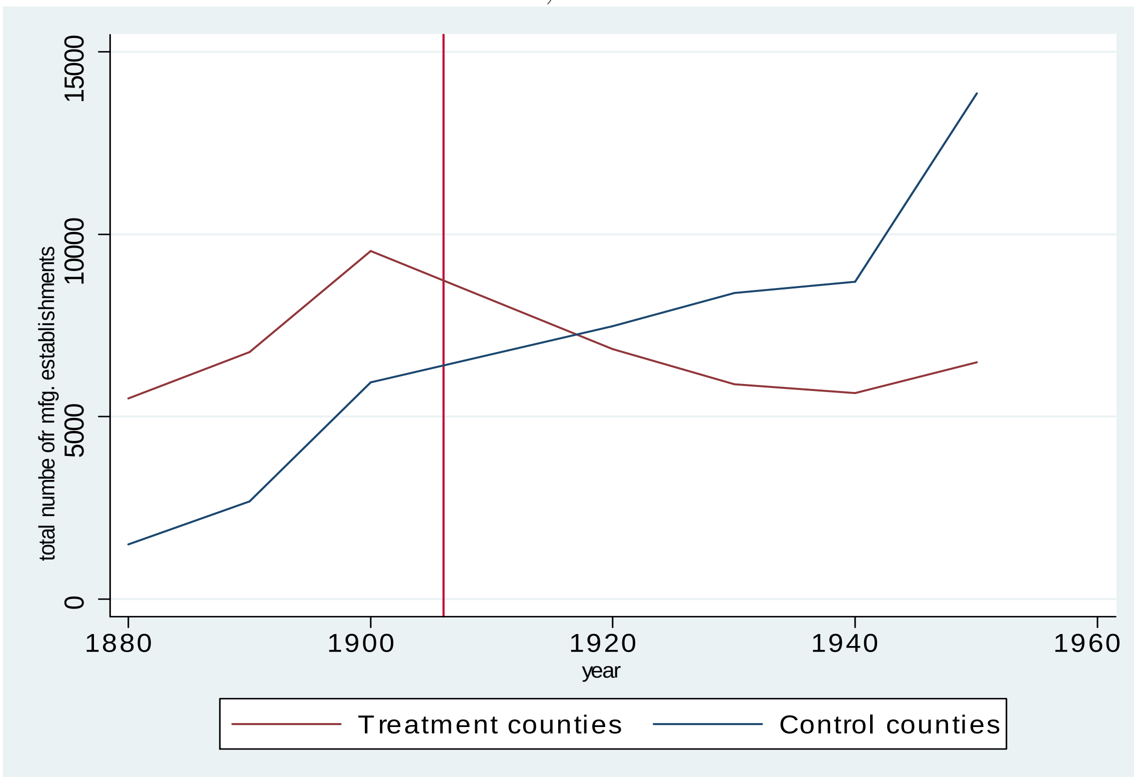


Panel B: Total population

Notes. These figures show the development of the total population and urban population, where we have grouped according to the treatment intensity measure used in the empirical analysis. Treatment is defined as the counties that were affected to some extent, whereas Control reflects the counties not affected at all. Panel A depicts the urban population and panel B the total population. The vertical indicates the year of the earthquake (1906).

Source: Haines (2010) and Boatwright and Bundock's (2005)

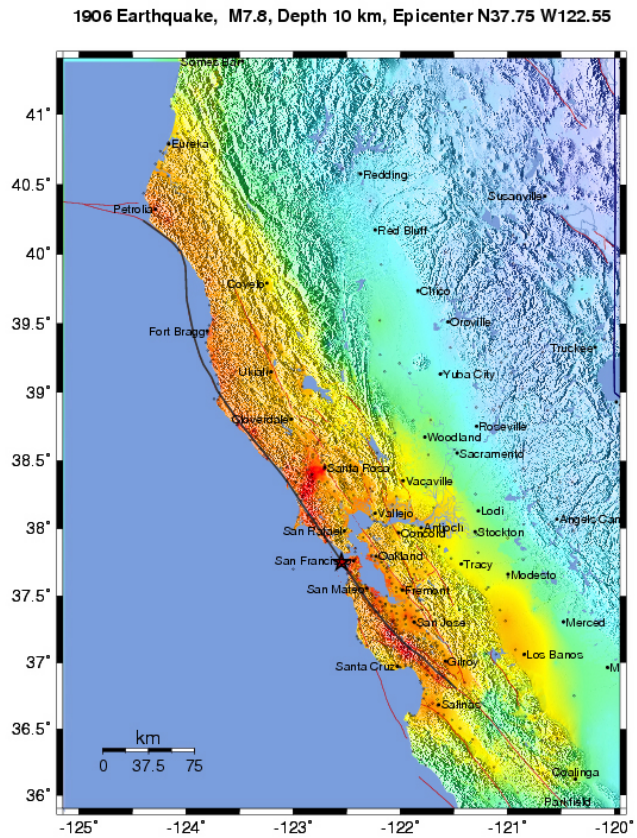
Figure IV: *Total number of manufacturing establishments in the affected and non-affected counties, 1880–1950*



Notes: These figures show the development in the total number of manufacturing establishments, where we have grouped according to the treatment intensity measure used in the empirical analysis. Treatment is defined as the counties that were affected to some extent, whereas Control reflects the counties not affected at all. The vertical indicates the year of the earthquake (1906). There are no data on manufacturing establishments in 1910.

Source: Haines (2010) and Boatwright and Bundock's (2005)

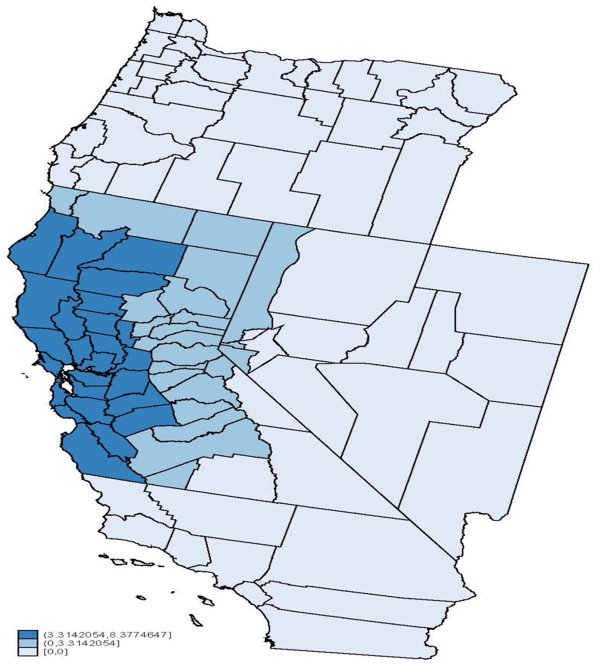
Figure V: *Shakemap of the 1906 San Francisco Earthquake*



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+

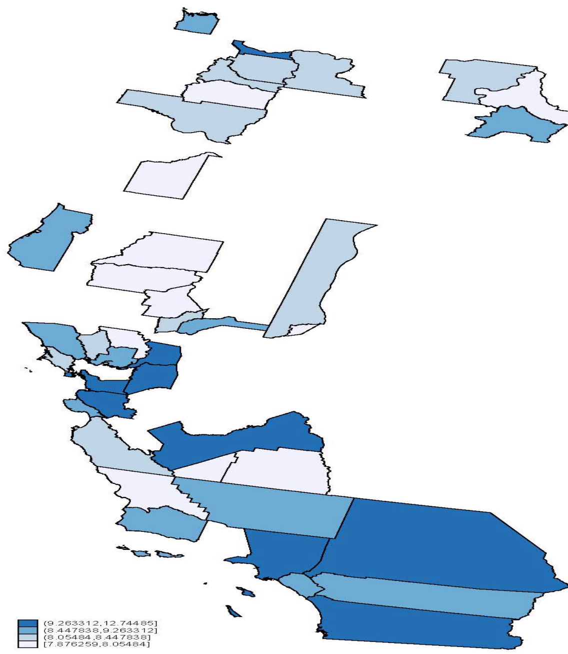
Notes: This figure show the earthquake intinsity of the 1906 San Francisco Earthquake.
Source: Boatwright and Bundock's (2005)

Figure VI: *Sample of affected and non-affected counties (shock intensity)*



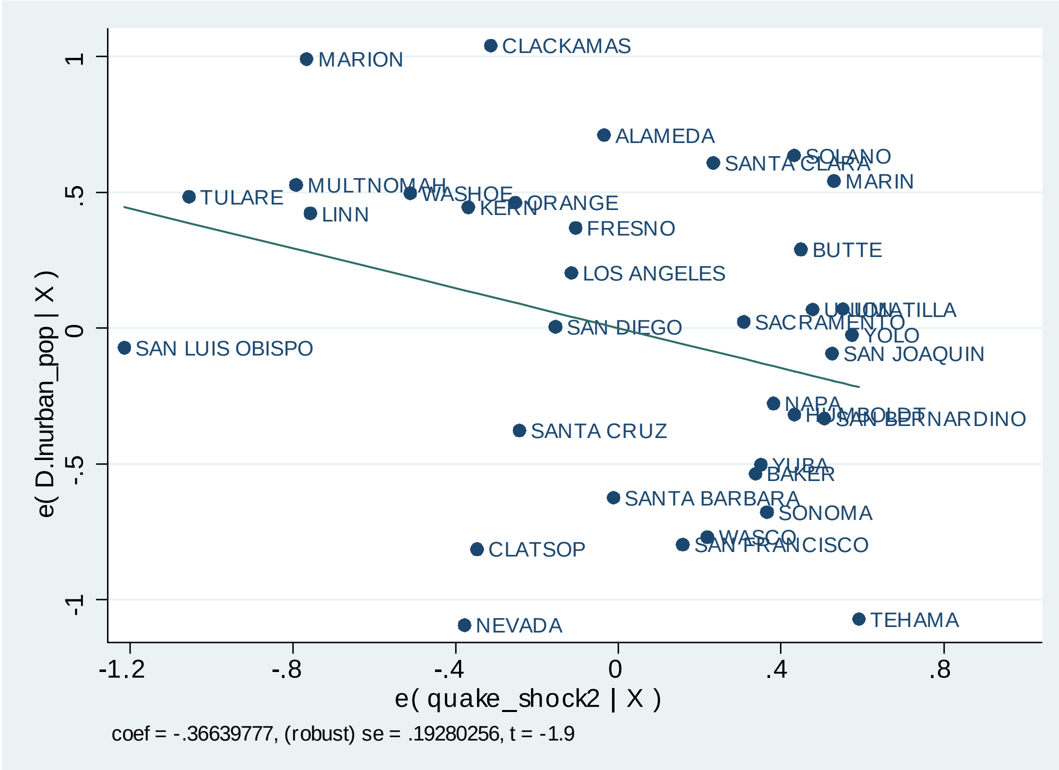
Notes: This is the sample of counties used in the empirical analysis, showing the earthquake intensity of the 1906 San Francisco Earthquake derived from Boatwright and Bundock's (2005) using Quantum GIS.

Figure VII: *Log urban population in 1900*



Notes: This is the reduced sample of counties used in the urban-population specification The criteria for a county be included is that is was urban, defined as a population larger than 2500, in 1890. The maps depicts log urban population size in 1900.

Figure VIII: *Partial correlation plot for log urban population*



Notes: Long difference specification for log urban population using the years 1900 and 1950.

Table I: Summary Statistics

	(1) N	(2) mean	(3) sd	(4) min	(5) max
Outcomes (in logs)**:**					
urban population	245	9.848	1.541	7.826	15.21
population	686	9.562	1.427	5.485	15.24
dwellings	588	8.099	1.336	4.159	13.78
mfg. establishments	511	3.840	1.387	0	9.156
mfg. workers	492	6.438	1.819	1.609	12.77
mfg. wages	459	13.53	2.143	8.300	20.82
mfg. value added	459	14.36	2.148	9.303	21.43
farms	588	6.592	1.319	0	9.446
land value per acre	686	3.621	1.136	0.693	9.700
agr. value added	588	14.24	1.491	5.030	17.96
quake damage intensity	686	1.939	2.690	0	8.377
distance epicenter	686	458.7	279.8	7.908	986.4
latitude	686	40.03	3.385	32.43	46.11
longitude	686	120.5	2.233	114.3	124.3
coast pacific	686	0.265	0.442	0	1
droughts	686	5.515	2.517	1.429	10.57
temperature	686	53.16	5.293	43.43	64.40
beaches	686	3.357	7.144	0	37
bays	686	5.724	9.216	0	46

Notes: The table reports summary statistics for the main variables used in the analysis. Notice that the geographical (cross-sectional) variables quake damage intensity, distance epicenter, ..., are time invariant, but in the empirical analysis they are interacted by time fixed effect in order not to be absorbed by county fixed effects.

Table II: Pre-Quake County Characteristics

		levels 1900		changes 1890-1900	
		geographical controls:			
		no	yes	no	yes
	# obs	(1)	(2)	(3)	(4)
Panel A: Population					
urban population	35	0.301	0.372	-0.0220	0.0775
		(0.234)	(0.324)	(0.0489)	(0.124)
population	98	0.505***	0.403**	-0.0312	-0.0123
		(0.122)	(0.170)	(0.0265)	(0.0441)
dwellings	98	0.465***	0.371**	-0.0223	0.000292
		(0.114)	(0.164)	(0.0267)	(0.0437)
Panel B: Manufacturing					
estabslishments	86	0.414***	0.383**	-0.0109	0.0450
		(0.153)	(0.192)	(0.0994)	(0.0886)
workers	83	0.551***	0.551***	0.0780	-0.0237
		(0.202)	(0.202)	(0.0814)	(0.140)
wages	77	0.495**	0.540**	0.0450	0.00629
		(0.213)	(0.259)	(0.0886)	(0.155)
value added	77	0.466**	0.549**	-0.0136	0.0140
		(0.201)	(0.235)	(0.0819)	(0.141)
Panel C: Agriculture					
land value per acre	98	0.370***	0.207*	-0.0980***	-0.0120
		(0.0956)	(0.116)	(0.0360)	(0.0572)
farms	98	0.302***	0.175	-0.0922**	-0.0497
		(0.113)	(0.181)	(0.0413)	(0.0498)
value added	98	0.470***	0.410**	-0.216***	-0.147**
		(0.110)	(0.192)	(0.0431)	(0.0585)

Notes: The table reports OLS estimates. All variables are reported in logs. Columns 1 and 2 report the pre-quake (1900) county-level characteristics by ln(1+quake-damage intensity) without and with geographical controls, respectively. Columns 3 and 4 report the pre-quake county-changes characteristics by ln(1+quake-damage intensity) without and with geographical controls, respectively. We refer to the data appendix for further details. Constants are not reported. Robust standard errors are reported in parentheses.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table III: Population and dwellings

	Dependent variable (in logs)					
	urban population		total population		dwellings	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Flexible estimates (OLS)						
ln(1+quake)×1890	-0.0775	-0.0395	0.0123	0.0373	-0.000292	0.0220
	(0.107)	(0.0930)	(0.0420)	(0.0437)	(0.0417)	(0.0435)
ln(1+quake)×1910	-0.246**	-0.279***	-0.109*	-0.124**	-0.107*	-0.115**
	(0.0991)	(0.0975)	(0.0587)	(0.0552)	(0.0598)	(0.0573)
ln(1+quake)×1920	-0.245	-0.277*	-0.0625	-0.0977	-0.0772	-0.0997
	(0.153)	(0.151)	(0.0869)	(0.0870)	(0.0846)	(0.0874)
ln(1+quake)×1930	-0.342**	-0.385**	-0.149	-0.198**	-0.153	-0.193*
	(0.160)	(0.161)	(0.101)	(0.0989)	(0.0998)	(0.102)
ln(1+quake)×1940	-0.392**	-0.434***	-0.209**	-0.259***	-0.221**	-0.257***
	(0.164)	(0.160)	(0.0997)	(0.0983)	(0.0897)	(0.0939)
ln(1+quake)×1950	-0.362**	-0.350**	-0.197	-0.270**		
	(0.147)	(0.154)	(0.121)	(0.117)		
Panel B: DiD estimates (OLS)						
ln(1+quake)×I^{post}	-0.279**	-0.325***	-0.151*	-0.208**	-0.139*	-0.177**
	(0.131)	(0.123)	(0.0916)	(0.0871)	(0.0814)	(0.0821)
Panel C: DiD estimates (2SLS)						
ln(1+quake)×I^{post}	-0.336*	-0.354**	-0.313***	-0.313***	-0.244**	-0.239**
	(0.176)	(0.174)	(0.121)	(0.119)	(0.115)	(0.111)
Kleibergen-Paap F-statistics	27.52	32.00	79.50	84.80	79.67	86.25
Controls ($\times \sum I^j$)						
geographical	Yes	Yes	Yes	Yes	Yes	Yes
initial outcome	No	Yes	No	Yes	No	Yes
observations	245	245	686	686	588	588
counties	35	35	98	98	98	98

Notes: The observations are reported at the county level over the period 1890-1950 (every decade). Panel A reports flexible estimates, where the comparison year is 1900 (OLS), Panel B reports DiD estimates (OLS), and Panel C reports DiD estimates using log 1 + distance to the epicenter interacted with a time indicator as an IV. All regressions include county and time fixed effects. The outcome variables are: log population for the sample of counties that were urban before the quake (columns 1 and 2), log population for all counties in the sample (columns 3 and 4), and log dwellings (columns 5 and 6). Quake is the potential damage intensity of the SF earthquake. I^{post} is the time indicator, which equals zero before 1910 and one after. The geographical controls are: latitude, longitude, coast-pacific dummy, # droughts, avg. temperature, # bays, and # beaches. Initial outcome is the outcome of interest in the year 1900. These cross-sectional measures are interacted with a full set of year fixed effects. Constants are not reported. Standard errors (in parentheses) account for arbitrary heteroskedasticity and are clustered at the county level.
*** p<0.01, ** p<0.05, * p<0.1.

Table IV: Manufacturing

	Dependent variable (in logs)							
	establishments		workers		wages		value added	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Flexible estimates (OLS)								
ln(1+quake)×1890	0.00489	0.0188	0.0238	0.0926	-0.00629	0.0492	-0.0118	0.0343
	(0.0946)	(0.0920)	(0.132)	(0.129)	(0.146)	(0.144)	(0.133)	(0.126)
ln(1+quake)×1920	-0.293***	-0.242**	-0.202	-0.173	-0.377*	-0.310	-0.413**	-0.358*
	(0.0923)	(0.0967)	(0.183)	(0.200)	(0.198)	(0.214)	(0.200)	(0.211)
ln(1+quake)×1930	-0.435***	-0.414***	-0.406*	-0.344	-0.649***	-0.536**	-0.724***	-0.650***
	(0.111)	(0.115)	(0.209)	(0.218)	(0.213)	(0.221)	(0.222)	(0.225)
ln(1+quake)×1940	-0.414***	-0.374***	-0.673***	-0.518**	-0.848***	-0.664**	-0.782***	-0.637**
	(0.116)	(0.120)	(0.228)	(0.238)	(0.255)	(0.261)	(0.254)	(0.258)
ln(1+quake)×1950	-0.530***	-0.464***	-0.587***	-0.414**	-0.663***	-0.460**	-0.687***	-0.500**
	(0.122)	(0.123)	(0.191)	(0.188)	(0.209)	(0.205)	(0.213)	(0.210)
Panel B: DiD estimates (OLS)								
ln(1+quake)×I^{post}	-0.421***	-0.383***	-0.478***	-0.408**	-0.631***	-0.517***	-0.645***	-0.554***
	(0.0938)	(0.0981)	(0.155)	(0.165)	(0.171)	(0.182)	(0.169)	(0.181)
Panel C: DiD estimates (2SLS)								
ln(1+quake)×I^{post}	-0.449***	-0.447***	-0.757***	-0.727***	-0.789***	-0.707***	-0.738***	-0.678***
	(0.146)	(0.144)	(0.222)	(0.224)	(0.220)	(0.221)	(0.213)	(0.219)
Kleibergen-Paap F-statistics	67.72	71.65	67.57	66.53	69.67	68.28	69.67	68.82
Controls ($\times \sum I^j$)								
geographical	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
initial outcome	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
observations	511	511	492	492	459	459	459	459
counties	86	86	83	83	77	77	77	77

Notes: The observations are reported at the county level over the periods 1890, 1900, and 1910-1950 (every decade). Panel A reports flexible estimates, where the comparison year is 1900, Panel B reports DiD estimates (OLS), and Panel C reports DiD estimates using log 1 + distance to the epicenter interacted with a time indicator as an IV. All regressions include county and time fixed effects. The outcome variables are: log establishments (columns 1 and 2), log workers (columns 3 and 4), log wages (columns 5 and 6), and log value added (columns 7 and 8). Quake is the potential damage intensity of the SF earthquake. I^{post} is the time indicator, which equals zero before 1910 and one after. The geographical controls are: latitude, longitude, coast-pacific dummy, # droughts, avg. temperature, # bays, # beaches, which are interacted with a full set of year fixed effects. Initial outcome is the outcome of interest in the year 1900 interacted with a full set of year fixed effects. Constants are not reported. Standard errors (in parentheses) account for arbitrary heteroskedasticity and are clustered at the county level.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table V: Agriculture

	Dependent variable (in logs)					
	farms		land value per acre		value added	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Flexible estimates (OLS)						
ln(1+quake)×1890	0.0497	0.0369	0.0120	0.0591	0.147***	0.136**
	(0.0475)	(0.0453)	(0.0545)	(0.0548)	(0.0557)	(0.0562)
ln(1+quake)×1910	-0.0793	-0.0923*	-0.283***	-0.299***	-0.159***	-0.210**
	(0.0516)	(0.0493)	(0.0681)	(0.0654)	(0.0614)	(0.0832)
ln(1+quake)×1920	-0.0187	-0.0553	-0.281***	-0.277***	-0.140	-0.255**
	(0.0748)	(0.0668)	(0.0922)	(0.0911)	(0.107)	(0.101)
ln(1+quake)×1930	-0.0185	-0.0646	-0.311***	-0.345***	-0.159	-0.247**
	(0.0819)	(0.0697)	(0.105)	(0.0994)	(0.103)	(0.0993)
ln(1+quake)×1940	-0.134	-0.179**	-0.301***	-0.312***	-0.250***	-0.331***
	(0.0879)	(0.0790)	(0.105)	(0.102)	(0.0885)	(0.0781)
ln(1+quake)×1950			-0.400***	-0.417***		
			(0.121)	(0.116)		
Panel B: DiD estimates (OLS)						
ln(1+quake)×I^{post}	-0.0874	-0.116*	-0.321***	-0.359***	-0.250***	-0.329***
	(0.0708)	(0.0642)	(0.0828)	(0.0798)	(0.0768)	(0.0747)
Panel C: DiD estimates (2SLS)						
ln(1+quake)×I^{post}	-0.110	-0.0287	-0.452***	-0.480***	-0.436***	-0.421***
	(0.0941)	(0.0747)	(0.123)	(0.116)	(0.132)	(0.118)
Kleibergen-Paap						
F-statistics	79.67	106.54	79.50	81.85	79.69	81.66
Controls ($\times \sum I^j$)						
geographical	Yes	Yes	Yes	Yes	Yes	Yes
initial outcome	Yes	Yes	Yes	Yes	Yes	Yes
observations	588	588	686	686	588	588
counties	98	98	98	98	98	98

Notes: The observations are reported at the county level over the periods 1890-1940 or 1890-1950 (every decade). Panel A reports flexible estimates, where the comparison year is 1900, Panel B reports DiD estimates (OLS), and Panel C reports DiD estimates using log distance to the epicenter interacted with a time indicator as an IV. All regressions include county and time fixed effects. The outcome variables are: log farms (columns 1 and 2), log land value per acre of farm land (columns 3 and 4), and log value added (columns 5 and 6). Quake is the potential damage intensity of the SF earthquake. I^{post} is the time indicator, which equals zero before 1910 and one after. The geographical controls are: latitude, longitude, coast-pacific dummy, # droughts, avg. temperature, # bays, # beaches, which are interacted with a full set of year fixed effects. Initial outcome is the outcome of interest in the year 1900 interacted with a full set of year fixed effects. Constants are not reported. Standard errors (in parentheses) account for arbitrary heteroskedasticity and are clustered at the county level. *** p<0.01, ** p<0.05, * p<0.1.

Table VI: Demographic Channels

	Dependent variable (in logs)					
	men 20-40 per capita	women 20-40 per capita	immi-grants per capita	child01/ women	child05/ women	married per capita
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Flexible estimates (OLS)					
ln(1+quake)×1880	-0.0234	0.114**	0.0282	0.0811	-0.0570	0.0278
	(0.0740)	(0.0520)	(0.0428)	(0.118)	(0.0706)	(0.0345)
ln(1+quake)×1910	0.0241	0.0273	-0.0827**	-0.0674	-0.154**	-0.00745
	(0.0647)	(0.0459)	(0.0410)	(0.0848)	(0.0716)	(0.0397)
ln(1+quake)×1920	0.0995	0.0152	-0.141***	-0.101	-0.128**	-0.0632*
	(0.0686)	(0.0611)	(0.0458)	(0.0811)	(0.0587)	(0.0375)
ln(1+quake)×1930	0.0172	-0.0865**	-0.0138	-0.0200	0.0454	-0.0368
	(0.0550)	(0.0417)	(0.0429)	(0.0931)	(0.0931)	(0.0297)
ln(1+quake)×1940	0.0354	-0.0138	-0.0877***	0.0435	-0.0600	-0.00122
	(0.0362)	(0.0213)	(0.0298)	(0.0931)	(0.0720)	(0.0180)
	Panel B: DiD estimates (OLS)					
ln(1+quake)×I^{post}	0.0557	-0.0715**	-0.0954**	-0.0767	-0.0455	-0.0411
	(0.0440)	(0.0330)	(0.0393)	(0.0538)	(0.0484)	(0.0273)
	Panel C: DiD estimates (2SLS)					
ln(1+quake)×I^{post}	0.0472	-0.0557	-0.159***	0.00791	-0.0623	-0.0425
	(0.0559)	(0.0480)	(0.0426)	(0.0761)	(0.0660)	(0.0287)
Kleibergen-Paap F-statistics	55.66	49.33	63.30	52.04	50.52	57.69
Controls ($\times \sum I^j$)						
geographical	Yes	Yes	Yes	Yes	Yes	Yes
initial outcome	Yes	Yes	Yes	Yes	Yes	Yes
observations	462	432	474	342	420	468
counties	77	72	79	57	70	78

Notes: The observations are reported at the county level over the period 1880 and 1900-1940 (every decade).Panel A reports flexible estimates, where the comparison year is 1900, Panel B reports DiD estimates (OLS), and Panel C reports DiD estimates using log distance to the epicenter interacted with a time indicator as an IV. The outcome variables are: log men aged 20-40 per capita (column 1), log women aged 20-40 per capita (column 2), log immigrants per capita (column 3), log children aged 0-1 per women of childbearing age (column 4), log children aged 0-5 per women of childbearing age (column 5), and log number of married individuals per capita (column 6). Quake is the potential damage intensity of the SF earthquake. I^{post} is the time indicator, which equals zero before 1910 and one after. The geographical controls are: latitude, longitude, coast-pacific dummy, # droughts, avg. temperature, # bays, # beaches, which are interacted with a full set of year fixed effects. Initial outcome is the outcome of interest in the year 1900 interacted with a full set of year fixed effects. Constants are not reported. Standard errors (in parentheses) account for arbitrary heteroskedasticity and are clustered at the county level.
*** p<0.01, ** p<0.05, * p<0.1.