13th World Conference on Earthquake Engineering
Vancouver, B.C., Canada
August 1-6, 2004
Paper No. 217

VULNERABILITY PARAMETERS FOR PROBABILISTIC RISK MODELLING – LESSONS LEARNED FROM EARTHQUAKES OF LAST DECADE

Alok KUMAR[1], Martin BERTOGG[2], Silvio TSCHUDI[3]

SUMMARY

Analysis of earthquake loss data is important for improving and validating the probabilistic risk assessment models widely used by the emergency planners and the insurance industry for estimating the future losses from earthquakes. In this paper important vulnerability parameters from the perspective of the probabilistic risk models are laid out. Published loss data from Northridge and Kobe earthquakes have been analyzed to quantify and illustrate these vulnerability parameters. Not all the parameters could be derived from the available loss data due to the various missing pieces of information. A few results that could be obtained are presented and uncertainties associated with those results are discussed. Limitations of each of the available loss data sets used in the analysis are identified and recommendations for improving the information content of the loss data are made. It is hoped that these recommendations would help to foster the data collection process for thorough loss analysis of any future event.

INTRODUCTION

Earthquakes pose a threat to life and property in most parts of the world. Building code developers and emergency planners in a community exposed to earthquake need to do an in-depth assessment of the risk in order to understand the ways of reducing the risk and to be better prepared for the emergency response and recovery following an earthquake. A general question they look to answer from such a risk assessment is what portion of the portfolio of risks might be affected.

Any (re)insurance company offering earthquake coverage also need to do such a detailed risk assessment to be able to answer the following key questions important for its long term business continuity and survival:

- What is the expected annual loss to a portfolio of risks?
- What is the maximum possible loss from any single event?

[1] Swiss Reinsurance Company, Zurich, Switzerland. Email: alok_kumar@swissre.com
[2] Swiss Reinsurance Company, Zurich, Switzerland.
[3] Swiss Reinsurance Company, Zurich, Switzerland.

Answer to the first question is needed to achieve long term equilibrium of losses and premium income over time. Answer to the second question can be crucial for the survival of the company when a portfolio of several thousands of policies is affected.

Today's state-of-the-art probabilistic earthquake risk models provide local emergency planners and insurance companies with the tools necessary to do an in-depth risk assessment for the region of interest or for the insurance portfolio. The outcome of such a detailed risk assessment can be used by local authorities for risk reduction and emergency preparedness and by the (re)insurance companies for better managing their exposure to catastrophic events and hence their business.

One of the key inputs in these probabilistic risk models is the impact of an earthquake on various objects in the built environment such as buildings, bridges, various utilities etc. There are several engineering methods available to estimate this impact. However, the most valuable and vital information to quantify the impact of ground shaking on built environment comes from an analysis of the real damage and loss data.

After a major earthquake event, typically a wealth of earthquake loss data is collected, compiled and analysed by the engineering and research community. However, very often some key information is found to be missing from these data, limiting the applicability of the results of loss analysis for risk modelling and answering above questions.

In this paper, as a starting point the important vulnerability parameters from the perspective of probabilistic risk assessment models are laid out. Published loss data-sets from the Northridge and the Kobe earthquakes, two of the biggest earthquakes in the last decade in terms of their regional impact and the losses to the insurance industry, are analyzed to illustrate the methods of quantifying these vulnerability parameters. These data-sets have also been used to bring to fore the common deficiencies in the loss data compiled after major earthquakes. The shortcomings with each of the data sets are discussed in this paper and recommendations are made for improving the data contents so that useful analysis can be carried out in future.

KEY VULNERABILITY PARAMETERS FROM RISK MODELLING PERSPECTIVE

The vulnerability Module in a risk assessment model incorporates parameters quantifying the impact of ground shaking on the built environment. These vulnerability parameters can be derived using loss data from past earthquakes. For application in risk assessment models, a loss analysis should ideally focus on deriving at least the following basic vulnerability parameters used in a typical risk model (Bertogg [1]):

Percentage of Policies[4] Affected (PPA)

For a given shaking intensity[5] level, the PPA figure denotes the ratio of number of policies (building or an insurance policy) affected by a loss to the total number of insurance policies exposed to the same shaking intensity. In other words, PPA would indicate number of policies or risks that would suffer some loss.

[4] When used for non-insurance purpose, word "policies" can be simply replaced by "buildings" or "risks".

[5] There are several measures of ground shaking available such as intensity, peak ground acceleration, peak ground velocity and spectral response. Due to the availability of intensity records for historical earthquakes and damage data correlated with it, intensity (Modified Mercalli Intensity, MMI) is widely used as the measure of ground shaking in risk assessment models.

Mean Damage Degree (MDD)
The MDD at a shaking intensity is the total loss to the affected policies as a ratio of total sums insured[6] for the affected policies only. This figure gives an indication of average size of loss to a policy that suffered loss.

Mean Loss Ratio (MLR)
The MLR per shaking intensity can be calculated as the sum of individual loss ratios for each of the affected policies divided by the total number of affected policies. MLR and MDD figures would be close to each other when there is not much variation in the sums insured for underlying policies. In case of large variations in the sums insured, MLR is a better indicator of average loss ratio than MDD. One disadvantage of MLR is that it can be calculated only if loss amount and sums insured is available for each individual policy, which is often not the case.

Mean Damage Ratio (MDR)
The MDR at a shaking intensity is the total loss to the affected policies as a ratio of total sums insured for the all the policies exposed to same shaking intensity. In contrast to the MDD, this figure also includes total sums insured for the policies which did not suffer any loss at the same level of shaking.

Loss size distribution
For application in risk modelling, it is essential to know not only the mean loss but also the distribution of the loss around mean. Knowledge about loss distribution function and parameters would help in modelling the impact of insurance conditions such as deductibles and limits.

Deductible credit curve
Deductible credit curve gives a relation between paid loss as a fraction of total incurred loss and the deductible as a percentage of sums insured. This relationship derived from loss data can be used to validate the deductible credit calculated by loss model using loss distribution parameters.

In addition, the impact of various risk features such as construction type, occupancy, size, height, age etc. on each of the vulnerability parameters is also of interest from risk modelling perspective. Increased quality and resolution of the portfolio data means insurers are trying to incorporate these features into their pricing considerations and give appropriate credit. So quantifying the impact of various risk features using the loss data would be of great value for application in the risk modeling for insurers.

VULNERABILITY PARAMETERS DERIVED FROM PAST EARTHQUAKES

Loss data from some of the past earthquakes have been analyzed with an aim to derive the key vulnerability parameters mentioned above. Not all the parameters could always be derived due to limitations with the available loss data. This section presents some of the results from loss analyses and identifies the limitations with each set of data considered. Table 1 below lists the earthquakes analyzed and brief comment on the data used for the analysis.

---

[6] In the context of this paper, total sums insured is used as the replacement value i.e. the cost for labor and material to rebuild the structure in new. This is strong contrast to the actual cash value, which includes time depreciation. Typically, insurance contracts protect for replacement value.

Table 1: Analysed earthquake events and loss data

Earthquake	Date	Magnitude	Data available
Northridge Earthquake	17.01.94	6.7	Published insurance loss data Published building safety inspection data
Kobe Earthquake	17.01.95	6.9	Insurance loss data

Northridge Earthquake – Insurance loss data

Data Analyzed

The insurance loss data published in a report by Comerio [2] that examines the state of earthquake recovery practice in California has been used in this analysis. This report contains the number of paid insurance claims and paid losses per policy type resulting from Northridge earthquake and reported by insurance companies to the California department of insurance as of March 1995. Summary of this data is presented in Table 2 below.

Table 2: Northridge insurance claims and losses by policy type (as of March 1995)

Policy Type	Claimed Losses	No. of Claims	Average Loss
Earthquake Policy	$412,274,270	6,687	$61,653
EQ rider on Homeowners Policy	$4,550,152,370	110,805	$41,065
EQ rider on Fire Policy	$246,988,189	9,169	$26,937
EQ rider on Mobile Home Policy	$102,308,519	5,494	$18,622
EQ rider on Condominium Policy	$272,432,177	20,717	$13,150
EQ Rider on Renters Policy	$80,636,029	9,652	$8,354
All Residential Policy	$5,664,792,554	162,524	$34,855

Apart from paid insurance losses, this report also contains claims ratio, insurance penetration, total number of single family housing and the median home value per zip code. Additionally the report contained the average MMI intensity observed in each zip code. These MMI intensities were based on intensity map developed by the USGS after the earthquake. However, the report did not provide some other key information on the total exposure at the time of earthquake such as total sums insured and number of policies exposed. Also missing were deductibles in force per policy and affected policies with losses below deductibles. Data was aggregated at zip code, so features such as age, size and height of each individual risk was not available.

Important comment regarding this data: Northridge insurance loss data used in this analysis represents only a sub-set of total residential claims paid. An estimate by Petak [3] in 2000 puts the residential claims paid at $9.88 billion as opposed to data corresponding to $5.66 billion of claims paid used in this analysis. Details regarding additional claims paid are not available to authors.

Data Analysis and Results

Focus of this analysis was to derive vulnerability parameters for single-family residential dwellings. These dwellings are insured for earthquake using earthquake rider on homeowner and fire policies or through separate earthquake policies. As such any data for mobile home, condominium and renters policies were not considered in this analysis.

First step in this loss analysis is to get an estimate of the total exposure, meaning number of earthquake insurance policies and total sums insured, at the time of earthquake. Total number of insurance policies with earthquake coverage is estimated using available information on claims ratio, number of paid claims and number of single family dwellings per zip code. The estimated number of exposed policies put the insurance penetration in the affected area to be 42%, consistent with widely reported 35-40% insurance penetration at the time of earthquake. Total sums insured at the time of earthquake is estimated using average building value obtained from Los Angeles County population and building inventory data available in Kircher [4]. Contents value of 50% of structure value have been assumed and added to the structure value to get total sums insured.

Losses below deductibles have been estimated by assuming an average deductible of 10%[7]. Missing from this estimate are losses to policies where no claims were paid because of total loss being less than the deductible. None of the available sources presented any indication about number of such policies. As such any reasonable estimate of this portion of total loss was not possible.

Table 3 presents per MMI summary of loss and exposure data estimated as described above for Northridge Earthquake. No adjustment has been made to average MMI assigned to each zip-code in the report.

Table 3: Summary of estimated exposure and loss per MMI for Northridge Earthquake

	Exposure data			Loss data		
Intensity (1)	Number of in force policies (2)	Total sums insured ($ bn) (3)	Sums insured for policies with paid loss (bn $) (4)	Reported no. of claims paid (5)	Reported Amount paid ($ bn) (6)	Estimated from ground-up losses[8] ($ bn) (7)
V	81,009	17.808	0.470	2,137	0.029	0.060
VI	196,904	43.285	1.811	8,237	0.177	0.298
VII	269,112	59.158	8.425	38,325	1.447	2.008
VIII	187,864	41.298	15.616	71,037	3.092	4.133
IX	10,115	2.224	1.522	6,925	0.464	0.566
Total	745,003	163.772	27.844	126,661	5.209	7.065

The data from Table 3 is used to get an estimate of likely PPA, MDD and MDR. The missing information regarding number of policies with losses below deductible and the amount of such losses mean that the derived parameter using the available data do not correspond to the actual definition of each of these parameters. In the present analysis, PPA is calculated as the ratio of number of paid claims to the number of in-force policies (column 5/2). MDD is calculated as the ratio of estimated from ground-up losses to the sums insured for policies with paid loss (column 7/4). MDR is calculated as the ratio of estimated from ground-up losses to the total sums insured (column 7/3). The relationships so obtained are displayed in Figure 1.

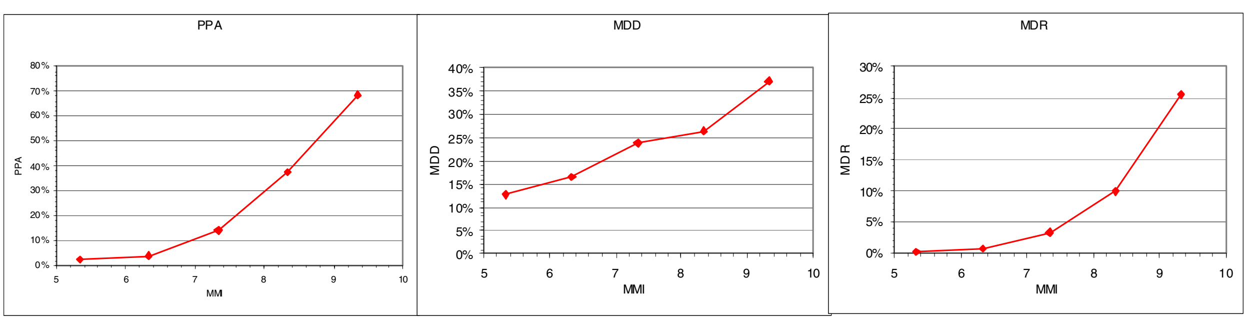

Figure 1: Vulnerability parameters derived using insurance loss data for Northridge. Assumptions regarding exposure and deductibles introduce large uncertainty in the results.

[7] Typical earthquake deductible in California at the time of earthquake.
[8] For policies where claims were paid. This does not contain any losses to policies which suffered losses that did not exceed deductible.

The vulnerability parameters derived above do not correspond to the true definition. They have been derived by ignoring the contribution of affected policies where losses remained below deductible. Additionally, assumptions involved in estimating the in-force insurance policies and total sums insured introduce large uncertainty in the results obtained. As such, the derived parameters cannot be directly used in any risk modeling. It is even not possible to say whether actual values of these parameters would be higher or lower than the relations derived here. Other vulnerability parameters namely, loss size distribution and influence of deductible could not be estimated due to aggregated nature of data and missing information on deductible per policy.

The grouping of the data per policy type revealed that average losses to earthquake policies were much greater than homeowner or fire policies, a very strong observation about which any insurance company would like to know more. However, lack of information regarding underlying exposure and deductibles per policy meant any meaningful analysis could not be carried out with this data.

Discussion
From loss analysis for risk modelling perspective, the available data was found to be insufficient due to key missing information such as number of policies in force, total sums insured for the in-force policies, number of affected policies where losses were not paid and losses below deductible. To derive the vulnerability parameters such as PPA, MDD and MDR a number of assumptions had to be made regarding missing information, which introduced substantial uncertainty in the results obtained.

Another deficiency with the data was its aggregated nature. Aggregation of data made it difficult to get any reasonable estimate of the loss size distribution and the influence of deductible on claims paid.

A very little additional effort to compile the missing information such as the number of in-force policies, the total sums insured of these policies and the losses below deductible would have reduced the number of assumptions and hence the uncertainties in the results of loss analysis. Grouping of data per policy type was a very positive feature of the data. Such a grouping clearly showed different behavior of claims paid under different policy types. A detailed per policy exposure and loss data with such a grouping would have been ideal for a more useful loss analysis.

Northridge Earthquake – Building inspection data

Data Analyzed
The Northridge earthquake building damage and inventory data in the affected area compiled in a database (called OES database from here on) by EQE Inc., a risk modeling consultant, and California Governor's Office of Emergency Services (OES) is used in this analysis. This database, publicly available in report [5], groups the building damage and inventory information per MMI based on key building features such as occupancy type, construction type, age, size and soil. Building damage data is based on estimates of damage ratio for buildings provided by local building safety inspectors during the building inspections carried out after the earthquake. Main purpose of these inspections was to assign a color-coding to the buildings based on their safety for occupation. Building inventory and information on key building characteristics is compiled from Los Angeles county tax assessor database. Soil and intensity associated with each building has been obtained by the OES by geocoding the address provided in the tax assessor database and using detailed soil and the USGS observed intensity map. A summary of information pertaining to single-family dwellings in the OES database is presented in Table 4

Table 4: Summary of single-family dwellings information in OES database

Number of Single Family Dwellings in affected area	1'756'857
Number of inspected buildings in LA County	105'019
Number of buildings with damage estimates	97'196
Total estimated damage	$2'562'352'452
Damage per affected building	$26'362

The OES database used in this analysis is found to be very comprehensive in terms of count of the residential buildings and classifying the building as per key characteristics such as construction type, age and size. However when compared to the insurance claims paid data, the estimates of building damage and even the count of damaged building as per the OES database were found to be severely understated. Compared to the estimates using the OES database shown in the table above, there were 126,661 claims paid to single-family dwellings amounting to a total of $5.2billion as of March 95. Average claim paid per single-family dwelling is estimated at $42'000, almost 60% higher than estimate from the OES database. Some of the reasons for this underestimation could be:

- The building safety inspectors estimated only the structural damage ignoring completely the damage to non-structural, contents and additional structures (such as garage, walk ways, chimneys etc).
- While severely damaged structures were easy to spot, lots of structures with small exterior damages were overlooked during the inspection. In the end these structures needed tens of thousands of dollars to repair.
- There was no systematic procedure in place to estimate the dollar damage to structures. The estimate varied depending on the experience of the building surveyor.

Data Analysis and Results

Due to clear underestimation of total losses by the OES damage data, this data was considered insufficient for estimating the parameters such as PPA, MDD and MDR. However, this data was found to be excellent for studying the relative impact of various building characteristics on its vulnerability. Using the OES data, the impact of building occupancy, size and age on its mean damage ratio (MDR) has been studied. Various MMI-MDR curves derived using this data are presented below in Figure 2 to Figure 4. Key findings from this analysis are:

- Multi-family dwellings are more vulnerable than single-family dwellings, especially at higher intensities. (Figure 2)
- Single-family buildings with larger floor areas are more vulnerable at higher intensities. Buildings with larger floor areas are perhaps the once with more than on stories, implying higher vulnerability of multi-storey single-family dwellings. (Figure 3)
- Very interesting: common notion that older buildings are more vulnerable could not be observed using this data (Figure 4). This is an important observation, especially for building code developer, as this implies buildings constructed with the new and improved building codes did not necessarily perform better than those constructed with the old building codes.

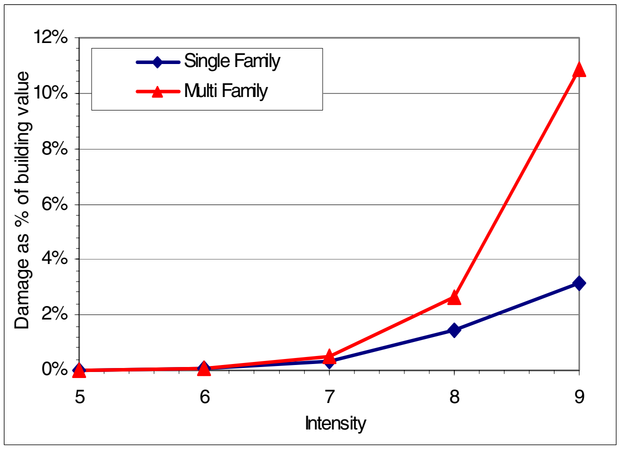

Figure 2: MMI-MDR relation derived for different occupancy using OES database. Multi-family dwellings are more vulnerable than single-family.

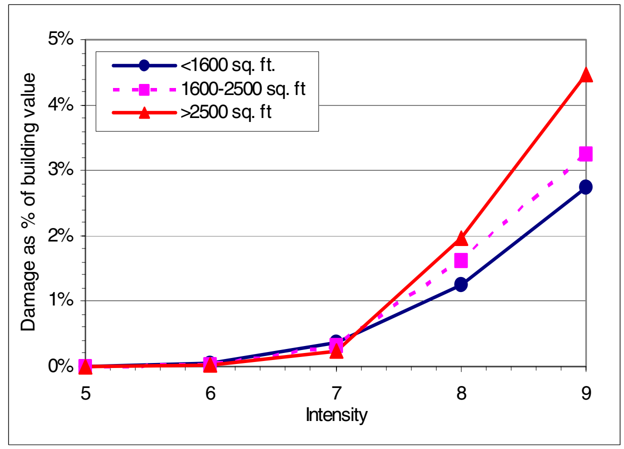

Figure 3: MMI-MDR relation derived for single-family dwellings with different floor area. Larger buildings are more vulnerable at higher MMI. (10 sq.ft. = 1 sq.m)

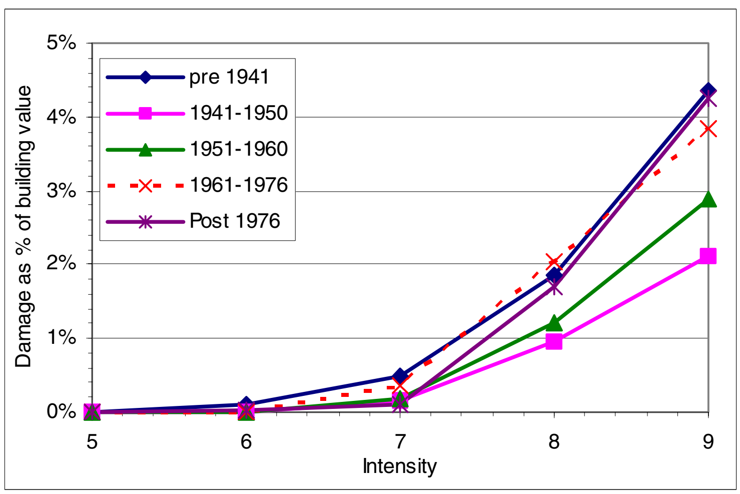

Figure 4: MMI-MDR relation for single-family housing in different age group. No clear trend regarding variation of MDR with age can be observed.

Discussion

Building damage estimates from building safety inspectors grossly understated the total losses. As such, this data was not considered a good data-set for deriving the vulnerability parameters that could be used in risk modeling.

The value of OES database lies in its comprehensiveness in terms classifying inventory and damage data as per key building characteristics. This compilation made it possible to study the relative impact of various building characteristics on its vulnerability.

Kobe Earthquake – Insurance loss data

Data Analyzed

Insurance loss data to the residential policies of one of the big insurance companies in Japan was used for the purpose of this analysis. This data, in the form it was presented, contained total number of in force policies, total exposure, number of claims paid and amount of claims paid per ward (city partitions) in the most heavily affected Hyogo Prefecture. The data was not linked to any estimate of the earthquake shaking intensity. Summary of this data is presented below in Table 5

Table 5: Kobe Earthquake – summary of residential exposure and paid loss data

Number of in-force policies in affected area	461,822
Total exposure (billion JPY[9])	2,281.85
Number of claims paid	73,227
Amount of claims paid (billion JPY)	99.56

The insurance company had very special insurance structure in place limiting the loss payout in the event of an earthquake. Information about the losses not paid out due to these insurance conditions was not available.

Data Analysis and Results

Since no estimate of shaking intensity was available with the loss data, the first step in this analysis involved estimating average shaking intensity per ward. The MMI intensity maps and PGA measurements available from RMS [6] and EQE [7] event reports were used as starting point. Assigning average intensity per ward proved to be a tricky task as the earthquake was characterized by an extreme directivity and very localized soil induced behavior. Especially the wards on the fault lines were exposed to an entire range of intensities from X down to VII, making it difficult to assign an average intensity value. The final estimation of average intensity per ward was based on combination of best judgment and available intensity maps and PGA readings.

With the average MMI per ward known, average paid loss per ward was calculated as the ratio of total paid loss to the total exposure in the ward. The resulting relation gives a lower bound of the MDR. The missing information about losses not paid due to insurance conditions means the actual observed MDR and other parameters such as PPA and MDD could not be calculated. Due to the aggregated nature of the data and missing information on insurance conditions per policy, it was not possible to get a reasonable estimate of the loss size distribution and the impact of deductibles.

[9] Japanese Yen

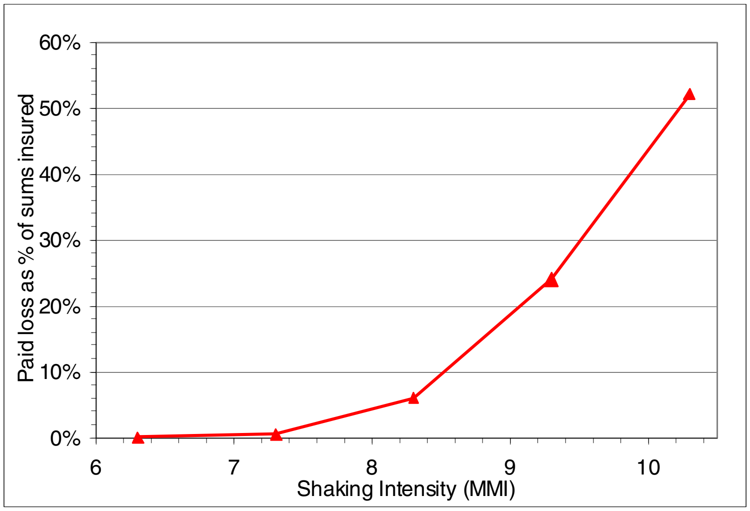

Figure 5: Lower bound of MDR derived from Kobe residential loss data. Actual MDR is expected be higher as it would also include losses not paid due to insurance conditions

Discussion

The Kobe loss data is considered comparatively better than the Northridge loss data. The data contains not only the information on the number and amount of claims paid, but also the total exposure in the affected area at the time of the event. As such it can be said with greater degree of confidence that the relation derived in this analysis represents at least lower bound of the MDR. If it can be assumed that the insurance condition did not change since the earthquake, the relationship derived has the potential for direct applicability in a risk assessment model.

Due to aggregated nature of the data and unavailable information on insurance conditions and unpaid losses, the data could not be used to derive all the meaningful vulnerability parameters.

Some additional effort in compiling and presenting the information on insurance conditions in place at the time of earthquake and losses which could not be paid due to these insurance conditions would have improved the data quality tremendously.

LESSONS LEARNED FROM LOSS ANALYSIS

From an insurance company's viewpoint, the insurance loss data include very valuable information regarding loss settlement, impact of insurance conditions and behavior of different kinds of insurance policies. Insurers can derive a set of vulnerability parameters from such a data that can be used in risk modeling tools. However, analysis of published insurance loss data from two big earthquakes of last decade indicates several limitations with the way insurance loss data is compiled and presented. It has been observed that the focus of published loss data is very often amount of claims paid, ignoring completely losses not paid due to insurance conditions and the unaffected portions of the portfolio. Due to these missing information, usability of the analysis results obtained from such published data can be severely limited.

Analysis of building damage estimates provided by the building safety inspectors and its comparison with insurance claims data clearly reveals that there is a big difference between a building surveyor's view of damage and damage from an insurance industry perspective. Focus of building damage survey carried out after an earthquake is usually to identify the unsafe buildings. So the buildings with very little exterior damage, but requiring substantial amount to repair, usually get overlooked during safety inspections. Additionally, shortage of experienced surveyor and building estimators would make it very difficult to assign an appropriate monetary value to losses sustained by structural, non-structural and contents in a wide area impacted by an earthquake. Due to these factors, vulnerability relationships derived using data from building safety inspectors would have large uncertainty and limited applicability in financial risk modeling.

Based on the observations made during loss analysis of published data, following recommendations are made for improving the value of assessed data from risk modelling perspective:

- As a first thing, the data should contain information on number of policies (risks) and total sums insured of these policies in the entire affected area, with geographic reference. This is easiest to obtain and one of the most important information needed for reasonable loss analysis.
- The loss information should include not only the total claims paid but also any filed claims, including the amount of loss that were not paid because of insurance conditions, namely deductible, limits and co-insurance.
- Information about at least the average deductible, limits and co-insurance would be very helpful.
- A very difficult to obtain but very useful information is the number and amount of loss to policies where losses remained below deductible. One source of this information could be number of claims reported but not paid. If available, this information should also be included for completeness of loss data. If not, we recommend insurance companies to store such data.
- If available, the data should be presented per policy. Geographically aggregating the data limits its usability, particularly in deriving loss distribution parameters and impact of insurance conditions.
- If available, the key building characteristics such as occupancy type, construction type, age, size etc. should also be indicated with the per policy data.

Summarizing, a major portion of policies or risks in an earthquake affected area typically do not suffer any loss. Given the importance of unaffected part the portfolio for building code developers, emergency planners or insurance industry, we propose a change in view point for future loss data collection efforts. Any future loss data collection effort should focus not only on the damaged risks but also provide answer to following questions in the order of importance:

1. How many risks did not suffer any loss?
2. What types of buildings did not suffer losses?
3. What types of structural systems/designs did not suffer losses?

CONCLUSIONS

Earthquake loss data contains very valuable information for emergency planners, building code developers and insurance industry. Results obtained by analyzing such loss data can help the local emergency planners be better prepared for eventuality. Building code developers can learn about how and why code improvements helped or not-helped in reducing the losses. And insurance industry can learn vital lessons which would help it in taking better decisions for its long term business continuity and survival.

Analysis of few loss data-sets from past earthquakes revealed their focus only on the damaged property and the paid losses. However, to help each of these agencies learn the desired lessons from the loss data, it is vital not to focus only on the damaged risks and paid loss amount, but also on the undamaged risks and the amount that was not paid out due to various reasons. With very little additional effort put in collecting such information, the value of loss data gathered after an earthquake can be multiplied.

REFERENCES

1. Bertogg M, Hitz L, Schmid E. “Vulnerability Functions derived from loss data for insurance risk modelling: findings from recent earthquakes”. 12^{th} European Conference on Earthquake Engineering, 2002.
2. Comerio MC, Landis JD, Firpo CJ. “Residential Earthquake Recovery – Improving California’s Post-Disaster Rebuilding Policies and Programs”. California Policy Seminar report, 1996.
3. Petak WJ and Elahi S. “The Northridge Earthquake, USA and its Economic and Social Impacts”. EuroConference on Global Changes and Catastrophe Risk Management. Earthquake Risks in Europe, IIASA, Laxenburg Austria, July 6-9, 2000.
4. Kircher CA, Reitherman RK, Whitman RV, Arnold, C. “Estimation of Earthquake Losses to Buildings”. Earthquake Spectra, November 1997, Volume 13, Number 4.
5. EQE International Inc. and GIS Group of Governor’s Office of Emergency Services “The Northridge Earthquake of January 17, 1994: Report of Data Collection and Analysis (PART A and B)”. Prepared for The Governor’s Office of Emergency Services of the State of California, 1995.
6. RMS Inc, Failure Analysis Associates Inc. “Event Report: Japan – The Great Hanshin Earthquake”. 1995
7. EQE International Inc. “The January 17, 1995 Kobe Earthquake – An EQE Summary Report”. April 1995.