2006

Spatial Distribution of Frequency and Severity of Water Claims in California

Gurbhag Singh
Policy Research Bureau, singhg@insurance.ca.gov

Max Tang
California Department of Health Services, mtang1@dhs.ca.gov

Don McNeill
Policy Research Division, McNeillD@insurance.ca.gov

Lyn Hunstad
Policy Research Bureau, lyn_hunstad@Yahoo.com

Follow this and additional works at: http://digitalcommons.unl.edu/joap

Part of the Accounting Commons, Business Administration, Management, and Operations Commons, Corporate Finance Commons, Finance and Financial Management Commons, Insurance Commons, and the Management Sciences and Quantitative Methods Commons

http://digitalcommons.unl.edu/joap/10

This Article is brought to you for free and open access by the Finance Department at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Journal of Actuarial Practice 1993-2006 by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.
Spatial Distribution of Frequency and Severity of Water Claims in California

Gurbhag Singh,* Max Tang,† Don McNeill,‡ and Lyn Hunstad§

Abstract

We examine the frequency and severity of water loss claims for homeowners insurance across the state of California for the experience years 2000, 2001, and 2002. The spatial distribution patterns of frequencies and severities are mapped and analyzed at the zip code level. The maps reveal the pockets of high frequencies and severities. The information provided in this paper will assist actuaries and policy makers in their quest to set accurate rates for homeowners insurance.

Key words and phrases: exposure, credibility, homeowners insurance, high risk locations

*Gurbhag Singh, Ph.D. (University of Cincinnati), is a research specialist with the Policy Research Bureau.
Dr. Singh's address is: Policy Research Bureau, California Department of Insurance, Sacramento CA 95814, U.S.A. E-mail: singhg@insurance.ca.gov
†Max Tang, Ph.D. (University of Hawaii), is an actuary with the California Department of Health Services.
Dr. Tang's address is: Department of Health Services, Sacramento CA 95814, U.S.A. E-mail: mtang1@dhs.ca.gov
‡Don McNeill, B.Sc., M.B.A., is a research program specialist with the Policy Research Division.
Mr. McNeill's address is: Policy Research Bureau, California Department of Insurance, Sacramento CA 95814, U.S.A. E-mail: McNeillD@insurance.ca.gov
§Lyn Hunstad, B.Sc., is a retired former bureau chief of the Policy Research Bureau.
Mr. Hunstad's address is: Lyn Hunstad, P.O. Box 1869, Willits, CA 95490, U.S.A. E-mail: lyn_hunstad@yahoo.com

We thank the anonymous reviewer and the editor for their helpful comments; Nettie Hoge, Deputy Commissioner, for allowing us to submit the article to this journal; and Brandt Stevens, Ben Gentile, Luciano Gobbo, Bob Loo, and Camilo Pizarro of the California Department of Insurance for their help.
1 Introduction

There has been widespread concern about the performance of the homeowners insurance market in California in recent years. These concerns are the result of a variety of factors such as non-renewals of many long standing customers, access to homeowners insurance, and high insurance rates. Of importance to us, however, is the impact of water loss, i.e., water damage and claims resulting from accidental discharge of water in a home. Water loss can be on account of leakage or overflow of water from a home’s plumbing system or from appliances used daily such as washing machines, water heaters, refrigerators, leaky faucets, and leaky hoses.

Though water losses constitute roughly a third of homeowner's insurance claims (Table 1), there is no water loss insurance line per se as is the case with earthquake line of insurance or fire insurance. A basic water loss coverage is generally included in most homeowner insurance policies. More comprehensive coverage can be acquired by an endorsement of an existing homeowner policy.

Table 1
Homeowner's Water Versus Non-Water Claims
In California for 2000, 2001, and 2002

<table>
<thead>
<tr>
<th></th>
<th>Water</th>
<th>Non-Water</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Exposure</td>
<td>21,760,364</td>
<td>25,550,121</td>
</tr>
<tr>
<td>Total Claims</td>
<td>478,728 (36%)</td>
<td>843,154 (64%)</td>
</tr>
<tr>
<td>Total Losses</td>
<td>$1,834,135,735 (34%)</td>
<td>$3,536,400,460 (66%)</td>
</tr>
<tr>
<td>Frequency</td>
<td>0.022</td>
<td>0.033</td>
</tr>
<tr>
<td>Severity</td>
<td>$3,831</td>
<td>$4,194</td>
</tr>
</tbody>
</table>

Notes: Frequency = Number of claims/Number of house years of exposure and Severity = Incurred property losses ($) / Number of claims.

Given the role water plays in homeowners insurance in California, we are surprised to have found that there has been no study of the spatial distribution of water claims across California. Our objective is to provide information about the frequency and severity of water loss claims in California. We do so by providing data on water claims at the zip code level and by identifying the geographic areas with high risk of water claims. While homeowners insurance of water peril is not, in general, priced separately from other perils, knowledge of water loss costs and the distribution of these costs across the state would significantly
facilitate the rate making process. As a result, this information will be especially helpful for actuaries and underwriters when evaluating risks and determining premium rates.

The paper is organized as follows: Section 2 describes the data and the methodology used. The main results are presented in Section 3, while Section 4 provides areas for further research.

2 The Data

The data source for this study is the earthquake and fire data call (EF-2002), a special data call, and an addendum data call, with each call for the experience years 2000, 2001, and 2002. The special data call was sent to over 535 insurance companies that write homeowners insurance in California. The companies provided data at the zip code level on total exposure, total incurred property loss, total claim counts, total losses for water damages, and claim counts for water damages. After the data were received from the insurance companies, the data were edited and checked for accuracy, which was a time consuming process. In order to expedite the completion of our project, the data from the 13 top (in terms of market share) homeowner insurance companies were used. These companies comprised a little over 77% of the homeowners insurance market. The data for three experience years 2000, 2001, and 2002 from the 13 companies are combined at zip code level to produce a statewide data set consisting of 1812 observations. Note that the data from different years were not adjusted for inflation. The data for the maps and the tables were collected through the data call. GIS (Geographic Information System) is used as our mapping tool.

As the zip code is the basic geographic unit used in this study, if the zip code data are sufficiently sparse, its data are adjusted for full credibility using limited fluctuation theory. This study uses the classical credibility approach (also known as limited fluctuation credibility) for adjusting the zip code data for credibility. The rationale for selecting this credibility approach is its simplicity: it uses relatively uncomplicated formulas and provides reasonable results. Additionally, many company actuaries use this approach in practice. Limited fluctuation credibility approach can be briefly described as follows.

Suppose we are interested in estimating the severity, i.e., the average water claim per unit of exposure. Let \( \bar{X} \) denote the severity in a single zip code with exposure base \( n \). Our problem is to find \( n \) such that

\[
\mathbb{P} \left[ (1 - k) \mathbb{E}[\bar{X}] \leq \bar{X} \leq (1 + k) \mathbb{E}[\bar{X}] \right] \geq p,
\]
where $\bar{X}$ has mean $\mu_X$ and variance $\sigma^2_X/n$. Again, using the normal approximation for $\bar{X}$ yields:

$$n \geq \left( \frac{\zeta_{1-\epsilon/2}}{k} \right)^2 \left( \frac{\sigma_X}{\mu_X} \right)^2,$$

which is the minimum exposure needed for full credibility. The popular standard for full credibility, which is based on $p = 90\%$ and $k = 5\%$ and a coefficient of variation of $1$, is a minimum exposure of $n = 1082$. However, following Mahler and Curtis (2001, pages 492-498) we use the value of $1,082$ claims corresponding to $p = 90\%$ and $k = 5\%$ as is commonly used in applications. In the rest of this paper we will use $1082$ claims as the standard for full credibility.

If the number of claims in a zip code is less than $1082$, i.e., not large enough to give full credibility, a credibility estimate of the severity ($\bar{X}$) is constructed. This is done by first placing zip codes into relatively homogeneous groups called regions. (Table 2 shows the 20 regions constructed for the state of California.) A credibility factor $Z (0 \leq Z < 1)$ and its complement $(1 - Z)$ are then used such that

$$\bar{X} = Z \bar{X} + (1 - Z) \mu_R,$$

where $\bar{X}$ is the sample mean from the zip code's experience and $\mu_R$ the collective sample mean from the region the zip code is assigned. We use the credibility factor

$$Z = \sqrt{\frac{\text{Number of Claims}}{1082}}.$$

As an example, consider the adjustment made to severities using Los Angeles Area for zip code 90717. This zip code has a severity of $4,399$ total water damages of $1,087,984$ and $249$ claims leading to a severity of $\bar{X} = 1,087,984/249 = 4,369$. This zip code belongs to Region 1, which has a severity of $\mu_R = 4,399$. Because the number of claims is less than $1082$, this zip code requires an adjustment for credibility. From equation (3), $Z = \sqrt{249/1082} = 0.47972$. Thus the adjusted severity is

$$\hat{X} = 0.47972 \times 4369 + (1 - 0.47972) \times 4399 = 4385.$$
For more on limited fluctuation theory and on credibility theory in general, see, for example, Goulet (1998), Mahler and Dean (2001), Klug­man, Panjer and Willmot (2004), and references therein. Boor (1996) provides a good treatment of the concept of complement of credibility and points out the basic principles that should be considered for selecting the information that receives the complement of credibility.

After the credibility adjustments, the zip codes and their corre­sponding counties are grouped into five geographic areas in the state: Northern California, Central California, Southern California, Bay Area (the San Francisco Oakland Bay), and Los Angeles Area. These geo­graphic areas are constructed based on member counties sharing com­mon characteristics, such as geographic location, degree of urbaniza­
Table 3
Grouping of California Counties into Geographic Areas

<table>
<thead>
<tr>
<th>Geographic Area</th>
<th>Counties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northern California</td>
<td>Napa, Sonoma, Lake, Mendocino, Humboldt, Trinity, Del Norte, Siskiyou, Yolo, Modoc, Lassen, Tehama, Shasta, Plums, Sierra, Nevada, Placer, El Dorado, Amador, Alpine, Butte, Yuba, Sutter, Sacramento, Colusa, Glenn</td>
</tr>
<tr>
<td>Central California (Non-Bay Area)</td>
<td>Tulare, Kings, Madera, Merced, Stanislaus, Fresno, San Joaquin, Calaveras, Monterey, San Benito, Tuolumne, Mariposa, Mono, Inyo</td>
</tr>
<tr>
<td>Bay Area Region</td>
<td>San Francisco, Santa Cruz, Santa Clara, Alameda, San Mateo, Contra Costa, Solano, Marin</td>
</tr>
<tr>
<td>Southern California</td>
<td>Kern, San Luis Obispo, Santa Barbara, Ventura, Orange, San Diego, Imperial, Riverside, San Bernardino</td>
</tr>
<tr>
<td>Los Angeles</td>
<td>Los Angeles County</td>
</tr>
</tbody>
</table>

tion, metro areas, and agricultural areas. Table 3 shows the counties grouped into geographic areas.

3 The Main Results

Table 4 displays the zip codes and corresponding cities with the highest ranges of frequency and severity. Note that the Southern California region has the most high frequency and high severity locations. Figures 1 to 6 show the water claims frequency for the entire state and the five geographic areas. Figures 7 to 12 show the water claims severities for the entire state and the five geographic areas.
Table 4
Highest Frequency and Severity Pockets of Water Claims in California by Region, City, and Zip Code

<table>
<thead>
<tr>
<th>Region</th>
<th>City</th>
<th>Frequency</th>
<th>Severity</th>
</tr>
</thead>
<tbody>
<tr>
<td>North. Cal.</td>
<td>Brentwood</td>
<td>94513</td>
<td>Chester</td>
</tr>
<tr>
<td></td>
<td>Echo Lake</td>
<td></td>
<td>96020</td>
</tr>
<tr>
<td></td>
<td>Markleeville</td>
<td></td>
<td>95721</td>
</tr>
<tr>
<td></td>
<td>Norden</td>
<td></td>
<td>96120</td>
</tr>
<tr>
<td></td>
<td>Portola</td>
<td></td>
<td>95724</td>
</tr>
<tr>
<td></td>
<td>Truckee</td>
<td></td>
<td>96122</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>96161</td>
</tr>
<tr>
<td>Centr. Cal.</td>
<td>Elk Grove</td>
<td>95758</td>
<td>Citrus Heights</td>
</tr>
<tr>
<td></td>
<td>Tracy</td>
<td>95304</td>
<td>95610</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>95619</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>95630</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>95667</td>
</tr>
<tr>
<td>Bay Area</td>
<td>Antioch</td>
<td>94509</td>
<td>Inverness</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>94937</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>94548</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>90265</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>90272</td>
</tr>
<tr>
<td>L.A. Area</td>
<td>Compton</td>
<td>90220</td>
<td>Agoura Hills</td>
</tr>
<tr>
<td></td>
<td>Carson</td>
<td>90745</td>
<td>91301</td>
</tr>
<tr>
<td></td>
<td>Gardena</td>
<td>90248</td>
<td>91302</td>
</tr>
<tr>
<td></td>
<td>Walnut</td>
<td>91789</td>
<td>91311</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>91316</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>92651</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>92211</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>90274</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>90290</td>
</tr>
<tr>
<td>South. Cal.</td>
<td>Colton</td>
<td>92324</td>
<td>Capistrano Beach</td>
</tr>
<tr>
<td></td>
<td>Laguna Niguel</td>
<td>92677</td>
<td>92624</td>
</tr>
<tr>
<td></td>
<td>Moreno Valley</td>
<td>92553</td>
<td>92629</td>
</tr>
<tr>
<td></td>
<td>Perris</td>
<td>92571</td>
<td>92612</td>
</tr>
<tr>
<td></td>
<td>Rancho Santa Marg.</td>
<td>92688</td>
<td>92037</td>
</tr>
<tr>
<td></td>
<td>Riverside</td>
<td>92507</td>
<td>92064</td>
</tr>
<tr>
<td></td>
<td>San Juan Capis.</td>
<td>92675</td>
<td>92270</td>
</tr>
<tr>
<td></td>
<td>Silverado</td>
<td>92676</td>
<td>92675</td>
</tr>
</tbody>
</table>

Notes: North. Cal. = Northern California; Centr. Cal. = Central California, which excludes the Bay Area; L.A. Area = Los Angeles Area; South. Cal. = Southern California, which excludes the Los Angeles Area.
Water Claim Frequencies Per 1000 Exposure Yrs.
State of California

Legend
- 0 - 0
- 1 - 19
- 20 - 23
- 24 - 27
- 28 - 31
- 32 - 36
- 37 - 42
- 43 - 49
- 50 - 61
- 62 - 83

Figure 1: Water Claim Frequencies for State of California

Notes: Water claim frequencies range from zero to 83 per 1,000 exposure years across California. In general, the Northern California Region has lower claim frequencies (0 to 42 per 1,000 exposure years) than the Southern California Region where a greater portion lies within the range 24 to 83 per 1,000 exposure years. In the Central California Region frequencies vary from 0 to 42 per 1,000 exposure years, around the Bay Area the range is from 1 to 36 per 1,000 exposure years while in the Los Angeles Area water claim frequencies range from 19 to 83 per 1,000 exposure years. Broadly speaking, water claim frequencies increase as we head south from Northern California to Southern California, with the highest claim frequencies at 83 per 1,000 exposure years.
Water Claim Frequencies Per 1000 Exposure Yrs.
Northern California

Figure 2: Water Claim Frequencies for Northern California

Notes: Water claim frequencies vary from 0 to 42 per 1,000 exposure years in this region with an average of 19 per 1,000 exposure years. The spatial pattern of distribution of water claim frequencies in Northern California shows that the water claim frequencies are higher (from 20 to 42 per 1,000 exposure years) in the south central areas and taper toward the surrounding coastal, northern, and Sierra Nevada areas where values vary from 0 to 19 per 1,000 exposure years. The reason for such a geographic distribution is not clear.
Water Claim Frequencies Per
1000 Exposure Yrs.
Central California

Figure 3: Water Claim Frequencies for Central California

Notes: This geographic region’s range of water claim frequencies is similar to the Northern California Region. Most of this region displays frequencies in the range of 0 to 36 per 1,000 exposure years with an average of 20 per 1,000 exposure years. With the exception of the Bay Area described in Figure 4, the highest range of water claim frequencies in the Central California Region is from 37 to 42 per 1,000 exposure years. Only a few zip codes have frequencies in this highest range, and these zip codes are located east of San Francisco Bay.
Figure 4: Water Claim Frequencies for Bay Area

Notes: The water claim frequencies distribution in the Bay Area generally varies from 0 to 49 per 1,000 exposure years with an average of 19 per 1,000 exposure years. Only a single zip code had no water claims. The frequencies increase from the coastal area and San Francisco Bay area near the cities of Berkeley, Oakland, San Leandro in Contra Costa, Alameda, and Solano counties inland. Around the city of San Leandro the water claim frequencies are highest, ranging from 43 to 49 per 1,000 exposure years. Compared to the state's highest water claim frequencies, however, the Bay Area has a moderate range.
Water Claim Frequencies Per 1000 Exposure Yrs.
Southern California

Figure 5: Water Claim Frequencies for Southern California

Notes: Southern California has the widest range of frequencies (0 to 83 per 1,000 exposure years) and an average of 22 per 1,000 exposure years. Most of the region has water claim frequencies over 23 per 1,000 exposure years, and there are several pockets in the highest frequency range of 62 to 83 per 1,000 exposure years. The spatial distribution of water claim frequencies is higher in the southern and southwestern part of the region compared to the north part of the region. The areas of very high concentration are in Orange, Riverside, and San Bernardino counties. Specifically, these pockets of highest water claim frequencies are in the following cities: Silverado, San Juan Capistrano in Orange County, Perris, Riverside in Riverside County, and Colton in San Bernardino County (Table 4).
Figure 6: Water Claim Frequencies for Los Angeles Area

Notes: Though water claim frequencies in the Los Angeles Area range from 0 to 61 claims per 1,000 exposure years with only a few zip codes below a frequency of 1 or less, the majority of the area has over 23 claims per 1,000 exposure years. About a dozen zip codes have frequencies in the range of 37 to 42 claims per 1,000 exposure years while six zip codes have claim frequencies between 43 and 49. Also, several zip codes have water claim frequencies between the ranges of 50 to 61. Only Southern California has a wider range of claim frequencies than the Los Angeles Area.
Figure 7: Water Claim Severities for State of California

Notes: The severity of claims varies considerably by zip code. The average severity across the state is $3,719. The highest severity values in California range from $6,503 to $11,138. Though the number of high severity pockets is small, Northern California, Los Angeles, and Southern California have more pockets of high severities than Central California and the Bay Area. In fact, Southern California and Los Angeles Area have close to 60% of these pockets, while the Bay Area has only 7%.
Figure 8: Water Claim Severities for Northern California

Notes: In the Northern California Region water claim severities range from $0 to $6,768 with an average of $3,480, though a sizeable proportion of the severities is under $4,128. The spatial pattern of distribution of water claim severities shows that the water claim severities are higher along the counties bordering Nevada in the Sierra area and lower in both the northern coastal areas and northern part of this region. Many parts of these Sierra counties have severities over $4,128. It seems that environmental conditions such as frost impact the claim severities in this part of California.
Figure 9: Water Claim Severities for Central California

Notes: Here severities range from $0 to $6,793 with an average of $3,480. A sizeable part of this area has severities in the $4,128 to $4,659 range with a small number of pockets in the $5,375 to $6,793 range. Claim severities in the range of $2,400 to $2,835 are predominant in a large portion of the central part of this region. A very small area with high severities is located in the southwestern portion of this region.
Singh et al: Spatial Distribution of Water Claims

Figure 10: Water Claim Severities for Bay Area

Notes: Water claim severities distribution in most of the Bay Area varies from $0 to $11,138 with an average of $3,702. The highest range of severities is in the range $6,503 to $11,138 and is located in Inverness, a city in Marin County. Only a single location has zero water claim severity and it is Burlingame, a city in San Mateo County. In general, severities increase from north to south in the Bay Area region.
Figure 11: Water Claim Severities for Southern California

Notes: This is the geographic region where the water claim severities are relatively higher than other regions (highest severity is $10,261 and average is $3,736). About one third of the total pockets of the highest severities in the state are concentrated in the Southern California Region. The pockets of highest severities are found near the coastal area where expensive homes are located. In general, the severities decline as we head inland from the coast. It appears that the proximity to water affects the pattern of distribution of severities for this region.
Figure 12: Water Claim Severities for Los Angeles Area

Notes: Though the range of water claim severities varies from almost from $0 to $7,591 with an average of $4,280. Next to Southern California, this region has the largest number of highest severity pockets. The major portion of this region has water claim severities over $3,247. The pockets of high severity are located in the coastal region. It appears that the proximity to water affects the pattern of distribution of severities for this region too.
4 Closing Comments

Overall we found that water claims frequencies are higher in the metropolitan areas such as Los Angeles, San Diego, San Francisco, and Sacramento, while lower in the rural areas. Also, the water frequencies are higher in the Southern California Region and lower in the Northern California Region. The largest concentration of the pockets of the highest frequencies is in the Southern California Region with 8 of the 16 pockets of highest frequencies. Southern California and Los Angeles Regions have 17 of the 29 pockets of highest severities.

A number of unanswered questions remain to be addressed: (i) identifying and analyzing the underlying factors that affect the spatial distribution pattern of water frequencies and severities; (ii) can additional understanding be gained about the distribution patterns of claim frequencies and severities by changing the geographic unit from zip code to CCD (Census County Division) or county; (iii) what other models can be used to adjust the data for credibility? Additional data will be required, however, to address these questions.

References


