Pricing Earthquake Exposure Using Modeling

Debra L. Werland
United Services Automobile Association, Debra.Werland@usaa.com

Joseph W. Pitts
GAINSCO, Inc, management@gainsco.com

Pricing Earthquake Exposure Using Modeling

Debra L. Werland* and Joseph W. Pitts†

Abstract‡

This paper demonstrates a practical methodology for determining a statewide rate level indication for the earthquake insurance and for determining more equitable territorial relativities within a state. The methodology is based on the output from a certain commercially available earthquake modeling software package. The methodology addresses some of the complex issues involved in pricing earthquake insurance exposure and potential regulatory acceptance. The paper also features a section dealing with the net cost of reinsurance in the proposed direct rates. A final consideration is the treatment of a model's output when it is believed the modeled results are less than fully credible.

Key words and phrases: *catastrophe modeling, reinsurance, target rate of return, zone relativities*

*Debra L. Werland, F.C.A.S., M.A.A.A., is executive director of homeowners and property pricing actuary for United Services Automobile Association. She has co-authored a paper entitled "Using a Geographic Information System to Identify Territory Boundaries" that appeared in *The Casualty Actuarial Society Forum* (Winter, 1996).

Ms. Werland's address is: United Services Automobile Association, USAA Building B-1-E, San Antonio TX 78288. Internet address: *Debra.Werland@usaa.com*

†Joe W. Pitts, F.C.A.S., M.A.A.A., is vice president and chief actuary for GAINSCO, Inc. He currently serves on the Casualty Actuarial Society Exam Committee.

Mr. Pitts' address is: GAINSCO, Inc., 500 Commerce, Fort Worth TX 76102. Internet address: *management@gainsco.com*

‡This paper is based on a previous paper entitled "Pricing Earthquake Exposure Using Modeling" that appeared in *The Casualty Actuarial Society Forum* (Winter, 1997), a nonrefereed publication of the Casualty Actuarial Society.

References to and descriptions of some of the inner workings of the earthquake computer simulation model developed by Applied Insurance Research, Inc. of Boston, Massachusetts, are done with their express written permission.

1 Introduction

Pricing hurricane and earthquake risk has never been an easy task. No insurer's loss history is adequate to cover the expectation of all possible type and size of events. Any ratemaking formula based on actual loss experience for such rare events will fail to capture the scope of possible events that could affect an insurer's financial results. Catastrophe hazard modeling represents a way of developing the scope of possible catastrophic events. The financial impact of these events is based on characteristics of the underlying peril and their interaction with the insured properties.

Actuaries are relying more than ever on the use of modeling in pricing catastrophic risks such as hurricanes and earthquakes. As a result, catastrophe hazard modeling has become an important tool for ratemaking in lines of business subject to low frequency, high severity type losses. Natural hazard events such as hurricanes and earthquakes rarely occur, but their devastation can be overwhelming when they do. Few insurance companies have enough historical loss data to sufficiently price these events.

In this paper we will focus on the earthquake peril and its pricing. The approach adopted is to use an earthquake computer simulation model. In particular we use an earthquake model developed by Applied Insurance Research, Inc. (AIR) of Boston, a leading computer simulation/modeling firm. While it is not necessary for one to completely understand the intricacies of all functions and assumptions used in the simulation model, it is important nonetheless to present an overview of the AIR model. Briefly, the AIR earthquake model is composed of three separate component models: an earthquake occurrence model, a shake damage model, and a fire-following model. The overall model uses sophisticated mathematical techniques to estimate the probability distribution of losses resulting from earthquakes anywhere in the 48 contiguous states. The AIR earthquake model is described in more detail later in the appendix.

For ratemaking purposes, the output from the model includes loss costs applicable to a specific location, type of construction, and policy form. Our interest is in a single family dwelling as covered under a typical homeowners policy. The loss costs generated by the model are the basic building blocks in the development of an appropriate rate.

We will discuss target underwriting profit provisions, reinsurance costs, and other components of developing an adequate rate per $1,000 of dwelling coverage for a typical book of homeowners business. The credibility of the results will be addressed in the derivation of the indicated rates, and the state will be partitioned into geographic zones based on the relative difference in loss costs determined from the modeled results.

2 Proposed Methodology

The goal of this paper is to present a methodology for developing a rate per $1,000 of earthquake coverage. We assume that the indicated rate is based on Coverage A (the dwelling limit of a typical homeowners single family dwelling). The modeled results include all coverages (dwelling, other structures, personal property, time element expenses), and the figures have been ratioed to Coverage A, in 1000s.

2.1 Statewide Indicated Rate

The statewide indicated rate is determined using the pure premium method. The losses are based on an insurer's own exposure distribution within the state. The first input into the methodology is the statewide modeled expected losses stated at a base deductible level. In this example the base deductible is 10 percent applicable to the dwelling limit. The expected annual losses represent the average annual amount of losses an insurer could expect from writing the earthquake line of business in state X if each insured had a 10 percent deductible.

The modeled results are generally available on an individual state basis as well as on a zip code or county basis within the state. The expected annual losses are trended (severity only) and adjusted for loss adjustment expense (LAE), then ratioed to the total trended value of insured dwellings to develop a projected pure premium which is used to determine the indicated rate as shown in Table 1. (A viable alternative would be to trend the insured values first and use these trended values as input to the catastrophe model, thus yielding an estimate of trended severity within the model results). In this example, the current rate is assumed to be $2.50 per $1,000 of dwelling coverage. The indicated rate is calculated by taking the projected pure premium and grossing it up to include reinsurance costs net of reinsurance recoveries, trended fixed expenses, and variable expenses. These calculations show that the indicated statewide rate is $3.77 per $1,000 of dwelling coverage.

Some of the rows of Table 1 are described in more detail as follows:

(1) This is the main output received from the modeling firm. It is an estimate of the expected annual losses at a base deductible for an

Table 1
Statewide Indicated Rate

(1)	Modeled Expected Annual Losses, 10% Deductible, 12/31/95	$19,500,000
(2)	Total Dwelling Coverage, 12/31/95	$10,965,281,000
(3)	Proposed Effective Date	7/1/96
(4)	LAE Factor	1.150
(5)	Loss Trend Factor Trended to 7/1/97	1.250
(6)	Exposure Trend Factor Trended to 7/1/97	1.190
(7)	State X Earthquake Share of Net Cost of Reinsurance	$7,592,703
(8)	Trended Fixed Expense Provision Per $1000 of Coverage	0.265
(9)	Pure Premium Per $1000 of Coverage	$2.99
(10)	Variable Permissible Loss and LAE Ratio	0.794
(11)	Indicated Rate: (9)/(10)	$3.77
(12)	Current Statewide Rate Per $1000 of Dwelling Coverage	$2.50
(13)	Indicated Percentage Change: (11)/(12) − 1	50.8%
(14)	Proposed Change	50.8%
(15)	Proposed Statewide Rate: (12) × [1 + (14)]	$3.77

insurer, given the current book of business within the state for the earthquake line of business;

(2) The total value of insured dwellings is provided to the modeling firm by the insurer and is used to determine the average expected annual losses per $1,000 of coverage in the pure premium method;

(3) The proposed effective date as selected by the insurer;

(4) The LAE factor is calculated based on a comparison of estimated ultimate loss adjustment expenses to estimated ultimate losses from the most recent earthquake events faced by the insurer;

(5) The modeled losses are trended using historical homeowners severity data. Earthquake loss trend data are not used because of their instability. Losses should not be trended for frequency, unless the insurer is confident there exists an increased period of seismicity in the future;

(6) The exposure trend is based on historical changes in the average amount of insurance for the earthquake line of business;

(7) The state X earthquake share of the expected net cost of reinsurance is calculated as described in Table 2;

(8) The trended fixed expense provision per $1,000 of coverage is calculated by trending fixed expenses to a point in time appropriate for the proposed effective date and dividing it by trended insured value, using an annualized fixed expense trend of 5 percent;

(9) The formula for Row (9) is:

$$\text{Pure Premium per } \$1{,}000 = (8) + \frac{[(1) \times (4) \times (5) + (7)] \times 1000}{(2) \times (6)},$$

which combines the modeled expected losses with the net cost of reinsurance for the state and line of business with the trended fixed expense provision to provide an estimate of the projected pure premium to be expected during the time the proposed rates are to be in effect; and

(10) The variable permissible loss and LAE ratio are calculated based on historical variable expenses and a consideration of the relative riskiness of the earthquake line of business compared to other lines being written and the overall required return on surplus. An 18.2 percent underwriting profit provision is used along with a 2.4 percent provision for variable expenses.

2.2 Net Cost of Reinsurance

An important component that we reflect in the rate indication is the net cost of reinsurance. An insurer should decide whether to include this component based on the costs and anticipated recoveries associated with its reinsurance program. The net cost of reinsurance should be included as a cost if the expected reinsurance recovery is less than the amount of premium paid to the reinsurer for reinsurance protection. This relationship generally will be the case due to the presence of transaction costs that include a margin for reinsurance risk load and profit.

The expected reinsurance recovery represents the average annual amount an insurer could expect to recover from the reinsurer(s) due to insured events and can be determined using catastrophe modeling. The expected reinsurance recovery needs to be calculated considering the attachment points or quota share percentages associated with an insurer's reinsurance program. An insurer's reinsurance program often is structured to provide protection against many types of hazards; however, some reinsurance contracts are designed to provide protection against only one hazard.

To accurately measure the net cost of reinsurance for a particular hazard, the reinsurance premium from all programs that provide protection for the hazard should be included. If other catastrophic hazards such as hurricanes are a large proportion of an insurer's exposure to catastrophe loss, the reinsurance premium for multihazard contracts could be segregated for each hazard. The reinsurance premium for each hazard then could be included with each net cost of reinsurance calculation for every line of business. In the example, however, the net cost of reinsurance is allocated to the earthquake line of business and to the appropriate state.

The allocation to line of business in the example shown in Table 2 is based on model results comparing expected earthquake reinsurance recovery to the total expected reinsurance recovery. This ratio is applied to the net cost of reinsurance to obtain the earthquake-only net cost of reinsurance. The allocation to a state level uses earthquake written premium. This allocation may introduce a distortion if the state in question has a different level of premium adequacy than countrywide premium adequacy. In addition, a premium base allocation may not adequately represent the riskiness of expected earthquake losses by state.

The rows of Table 2 are described in more detail as follows:

(1) This is the total of all reinsurance premium paid for reinsurance contracts that provide protection for earthquake losses;

(2) This is a model output number. It is determined based on the attachment point or quota share arrangement an insurer has with its reinsurer(s);

(3) The net cost of reinsurance is the difference between the reinsurance premium paid for contracts providing earthquake protection and the expected total reinsurance recovery;

(4) Model results are used to determine what portion of the expected recovery is due to earthquake;

(5) The earthquake proportion of the total expected reinsurance recovery is expressed as a factor to be applied to the total net cost of reinsurance;

Table 2
Estimated Net Cost of Reinsurance

(1)	1995 Countrywide Reinsurance Premium for Contracts Covering the Earthquake Peril	\$37,890,000
(2)	Expected Reinsurance Recovery	\$17,481,970
(3)	Net Cost of Reinsurance: (1) − (2)	\$20,408,030
(4)	Expected Earthquake Reinsurance Recovery	\$9,154,600
(5)	Proportion of Earthquake Recovery to Total Recovery: (4)/(2)	52.4%
(6)	Earthquake Share of Net Cost of Reinsurance: (3) × (5)	\$10,693,808
(7)	1995 State X Earthquake Written Premium	\$27,271,677
(8)	1995 Countrywide Earthquake Written Premium	\$38,551,154
(9)	State X Earthquake Share of Net Cost of Reinsurance: [(7)/(8)] × (6)	\$7,592,703

(6) The earthquake share of the net cost of reinsurance is the proportion of the earthquake recovery to the total recovery multiplied by the total net cost of reinsurance;

(7) The latest year state X earthquake written premium is used to allocate the earthquake share of the net cost of reinsurance to a state level; and

(8) The latest year countrywide earthquake written premium is used to find what proportion is represented by state X. Each state's written premium is first adjusted to current rate levels, if applicable.

The concept of including the net cost of reinsurance in a rate indication is relatively new and likely will be challenged or subjected to additional scrutiny by regulatory agencies. It does represent a cost of doing business, however; therefore, we include its net costs. Reinsurance costs also may be considered in conjunction with the selected rate of return.

2.3 Target Rate of Return

To develop an underwriting profit provision, we choose a total rate of return methodology. We are not proposing one method over another; we have selected this particular method for the development of a reasonable profit target for the earthquake line of business. The target rate of return on GAAP equity is developed using a discounted cash flow (dividend yield) method and the capital asset pricing model (CAPM). The selected rate of return, averaged from the results of these two methods, is 13.0 percent. From this selected rate of return we have subtracted 8.0 percent (which represents the post-tax investment rate of return from all investable funds). Table 3 converts this difference to a pre-tax basis, using a corporate tax rate of 35 percent. For an insurer's total book of business this percentage is divided by the company's premium-to-surplus ratio to convert the target underwriting profit provision to a percentage of premium. Although we do not endorse the divisibility of surplus or leverage ratios, we propose this method for calculating a reasonable earthquake underwriting profit provision.

We have selected a company whose underwriting results resemble the years 1985-1994 for all property and casualty insurers writing personal lines automobile, homeowners multiperil, and earthquake coverages. (It would be appropriate for more years to be used; however, the earthquake line of business was not segregated prior to 1985). The data are from *Best's Aggregate and Averages*. A company's own data also can be used for this purpose.

Table 3
Target Underwriting Profit Provision

A. Target Rate of Return (% of GAAP Surplus)		
1.	Dividend Yield Model	12.0%
2.	Capital Asset Pricing Model	14.0%
3.	Selected Target Rate of Return	13.0%
B. Target Underwriting Rate of Return (% of GAAP Surplus)		
1.	Investment Rate of Return After Tax	8.0%
2.	Target U/W Return After Tax $(A3) - (B1)$	5.0%
3.	Target U/W Return Before Tax $(B2)/(1 - 0.35)$	7.7%
C. Target Underwriting Profit Provision (% of Direct Earned Premium)		
1.	Net Written Premium/GAAP Surplus Ratio	1.30
2.	Indicated U/W Profit Provision $(B3)/(C1)$	5.9%
3.	Selected U/W Profit Provision	5.9%

Note: Insurers are chosen that resemble the mix of business written by the filing insurer. Company betas and projected dividend yields are from Value Line. Both the dividend yield method and CAPM are used in determining an appropriate rate of return. The selected target rate of return is a straight average of the two methods.

A company's underwriting profit provision should vary based on the riskiness of the line of business. A measure of risk we have chosen is the coefficient of variation (measured as standard deviation/mean) of a series of underwriting results for each line. Alternatively, combined ratios could be used, where a 100.0 combined ratio reflects a 0 percent underwriting result. Because the selected period includes the effects of Hurricane Andrew and the Northridge Earthquake, we adjust the losses so that Andrew reflects a 1-in-30 year event and Northridge a 1-in-50 year event. We did not adjust for Hurricane Hugo.

Table 4 shows the industry's yearly (1985-1994) underwriting gains and losses as a percent of net earned premium. Table 5 shows the coefficient of variation of each line, the weighted average of the coefficients of variation using the latest ten years of premium, and a risk index (the ratio of each line's coefficient of variation to the weighted coefficient of variation).

Table 4
Annual Underwriting Results as a Percentage of Premium

Year	Private Passenger Automobile	Homeowners Multiperil	Earthquake
1985	-11.0%	-11.7%	60.0%
1986	-8.3%	-3.5%	58.0%
1987	-6.0%	3.3%	44.2%
1988	-6.8%	0.0%	57.5%
1989	-8.9%	-13.9%	-42.1%
1990	-9.1%	-12.9%	43.8%
1991	-4.6%	-17.7%	55.3%
1992	-1.9%	-58.4%	61.4%
1993	-1.8%	13.5%	68.0%
1994	-1.3%	-18.4%	-222.2%

Assume the company's premium-to-surplus ratio corresponds to the industry's at 1.30, so that its inverse is 0.77. The risk indices are used to adjust each line's surplus ratio (surplus-to-premium) in the total rate of return methodology, resulting in target underwriting profit provisions that reflect the risk of each line of business. The resulting earthquake profit provision will be used in the derivation of the variable permissible loss and loss adjustment expense provision. Table 6 summarizes this information.

Table 5
Coefficient of Variation (CV) and Risk Index (RI)

Line of Business	PD	CV*	RI
Private Passenger Automobile	80.1%	0.550	0.92
Earthquake	0.5%	1.854	3.09
Homeowners Multiperil	19.4%	0.780	1.30
Total	100.0%	0.600	1.00

Notes: PD = Premium Distribution; *Absolute value.

Table 6
Target Underwriting Profit Provision

Line of Business	RI	S/P	TUPP
Private Passenger Automobile	0.92	0.71	5.4%
Earthquake	3.09	2.38	18.2%
Homeowners Multiperil	1.30	1.00	7.7%
Total	100.0%	0.77	5.9%

Notes: RI = Risk Index; S/P = Implied Surplus Ratio; TUPP = Target Underwriting Profit Provision.

In this example industry net underwriting results are used to determine an appropriate underwriting profit provision for the earthquake line of business. A larger earthquake underwriting profit provision would result if direct results were used. The variability of net underwriting results is removed by the stabilization of reinsurance. Using our methodology it is reasonable to conclude that part of the difference between underwriting profit provisions calculated using net or direct underwriting results would be due to reinsurance costs. An insurer should expect a lower net cost of reinsurance if part of the reinsurance cost is reflected in the earthquake underwriting profit provision calculated using direct underwriting results. Efforts could be made to quantify what portion of the net cost of reinsurance is contained in an earthquake underwriting profit provision based on direct underwriting results. One possible approach would be to compare the difference in earthquake underwriting profit provisions calculated using net and direct underwriting results to a net cost of reinsurance as calculated in this example.

2.4 Zone Relativities

Model results also can be used to determinc revised earthquake zone definitions and earthquake zone relativities. The data used to establish earthquake zone definitions are model results at a five digit zip code level. The sum of all the five digit zip code modeled losses and dwelling insured values should balance to the statewide totals used to determine the statewide indicated rate.

Table 7
State X Earthquake Model Results, Zip Code Level

Zip	DIV (in $000)	EAL at 10% Deductible	Loss Cost
1	$ 921,339	$ 2,303,348	$2.50
2	1,096,528	1,644,792	1.50
3	258,481	387,722	1.50
4	548,264	603,090	1.10
5	922,272	830,045	0.90
6	79,839	98,897	1.24
7	722,114	902,643	1.25
8	103,211	232,225	2.25
9	803,112	3,011,670	3.75
10	801,247	721,122	0.90
11	552,322	359,009	0.65
12	402,178	623,376	1.55
13	700,659	1,156,087	1.65
14	1,102,321	2,369,990	2.15
15	200,321	490,786	2.45
16	402,111	1,105,805	2.75
17	727,727	1,928,477	2.65
18	202,001	490,786	1.03
19	112,007	123,768	1.11
20	307,227	399,088	1.30
Total	$10,965,281	$ 19,500,000	$1.78

Notes: Zip = Five Digit Zip Code Area; DIV = Dwelling Insured Value; and EAL = Expected Annual Loss.

In the example we assume the state comprises 20 distinct five digit zip codes. Table 7 shows the data segregated by five digit zip code. We use a SAS clustering program to determine the new earthquake zone definitions and zone relativities. The SAS procedure we used is described in the SAS user's manual (1989).

PROCFASTCLUS performs a joint cluster analysis on the basis of Euclidean distances computed from one or more quantitative variables. The observations are divided into clusters such that every observation belongs to one and only one cluster. The procedure is intended for use with large data sets, from approximately 100 to 100,000 observations. With small data sets the results may be highly sensitive to the order of the observations in the data set.

PROCFASTCLUS uses a method referred to as *nearest centroid sorting*. A set of points called *cluster seeds* is selected as a first guess of the means of the clusters. Each observation is assigned to the nearest seed to form temporary clusters. The seeds are replaced by the means of the temporary cluster, and the process is repeated until no further changes occur in the cluster.

Specifying the desired number of earthquake zones and using the SAS procedure yields the results in Table 8. The number of zones to be used in a real application will depend on the size of the insurer's earthquake book of business, geographic spread, and the level of seismic variation within the state. The proposed earthquake zones probably will not be contiguous because five digit zip codes from different parts of the state will fall into the same cluster in the SAS procedure. We only use 20 zip codes in our example; however, the SAS procedure has the capability to handle a much larger number of zip codes. The relativities shown in Table 8 are applied to the statewide indicated rate previously calculated to determine each zone's earthquake rate.

The resultant earthquake zone rates should display a wider variance, as it could be argued that risk margins should vary by geographic location for the earthquake peril. We view this as another area deserving further consideration and an important aspect of determining adequate earthquake rates.

3 Shortcomings Inherent in Modeling

3.1 Data Problems

Modeled results can be understated for many reasons, most of which can be attributed to company issues or to adjustments not made within the models. We first will discuss company shortcomings and then follow with model shortcomings. Where appropriate, we will make sug-

Table 8
State X Earthquake Zone Relativities

Zone	(1)	(2)	(3)	(4)	(5)
1	$552,322	$359,009	$0.65	0.37	$1.38
2	3,694,971	3,886,713	1.05	0.59	2.23
3	3,560,167	6,181,967	1.74	0.98	3.68
4	2,354,709	6,060,641	2.57	1.45	5.46
5	803,112	3,011,670	3.75	2.11	7.95
Statewide	$10,965,281	$19,500,000	$1.78	1.00	$3.77

Notes: Zone = Earthquake Zone; Column (1) = Dwelling Insured Value in ($000); Column (2) = Expected Annual Loss at 10% Deductible; Column (3) = Loss Cost = (2)/(1); Column (4) = Indicated Relativity to Statewide = (3)/1.78; and Column (5) = Indicated Earthquake Zone Rate = (4) × 3.77.

gestions on how to handle quantifiable and supportable adjustments to the modeled input or output. The following list is not meant to be exhaustive, but is typical of company issues. Company shortcomings include:

- Underinsurance (homes insured less than their value) or overinsurance (homes insured more than their value);
- Demand surge for labor and materials after a catastrophic event;
- The need for extra claims adjusters following catastrophic events;
- No data collecting or coding for retrofitting safety features; and
- Invalid or incomplete data.

The major company shortcoming may be the problem of underinsurance. Expected loss to a particular structure in a particular area is based on applying an average damage ratio (defined as the ratio of the repair cost of a building to its total replacement value) to the total insured value of the structure. It is assumed that the insured value of a building represents its true replacement cost. A company should estimate its underinsurance (or overinsurance) problem before providing data to a modeling firm. If, on average, it is determined that a book of business is underinsured by 10 percent, then all limits should be adjusted before the model is run.

The effects of demand surge can be significant and should be factored into all modeled results. (It is not clear whether this adjustment should be made by the insurer or by the modeler.) The demand for labor and materials will vary depending on the location and magnitude of each earthquake. The additional cost probably varies between 0 percent and 30 percent, but the highest demand is associated with events that have the lowest expected probability; therefore, the effect on average annual aggregate losses should be minimal (albeit the effect could be substantial for large catastrophic events). We believe this adjustment to the modeled loss costs is important, yet is an uncertain aspect of the process. Studies should be conducted to determine the impact of demand surge factors, perhaps by studying the payout of events such as Loma Prieta and Northridge, if data are available. Either overall average demand surge factors should be applied to the resultant loss costs or variable demand surge factors should be determined and applied by location and event.

The need for independent claims adjusters is a real cost of settling claims following large catastrophic events. It is not clear which loss adjustment expense (LAE) factors should be applied to the modeled expected loss costs—there has not been enough loss experience to determine appropriate factors. We suggest using either the ratio of LAE to losses of past events (which may understate the true ratio) or the underlying policy average LAE factor, given earthquake coverages are normally endorsed to a homeowners or dwelling fire program.

Modeled results should account for retrofitting safety features of an insured structure. This is especially applicable to buildings made of unreinforced masonry. Average damage ratios should be adjusted for these features. It is not clear how the effects of retrofitting can be measured, but research should be conducted and insurers should encourage their installation. A strongly built and reinforced home should withstand the initial impact and aftershocks of an earthquake, as opposed to a home whose frame is not bolted to the foundation, for example. Most insurance companies do not request information on retrofitting mechanisms, nor do they store the data. We would encourage the Institute for Business and Home Safety (IBHS)[1] to study the effects of such safety features and simulate an earthquake under monitored laboratory conditions to determine the extent of damage on the structure and its contents. The Institute for Business and Home Safety is a nonprofit organization sponsored by the insurance industry. The mission of IBHS is "to reduce injuries, deaths, property damage, economic losses and human suffering caused by natural disasters."

[1]Institute for Business and Home Safety, 73 Tremont Street, Suite 510, Boston MA 02108.

Finally, there is always the possibility of invalid data, incomplete data, or no data at all. Invalid data are most prominent if zip code, county, or street address is not validated before being stored on the insurer's database. Either the data should be cleaned before the input files are created or the data should be eliminated from analysis. Alternatively, invalid data could be proportionally distributed throughout the state by county or zip code based on the distribution of the insurer's valid data. Most companies do not have enough insureds located in all areas of the state. Therefore, there will be many locations with no modeled loss costs. In these situations, modeling firms have access to an inventory of typical building structures by location, average dwelling limit, type of construction, average year of construction, building height, etc. Modeled loss costs from this generic inventory can supplement an insurer's results where few or no insureds reside.

There will also be locations with insufficient data. Assume for a moment that an insurer's book of business is mapped to the geographic zip code centroid of each zip code within the state. Although modeled results are assumed to be 100 percent credible by location, the reader could question whether one, ten, or even 100 exposures are enough to deem the results credible. An insurer's database could be complemented with the results of the generic inventory. The authors have chosen to consider data 100 percent credible by zip code with more than 100 exposures; otherwise, the generic inventory is given full credibility.

3.2 Inadequate Information

These brief remarks are not intended to criticize any model or modeler, but to highlight the importance of their impact on modeled results. The following list is also not meant to be exhaustive, but does represent typical shortcomings:

- Factor for unknown faults;
- Inclusion of debris removal expenses;
- Effects of aftershocks; and
- Parameter risk within the model.

The 1994 Northridge Earthquake is a perfect example of an unknown fault, a blind thrust fault that does not break the earth's surface. Not even seismologists know the extent of undiscovered fault lines beneath the earth's surface. How understated could the modeled results be? No one knows for sure, and we propose no solution to handle this uncertainty. Although the models account for possible earthquakes in all historical seismic source zones, it is questionable if distributions in the model account for all potential seismicity. With the passage of time and with advancing technology, perhaps these models may account for all possible faults some day. For now we must assume that a model's results may understate expected average annual losses and, hence, expected loss costs per $1,000 of coverage.

Debris removal expenses, although small, should be added to the model's expected loss costs. More prominent would be the effects of aftershocks that follow moderate to large earthquakes. Claims often are reopened months later due to weakened structures repeatedly damaged from aftershocks. Future modifications to catastrophe models should account for this possibility.

Because catastrophe modeling is based on incomplete distributions developed from historical information, parameter risk always will exist. This risk may lead to gross understatement (or overstatement) of potential insured losses and represents a potential shortcoming of modeling.

3.3 Additional Considerations

There will always exist areas that deserve further consideration. While we have presented a practical procedure for developing adequate earthquake rates, some areas deserve additional research and attention. We will divide these topics into four categories: (1) shortcomings of models, (2) credibility of data, (3) necessary target rate of return, and (4) net reinsurance costs.

We devote an entire section of this paper to model shortcomings and company data issues. We repeat them to emphasize their importance and the need for further study. The cooperation of the insurance industry, modeling firms, and the IBHS is necessary to quantify the impact of outstanding issues on expected loss costs. Perhaps special data calls or cooperative studies can be conducted and the results shared with all interested parties.

Computer modeling simulates thousands of possible events, and its results are generally considered credible. The earthquake peril is unique by location, especially in California, so a feasible complement of credibility to augment a local result does not exist. Perhaps a regional complement could be used, but its applicability is questionable, given local soil conditions and proximity to fault lines. We believe that an industry inventory database represents the best alternative for a complement.

Insuring the earthquake peril is much riskier than insuring auto physical damage coverages. Due to the relationship between risk and return, a higher rate of return (and therefore a higher underwriting profit and contingency provision) should be allowed to cover a company's earthquake exposure. This provision also should vary by location. We have presented a simplified method for deriving a reasonable profit provision, but we encourage more research in this important area.

Should rates include the costs of reinsurance on an insurer's book of business? Their inclusion could be viewed as a pass-through to the consumer. Also, in the long run neither the insurer nor the reinsurer(s) should be worse off for engaging in a reinsurance program; otherwise, neither party would enter the contract. In the short run, however, reinsurance costs are a legitimate expense of doing business, and we believe that all parties should share in that expense, including policyholders. Policyholders benefit from financially strong companies.

4 Summary

Catastrophe hazard modeling has become an integral part of the ratemaking process. Casualty Actuarial Society ratemaking principles (1988) state that "other relevant data may supplement historical experience. These other data may be external to the company or to the insurance industry." We have entered the realm of that other relevant data. *Actuarial Standard of Practice (SOP) No. 9* (1991) states that "an actuary should take reasonable steps to ensure that an actuarial work product is presented fairly ... if it describes the data, material assumptions, methods, and material changes in these with sufficient clarity that another actuary practicing in the same field could make an appraisal of the reasonableness and the validity of the report." With the advent of modeling, however, the actuary must rely on the work of another person. *SOP No. 9* states that "reliance on another person means using that person's work without assuming responsibility therefore." These other persons now include experts in the fields of geology, seismology, and structural engineering, to name a few. Actuaries, however, can play a key role in contributing to the development of the models and, more importantly, the interpretation and communication of their valuable results.

Catastrophe hazard modeling has become a necessary tool for the pricing of large catastrophic events such as hurricanes and earthquakes.

Their frequency is so low and their severity so potentially high that not even all of the property and casualty companies in a state could have enough loss history upon which to base rates. Despite any shortcomings models may have, they hold the key to the future and the pricing of nature's perilous attacks.

References

Actuarial Standards Board. *Documentation and Disclosure in Property and Casualty Insurance Ratemaking, Loss Reserving, and Valuations.* Washington, D.C.: American Academy of Actuaries, January 1991.

Statement of Principles Regarding Property and Casualty Insurance Ratemaking. Arlington, Va.: Casualty Actuarial Society, May 1988.

Best's Aggregates and Averages, Property-Casualty United States. Oldwick, N.J.: A.M. Best Company, Inc., (1995): 174, 176.

SAS/STAT User's Guide, Version 6, Fourth Edition, 1. Cary, N.C.: SAS Institute Inc., (1989): 823–824.

Appendix: The Applied Insurance Research Model

Overview

The model developed by Applied Insurance Research uses sophisticated mathematical techniques to estimate the probability distribution of losses resulting from earthquakes anywhere in the 48 contiguous states. The earthquake model is composed of three separate components: an earthquake occurrence model, a shake damage model, and a fire-following model. The earthquake occurrence component of the model uses a probabilistic simulation to generate a synthetic catalog of earthquake events that is consistent with the historical record. The shake damage estimation component uses analytical numerical techniques to calculate the distribution of losses for individual buildings given the characteristics of the event. The fire-following component uses simulation to estimate fire losses following an earthquake. Together these techniques allow the estimation of a wide range of information about potential earthquake losses in the United States.

The earthquake simulation model incorporates descriptions of a large number of variables that define both the originating event (the earthquake) and its effect on structures. Some of these variables are random and others are deterministic. We will describe the key aspects of the model, the main variables affecting the outcomes, and the relationships between the primary variables in the rest of this appendix.

Earthquake Occurrence in the USA

For earthquakes there are three key types of variables that describe the physical phenomenon. In broad terms, these variables describe where earthquakes can occur, the size of the earthquake, and the likelihood of seeing an earthquake of a particular size. In other words, the variables describe where, how big, and how often earthquakes occur.

The issue of where earthquakes occur is handled by identifying *faults* or *seismic zones* where actual earthquakes have been observed. On the West Coast earthquakes tend to occur along well-defined geological features called faults, which are places where the surface of the earth has been ruptured by past earthquakes and which are observable at the ground surface or by subsurface sounding techniques.

Not all faults are active, i.e., not all faults are believed capable of rupturing in the present, although they have ruptured in the distant past. Where faults are observed and where the historical catalog (record) of earthquakes indicate that the faults are still capable of rupturing, the surface trace of the fault defines a possible location for future earthquakes.

Not all earthquakes occur on identifiable faults, however. Many earthquakes, especially those east of the Rocky Mountains, occur on faults that are not visible at the surface. Such faults are inferred from the occurrence of actual earthquakes in the historical record. For these areas, a source zone is created, which is an area with fuzzy boundaries within which future earthquakes are possible.

The AIR model contains approximately 250 seismic source zones covering the 48 contiguous states. Each source zone is defined by a line on the surface of the earth with probability distributions describing the variability of potential epicenters both along and perpendicular to that line. A potential earthquake is not limited to occur along a known fault line, but can occur anywhere in the vicinity of a fault or anywhere within a seismic source zone, depending on the degree of uncertainty associated with the historical record of earthquakes in that area. The central line of the source zone does define the dominant direction of faults in the area and characterizes the orientation of the rupture surface.

The size of an earthquake is usually measured by one of several magnitude scales. In the AIR model the surface wave magnitude Ms scale[2] is used to characterize the earthquake magnitude. For every fault and source zone the frequency of earthquakes of different magnitudes must be described. Seismologists generally agree that, over a considerable magnitude range, the logarithm of the number of historic earthquakes that exceed a given *magnitude scales* varies linearly with magnitude. This indicates that the frequency-magnitude relationship is approximately exponential.

Additionally, prehistoric seismologic data have been interpreted by some researchers to indicate that the frequency-magnitude relationship for large earthquakes differs from exponential scaling, leading to the notion of characteristic earthquakes in certain geographic areas. The AIR model incorporates a truncated exponential distribution, or truncated Gutenberg-Richter relationship, to represent potential seismicity in each source zone. Where appropriate we incorporate a characteristic earthquake model.

The AIR earthquake model is calibrated to a catalog of historical earthquakes that covers the historical record from the mid-1600s to the present. Because the completeness of the catalog varies both in time and as a function of magnitude (larger earthquakes are more likely to be included in the historical record), the fitting of the frequency-magnitude distribution is adjusted to account for the variation in historical completeness.

Earthquake Attenuation

After earthquakes are simulated using the probability distributions of the different earthquake parameters, the shaking intensity of the earthquake at every location affected by the earthquake is calculated using a relationship called an *attenuation function*.[3] The local intensity is corrected to reflect local soil conditions, as some types of soil amplify the shaking intensity relative to other soil types. This section discusses the variable interrelationships required to calculate the local shaking intensity.

From the characteristics of the earthquake the local shaking intensity is calculated using an attenuation relationship. The attenuation relationship depends on the location of the source zone, as earthquake shaking attenuates more quickly in the western U.S. than in the east-

[2]The Ms scale measures the strength of an earthquake as determined by observations of its local surface waves.

[3]This function measures the reduction in the shaking intensity as we move away from the epicenter of the earthquake.

ern part of the country. The same magnitude earthquake will affect a smaller area in California than in the northeast.

The attenuation calculation starts by spreading the energy released by the earthquake over the rupture surface and integrating over the entire rupture surface to calculate the total effect of the earthquake. In effect, energy is assumed to be released uniformly over the rupture, and each incremental piece of energy is attenuated separately to obtain the effect at some distant point. This results in contours of equal intensity that are elongated along the orientation of the rupture.

The calculation of local shaking intensity consists of two parts. First, a basic intensity is calculated that assumes uniform soil conditions at every location. This intensity (called a *Rossi-Forel intensity*) depends on the distance of the site from the earthquake rupture, the orientation of the rupture, and the earthquake magnitude and focal depth. The rupture length is calculated from the basic earthquake parameters. Second, the Rossi-Forel intensity is modified to reflect the soil conditions at the site. Soil conditions for the entire country are digitized on grids varying from 0.1 degree latitude/longitude squares to 0.5 minute latitude/longitude squares. The local soil condition can significantly affect shaking intensity. The final intensity is identified as a *modified Mercalli intensity* (MMI).

The MMI is a generally accepted unit of shaking intensity. It describes, in general terms, the type of damage that might be expected to buildings of usual design and other effects of earthquakes that would be expected at a particular location. The MMI is a good metric for estimating damages to structures.

Exposure Characterization

In order to calculate damages from an earthquake, the AIR model incorporates an extensive description both of the structural characteristics of an exposure and of the policy conditions describing the treatment of deductibles and other factors.

The seismic performance of a building depends primarily on the structural system resisting the lateral loads, but is also affected by other factors (including, in the AIR model, the age of the building and the height of the building). The age of the building is used to determine the likely code provisions under which the building was designed and constructed. Newer buildings, which may have been built to more exacting code provisions for seismic performance, are expected to perform better than older buildings.

The AIR model incorporates damageability relationships for many different classes of exposures, with up to three height categories in each class. There are 42 different damage relationships for each coverage type, plus several different age categories. The categories of structural types are based in part on the structural types defined in ATC-13,[4] although the actual damage relationships are modified and extended beyond those covered in that reference.

The exposures are characterized by policy limits for four different coverages:

- Coverage A refers to the dwelling limit;
- Coverage B refers to the appurtenant structures;
- Coverage C refers to personal property; and
- Coverage D refers to additional living expense.

Most commonly, Coverage B is combined with Coverage A for calculation purposes and is assumed to apply to the same structural type as Coverage A. The policy limit for each coverage may be defined by both a replacement value and a policy limit. The replacement value may rise in time without the policy limit being adjusted to reflect inflation. Damage is always calculated with respect to replacement value and then is capped at the policy limit if appropriate.

The location of the risk can be defined by a latitude and longitude point or by the five digit zip code in which the risk is located. The risk also can be associated with a line of business (homeowners, renters, commercial multiperil, etc.) in order to report losses separately in categories meaningful to the insurer.

Damage Estimation

Given the local shaking intensity in MMI units, damages to structures at a particular location can be calculated if sufficient information is available about the structure. Two types of damages are calculated by AIR: shake damage due to the lateral and vertical motions of the ground and fire damage due to earthquake-induced fires.

In order to calculate shake damage, the exposure information is combined with the level of shaking intensity at the building. Information on the structural characteristics of the properties at risk is used to

[4] The Applied Technology Council is a 13 member advisory project engineering panel established in 1982 to develop earthquake damage/loss estimates for facilities in California.

select an appropriate damageability relationship (also sometimes called a *damage function* or a *fragility curve*) relating the probability of different levels of damage to the local shaking intensity (MMI). The damageability relationship is a complete probability distribution of damage, ranging from no damage to complete destruction (0 to 100 percent damage), with a probability corresponding to every level of damage. Thus the probability distribution is a continuous function of the local MMI level.

The earthquake damageability relationships have been derived and refined over a period of several years. They incorporate well-documented engineering studies by earthquake engineers and other experts both within and outside AIR. These damageability relationships also incorporate the results of post-earthquake field surveys performed by AIR engineers and others as well as detailed analyses of actual loss data provided to AIR by its client companies. These relationships are continually refined and validated.

Fire-Following Loss Estimation

Once the shake damages have been calculated for a particular earthquake, fire-following losses are estimated. This part of the model uses a separate simulation to estimate fire losses for each event.

First, the number of fires spawned by the earthquake is generated. The fire ignition rate is based on the local MMI intensity and the total population in the area. A number of fires is simulated for each affected zip code. The mean ignition rate increases as the MMI increases. The probability distribution of ignition rates is assumed to be uniform in some interval around the mean rate. Once the number of fires is simulated, each fire is randomly placed within a zip code and is assigned to affect either residential properties, commercial properties, and/or mobile homes.

The fire simulation then simulates the spread of the fires as well as the actions taken by local fire departments to control the fires. The fire spread rate is affected by a randomly selected wind speed appropriate for the location of the earthquake. Higher wind speeds increase the rate of spread of the fire.

Some of the factors included in the fire simulation are the time to report the fire, the time for one or more fire engines to reach the fire, and the availability of water to fight the fire. All of these factors are affected by the local MMI, as areas experiencing high shaking intensity are more likely to have obstructed roads and broken water mains. Also, the influence of fire breaks—wide roads or other natural impediments to fire spread—is included in the simulation. Fire engines can move from fire to fire as fires are controlled.

Because the fire losses are determined by simulation, different levels of fire loss can be calculated for a given earthquake. Typically, the variability of fire losses is large, at least for the larger earthquakes, such that fire losses can vary by at least a factor of two if the same earthquake is simulated several times. This reflects the uncertainty in fire losses for larger earthquakes.