A Non-Random Walk in the City: The 1906 San Francisco Disaster and Industry Localization

James Siodla*

September 4, 2015

Abstract

The geographic concentration of business activities in a city is a dynamic process that is sensitive to the way industries evolve and change over time. Is localization also sensitive to large, temporary shocks? To answer this question, this paper studies the impact of the 1906 San Francisco earthquake and fire on the localization of industries in the city. The disaster disrupted normal business activity through the destruction of over 28,000 buildings on more than 500 city blocks. Using location data gathered from city business directories in various years between 1900 and 1930, this study quantifies localization at the industry level and tests whether it follows a random walk through time. Evidence suggests that the disaster had a significant, but temporary, effect on localization in San Francisco. Thus, the effect of a large shock on localization is mean-reverting at the level of business locations within a city.

JEL codes: N91, R12.
Keywords: urban disaster, shock, agglomeration, San Francisco.

*Department of Economics. Colby College. Email: jrsiodla@colby.edu. Web: https://jamessiodla.wordpress.com. I am extremely grateful to Dan Bogart and Jan Brueckner for their advice and comments. I gratefully acknowledge research support from the UC Irvine Department of Economics and the All-UC Group in Economic History, as well as support from the Institute for Humane Studies.
1 Introduction

Large disasters can have big consequences for cities and regions. Their destructive nature disrupts normal business activity and temporarily redirects it toward other pursuits, such as relief. In the reconstruction phase, however, they provide opportunities to reorganize economic activity. The goal of this paper is to estimate the impact of the 1906 San Francisco earthquake and fire on industry localization (i.e., agglomeration) in the city. It was the largest urban calamity in U.S. history up to that point, far exceeding the total damage inflicted by previous urban disasters such as the fires in Baltimore, Boston, and Chicago. Did the disaster then have an effect on industry localization in the city? If so, was it temporary or permanent? In other words, does localization follow a random walk over time so that it can be greatly perturbed by a large shock, or does it follow a more intentional path through time?

There are two distinct theories regarding the impact of temporary shocks on agglomeration. The theory of natural advantages suggests a single equilibrium outcome in location patterns, and hence a return to the previous state following a large shock. On the other hand, models in the New Economic Geography (NEG) tradition, with their focus on self-reinforcing processes, suggest instability and thus multiple equilibria. In this case, a city that suffers a large shock may not return to a previous state and instead adjust to a new equilibrium. If the 1906 disaster had a temporary and non-disruptive impact on the spatial distribution of businesses in San Francisco, then natural advantages and the single equilibrium hypothesis is supported. However, if the city’s business concentration patterns changed significantly and persist as a result of the disaster, then there is evidence in favor of the NEG hypothesis. Empirical work has often focused on the impact of temporary shocks on urban systems and agglomeration at the regional level, thus pitting these two theories against one another (Davis and Weinstein (2002); Brakman et al. (2004)). However, little research has been conducted focusing on the impact of large shocks on outcomes within a city. This pa-
per aims to fill this gap. If industry localization follows a random walk, such that the 1906 disaster was sufficient to permanently alter agglomeration patterns, then this lends support for the NEG hypothesis. However, if the disaster only had a temporary effect (or no effect at all) on localization, there is evidence in favor of the single-equilibrium hypothesis.

Urban disasters have had many implications for cities. After the 1872 Fire, Boston experienced higher land values and was rebuilt with higher quality buildings (Hornbeck and Keniston (2014)). Also, changes in industry agglomeration after the fire were slight and were not the primary drivers of the higher post-disaster land values. After the 1906 Fire, San Francisco was rebuilt at much higher residential density, resulting in a density gap between burned and unburned areas that still exists today (Siodla (2015)). Fires in Baltimore and Chicago seem to have resulted in the movement of businesses in general, although it is not clear whether agglomeration patterns were impacted (Rosen (1986)). Other studies of disastrous events in cities have focused largely on short-run outcomes rather than causal effects and long-run implications (Fales and Moses (1972); Douty (1977)).\footnote{For instance, Fales and Moses (1972) study the distribution of population and industry in Chicago only after the 1871 Fire. In a study on the 1906 disaster in San Francisco, Douty (1977) looks at various aspects of the city’s redevelopment and reconstruction.} Thus, much can be learned from studying the impact of a large disaster on industry localization in a city.

Using 1906 as a break point, this paper compares the geographic distribution of businesses in San Francisco between 1900 and 1930. The focus is on two pre-disaster periods in 1900 and 1905, and three post-disaster periods in 1906, 1915, and 1930. Industry localization is first measured using an index developed by Mori et al. (2005), which is then used to test whether localization follows a random walk in a city. Evidence presented here suggests that the fire indeed had an impact on localization in San Francisco, but that the effect was only temporary and had mostly dissipated by 1930.
2 Agglomeration, Cities, and Shocks

Business location decisions ultimately determine agglomeration patterns. Foundational studies on business location patterns include Alao (1974) and Moomaw (1980), which are primarily concerned with initial firm location decisions. However, as cities develop over time, businesses must weigh the benefits and costs of relocating. Three primary determinants of firm relocations include changes in demand, initial plant size, and changes in transport costs (Cooke (1983)). Changes in demand and transport costs are likely to occur over time, and thus firms must adjust accordingly. Initial plant size acts as a fixed investment, so that larger plants are relatively more anchored to initial locations. Due to fixed capital investments, a firm’s choice of a particular location has an impact on future location decisions, and thus past commitments play a role in the dynamic process of business location patterns. In this way, it is useful to understand historical location decisions in order to gauge firm behavior (Stahl (1987)).

Closely related to firm location decisions is the concept of agglomeration economies in production. Mills and Hamilton (1994, p. 118) describe agglomeration economies as weakening with distance to the CBD, thus encouraging firms to locate near the center of the city. In nineteenth century cities, the high costs of moving goods relative to moving people explains much business clustering (Moses and Williamson (1967)). Other reasons for agglomeration include labor pooling and the sharing of ideas. Whatever the sources and determinants of agglomeration, attempts to measure it and explain its patterns are an important part of the analysis. Various indexes have been developed to do so, with each index suggesting strong tendencies for industries—primarily manufacturing—to agglomerate in the U.S. (Ellison and Glaeser (1997), U.K. (Duranton and Overman (2002)), and Japan (Mori et al. (2005)). For the U.S., Ellison and Glaeser (1997) find evidence that extreme cases of agglomeration are likely due to natural advantages, although they assert that there is much concentration left
to explain. Other studies have focused on the dynamic process of localization and the role of history in determining current patterns. Kim (1995) studies the long-run trends in U.S. manufacturing and finds evidence of a general decline in specialization and localization from 1860-1987, which is largely due to changes in resource use and scale economies. Dumais et al. (2002) show stability among many industries from 1972-1992, but with a decline in concentration in the last half of the period. The authors decompose the changes in concentration and find evidence that location choices of new firms tend to de-concentrate an industry, while plant closures reinforce agglomeration. They further find that historical accidents, while important in certain industries, generally do not have long-lasting effects on many other industries. This phenomenon is supported by evidence that shows concentrated industries tend to be as mobile as unconcentrated industries over time.

To what extent are large shocks important in explaining geographic concentration patterns over time? Various studies have answered this question in the context of urban systems. Following World War II, Japanese cities that were bombed by the U.S. had suffered population losses during the war, but had recovered lost growth within a couple of decades (Davis and Weinstein (2002)). In this case, the bombings had a significant impact on city size and population growth, but the effect was only temporary. The same is true in the context of industries in Japan, which had recovered in the same cities in which they existed prior to the war (Davis and Weinstein (2008)). The same result holds in the case of Germany overall, although East Germany experienced a more permanent impact due to institutional reasons (Brakman et al. (2004)). Overall, the evidence regarding the impact of war-time destruction on urban systems supports the idea that patterns of agglomeration are determined by fundamentals. However, other cities may thrive due to self-reinforcing agglomeration economies that are generated over time. For instance, many cities in the U.S. are still centered on old portage sites that are no longer used (Bleakley and Lin (2012)), and cities in Kenya are still located near old colonial railroads (Jedwab et al. (2015)). The literature is thus mixed on
the role of fundamentals and historical accidents in determining the location of economic activity in urban systems.

Cities exhibit enormous fixed investments that are difficult to alter and adapt, and thus there is great potential for history and shocks to play a large role in their development at very micro levels. For instance, persistence of land use is evident in urban density patterns. (Brooks and Lutz (2013) show that high density patterns around now-defunct streetcar stops are still evident in Los Angeles, while Siodla (2015) shows a long-lasting impact of the 1906 San Francisco fire on residential density. In these cases, evidence supports the idea that micro-level outcomes like land use in cities are not mean-reverting in the face of shocks.

Many forces may be at work in explaining location patterns as cities develop over time. For instance, improvements in urban transport networks can have either agglomerating or de-agglomerating effects. Regulations such as zoning might bear responsibility for the location patterns of businesses within a city. Thus, patterns emerge over time in how businesses locate. The impact of a large-scale disaster on such patterns is the focus of this paper, which estimates this effect at a very micro level. Did the 1906 disaster have an impact on business location patterns in San Francisco? If so, was the impact temporary or permanent? The remainder of the paper is focused on answering these questions.

3 The History and the Data

3.1 Historical Background

At the time of the disaster in 1906, San Francisco was experiencing the type of growth witnessed in many cities during this time of heavy U.S. industrialization. With nearly 343,000 people in 1900, the city’s population grew an average of 20 percent each decade between 1900 and 1930 (Issel and Cherny (1986, p. 24, Table 1)). The city’s manufacturing sector was also experiencing growth leading around the time of the disaster, with 1,748
establishments in 1899 and 2,251 in 1904 (Douty (1977, p. 366, Table 29). Over this time period, there were also large increases in the number of wage earners employed in the manufacturing sector, as well as increases in total wages, capital, and output. While the disaster temporarily impacted these numbers for the worse, they had greatly recovered by 1914 and manufacturing remained a large part of the city’s economic activity after 1906.

The bulk of total damage was caused by the fire. In total, 28,188 buildings were consumed across 2,831 acres of land. Figure 1 shows the coverage of the fire, which is represented by the darkest area on the map. The buildings in unburned areas suffered from the earthquake, but were often repaired as they were much less compromised relative to those that burned. It is estimated that the fire’s destruction represents at least 80 percent of the total damage inflicted by the disaster.\(^2\) Thus, the fire will be treated throughout the study as the most important component of the disaster.

Reconstruction was fairly rapid, one indicator of which is the number of building permits issued in each year following the fire itself. Data gathered from municipal reports show that city-wide building permit issues returned to their pre-disaster level by 1914 (SFMR (1904-1916)). Furthermore, reconstruction was largely initiated by business owners, and the city supported their efforts in an attempt to quickly re-establish a significant part of the tax base.

Overall, private interests largely guided building during the reconstruction phase. Even a plan developed just prior to the disaster to reorganize the city’s layout was ignored upon reconstruction. Although attempts were made to implement new building codes, most aspects of the new codes were ignored in the rush to rebuild, including height limitations and requirements for fire-resistant walls (Fradkin (2005)). The most significant—and generally followed—changes were a moderate expansion of the city’s fire limits (where buildings were required to be largely non-combustible), a new fireproof roof area, and the legal permissibility of concrete in buildings (Tobriner (2006)). The city’s first zoning code was implemented

\(^2\)Tobriner (2006) suggests that the fire accounts for 95 percent of total property damage.
in 1921, but it did not play a significant role in determining land use in the city until after the Great Depression (Weiss (1988)). Overall, aside from relatively minor changes to the fire limits and the new fireproof roof area, the regulatory environment was largely unchanged after the disaster and should not greatly impact the post-disaster patterns of localization among industries.

3.2 Data Construction

The primary data sources for this study are the San Francisco city directories published by Crocker-Langley. These directories were produced annually in the nineteenth and twentieth centuries and conveniently list business names and street addresses by trade. Figure 2 shows a sample page from the 1915 directory, from which data on cigar manufacturers were gathered.

Address listings were gathered for 55 randomly chosen business categories in the following years: 1900, 1905, 1906, 1915, and 1930 (Langley (1900); Langley (1905); Langley (1906); Langley (1915); Polk (1930)). The directory for 1906 was developed after the disaster, showing business locations as of December 31st of that year. Addresses were linked to city blocks, which remained relatively constant in size and shape over the time period under study. These block locations were determined using Sanborn maps produced in 1900, 1905, 1914, and 1931 (Sanborn (1899–1900); Sanborn (1905); Sanborn (1928–1931)). These maps provide address ranges for city blocks in each year under study. Since street names and addresses ranges change over time, the data produced in this study are more historically accurate than they would be using a modern geocoder, which relies on today’s information to identify business locations.

The SIC (Standard Industrial Classification) system, developed in 1937, was used to classify business categories from the directories into the following major divisions: manufacturing, wholesale trade, retail trade, FIRE (i.e, fire, insurance, and real estate) services,
and general services. Manufacturing features the largest number of industries with 30, while there are 25 industries across the rest of the divisions. Each business category was assigned to one of these divisions. A large focus in the study is on manufacturing due to the established economic benefits of agglomeration among these firms, and also because it is the primary focus of other studies in agglomeration. Table 1 reports the SIC divisions and business categories used in the study. The categories are an assortment of important industries in San Francisco. While businesses enter and exit over time, the dataset is longitudinal in business category and balanced featuring five time periods.

Other important facets of the study include the determination of the blocks that were razed in the disaster and the spatial area of focus. Figure 1 shows the map used to determine which blocks were burned and which were left unscathed, which is used in the study to determine the degree to which industries were impacted by the disaster. The darker portion of the map refers to the city blocks that burned in 1906. Most businesses were located in or near the core of the city, which was mostly burned by the fire. This study will focus on the primary areas of the city, such as Downtown, Western Addition, South of Market, and Mission District. These areas are depicted in Figure 3. The relatively few firms located in the outskirts of the city in any given year, such as the Richmond, Sunset, and Outer Mission areas, were coded as being located in the suburbs rather than a specific city block.

4 Estimation Methods and Results

4.1 Index of Localization

Localization is measured using an index developed by Mori et al. (2005), which is described in detail in the appendix. The data requirements for the index are relatively meager, requiring only a count of firms for each industry in each spatial unit, which in this case is a city block. The index, which is independent of sample size, relates the degree of localization found
among particular industries across blocks to a reference distribution that assumes complete 
spatial dispersion of firms. In this study, the reference distribution is assumed to be uniform 
across city blocks. Each industry has an index value with a unique minimum value of zero, 
with a larger value indicating a greater level of firm localization relative to the reference 
distribution. An index was calculated for each of the 55 industries in each year of the study.

A summary of the index is given in Table 2, which compares index values across years 
and SIC divisions. The disaster seems to have had an impact in all divisions, with average 
index values in 1906 substantially different from those in 1905. However, by 1930, localiza-
tion returns to 1900 levels in manufacturing and FIRE, while it rises in the wholesale, retail, 
and services divisions. Only in retail does it appear that the fire itself caused a substan-
tial redirection in localization, such that the mean index values in 1906 and 1930 are very 
similar, and substantially different from the pre-disaster values. Overall, retail and services 
experienced the largest percentage increases in localization between 1900 and 1930, rising 
almost 30 percent across the two divisions. Wholesale trade experienced a relatively small 
percentage increase in localization over this time.

Table 3 shows the correlation of the index values across time and SIC divisions. The 
table illustrates that localization in each division was stable leading up to the disaster, with 
correlations of at least 0.85 from 1900 to 1905. In each division, the correlation of indexes 
is weaker between 1900 and 1930, and considerably so in the retail and services divisions. 
Retail trade is the only division where correlation is low between 1900 and 1930, but high 
between 1906 and and years thereafter, again suggesting a large impact of the fire. The 
manufacturing, wholesale, and FIRE divisions display higher degrees of correlation between 
1900 and 1930 than other divisions, suggesting stability in the face of temporary shocks. As 
a comparison, in their study of U.S. manufacturing, Dumais et al. (2002) find a correlation 
of the localization index of 0.92 between 1972 and 1992 for manufacturing industries, which 
suggests stable agglomeration over time. In a more long-run view, Kim (1995) finds a 0.64
correlation of localization index values in manufacturing industries between 1860 and 1987. Thus, the stability seen in manufacturing in San Francisco is in line with previous studies of long-run industry agglomeration patterns, even in the face of the 1906 disaster.

Figure 4 shows the changes in localization that occur across time periods within each SIC division. The change in the index value for each industry in a particular period is displayed relative to the comparison year’s index value. In all five divisions, there are very small changes in localization leading up to the fire, as most industries experienced very little change between 1900 and 1905. However, the graphs in the second column show that localization changed considerably after the disaster. There are fewer observations near the zero-change line, thereby suggesting that the disaster had an impact on industry localization. Substantial changes in localization also occur in later time periods, presumably as industries recover and reestablish location patterns after the shock. For instance, in manufacturing, the disaster caused temporarily high levels of localization, with downward adjustments occurring in later periods. Retail, on the other hand, experienced a localizing impact of the disaster and reinforced it with little change occurring between 1906 and 1915 and greater localization between 1915 and 1930.

Overall, this preliminary analysis suggests that, in most cases, the fire likely had only a short-lived impact on localization in the city. If this is true, how long did it take for industries to recover their pre-disaster localization paths? The remainder of the paper will focus on more clearly identifying whether the fire indeed had an impact, and whether it was only temporary or permanent. In other words, does industry localization in a city follow a random walk?

4.2 The Disaster and Agglomeration

The technique used in identifying the disaster’s impact and its persistence follows the methodology developed by Davis and Weinstein (2002), which has been subsequently used
in other studies (Brakman et al. (2004); Davis and Weinstein (2008)). The goal of this section is to test whether industry localization follows a random walk in a city. If so, then a temporary shock the size of the 1906 disaster may be enough to redirect the evolution of localization in San Francisco. If localization does not follow a random walk, but is instead mean-reverting, then the effect of the shock dissipates over time. Let $D$ represent the index value (in logs), so that the following equation represents localization for industry $i$ at time $t$:

$$D_{it} = \Omega_i + \varepsilon_{it},$$

(1)

where $\Omega_i$ is an initial localization index value for industry $i$ and $\varepsilon_{it}$ represents industry-specific shocks. The persistence of a shock in period $t$ can then be modeled as

$$\varepsilon_{i,t+1} = \rho \varepsilon_{it} + \nu_{i,t+1},$$

(2)

where $\nu_{it}$ is independently and identically distributed (i.i.d.). It is assumed that $0 \leq \rho \leq 1$.

First-differencing (1) and using (2) yields the following equation:

$$D_{t,t+1} - D_{it} = (\rho - 1)\nu_{it} + [\nu_{i,t+1} + \rho(1-\rho)\varepsilon_{i,t-1}].$$

(3)

The term in brackets is the error term, and is uncorrelated with the shock in period $t$. A shock is permanent if $\rho = 1$, while it dissipates over time with varying degrees of persistence if $0 < \rho < 1$. If $\rho = 0$ the shock has no persistence at all.

A unit root test can be employed to determine the value of $\rho$, as mentioned by Davis and Weinstein (2002, p. 1280). Harris and Tzavalis (1999) develop a unit root test for panel data with relatively high $N$ and fixed $T$. The null hypothesis is that the panels contain a unit root, while the alternative hypothesis is that they do not. The test statistic is highly significant, with an estimate of $\rho = 0.19$ across all divisions. This outcome suggests that
localization does not follow a random walk in a city. But while this test is informative, its power is low. The parameter $\rho$ can also be estimated through a more convincing means using a variety of time horizons beyond the disaster.

The key to estimating the value of $\rho$ is in identifying the shock, $\nu_\delta$. In the present case, one proxy for the shock is the change in localization between 1905 and 1906, years which represent time periods in the data that are directly before and after the disaster. Using (3), the following equation can then be estimated (in logs):

$$D_{i,1906+\delta} - D_{i,1906} = \alpha(D_{i,1906} - D_{i,1905}) + \beta_0 + \text{error}_i,$$

where $\alpha \equiv \rho - 1$. This equation tests whether localization follows a random walk with drift, where drift is captured by $\beta_0$. This parameter captures long-run trends in localization within the city that may be due to a variety of changes at the beginning of the twentieth century, such as the introduction of zoning or changes in transport costs. In this setup, localization follows a random walk if $\alpha = 0$ ($\rho = 1$). If $\alpha = -1$ ($\rho = 0$), then the shock had no effect at all. If $-1 < \alpha < 0$, then the disaster had a temporary (i.e., mean-reverting) effect on localization paths in San Francisco.

The shock itself is measured with error since it contains information about past localization changes and is thus correlated with the error term in equation (3). In order to identify the innovation $\nu_\delta$, it is necessary to use valid instruments for the shock in 1906. Two variables that reflect the destruction experienced in the disaster are the proportion of firms burned by the fire within a particular industry and the percentage change in the number of firms in the industry between 1905 and 1906. These variables proxy for the degree of the fire’s impact on particular industries. Industries that were more greatly impacted by the disaster—as reflected by a greater proportion of firms burned out and relatively fewer firms remaining in the city by 1906—potentially experienced more substantial changes in
Finally, it is necessary to choose $t$ for the left-hand side variable. This choice of year should reflect the long-run adjustment in recovery from the shock. In the case of San Francisco, the city declared itself “new” and largely recovered by the time of the Pan-Pacific International Exposition, which took place in 1915. Thus, $t = 9$ in this case. Analysis is also conducted using a longer horizon out to 1930, in which case $t = 24$. These values sandwich the time horizons used by Davis and Weinstein (2002) and Brakman et al. (2004), which are between 13 and 18 years.

Table 4 reports the first-stage regression results where the change in localization between 1905 and 1906 (in logs) is regressed on the two instrumental variables described. For reasons alluded to in Section 4.1 and described below, the analysis is carried out separately for all divisions in the sample, manufacturing alone, and non-retail divisions. The table reveals that the instruments are highly correlated with the dependent variable and explain much of its variation, namely 56 percent for all divisions, 61 percent for manufacturing, and 51 percent for non-retail divisions. The $F$-statistics in these regressions are between 13 and 23. Burned-out firms had an effect in each case, except for manufacturing.

Table 5 reports the results of estimating equation (4) using the two instruments, which is the test of whether temporary shocks have permanent effects on localization patterns. The coefficient of interest is on the change in localization between 1905 and 1906, which corresponds to $\alpha = \rho - 1$. Aside from analyzing all divisions in a single specification, analysis is conducted for manufacturing alone since previous studies of agglomeration have focused on manufacturing. Also, since retail seemed to have experienced a differential impact of the shock relative to other divisions (see Section 4.1), the last specification drops this division from the analysis.

First consider columns 1 through 3, which show results for the 9-year time horizon. For all divisions and manufacturing, the coefficient on the 1905–1906 change in localization is
nearly \(-0.50\), which means that \(\rho\) is approximately 0.50. When dropping retail from the analysis, the coefficient rises to \(-0.56\), thereby suggesting \(\rho = .44\). Excluding retail causes \(\rho\) to decrease, which means the fire's effect is closer to zero for non-retail divisions. In the nine years following the disaster, the typical industry had halfway recovered its former localization path, so that the fire's effect was only temporary.

The fact that industries had only halfway recovered their previous localization paths by 1915 suggests that more time was needed to completely recover from the disaster. Thus, an informative exercise is to extend the time horizon. Columns 4 through 6 report the results of this exercise for a 24-year horizon that extends out to 1930, which covers the exuberant building period of the 1920s (Field (1992)) and leads up to the Great Depression. The results suggest that industries had gotten closer to their pre-disaster localization paths after 1915. The coefficient for all divisions and manufacturing alone is at least \(-0.61\) \((\rho \approx 0.39)\) on the 1905–1906 change in localization. Dropping retail from the analysis decreases the coefficient considerably to \(-0.81\) \((\rho \approx 0.19)\), which implies that the \(\rho\) value for retail is closer to zero than that of other divisions. Thus, by 1930, industries had mostly recovered their pre-1906 localization paths, especially non-retail industries.

Overall, these tests establish that the destruction from the 1906 disaster had only a temporary impact on the localization patterns of industries in San Francisco. The city's industries had halfway returned to their pre-disaster localization paths by 1915 and had gotten even closer by 1930. This strongly rejects the hypothesis that the path of localization among industries in a city follows a random walk. These outcomes broadly support the natural advantages literature in the context of industry localization: the clustering of businesses in a city does not follow a random walk and shock persistence dissipates over time. These results are in line with previous studies on urban population growth, which suggest that large shocks had only temporary effects on urban systems and structures in Japan and Germany as a whole (Davis and Weinstein (2002); Davis and Weinstein (2008); Brakman
et al. (2004)). Furthermore, studies of agglomeration in manufacturing industries suggest remarkable resilience in localization over long periods of time (Dumais et al. (2002); Kim (1995)), as shown by a high degree of correlation in localization index values across long periods of time.

4.3 Discussion

The analysis was carried out for three separate samples: all divisions, manufacturing, and non-retail divisions. In all specifications and time horizons, the most substantial return to pre-disaster paths was seen when excluding retail from the analysis. Furthermore, the discussion in Section 4.1 suggests that the fire had a differential effect on retail relative to other divisions. Why might this be the case? In other words, why might the effects of large, temporary shocks persist in some cases and not in others? One might attribute this outcome to various institutional factors, such as land-use regulations that favor certain industries or the political power of industry magnates. These factors notwithstanding, economic forces likely play a role as well.

For instance, manufacturing industries tend to place such a high value on clustering due to the economic benefits accrued that they tend to outbid residential users for valuable land. In this case, a shock would not be expected to alter the trends or patterns of localization that a thriving industry had established in the past. In the present study, this might also be true for other non-retail divisions, whose mean localization index values in 1906 were fairly similar to those in 1905. But what about the case of retail, which seemed to have been affected greatly by the shock by becoming more concentrated after 1906 and remaining so at least through 1930?

One reason why the retail sector may have experienced higher concentration in the years after the fire might have to do with the development of the streetcar system, which was expanding prior to the disaster. Localization in many industries, such as manufacturing
and FIRE services, likely experienced little or no change from this innovation since it only meant shorter commute times for workers. In fact, Walker (2001) asserts that manufacturing industries in San Francisco were forming new districts and changing locations well before World War I and the advent of more modern transport technology. The financial district—where most FIRE services were located—also existed well before the fire, and was only reinforced after 1906. Thus, a temporary shock would not likely perturb the seemingly beneficial localization paths followed by many industries even years before the disaster.

But retail may be a different story. With the development of the streetcar system, which reduced the cost of moving people across the city, and a clean slate following the fire, retail businesses had more freedom to change location patterns. Increasingly, these businesses could locate farther from their customers and form districts to gain from inter-store externalities. This seems to have occurred in San Francisco after the disaster. New retail districts had emerged within a few short years after the fire. One example is the change experienced on Market Street, the central thoroughfare in San Francisco that became a shopping district unto itself after 1906. Upon visiting the city in 1909, one writer observed, "Market street [sic], which, previous to 1906, was strictly a wholesale thoroughfare...is now lined with retail stores..." (Grant (1909, p. 370)). The emergence of retail shopping on Market Street would not have been possible without adequate transportation to move people to the city center from their increasingly suburban residential locations. Additionally, as it pertains to this study, the process of moving retail downtown would likely have taken much longer in the absence of the fire. For this reason, more confidence is placed in the study’s results that exclude the retail industry, which may have been subject to shocks outside of the disaster during this time.
5 Conclusion

Industry localization is a dynamic process that is susceptible to a variety of exogenous events that occur within industries over time. Evidence presented in this paper suggests that localization in a city is susceptible to large shocks, but that this effect is only temporary. In this case of San Francisco, the 1906 disaster had an impact on industry localization that had mostly dissipated by 1930. This outcome supports the hypothesis that localization in a city does not follow a random walk, but is instead a process that is mean-reverting in the face of a large and temporary shock. This further suggests that localization may not be characterized by multiple equilibria, since an enormous shock the size of the 1906 fire was not enough to alter the paths of localization experienced by industries in San Francisco in the early twentieth century.
Appendix

The localization index used in the study, developed in Mori et al. (2005), is a divergence statistic that is independent of sample size. Following the authors, for each industry $\iota$, the theoretical index value $D$ is given as

$$D(p_i \mid p_0) = \sum_{i=1}^{B} p_{ii} \ln \left( \frac{p_{ii}}{p_{0i}} \right),$$

(5)

where industries $\{\iota\}$ are located on $B$ city blocks. The probability of a randomly sampled industry-firm being located on block $i$ is given by $p_{ii}$ and the reference distribution, $p_{0i}$, is the probability that a randomly sampled industry-firm is located on block $i$ under spatial dispersion. In this study, the reference distribution is a uniform distribution. Since $p_{ii}$ is not directly observable, a sample estimate is given as $\hat{p}_{ii} = \frac{N_{ii}}{\sum_{j=1}^{B} N_{ij}}$, where industry $\iota$ has $N_{ii}$ firms on block $i$. This yields the following statistic:

$$D(\hat{p}_i \mid p_0) = \sum_{i=1}^{B} \hat{p}_{ii} \ln \left( \frac{\hat{p}_{ii}}{p_{0i}} \right) \approx D(p_i \mid p_0).$$

(6)

Thus, greater relative localization of firms (so that $\hat{p}_{ii}$ is high relative to $p_{0i}$) implies larger values of $D$. 
References


SEIC. State earthquake investigation commission. map of the city of san francisco showing the streets and the burnt area, 1906. Technical report, Carnegie Institution of Washington, 1908.


Figure 1: Fire Coverage from the 1906 San Francisco Disaster
Source: SEIC (1908); David Rumsey Historical Map Collection (www.davidrumsey.com).
Figure 2: Sample Page from the 1915 Business Directory
Source: Internet Archive (www.archive.org).
Figure 3: Primary Neighborhoods

Figure 4: Changes in Index Over Time

*Source:* See text.

*Notes:* FIRE refers to finance, insurance, and real estate.
### Table 1: SIC Divisions and Directory Classifications

<table>
<thead>
<tr>
<th>SIC division</th>
<th>Directory classifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manufacturing</td>
<td>Belting manufacturers; boiler makers; book printers; bookbinders; brass foundries; cabinet makers; canneries; chemical works; cigar manufacturers; confectioners; flour mills; furniture manufacturers; iron foundries; jewelry manufacturers; machinists; paper manufacturers; planing mills; sausage makers; ship builders; shirt manufacturers; soap makers; stationery manufacturing; steam and hydraulic packing; stove manufacturers; sugar refiners; tank manufacturers; tool manufacturers; watch case makers; wire manufacturers; wire cloth, netting, and rope manufacturers</td>
</tr>
<tr>
<td>Wholesale trade</td>
<td>Cigars and tobacco; engineer supplies; hardware; junk dealers; machinist supplies; stationery; tobacco leaf</td>
</tr>
<tr>
<td>Retail trade</td>
<td>Booksellers; cigars and tobacco; stationers; variety stores</td>
</tr>
<tr>
<td>FIRE</td>
<td>Banks; insurance companies—accident, fire, life, marine</td>
</tr>
<tr>
<td>Services</td>
<td>Blacksmiths; drayage; engineers—civil, consulting, electrical, mechanical, mining; horseshoers; upholsterers</td>
</tr>
</tbody>
</table>

*Notes: FIRE refers to finance, insurance, and real estate.*

### Table 2: Mean Levels of Geographic Localization, 1900-1930

<table>
<thead>
<tr>
<th>SIC division</th>
<th>1900</th>
<th>1905</th>
<th>1906</th>
<th>1915</th>
<th>1930</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manufacturing</td>
<td>4.93</td>
<td>4.82</td>
<td>5.25</td>
<td>4.96</td>
<td>4.87</td>
</tr>
<tr>
<td>Wholesale trade</td>
<td>4.93</td>
<td>4.78</td>
<td>5.24</td>
<td>5.09</td>
<td>5.44</td>
</tr>
<tr>
<td>Retail trade</td>
<td>2.79</td>
<td>2.85</td>
<td>3.82</td>
<td>3.46</td>
<td>3.75</td>
</tr>
<tr>
<td>FIRE</td>
<td>5.29</td>
<td>5.04</td>
<td>4.55</td>
<td>4.91</td>
<td>5.33</td>
</tr>
<tr>
<td>Services</td>
<td>4.00</td>
<td>4.02</td>
<td>4.28</td>
<td>4.10</td>
<td>4.92</td>
</tr>
</tbody>
</table>

*Notes: FIRE refers to finance, insurance, and real estate.*
The number of firms is reported in brackets.
Table 3: Correlation of Index Over Time, 1900-1930

<table>
<thead>
<tr>
<th>Manufacturing</th>
<th>1900</th>
<th>1905</th>
<th>1906</th>
<th>1915</th>
</tr>
</thead>
<tbody>
<tr>
<td>1905</td>
<td>0.95</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1906</td>
<td>0.60</td>
<td>0.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1915</td>
<td>0.79</td>
<td>0.80</td>
<td>0.80</td>
<td></td>
</tr>
<tr>
<td>1930</td>
<td>0.66</td>
<td>0.72</td>
<td>0.59</td>
<td>0.78</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Wholesale trade</th>
<th>1900</th>
<th>1905</th>
<th>1906</th>
<th>1915</th>
</tr>
</thead>
<tbody>
<tr>
<td>1905</td>
<td>0.97</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1906</td>
<td>0.68</td>
<td>0.59</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1915</td>
<td>0.69</td>
<td>0.72</td>
<td>0.88</td>
<td></td>
</tr>
<tr>
<td>1930</td>
<td>0.87</td>
<td>0.85</td>
<td>0.84</td>
<td>0.90</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Retail trade</th>
<th>1900</th>
<th>1905</th>
<th>1906</th>
<th>1915</th>
</tr>
</thead>
<tbody>
<tr>
<td>1905</td>
<td>0.85</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1906</td>
<td>0.31</td>
<td>0.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1915</td>
<td>0.34</td>
<td>0.76</td>
<td>0.98</td>
<td></td>
</tr>
<tr>
<td>1930</td>
<td>-0.10</td>
<td>0.38</td>
<td>0.85</td>
<td>0.89</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FIRE</th>
<th>1900</th>
<th>1905</th>
<th>1906</th>
<th>1915</th>
</tr>
</thead>
<tbody>
<tr>
<td>1905</td>
<td>0.95</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1906</td>
<td>0.46</td>
<td>0.38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1915</td>
<td>0.90</td>
<td>0.98</td>
<td>0.45</td>
<td></td>
</tr>
<tr>
<td>1930</td>
<td>0.65</td>
<td>0.67</td>
<td>0.72</td>
<td>0.78</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Services</th>
<th>1900</th>
<th>1905</th>
<th>1906</th>
<th>1915</th>
</tr>
</thead>
<tbody>
<tr>
<td>1905</td>
<td>0.97</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1906</td>
<td>0.76</td>
<td>0.82</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1915</td>
<td>0.77</td>
<td>0.86</td>
<td>0.82</td>
<td></td>
</tr>
<tr>
<td>1930</td>
<td>0.28</td>
<td>0.17</td>
<td>0.00</td>
<td>-0.01</td>
</tr>
</tbody>
</table>

Notes: The table reports the correlation between the MNS index values for different years.
Table 4: First-Stage Regression Results

<table>
<thead>
<tr>
<th></th>
<th>All divisions</th>
<th>Manufacturing</th>
<th>All divisions except retail</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>% firms burned</td>
<td>-0.408***</td>
<td>-0.077</td>
<td>-0.296***</td>
</tr>
<tr>
<td></td>
<td>(0.122)</td>
<td>(0.095)</td>
<td>(0.089)</td>
</tr>
<tr>
<td>% change in number of firms</td>
<td>-0.290***</td>
<td>-0.344***</td>
<td>-0.271***</td>
</tr>
<tr>
<td></td>
<td>(0.048)</td>
<td>(0.069)</td>
<td>(0.045)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.317***</td>
<td>-0.010</td>
<td>0.214**</td>
</tr>
<tr>
<td></td>
<td>(0.112)</td>
<td>(0.085)</td>
<td>(0.080)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.557</td>
<td>0.609</td>
<td>0.511</td>
</tr>
<tr>
<td>Number of observations</td>
<td>55</td>
<td>30</td>
<td>51</td>
</tr>
</tbody>
</table>

Notes: The dependent variable is the change in localization (in logs) between 1905 and 1906, time periods directly before and after the disaster. Robust standard errors are given in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels.
Table 5: Two-Stage Least-Squares Estimates

<table>
<thead>
<tr>
<th>Change in localization between 1905 and 1906</th>
<th>Dependent variable = change in localization between 1906 and 1915</th>
<th>Dependent variable = change in localization between 1906 and 1930</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All divisions (1)</td>
<td>Manufacturing except retail (2)</td>
</tr>
<tr>
<td>Change in localization between 1905 and 1906</td>
<td>-0.477***</td>
<td>-0.451***</td>
</tr>
<tr>
<td></td>
<td>(0.089)</td>
<td>(0.104)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.004</td>
<td>-0.019</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.015)</td>
</tr>
<tr>
<td>Number of observations</td>
<td>55</td>
<td>30</td>
</tr>
</tbody>
</table>

Notes: Robust standard errors are given in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels.