Annual Report:

Trimming the Hazard Logic Tree, Phase 2

Prepared for
Southern California Earthquake Center
Zumberge Hall of Science 169
University of Southern California
3651 Trousdale Parkway
Los Angeles, CA 90089-0742

By
Keith Porter and Charles Scawthorn
SPA Risk LLC
Denver CO, USA

16 December 2009

Abstract. This document summarizes 2nd-year progress on an effort to test and depict the sensitivity of societal risk estimates to branches in the UCERF hazard logic tree. The work is not yet complete. In this phase, probabilistic seismic vulnerability functions have been created that relate building repair costs to shaking intensity, by structure type and occupancy classification. Intensity is measured using a vector measure: 5%-damped elastic spectral acceleration response at 0.3-sec and 1.0-sec periods, also conditioned on magnitude, distance, site soil classification, and tectonic regime. Casualty-rate seismic vulnerability functions were previously created for another (USGS) project and both mean and probabilistic seismic vulnerability functions of repair cost were created for SCEC under the SCEC 2008 year. In the current year, a portfolio of assets exposed to seismic risk was also estimated, in work for this 2009 SCEC project and another USGS project. (The USGS work quantified indoor occupants; the SCEC work added square footage and replacement costs, by census tract, occupancy classification, and structure type.) A component of the OpenRisk software, designed for SCEC in previous work and developed in collaboration with USGS programmers, will be used to carry out the loss calculations. The sensitivity analysis, not yet begun, will employ a tornado-diagram-analysis approach developed for decision analysis and applied and extended in the last 10 years by SPA personnel and others for use in earthquake engineering loss estimation.

Table of Contents

Index of Figures

Index of Tables

1   INTRODUCTION AND OBJECTIVES

The goal of the present project is to assess the sensitivity of selected measures of societal risk to selected branches in the hazard model, in particular to the branches of the Uniform California Earthquake Rupture Forecast, either version 2 or the just-beginning UCERF 3. By "measures of societal risk" is meant expected annualized loss—either economic or human deaths and injuries—or aspects of the loss-exceedance curve for the same measures. A loss-exceedance curve here means a relationship between loss and mean exceedance frequency. The value of such a capability is that it will provide insight into the relative importance of scientific inquiry into this or that aspect of the ERF in particular or the hazard model in general.

The insight can help in two ways: first, by identifying branching points that do not matter much to important measures of risk, this "tree-trimming" capability can help to reduce computational effort expended on risk modeling. One could in principle pick a single branch and not iterate or simulate over several. A second, possibly less important and more controversial use is to identify uncertainties that warrant, because of their effect on societal risk, additional study. For example, if an uncertainty in seismic hazard can potentially be reduced by further study, this might potentially reduce uncertainty in a practical risk model, such as the loss-exceedance curve of an insurer, which in turn affects reinsurance needs and costs.

This is important: we do not wish to imply that scientific inquiry should be judged solely by its effects on uncertainty in social or economic risk. Scientific inquiry into any particular topic of the hazard model has its own internal values and drivers: curiosity of the investigators, relevance to other aspects of seismic hazard, broader relevance outside of seismology, etc. But nonscientists affected by science may wish to direct research resources to topics with a potentially important practical benefit to them, and this tree-trimming capability can potentially aid in those decisions. Others for whom computational expense can be significantly reduced by eliminating relatively immaterial branches to an uncertain hazard model may also value the capability aimed for here. To achieve this goal requires a few developments, begun in the SCEC 2008 year and continued here.

1. An estimated portfolio of assets exposed to shaking—building occupants, building value, etc. The "portfolio" comprises an estimate, by relatively small geographic area such as census block or census tract, of the quantity of people and building replacement cost by structure type and possibly occupancy class (if breakout of risk by occupancy class is important).

2. A computer model capable of estimating the expected annualized loss (EAL) for a portfolio of assets exposed to seismic shaking. The model must be capable of estimating EAL for single paths along the hazard logic tree.

3. A similar model capable of estimating the portfolio loss exceedance curve (LEC).

4. A set of (mean) seismic vulnerability functions for deaths and nonfatal injuries of indoor occupants in each structure type in the portfolio. A mean seismic vulnerability function here refers to a relationship between the expected number of deaths or injuries (typically as a fraction of occupants, i.e., the mean damage factor) and a scalar or vector shaking intensity, whether shaking is measured in terms of damped elastic spectral acceleration response at some index period, or instrumental intensity, or some combination of these with one or more of magnitude, distance, site class, etc.

5. A set of mean seismic vulnerability functions for the economic repair costs for each structure type in the portfolio. Again, this is a relationship between the expected value of loss—this time as a fraction of replacement cost, new—and shaking intensity.

6. For an LEC of repair costs, the repair-cost seismic vulnerability functions must be probabilistic, i.e., providing an estimate of the probability distribution of repair-cost damage factor of the asset class as a function of shaking intensity.

7. For an LEC of casualties, the casualty-rate seismic vulnerability functions must be probabilistic, i.e., providing an estimate of the probability distribution of casualty rate for each injury level and structure type, as a function of shaking intensity.

8. A methodology for quantifying sensitivity of the loss estimate to each logic-tree branch in question.

2   PROGRESS OF THE WORK

2.1   ESTIMATED PORTFOLIO OF ASSETS: DONE; DETAILS IN APPENDIX

The work has begun somewhat out of order. In work for the USGS's PAGER project, the number of indoor occupants was estimated and tabulated by census tract, type of occupancy, and structure type. In 2009-2010 work for SCEC, the replacement cost of buildings and contents was estimated and tabulated alongside indoor occupants. The work amounted to extracting the inventory from HAZUS-MH, though it was a little more difficult than it sounds. The HAZUS-MH "inventory" is contained in a normalized database: it requires a sequence of queries of up to 15 tables to extract sufficient inventory information to compile the basic unit of the portfolio required for the present work. That is, there is no single table in the HAZUS-MH database that can answer the question: How many people are there buildings in census tract W, at time of day X, in structure type Y, and how much building square footage, building value and content value in those buildings in occupancy classification Z? To extract this inventory information required denormalizing the HAZUS-MH database, an effort documented in detail in Porter (2009a), copied to Appendix 1. A sample of the database is shown in Table 1. The fields are defined in Table 2.

Table 1. Sample of California portfolio

ID	Tract	OccLabel	SsType	DesignLevel	A	Vb	Vc	PopDay	PopNight	PopCommute
349	06001400100	RES1	W1	MC	1482	249794	124905	247	1544	546
350	06001400100	RES3A	C2L	HC	2.00	184	92	0	2	1
351	06001400100	RES3A	C2L	LC	0.65	59	30	0	1	0
352	06001400100	RES3A	C2L	MC	1.94	178	90	0	2	1
353	06001400100	RES3A	C3L	MC	0.65	59	30	0	1	0

Table 2. Layout of inventory table

Field name	Data type	Description	Comment
ID	Autonumber	An index	1, 2, …
Tract	Text, 11	Census tract number	e.g., 06001400100, from hzTract.Tract
BldgSchemesId	Text, 5	Identifies scheme to distribute from occupancy to material, structure type and design level	In California, distinguishes L/M/H hazard
OccLabel	Text, 5	HAZUS-MH occupancy label	RES1, RES2, … EDU2
SsType	Text, 10	FEMA earthquake structure type	W1, W2, … or MH
DesignLevel	Text, 2	HAZUS-MH seismic design level	PC, LC, LS, MC, MS, HC, or HS
A	Double	1000 sq ft	See Porter (2009a) for deriv
ValYr	MM/DD/YYYY	Year in which dollar valuation is made	Default = 1/1/2003
Vb	Double	Building replacement cost, $1000s	See Porter (2009a) for deriv
Vc	Double	Content replacement cost, $1000s	See Porter (2009a) for deriv
PopDay	Long integer	Daytime population, people at 2 PM	See Porter (2009a) for deriv
PopNight	Long integer	Nighttime population, people at 2 AM	See Porter (2009a) for deriv
PopCommute	Long integer	Commute-time population, people at 5 PM	See Porter (2009a) for deriv

2.2 PORTFOLIO EAL CALCULATOR: DONE; NEEDS REVISION

In previous work for the US Geological Survey and SCEC, we addressed item 2 and designed OpenRisk (Porter and Scawthorn 2007, 2009), open-source software that extends USGS and SCEC's suite of open-source seismic hazard analysis software OpenSHA to calculate damage and loss. We specified a portfolio data file format, and in collaboration with USGS, SCEC, and Instrumental Software Technologies, Inc., created a portfolio import tool and a portfolio EAL calculator. The OpenRisk Portfolio EAL calculator is currently available at www.risk-agora.org and its source code is available within the OpenSHA software repository. The software currently does not include the suite of seismic vulnerability functions referred to in items 4 and 5.

2.3 PORTFOLIO LEC CALCULATOR: 10% COMPLETE

In meetings with Field and others, we have discussed the algorithm required to perform a probabilistic portfolio risk calculation, but coding and testing has not yet begun. A single-site LEC calculator has been completed in collaboration with a USGS programmer, but this tool is inappropriate to the present task. It can be used to estimate the sensitivity of loss at a single point to branches of the hazard logic tree, but not societal risk.

2.4 CASUALTY RATE VULNERABILITY FUNCTIONS: DONE, PUBLISHED

In the 2008-2009 work for USGS and SCEC, we created the (mean) seismic vulnerability functions called for in item 4. In Porter (2009b) we created and published a set of mean seismic vulnerability functions for human deaths and injuries. The work on fatalities was performed largely for USGS, but is relevant to the SCEC 2008-2009 work as well, along with the seismic vulnerability functions for nonfatal injuries. These seismic vulnerability functions essentially extract functional relationships that are implicit in the HAZUS-MH model; see especially Kircher and Whitman (1997) and NIBS and FEMA (2003). They honor all the HAZUS-MH data and methods, while avoiding the iteration required by the structural analysis component of the method, and allow for risk analysis outside of HAZUS-MH. A sample seismic vulnerability function for mean indoor fatality rate is shown in Figure 1. It shows mean fatality rate for a high-code W1 (woodframe building < 5000 square feet), western US site, NEHRP site soil classification D, medium-duration shaking (magnitude 6.5 $\leq M < 7.5$), at a distance 15 km $\leq R < 30$ km. See Porter (2009b) for derivation of the curve and a sample calculation at the point highlighted by an open circle in the figure.
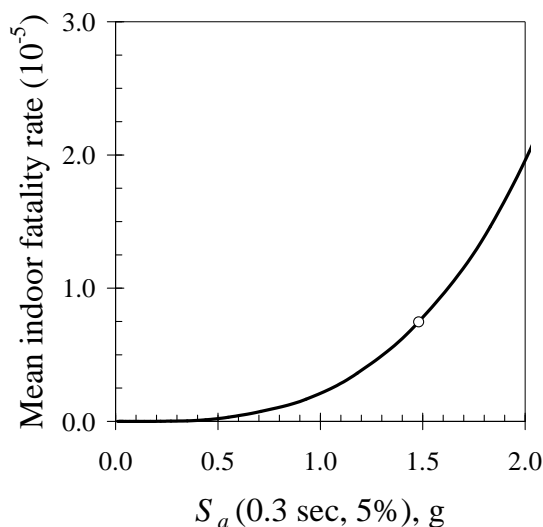

Figure 1. Sample vulnerability function

2.5 REPAIR COST VULNERABILITY FUNCTIONS: DONE, PUBLISHED

In 2008 work for SCEC, we created mean seismic vulnerability functions called for in item 5: repair cost as a fraction of replacement cost (new) for each building type in the portfolio. In Porter (2009c) we created and published a set of mean seismic vulnerability functions for repair cost as a fraction of replacement cost (new) for each structure type treated by HAZUS-MH. The work has been peer reviewed and duplicated by several other engineers. A sample of the resulting seismic vulnerability functions is shown in Figure 2. It depicts the mean damage factor (repair cost divided by replacement cost, new) of a W1 building, high-code design level, single-family dwelling (RES1) occupancy, NEHRP site soil classification D, medium-duration shaking (magnitude $6.5 \leq M < 7.5$), at a distance of 15 km $\leq R < 30$ km. Dot shows results of a sample calculation in Porter (2009c).
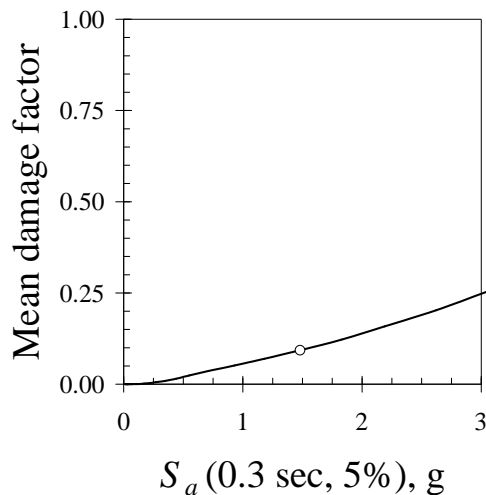
Figure 2. Sample repair-cost seismic vulnerability function

2.6 PROBABILISTIC REPAIR-COST VULNERABILITY: DONE, IN PRESS

This work amounts to evaluating a standard deviation or coefficient of variation of repair cost for each structure type at each of many intensity levels. Together with the (mean) repair-cost seismic vulnerability functions discussed above, and an assumption of the conditional probability distribution of repair cost conditioned on structure type and intensity. The work, performed for the 2009 SCEC year, is documented in a manuscript that has been accepted for publication in *Earthquake Spectra* (Porter ND) and included in Appendix 2. Sample results are illustrated in Figure 3, which shows that coefficient of variation of repair-cost damage

factor tends to decrease with increasing mean damage factor, while the standard deviation tends to increase with mean damage factor.

Figure showing plot (a) with y-axis "Coefficient of variation" (0 to 5) and x-axis "Mean damage factor" (0.00 to 0.20), with equation $y = 0.24x^{-0.48}$, $R^2 = 0.85$. Plot (b) with y-axis "Standard deviation" (0.00 to 0.25) and x-axis "Mean damage factor" (0.00 to 0.20), with equation $y = 0.24x^{0.52}$, $R^2 = 0.87$.

(a)   (b)

Figure 3. Trends in damage-factor uncertainty versus mean damage factor

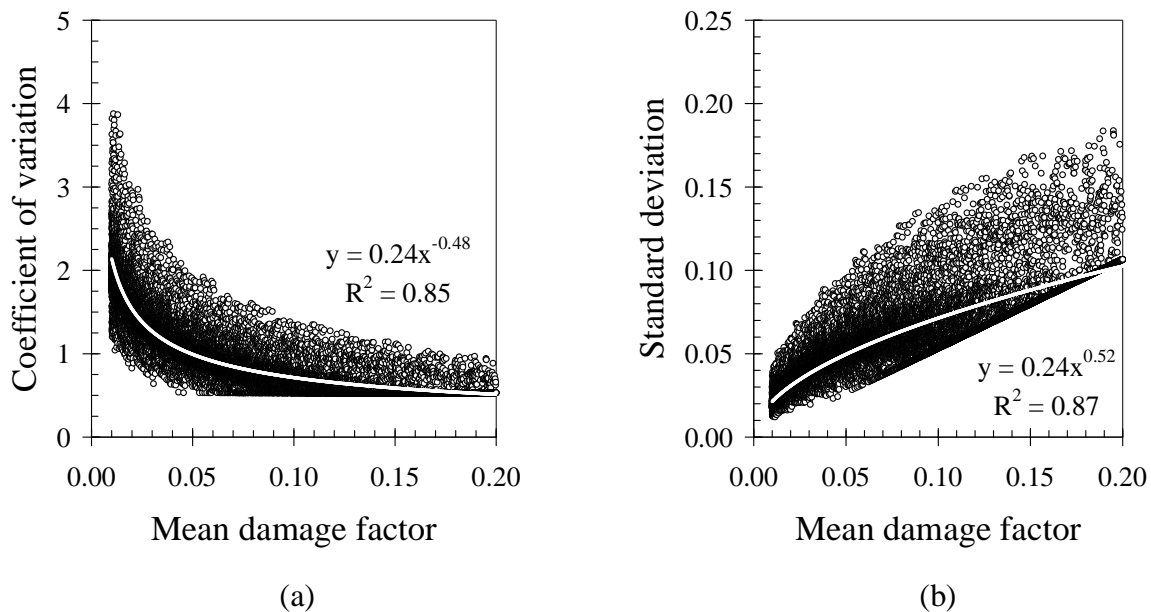## 2.7   PROBABILISTIC CASUALTY-RATE VULNERABILITY: NOT BEGUN

This work amounts to evaluating a standard deviation or coefficient of variation of casualty rate for each injury severity level and each structure type at each of many intensity levels. Together with the (mean) casualty-rate seismic vulnerability functions discussed above, and an assumption of the conditional probability distribution of casualty rate conditioned on structure type and intensity. The work has not yet begun.

2.8   SENSITIVITY METHODOLOGY: DONE, PUBLISHED

We will use tornado-diagram analysis to assess the sensitivity of societal loss to branches in the hazard logic tree. The method appears to have been developed for use in decision analysis. Described in Howard (1988), it is a method to test and depict the sensitivity of a scalar function to uncertainty in its arguments. In brief, one evaluates the function with all its arguments set to their expected value, except for one, which is set first to a lower bound, the function evaluated, and then the argument is set to its upper bound and the function evaluated again. The difference between the two results is a measure of the sensitivity of the function to

the varied argument, and referred to as the swing of the function with respect to that argument. The process is repeated for all the other arguments, which are then sorted in order of decreasing swing. (The parameters already tested are set to their expected value in all the subsequent tests; i.e., only one parameter is varied at a time.) The results are plotted in a horizontal bar chart where the argument with the greatest swing is the topmost item on the y-axis of the bar chart, the argument with the second-largest swing beneath that, etc. Beside each argument is a bar: a thin rectangle with its left and right edges spanning from between the two values of the function at the two bounds. The x-axis depicts the value of the function. Since the arguments are arranged in from top down in order of decreasing swing, the resulting chart resembles a tornado in profile. The methodology has been applied to earthquake engineering problems, e.g., in Porter et al. (2002) or MMC (2005), the latter of which shows how to select the lower and upper bounds and combine the results to estimate the first several moments of the probability distribution of the function, using a method presented in Ching et al. (2003, 2008). Figure 4 shows a sample tornado diagram, from Porter et al. (2002). It depicts the sensitivity of building repair cost for a particular building to major uncertain variables in a performance-based earthquake engineering analysis. It shows that, for the particular building under consideration, building repair cost is most sensitive to the capacity of the individual building components to resist damage. This variable is estimated to be more important than the uncertainty in the maximum shaking intensity that the building is expected to experience in 50 years (denoted by Sa in the figure).
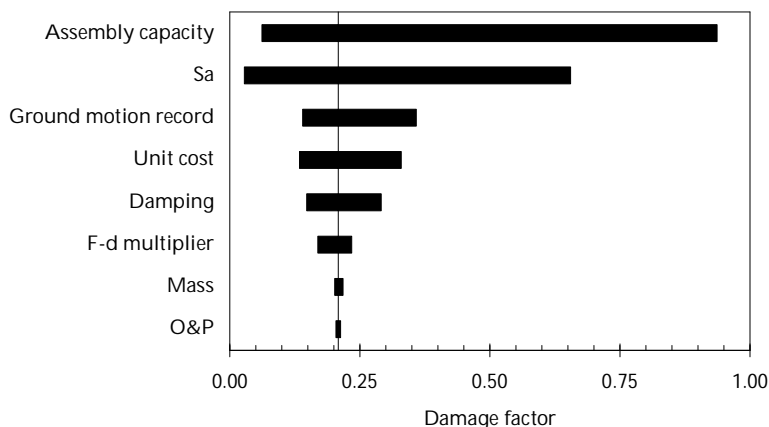

Figure 4. Sample tornado diagram

3   CONCLUSIONS

Tornado-diagram analysis (Howard 1988) will be used to test and depict the sensitivity of societal risk measures to branches in the hazard logic tree. The risk measures can be the expected annualized loss (EAL) in terms of human deaths or injuries, points on the loss exceedance curve (LEC) for human deaths or injuries such as the value with 2% exceedance probability in 50 years, or similar EAL or LEC measures of building repair cost. Tornado diagrams have been used in earthquake engineering before, such as to quantify the sensitivity of building repair cost to major uncertain variables in a performance-based earthquake engineering analysis (Porter et al. 2002) or to quantify the first several moments in the probability distribution of societal benefits resulting from FEMA's multihazard risk mitigation efforts between 1993 and 2003 (MMC 2005). The latter uses a quadrature method presented by Ching et al. (2003, 2008).

To carry out the analysis will require an extension of the OpenRisk software (Porter and Scawthorn 2007, 2009) begun for SCEC and USGS in collaboration with Field and others. A portfolio EAL calculator is currently capable of estimating sensitivity of portfolio EAL to branches in the logic tree, but it needs to be modified to accept new seismic vulnerability functions. We are discussing those modifications with Field, with the expectation that a USGS or SCEC programmer will perform the modifications. We have begun the design of a portfolio LEC calculator.

The analysis will also require mean and probabilistic seismic vulnerability functions of casualties and building repair costs. These have been completed and have either been published (Porter 2009b, c) or accepted for publication (Porter ND), with the exception of probabilistic seismic vulnerability functions for casualties, which have not yet begun development.

Finally the analysis requires an estimated portfolio of assets exposed to seismic risk: number of building occupants by time of day, building square footage, building replacement cost, and contents replacement cost, by geographic area (e.g., census tract), occupancy classification (e.g., single-family dwelling), and structure type (e.g., small woodframe dwelling, pre-code construction). This inventory work has been completed for California, and is also available for the rest of the United States (Porter 2009a).

4 REFERENCES CITED

Ching, J.Y., K.A. Porter, and J.L. Beck, 2003. *Uncertainty Propagation and Feature Selection for Loss Estimation in Performance-Based Earthquake Engineering*, EERL Report No. 2003-03, California Institute of Technology, Pasadena, CA http://caltecheerl.library.caltech.edu/443/01/EERL-2004-02.pdf

Ching, J.Y., K.A. Porter, and J.L. Beck, 2008. Propagating uncertainties for loss estimation in performance-based earthquake engineering using moment matching. *Structure and Infrastructure Engineering,* http://www.informaworld.com/10.1080/15732470701298323

Howard, R.A., 1988. Decision analysis: practice and promise. *Management Science* 34 (6), 679-695.

Kircher, C.A. and R.V. Whitman, 1997. Estimation of earthquake losses to buildings, *Earthquake Spectra*, **13** (4) 703-720

(MMC) Multihazard Mitigation Council, 2005. *Natural Hazard Mitigation Saves: An Independent Study to Assess the Future Savings from Mitigation Activities*, National Institute of Building Sciences, Washington, DC.

(NIBS and FEMA) National Institute of Building Sciences and Federal Emergency Management Agency, 2003. *Multi-hazard Loss Estimation Methodology, Earthquake Model, HAZUS$^{®MH}$ Technical Manual,* Federal Emergency Management Agency, Washington, DC, 690 pp.

Porter K.A. and C. Scawthorn, 2007. OpenRisk: open-source risk software and access for the insurance industry. *1st International Conference on Asian Catastrophe Insurance (ICACI)*, Kyoto University, Dec 3-4, 2007

Porter, K.A. and C. Scawthorn, 2009. *Development of Open Source Seismic Risk Modeling Framework*. Final Technical Report, award 07HQAG0002, US Geological Survey, Reston VA, http://earthquake.usgs.gov/research/external/

Porter, K.A., 2009a. Cracking an open safe: extracting HAZUS's building inventory to a single table. Alliance for Global Open Risk Analysis, www.risk-agora.org

Porter, K.A., 2009b. Cracking an open safe: HAZUS vulnerability functions in terms of structure-independent spectral acceleration. *Earthquake Spectra,* 25 (2), 361-378

Porter, K.A., 2009c. Cracking an open safe: more HAZUS vulnerability functions in terms of structure-independent spectral acceleration. *Earthquake Spectra,* 25 (3), 607-618.

Porter, K.A., J.L. Beck, and R.V. Shaikhutdinov, 2002. Sensitivity of building loss estimates to major uncertain variables. *Earthquake Spectra*, 18 (4), 719-743

Porter, K.A., ND. Cracking an open safe: uncertainty in HAZUS-based seismic vulnerability functions. Accepted for publication in *Earthquake Spectra*, Nov 2009

APPENDIX 1: BUILDING INVENTORY

Cracking an Open Safe: Extracting HAZUS's Building Inventory to a Single Table

Keith Porter

It can be valuable to perform seismic risk studies of buildings at the societal level, which requires an inventory of the building stock exposed to damage. HAZUS-MH offers a nationwide default building-stock inventory, but it can be difficult to use outside of HAZUS-MH (as is sometimes desirable): it is encoded in 15 tables in 2 Microsoft Access databases for each of 50 states. A new Microsoft Access database is developed here that links to the HAZUS-MH tables, and includes 54 scripted query language (SQL) queries, a macro to perform them all, and a number of supporting tables, for the purpose of extracting the HAZUS-MH inventory to a single de-normalized table, which is more practical for use by PAGER and some other applications. The table shows by census tract, county or state: square footage of construction, building value, content value, and number of indoor occupants at 2 PM, 2 AM, and 5 PM. The quantities are distributed by HAZUS-MH occupancy classification, structure type, and design level. The database and results for several states are available at www.risk-agora.org.

INTRODUCTION

To perform a seismic risk analysis at the societal level requires a database of the assets exposed to risk (people, property, or both), the seismic hazard to which they are exposed, their vulnerability (loss—dollars, deaths, or downtime—as a function of seismic excitation), and an algorithm to integrate hazard and vulnerability to calculate the desired loss metric. Seismic vulnerability functions are readily available for a variety of asset types, e.g., empirically derived functions such as Algermissen et al. (1972), ones based on expert opinion such as ATC (1985), the largely analytical models of HAZUS-MH (Kircher et al. 1997, NIBS and FEMA 2003, Porter 2009a,b), and building-specific performance-based vulnerability functions such as Porter et al. (2002, 2004) or Goulet et al. (2007). Seismic hazard information is also readily available for several countries; for example, in the United States see Petersen et al. (2008) or Field (2005).

But the first ingredient, societal-level inventories of buildings with seismic attributes, can be difficult to acquire. There is a nationwide database of the estimated building stock encoded in

HAZUS-MH, developed explicitly for emergency planning, preparedness, and mitigation decision-making. However, there are sometimes reasons to perform such a study outside of HAZUS-MH, e.g., to employ an alternative depiction of seismic hazard, to perform the loss calculations more quickly, or to integrate the loss estimate with other functions. All three reasons are present in the US Geological Survey's Prompt Assessment of Global Earthquakes for Response (PAGER), which aims to estimate fatalities within minutes after an earthquake, and notify decision-makers involved in humanitarian assistance. PAGER already comprises software to rapidly calculate shaking intensity and evaluate vulnerability functions; it needed for one of its methods a simple inventory of US building stocks.

Developing such an inventory (or rather, extracting it from HAZUS-MH data) is the subject of this work. As used here, an inventory comprises an estimate of people, building area, and property replacement cost by structure type, occupancy classification, and geographic area. Structure type must be defined at the same level of detail as the property vulnerability functions. Occupancy classification must be defined at the same level of resolution as the business interruption model requires. And geographic area must be small enough to capture spatially varying effects of fault distance and site soil classification.

SOME RELATED INVENTORY WORK

In the United States, county tax assessors collect a variety of building data for purposes of calculating real estate taxes. These data tend to be closely held by the individual county tax assessors, each with its own idiosyncratic content and format, and usually lacking earthquake vulnerability attributes (lateral force resisting system, often height category, number of occupants, replacement cost new, etc.).

ATC-13 (1985) offers a methodology (though not an actual database) for estimating buildings stocks in terms of square footage, replacement cost, and number of occupants at three times of day, by structure type, occupancy classification, and census area. The methodology uses census proxies, especially the population and housing census (which enumerates residents by census area) for residential occupancies and the economic census (which gives employment by economic sector by census area) for commercial and industrial ones. In general, one factors number of people by square footage per person and replacement cost per square foot to arrive at building area and value by census area and occupancy classification. Then one distributes these totals among a number of structure types, with the fraction of the population in each structure

type varying by occupancy classification. These fractions are arrived at by judgment. The HAZUS-MH developers (Kircher et al. 1997, NIBS and FEMA 2003) extended the ATC-13 methodology and applied the new methodology to estimate building stocks by census tract for the entire United States. More on the HAZUS-MH approach later.

A few attributes of durable housing have been compiled by the United Nations Human Settlements Program for some 60 countries (UN-HABITAT 2007), including floor, wall, and roof material. The United Nations has also compiled housing census data from 170 countries; data from 73 of these countries are readily available and include an indication of the exterior wall material. The United States Census Bureau compiles at the level of census blocks data about the area, age, replacement cost, and other attributes of US housing, though little information about engineering features. The World Housing Encyclopedia has compiled profiles of the structural engineering characteristics of more than 100 buildings that are typical of common housing in 40 countries. Jaiswal et al. (ND) have synthesized these and other sources to estimate country-level distribution of building stocks. The work required a good deal of guesswork, but seems to be the first-ever global database of earthquake-resistance structure types.

Some cities and countries have created building censuses where buildings are individually examined and their seismic attributes tabulated. Australia has completed a census of residential buildings (Edwards et al. 2004) and is in the process of creating a census of commercial buildings (Edwards 2009). Istanbul's building stock has been inventoried through remote sensing and field reconnaissance by survey workers (Erdik et al. 2003). The Extremum software (Shakhramanian et al. 2000) contains a global database of buildings by settlement, but the data are not publically available.

For the United States, the only complete and geographically detailed inventory for earthquake risk-modeling purposes seems to be the HAZUS-MH inventory. The HAZUS-MH default inventory is contained in a relational database that is derived similarly to ATC-13 (1985). The user is free to replace or enhance the tables to improve the estimated building stock. In keeping with standards of conventional normalized databases, the HAZUS-MH inventory does not contain a single table that comprises the inventory in the form desired here: people, area, and replacement cost, by structure type, occupancy classification, and census area. Instead, it contains for each state a number of tables in two database files that are summarized next. From these, the present work derives the desired inventory table.

HAZUS-MH INVENTORY DATA

The HAZUS-MH tables are laid out somewhat inconveniently for present purposes, so in several places in this discussion, duplicate tables are introduced that contain the same information as the HAZUS-MH tables but in a more convenient format for the present objective. In general, the redesign aims at combining HAZUS' smaller relational tables into fewer (but bigger) tables. The objective is to reduce the need to join and query numerous small tables at the time when a risk calculation is performed.

In database terminology this process is referred to as denormalizing the data, which according to Wikipedia means, "the process of attempting to optimize the performance of a database by adding redundant data or by grouping data…. A relational normalized database imposes a heavy access load over physical storage of data even if it is well tuned for high performance." A normalized database is one with several small tables in which each fact appears only once. Many researchers are demonstrating that while denormalized databases can be weaker than normalized ones at data integrity (more onus is on the application developer), they can be stronger at delivering results with speed and scalability. See, e.g., Bock and Schrage (1996) or Shin and Sanders (2006) for a discussion of the costs and benefits of denormalized databases. For present purposes, the benefits of simplicity and scalability outweighed the costs.

For each state, there are two database files that are relevant here: bndrygbs.mdb and MSH.mdb. The former quantifies the square footage, replacement cost, and number of occupants by census area and occupancy classification, but not structure type. The latter contains, among other things, the information needed to distribute these quantities among five general construction materials, and thence to detailed structure types. It seems that the developers intend to produce a public HAZUS-MH data dictionary (see, e.g., a reference in the HAZUS-MH earthquake user manual to an Appendix F), but have not yet done so, and of course the developers' internal documentation for the HAZUS-MH databases (the schema and data dictionary) are not publicly available. It can therefore be difficult for the (perhaps rare) outside user to understand the meaning of the tables and their fields. For the present effort it was helpful to have a summary of the relevant tables, so the following table descriptions are offered both for the purpose of understanding where the data come from to create the inventory desired here, and for general reference. These descriptions are based on the author's examination of the database tables, in particular those published with HAZUS-MH MR3 (v1.3), dated September 2007.

BNDRYGBS.MDB: QUANTITY BY CENSUS AREA AND OCCUPANCY CLASS

Let us first consider the HAZUS-MH data file bndrygbs.mdb. This database contains tables that provide quantities (number of buildings, people, households, square footage, building value, and content value) by census block and tract. Where a table name ends in a capital T, it contains tract-level information. Tables ending in B contain the same sort of information at the census-block level. The present work deals only with tract-level data. The relevant tables are as follows.

hzDemographicsT: population by census tract. This table contains number of people in various demographic categories living or working in each census tract. For example, in census tract 06001400100 (Alameda County, in the San Francisco Bay Area), there are 1145 households, 2498 residents, most of whom live in single-family dwellings, 746 people in residences in the day, 2474 at night, 590 workers in commercial occupancies, 31 industrial employees, and 1,034 single family dwellings. Other demographic data in that table that are irrelevant here, such as household income. Absent from in the database but relevant here are the contents of NIBS and FEMA (2003) Table 13.2, duplicated here in **Error! Reference source not found.**, which shows how the demographics data are used by HAZUS-MH to estimate number of occupants in various categories at various times of day.

Table 3. Default relationships for estimating population distribution (NIBS and FEMA 2003 Table 13.2)

Distribution of People in Census Tract			
Occupancy	**2:00 a.m.**	**2:00 p.m.**	**5:00 p.m.**
Indoors			
Residential	(0.999)0.99(NRES)	(0.70)0.75(DRES)	(0.70)0.5(NRES)
Commercial	(0.999)0.02(COMW)	(0.99)0.98(COMW) + (0.80)0.20(DRES) + 0.80(HOTEL) + 0.80(VISIT)	0.98[0.50(COMW) + 0.10(NRES)+ 0.70(HOTEL)]
Educational		(0.90)0.80(GRADE) + 0.80(COLLEGE)	(0.80)0.50(COLLEGE)
Industrial	(0.999)0.10(INDW)	(0.90)0.80(INDW)	(0.90)0.50(INDW)
Hotels	0.999(HOTEL)	0.19(HOTEL)	0.299(HOTEL)
Outdoors			
Residential	(0.001)0.99(NRES)	(0.30)0.75(DRES)	(0.30)0.5(NRES)
Commercial	(0.001)0.02(COMW)	(0.01)0.98(COMW) + (0.20)0.20(DRES) + (0.20)VISIT + 0.50(1-PRFIL)0.05(POP)	0.02[0.50(COMW) + 0.10(NRES) + 0.70(HOTEL)] + 0.50(1-PRFIL) [0.05(POP) + 1.0(COMM)]
Educational		(0.10)0.80(GRADE) + 0.20(COLLEGE)	(0.20)0.50(COLLEGE)
Industrial	(0.001)0.10(INDW)	(0.10)0.80(INDW)	(0.10)0.50(INDW)
Hotels	0.001(HOTEL)	0.01(HOTEL)	0.001(HOTEL)
Commuting			
Commuting in cars	0.005(POP)	(PRFIL)0.05(POP)	(PRFIL)[0.05(POP) + 1.0(COMM)]
Commuting using other modes		0.50(1-PRFIL)0.05(POP)	0.50(1-PRFIL) [0.05(POP) + 1.0(COMM)]

The quantities referred to in **Error! Reference source not found.** are given in fields in the HAZUS-MH database with slightly different names, in table hzDemographicsT, as shown in **Error! Reference source not found.**.

Table 4. Relationship between terms in NIBS and FEMA (2003) Table 13.2 and fields in hzDemographicsT

Table 13.2	Field in hzDemographicsT	Definition
POP	Population	Census tract population taken from census data
DRES	ResidDay	Daytime residential population inferred from census data
NRES	ResidNight	Nighttime residential population inferred from census data
COMM	Commuting5PM	Number of people commuting inferred from census data
COMW	WorkingCom	Number of people employed in the commercial sector
INDW	WorkingInd	Number of people employed in the industrial sector
GRADE	SchoolEnrollmentKto12	Number of students in grade schools (K-12)
COLLEGE	SchoolEnrollmentCollege	Number of students on college and university campuses in the census tract
HOTEL	Hotel	Number of people staying in hotels in the census tract
PRFIL		A factor representing the proportion of commuters using automobiles, inferred from profile of the community (0.60 for dense urban, 0.80 for less dense urban or suburban, and 0.85 for rural). The default is 0.80.
VISIT	Visitor	Number of regional residents who do not live in the study area, visiting the census tract for shopping and entertainment. Default is set to zero.

hzSqFootageOccupT: building square footage (1000s) by census tract and occupancy classification. This table contains building square footage (apparently in 1000s) by census tract and occupancy classification (RES1, RES2, … EDU2). Let us denote an entry in the table as $A_{tract,occ}$ where A denotes area, *tract* refers to the census tract (e.g., 06001400100) and *occ* refers to occupancy classification (e.g., RES1). For example, in census tract 06001400100, there appears to be 2,268,789 sq ft of building area in single-family dwellings (RES1). This figure seems slightly higher than average but in the right order of magnitude: dividing 2.3 million square feet among 1,034 single family dwellings suggests an average single family dwelling of slightly more than 2,000 sf; 1,500 sf would be more typical, but in an affluent community 2,000 sf seems realistic.

The table is laid out with tract values for each of the 33 occupancies RES1-EDU2 in separate columns, which will prove inconvenient later. As hinted at earlier, a duplicate table named hzSqFootageOccupT2 was created with 2 fields replacing the 33 fields: one named OccLabel

(and containing one of the values "RES1," "RES2," etc.) and another labeled A and containing the area of the appropriate occupancy type in hzExposureOccupT.

Occupancy groups. The new table hzSqFootageOccupT2 also contains a new field labeled Frac, indicating the fraction of area in the general occupancy group. The term "occupancy group" is introduced here for use in distributing the demographic data to the detailed occupancy classes. The groups are listed in **Error! Reference source not found.**.

Table 5. Occupancy groups used here

Group	Occupancies
RES	RES1, RES2, RES3A-RES3F, RES5, RES6
HOTEL	RES4
COM	COM1-COM10
IND	IND1-IND6
GOV	GOV1, GOV2
EDU1	EDU1
EDU2	EDU2
REL	REL1
AGR	AGR1

hzExposureOccupT: *building replacement cost ($1000s) by census tract and occupancy classification*. This table contains building replacement cost (apparently in $1000s) by census tract and occupancy classification (RES1, RES2, … EDU2). Let us denote an entry in the table as $Vb_{tract,occ}$ where Vb denotes building value, *tract* means census tract (e.g., 06001400100) and *occ* means occupancy classification (e.g., RES1). For example, in census tract 06001400100, there appears to be $382,298,000.00 in building replacement cost in single-family dwellings (RES1). Dividing Vb by A for this tract and occupancy classification suggests $170/sq ft replacement cost, which is reasonable—perhaps somewhat low—for California residential construction in Alameda County in 2003.

The table is laid out with tract values for each of the 33 occupancies RES1-EDU2 in separate columns, which will prove inconvenient later, so a duplicate table hzExposureOccupT2 was created with the 33 fields replaced by two, one named OccLabel (and containing one of the values "RES1," "RES2," etc.) and another labeled V and containing the quantity for the appropriate occupancy type.

hzExposureContentOccupT: content replacement cost ($1000s) by census tract and occupancy classification. This table contains content value (apparently in $1000s) by census

tract and occupancy classification (RES1, RES2, … EDU2). Let us denote an entry in the table as $Vc_{tract,occ}$ where Vc denotes content value. For example, in census tract 06001400100, there appears to be \$191,161,000 of content value in single-family dwellings (RES1), equivalent to half the value of buildings in the same group, which is reasonable for California residential construction.

The table is laid out with tract values for each of the 33 occupancies RES1-EDU2 in separate columns, which will prove inconvenient later, so a duplicate table hzExposureContentOccupT2 was created with the 33 fields replaced by two, one named OccLabel (and containing one of the values "RES1," "RES2," etc.) and another labeled V and containing the quantity for the appropriate occupancy type.

MSH.MDB: FROM OCCUPANCY TO STRUCTURE TYPE AND DESIGN LEVEL

This database contains tables that use as input quantities in a census area that are already distributed by occupancy class to structure type and design level, and distribute them first among materials (wood, steel, concrete, masonry, and mobile home), and from materials to detailed structure types.

hzGenBldgScheme: building material distribution by occupancy. This table provides the distribution of building quantity (sq ft or replacement cost) among five building materials (concrete, steel, masonry, wood, and mobile homes), by detailed occupancy classification. It provides distribution schemes labeled BldgSchemeId. In California, there are three: "CA1," "CA2," and "CA3". For each BldgSchemeId and occupancy (33 types mentioned above), there is one record, for a total of 99 records in California. Each record has a unique label, in the field GenBldgSchemeId, from "CA1" to "CA99." For example, under BldgSchemeId = "CA2," one record (GenBldgSchemeId = "CA34") distributes building quantities in RES1 occupancy among the five building materials as follows: 99% wood, 1% masonry, 0% everything else.

For example, given BldgSchemeId = "CA2," we can calculate the square footage of single-family dwellings (RES1) in census tract 06001400100 made of wood. The census tract has an estimated 2,268,789 sf of single-family dwellings, so 99% of it, or 2,246,101 sf, are of woodframe buildings. The native table hzGenBldgScheme contains percentage of census-tract square footage laid out in separate columns, one for each material. This layout will be somewhat inconvenient layout for later calculations, so for present purposes a duplicate table named hzGenBldgScheme2 was created with a slightly different layout: two field labeled "Material"

and "Pct" replace the 5 "WPct," "CPct," etc., and each record contains area fraction for one census tract and material, so it takes 5 records to reflect the material breakout for one census tract. Though this creates a longer table, it is more flexible, and can be expanded to reflect more materials or shortened to contain only records with nonzero values of "Pct."

eqWBldgTypeMp: wood structure type and design level distribution. This table provides the distribution of quantities of wood construction by woodframe building type and design level. That is, given for example 100 sq ft of woodframe buildings, this table gives the fraction of the 100 sf that are in each of 2 detailed woodframe structure types (W1 and W2, referring to woodframe buildings of < 5000 sf and ≥ 5000 sf, respectively) and 7 design levels. The design levels are PC, LC, LS, MC, MS, HC, and HS, where PC means pre-code, LC, MC, and HC mean low- moderate- and high-code, and S means superior, i.e., stronger construction for the given code era.

This table is laid out with separate columns for W1 and W2 percentages, which will be inconvenient for later use, so a duplicate table was created, named eqXBldgTypeMp. The new table replaces the W1 and W2 fields with three fields, named Material, SsType, and Pct, where Material = "W" (for wood), SsType = "W1" or "W2," and Pct = the value previously contained in the W1 or W2 fields. Thus, where eqWBldgTypeMp has a single record for every combination of eqWBldgTypeMpId, eqSpcBldgSchemeId, and DesignLevel, the new table has two: one for each structure type. The new table is longer than the original, but more flexible for present purposes.

eqCBldgTypeMp: concrete structure type and design level distribution. This table is like eqWBldgTypeMp but for 13 concrete building types in each of 7 design levels. Its contents can be appended to the new table eqXBldgTypeMp with entries whose material code = "C."

eqHBldgTypeMp: mobile home design level distribution. Like eqWBldgTypeMp, but for the 7 design levels of mobile home.

eqMBldgTypeMp: masonry structure type and design level distribution. Like eqWBldgTypeMp, but for 7 design levels of 7 masonry types.

eqSBldgTypeMp: steel structure type and design level distribution. Like eqWBldgTypeMp but for 13 steel building types in each of 7 design levels.

eqSpcBldgSchemes: labels for structure type mappings. This table provides a name and description for each eqSpcBldgSchemeId. For example, for eqSpcBldgScheme = "CA166," one table entry provides the SchemeName = "CA2RES1W" and Description = "CA2RES1W," along with the information that it is for construction material BldgTypeSource = "W." Note that SchemeName for California is always of the form CA1*OOOOX*, CA2*OOOOX*, and CA3*OOOOX*, where *X* refers to construction material (W, C, S, H, or M), and *OOOO* refers to a detailed occupancy class, e.g., "RES1." To unpack the SchemeName and make it more useful later, a duplicate table named *eqSpcBldgSchemes2* was created with the same fields, except renaming the field BldgTypeSource to a new field named Material, and adding the fields OccLabel (to contain just the occupancy label, e.g., "RES1") and BldgSchemesId to contain the 1st 3 characters ("CA1," "CA2," or "CA3"), as in table hzBldgScheme.

eqrlnSchemes: structure type and design level distribution by material. This table links the GenBldgSchemeId to a mapping scheme that will distribute quantity by building material and occupancy among the detailed structure types of HAZUS-MH. Each record contains a combination of GenBldgSchemeId and eqSpcBldgSchemesId, the latter of which will be used later to distribute building quantity from construction material to structure type. For each GenBldgScheme there are 5 records, one for each material. For example, there are 5 records with GenBldgSchemeId = "CA34," one each with eqSpcBldgSchemesId = "CA166," "CA199," "CA232," "CA265," and "CA298." For example, as will be shown shortly, eqSpcBldgSchemesId = "CA166" appears in a separate table that distributes quantity of wood construction to each of 14 combinations of structure type (W1 and W2) and design level (pre-code, low code, low-code superior construction, moderate code, moderate-code superior, high code, and high-code superior).

CREATING A CENSUS-TRACT INVENTORY FOR EARTHQUAKE RISK

We now have all the data and tables required to create and populate a census-tract inventory for earthquake risk modeling. In particular, we can now create a table containing square footage, building value, content value, daytime occupants, nighttime occupants, and commute-time occupants by census tract, occupancy class, structure type, and design level. Let us denote the quantities of square footage, building value, content value, and indoor occupants at 2 PM, 2 AM and 5 PM by A, Vb, Vc, PopDay, PopNight, and PopCommute, respectively. We begin by distributing square footage, building value, content value, and indoor occupants by census tract, occupancy class, and material, as follows:

$$A_{tract,occ,matl} = A_{tract,occ} \cdot Pct_{matl|occ} \qquad (1)$$

$$Vb_{tract,occ,matl} = Vb_{tract,occ} \cdot Pct_{matl|occ} \qquad (2)$$

$$Vc_{tract,occ,matl} = Vc_{tract,occ} \cdot \frac{Pct_{matl|occ}}{100} \qquad (3)$$

$$PopDay_{tract,occ,matl} = \sum_i F_i \cdot N_{tract,i} \cdot Q_{occ|group} \cdot \frac{Pct_{matl|occ}}{100} \qquad (4)a$$

$$PopNight_{tract,occ,matl} = \sum_j F_j \cdot N_{tract,j} \cdot Q_{occ|group} \cdot \frac{Pct_{matl|occ}}{100}$$

Error! Reference source not found.b

$$PopCommute_{tract,occ,matl} = \sum_k F_k \cdot N_{tract,k} \cdot Q_{occ|group} \cdot \frac{Pct_{matl|occ}}{100}$$

Error! Reference source not found.c

In these equations, $Pct_{matl|occ}$ is used to distribute square footage, value, and people equally. There does not appear to be any documentation to indicate whether the Pct values in the HAZUS-MH tables refer to square footage, building value, content value, etc., but HAZUS-MH does not appear to have any data to distinguish the percentage of one quantity, such as square footage, from that of another, such as building value. The same percentages are used regardless of which quantity is being distributed. This may be a somewhat crude assumption: fraction of area may be very different from fraction of value in highly disparate occupancies. However, it seems sufficient considering other sources of error in risk analyses.

In Equations **Error! Reference source not found.**a-c, the summations are over the demographic groups that contribute to each occupancy type according to **Error! Reference source not found.**. F denotes the fraction of the occupancy group's square footage in the given the occupancy type, N is the number of people in that demographic group in that tract. For example, referring to **Error! Reference source not found.**, indoor occupants of COM buildings at 2 PM comprise 97% of commercial workers, 16% of daytime residents, 80% of hotel occupants, and 80% of visitors in the census tract, so in Equation **Error! Reference source not found.**a, the summation is over these four demographic groups. The F values are 0.97, 0.16, 0.80, and 0.80, respectively, and the N values are the number of people in those demographic groups in a given tract. Q denotes the fraction of the occupancy

group's square footage represented by the occupancy class (the field "Frac" mentioned previously during the discussion of the new table hzSqFootageOccupT2). For example, in a particular tract, RES1 might comprise 96% of the square footage of all residential occupancies (except for hotel, which is counted separately).

Once the area, values, and indoor occupants have been distributed to the level of tract, occupancy class, and material, they are then further distributed to the level of structure type and design level as follows:

$$A_{tract,occ,type,design} = A_{tract,occ,matl} \cdot Pct_{type,design|matl} \tag{5}$$

$$Vb_{tract,occ,type,design} = Vb_{tract,occ,matl} \cdot Pct_{type,design|matl} \tag{6}$$

$$Vc_{tract,occ,type,design} = Vc_{tract,occ,matl} \cdot Pct_{type,design|matl} \tag{7}$$

$$PopDay_{tract,occ,type,design} = PopDay_{tract,occ,matl} \cdot \frac{Pct_{type,design|matl}}{100} \tag{8}a$$

$$PopNight_{tract,occ,type,design} = PopNight_{tract,occ,matl} \cdot \frac{Pct_{type,design|matl}}{100}$$

Error! Reference source not found.b

$$PopCommute_{tract,occ,type,design} = PopCommute_{tract,occ,matl} \cdot \frac{Pct_{type,design|matl}}{100}$$

Error! Reference source not found.c

where $Pct_{type,design|matl}$ denotes the fraction of square footage in the given material represented by the given particular structure type and design level.

IMPLEMENTATION IN SQL CODE

These calculations have been encoded in queries and tables in a Microsoft Access database and placed for public use at www.risk-agora.org. The relevant tables are listed and described in **Error! Reference source not found.**; the queries in **Error! Reference source not found.**. Results of Equation **Error! Reference source not found.** are stored in an intermediate table named raInventory1. (The "ra" prefix is intended to distinguish new helper tables from the native HAZUS-MH tables; it refers to risk-agora.org.) Equation **Error! Reference source not found.** results are placed in a table called raInventory2; Equation **Error! Reference source not found.** in raInventory3, and Equations **Error! Reference source not found.**a-c in raInventory4. Results

of Equations **Error! Reference source not found.**, **Error! Reference source not found.**, and **Error! Reference source not found.** are stored in intermediate tables called raInventory5, raInventory6, and raInventory7. Results of Equations **Error! Reference source not found.**a-c are stored in an intermediate table named raInventory8. These last 4 tables are then merged into a single table named raInventory, with the layout shown in **Error! Reference source not found.**. County and state-level aggregate tables are also created, named raInventoryCounty and raInventoryState, respectively, where county and state are indicated by their FIPS codes (the left-hand 5 characters of tract, and the left-hand 2 characters of tract, respectively). To use the database, do as follows:

1. Copy the blank database to a directory on your computer

2. Open the blank database and start the linked table manager (in Access 2003, run Tools | Database utilities | Linked table manager | Select all | Ok.

3. Place the HAZUS-MH data DVD containing the state of interest in an accessible drive. It is important to do this after step 2. Within the linked table manager, point Access to the HAZUS-MH directory, file and table for the state of interest, for each table labeled "HAZUS" in **Error! Reference source not found.**.

4. Run the macro "raDenormalizeHAZUSInventory," which performs the queries 02 through 13b in sequence. The queries take between 5 and 30 minutes in a 2007 Dell Latitude D620, depending on the size of the state. For large states like California, the 2 GB file-size limit in MS Access prevents the macro from completing and the database must be split in two in a process that is too complicated to describe here. For smaller states the macro can be run as-is. It helps to turn off confirmations first.

5. Use or export the tables raInventory, raInventoryCounty, and raInventoryState, as desired. Partly to testing this algorithm, the author has exported the inventories of a number of states to comma-and-quote delimited text files named, e.g., raInventoryTractCA.txt, raInventoryCountyCA.txt, raInventoryStateCA.txt, etc., and placed the resulting data files online at www.risk-agora.org for free download.

Table 6. Tables used in the calculations

Table	Source[1]	Contents that are relevant here
eqcbldgtypemp	msh.mdb	distributes concrete to structure type and design level
eqhbldgtypemp	msh.mdb	ditto, mobile homes
eqmbldgtypemp	msh.mdb	ditto, masonry
eqsbldgtypemp	msh.mdb	ditto, steel
eqwbldgtypemp	msh.mdb	ditto, wood
eqxbldgtypemp	new	reorganizes and combines eqcbldgtypemp, eqhbldgtypemp, eqmbldgtypemp, eqsbldgtypemp, and eqwbldgtypemp for convenience here
eqspcbldgschemes	msh.mdb	schemes to distribute material to structure type and design level
eqspcbldgschemes2	new	ditto, reorganized for convenience here
eqrlnschemes	msh.mdb	relates general building schemes (distributing area among materials) to earthquake-specific schemes that distribute area from construction material to earthquake structure type and design level. [not used here]
hzbldgcountoccupt	bndrygbs.mdb	number of buildings by occupancy type and tract [not used here, may be later.]
hzcounty	bndrygbs.mdb	county names and fips codes [not used here.]
hzdemographicst	bndrygbs.mdb	number of people by tract and demographic group
hzexposurecontentoccupt	bndrygbs.mdb	content replacement cost (2003 $1000s) by tract and occupancy
hzexposurecontentoccupt2	new	reorganizes hzexposurecontentoccupt for convenience here
hzexposureoccupt	bndrygbs.mdb	building replacement cost (2003 $1000s) by tract and occupancy
hzexposureoccupt2	new	reorganizes hzexposureoccupt for convenience here
hzgenbldgscheme	msh.mdb	distributes occupancy among material (pctmatl|occ)
hzgenbldgscheme2	new	reorganizes hzgenbldgscheme for convenience here
hzgenbldgschemes	msh.mdb	lists schemes for distributing occupancy among materials
hzsqfootageoccupt	bndrygbs.mdb	building square footage by tract and occupancy
hzsqfootageoccupt2	new	reorganizes hzsqfootageoccupt for convenience here
hzsqfootageoccupt3	new	sum of square footage by tract and occupancy group
hztract	bndrygbs.mdb	identifies which building scheme to use in each tract for distributing from occupancy among materials
rainventory	new	final tract-level inventory: square footage, building value, content value, and indoor occupants by structure type, design level, and occupancy
rainventory1	new	square footage by tract, occupancy, and material. from equation

Table	Source[1]	Contents that are relevant here
		Error! Reference source not found.
rainventory2	new	building value by tract, occupancy, and material. from equation **Error! Reference source not found.**
rainventory3	new	content value by tract, occupancy, and material. from equation **Error! Reference source not found.**
rainventory4	new	indoor occupants by tract, occupancy, and material. from equations **Error! Reference source not found.**a-c
rainventory5	new	square footage by tract, occup., structure type, and design level. from equation **Error! Reference source not found.**
rainventory6	new	building value by tract, occup., structure type, and design level. from equation **Error! Reference source not found.**
rainventory7	new	content value by tract, occup., structure type, and design level. from equation **Error! Reference source not found.**
rainventory8	new	indoor occs by tract, occup., structure type, and design level. from equations **Error! Reference source not found.**a-c
rainventorycounty	new	final county-level inventory: square footage, building value, content value, and indoor occupants by structure type, design level, and occupancy
rainventorystate	new	final state-level inventory: square footage, building value, content value, and indoor occupants by structure type, design level, and occupancy
ramaterial	new	helper table: lists materials
raoccupancy	new	helper table: lists occupancies and their occupancy group
raoccupancy2	new	helper table: lists occupancy groups
raoccupancy3	new	indoor occupants by tract and occupancy group
rasstype	new	helper table: lists structure types and their material
ratimes	new	helper table: lists times of day

1. MSH.mdb and bndrygbs.mdb are Microsoft Access database files distributed on HAZUS-MH data DVDs.

Table 7. Summary of SQL queries to carry out the foregoing calculations

Query	Purpose
01	Not used; reserved for future use
02	Populates the table hzGenBldgScheme2, the reorganized hzGenBldgScheme
02a	Updates hzGenBldgScheme2.Pct for wood material
02b	Ditto, concrete
02c	Ditto, steel
02d	Ditto, masonry
02e	Ditto, mobile home
03a	Populates the table hzExposureOccupT2, the reorganized hzExposureOccupT
03b	Update Vb for RES1-RES3F occupancies
03c	Ditto, RES4-RES6
03d	Ditto, COM1-COM5
03e	Ditto, COM6-COM10
03f	Ditto, IND1-IND5
03g	Ditto, IND6-GOV2
03h	Ditto, EDU1-EDU2
04a	Populates the table hzSqFootageOccupT2, the reorganized hzSqFootageOccupT
04b	Updates A for RES1-RES6 occupancies
04c	Ditto, COM1-COM10
04d	Ditto, IND1-EDU2
04e	Updates the occupancy-group label in hzSqFootageOccupT2 from raOccupancies
05a	Populates hzExposureContentOccupT2, the reorganized hzExposureContentOccupT
05b	Update Vc for RES1-RES6 occupancies
05c	Ditto, COM1-COM10
05d	Ditto, IND1-EDU2
06a	Populates wood-building entries for eqXBldgTypeMp, the reorganized and merged version of eqWBldgTypeMp, eqSBldgTypeMp, eqCBldgTypeMp, eqMBldgTypeMp, and eqHBldgTypeMp
06b	This query updates the Pct value in eqXBldgTypeMp for wood structure types
06c, d	Like 06a and 06b, for steel
06e, f	Ditto, concrete
06g, h	Ditto, masonry
06i, j	Ditto, mobile home
07	Populates the reorganized table eqSpcBldgSchemes2 from eqSpcBldgSchemes
08a	Sum square footage by tract and OccGroup. Places the sum in hzSqFootageOccupT3
08b	Calculates the area fraction Frac in hzSqFootageOccupT2.
08c	Updates null area fractions in hzSqFootageOccupT2 to zero
09a	Populates an intermediate table raOccupancy3 of people by occupancy class
09b	Calculates residential people by tract and time in table raOccupancy3
09c	Calculates commercial people by tract and time in raOccupancy3
10a	Calculates sq ft by tract, occupancy and material in raInventory1
10b	Calculates Vb by tract, occupancy and material in raInventory2
10c	Calculates Vc by tract, occupancy and material in raInventory3
10d	Calculates indoor occupants by tract, occupancy, and material in raInventory4
11a	Calculates sq ft by tract, occupancy, structure type, design level in raInventory5

11b	Calculates Vb by tract, occupancy, structure type, design level in raInventory6
11c	Calculates Vc by tract, occupancy, structure type, design level in raInventory7
11d	Calculates indoor occupants by tract, occupancy, structure type, design level in raInventory8
12a	Populates the final tract-level inventory table raInventory and inserts sq ft
12b	Updates Vb in raInventory
12c	Updates Vc in raInventory
12d	Updates indoor occupants in raInventory
13a	Aggregates all quantities by county FIPS code in table raInventoryCounty
13b	Aggregates all quantities by state FIPS code in table raInventoryState

Table 8. Layout of table raInventory, containing census-tract-level inventory for earthquake risk

Field name	Data type	Description	Comment
ID	Autonumber	An index	1, 2, …
Tract	Text, 11	Census tract number	e.g., 06001400100, from hzTract.Tract
BldgSchemesId	Text, 5	Identifies scheme to distribute from occupancy to material, structure type and design level	In California, distinguishes L/M/H hazard
OccLabel	Text, 5	Occupancy label	RES1, RES2, … EDU2
SsType	Text, 10	Earthquake structure type	W1, W2, … or MH
DesignLevel	Text, 2	Seismic design level	PC, LC, LS, MC, MS, HC, or HS
A	Double	1000 sq ft	See Equation **Error! Reference source not found.**
ValYr	MM/DD/YYYY	Year in which dollar valuation is made	Default = 1/1/2003
Vb	Double	Building replacement cost, $1000s	See Equation **Error! Reference source not found.**
Vc	Double	Content replacement cost, $1000s	See Equation **Error! Reference source not found.**
PopDay	Long integer	Daytime population, people at 2 PM	See Equation **Error! Reference source not found.** a
PopNight	Long integer	Nighttime population, people at 2 AM	See Equation **Error! Reference source not found.** b
PopCommute	Long integer	Commute-time population, people at 5 PM	See Equation **Error! Reference source not found.** c

CONCLUSIONS

It can be desirable for seismic risk studies to possess a single table of the building stock exposed to shaking. Although the HAZUS-MH methodology and data DVDs offer such a database, it is contained in 15 separate (normalized) tables in 2 files and can therefore be difficult to use outside of HAZUS-MH, as is sometimes desirable. To inform such studies, a Microsoft Access database was created here that links to tables in the HAZUS-MH inventory database files bndrygbs.mdb and MSH.mdb, and includes supporting tables, 54 queries, and a macro to run them in sequence. Running the macro produces a single (denormalized) table of the HAZUS-MH building stock: area, building value, content value, and indoor occupants day, night, and commute time, by census tract, building type, design level, and occupancy class. Additional tables are created that sum the inventory by county and state. The database is made available for free download at www.risk-agora.org.

ACKNOWLEDGMENTS

REFERENCES

Algermissen, S.T., W.A. Rinehart, J. Dewey, K.V. Steinbrugge, H.J. Degenkolb, L.S. Cluff, F.E. McClure, R.F. Gordon, S. Scott, and H.J. Lagorio, 1972. *A Study of Earthquake Losses in the San Francisco Bay Area*, Office of Emergency Preparedness, National Oceanic and Atmospheric Administration, Washington, D.C., 220 pp.

(ATC) Applied Technology Council, 1985. *ATC-13, Earthquake Damage Evaluation Data for California,* Redwood City, CA, 492 pp.

Bock, D.B., and J.F. Schrage, 1996. Benefits of denormalized relational database tables. *1996 Proceedings of the Decision Sciences Institute*, Orlando, Florida, November.

Edwards, M.R., D. Robinson, K.J. McAneney, K.J., and J. Schneider, 2004. Vulnerability of residential structures in Australia. *Proc. 13th World Conference on Earthquake Engineering*, Vancouver, B. C., Canada.

Edwards, M.R., 2009. Personal communication.

Erdik, M., N. Aydinglu, Y. Fajhan, K. Sesetyan, M. Demircioglu, B. Syahi, E. Durukal, C. Ozbey, Y. Biro, H. Akman, and O. Yuzugullu, 2003. Earthquake risk assessment for Istanbul metropolitan area. *Earthquake Engineering and Engineering Vibration*, 2 (1), 1-23.

Field, E.H., N. Gupta, V. Gupta, M. Blanpied, P. Maechling, and T.H Jordan, 2005. Hazard calculations for the WGCEP-2002 forecast using OpenSHA and distributed object technologies. *Seismological Research Letters* 76, 161-167.

Goulet, C., C. Haselton, J. Mitrani-Reiser, J. Beck, G. Deierlein, K. Porter, and J. Stewart. Evaluation of the seismic performance of a code-conforming reinforced-concrete frame building - from seismic hazard to collapse safety and economic losses. 2007. *Earthquake Engineering and Structural Dynamics*. 36 (13), 1973-1997.

Jaiswal, K., and D.J. Wald, 2008. Creating a Global Residential Building Inventory for Earthquake Loss Assessment and Risk Management. Open File Report 2008-1160. US Geological Survey, Reston VA.

Kircher, C.A., A.A. Nassar, O. Kustu, and W.T. Holmes, 1997. Development of building damage functions for earthquake loss estimation. *Earthquake Spectra*, **13** (4) 663-682,

(NIBS and FEMA) National Institute of Building Sciences and Federal Emergency Management Agency, 2003. *Multi-hazard Loss Estimation Methodology, Earthquake Model, HAZUS$^{®MH}$ Technical Manual,* Federal Emergency Management Agency, Washington, DC, 690 pp.

Petersen, M.D., A.D. Frankel, S.C. Harmsen, S.S. Mueller, K.M. Haller, R.L. Wheeler, R.L. Wesson, Y. Zeng, O.S. Boyd, D.M. Perkins, N. Luco, E.H. Field, C.J. Wills, and K.S. Rukstales, 2008. *Documentation for the 2008 Update of the United States National Seismic Hazard Maps*. Open File Report 2008-1128, US Geological Survey, Reston VA, 61 pp.

Porter, K.A., J.L. Beck, H.A. Seligson, C.R. Scawthorn, L.T. Tobin, and T. Boyd, 2002. *Improving Loss Estimation for Woodframe Buildings,* Consortium of Universities for Research in Earthquake Engineering, Richmond, CA, 136 pp.

Porter, K.A., 2004. *Evaluating Premium Incentives for the California Earthquake Authority*. California Institute of Technology, Pasadena, CA, 48 pp., http://www.sparisk.com/publications/

Shakhramanian, M.A., V.I. Larionov, G.M. Nigmetov, and S.P. Sutschev, 2000. *Assessment of the Seismic Risk and Forecasting Consequences of Earthquakes While Solving Problems on Population Rescue (Theory and Practice)*. Ministry of Russian Civil Defense and Disaster Management Research Institute, Moscow, Russia, 180 pp.

Shin, S.K., and G.L. Sanders, 2006. Denormalization strategies for data retrieval from data warehouses. Decision Support Systems 42 (1), 267-282.

UN-HABITAT, 2007. Housing in the World-Demographic and Health Survey

APPENDIX 2
UNCERTAINTY IN REPAIR COST SEISMIC VULNERABILITY FUNCTION

Cracking an Open Safe: Uncertainty in HAZUS-Based Seismic Vulnerability Functions

Keith Porter

The "cracking an open safe" methodology has been used to tabulate HAZUS-based seismic vulnerability as functions of structure-independent intensity, while avoiding iteration in the structural analysis. The vulnerability functions give mean damage factor (MDF, defined here as mean repair cost as a fraction of replacement cost) versus 5%-damped elastic spectral acceleration response at 0.3-second and 1.0-second periods, for every combination of occupancy type, model building type, design level, magnitude, distance, site soil classification, etc. Like HAZUS-MH, these prior seismic vulnerability functions give no estimate of uncertainty in damage factor. The coefficient of variation (COV) of damage factor is readily calculated by taking advantage of the fact that that at any level of excitation there is a probability mass function of damage state and an implicit distribution of repair cost conditioned on damage state. COV is calculated here for each combination of occupancy type, model building type, etc., tabulated alongside MDF, and the tables presented for public use at www.risk-agora.org. It is found that a HAZUS-based COV generally decreases with increasing MDF (as has been observed using other analytical vulnerability methods), and the standard deviation of damage factor generally increases with increasing MDF.

INTRODUCTION

In prior work (Porter 2009a, b), the open safe of the HAZUS-MH seismic vulnerability methodology was "cracked" to calculate tables of mean damage factor versus structure-independent 5%-damped spectral acceleration response at 0.3-sec and 1.0-sec periods. The open-safe metaphor refers to the fact that the HAZUS-MH methodology has been thoroughly documented and all its parameter values made available for public use; see especially Kircher and Whitman (1997) and NIBS and FEMA (2003). However, the HAZUS-MH developers do not offer the resulting seismic vulnerability functions in a tabular or graphical form plotted against a structure-independent intensity measure, which can be very inconvenient if one wishes to perform societal risk estimation outside of the HAZUS-MH software. Hence the need to crack

the safe. Furthermore, the HAZUS-MH methodology typically involves iteration to estimate structural response by the capacity spectrum method. The cracking-an-open-safe methodology offers a technique to avoid iteration while still honoring the underlying methodology for structural analysis.

Here, "mean damage factor" refers to the expected value of repair cost as a fraction of replacement cost. In calculating this expected value, the HAZUS-MH methodology propagates (in an approximate way) three sources of uncertainty: structural response given intensity (not in the sense of Modified Mercalli Intensity, but a vector-valued structure-independent spectral acceleration response), damage given structural response, and loss given damage.

Having the mean damage factor available for a particular building in a particular location is clearly very useful and often sufficient, especially when one is dealing with an expected-value problem such as the benefit-cost ratio for a risk-mitigation policy. But what if one wants to know something about the loss exceedance curve (frequency with which various levels of loss are exceeded for a given property), which requires accounting for the uncertainty in loss conditioned on intensity? Estimating the mean damage factor is not enough. The problem addressed here is to estimate the coefficient of variation (denoted here by COV) of damage factor at the same intensity values as the mean damage factor. COV refers to the standard deviation divided by the mean value.

RECAP METHODOLOGY FOR EVALUATING MEAN DAMAGE FACTOR

Before addressing the question of calculating COV versus intensity it is useful to recap the cracking-an-open-safe method. The reader is referred to Porter (2009a, b) for details. Both the HAZUS-MH and cracking-an-open-safe methodologies use the capacity spectrum method to estimate structural response (CSM, Freeman 1998). This pseudostatic nonlinear procedure idealizes a building as a single-degree-of-freedom nonlinear damped harmonic oscillator with a 3-part pushover curve: linear up to a yield point (D_y, A_y), perfectly plastic past an ultimate point (D_u, A_u), and a portion of an ellipse between the two. The effective damping ratio is estimated as the sum of an elastic damping ratio and a fraction of the hysteretic damping, where the fraction, κ, is a function of earthquake magnitude (and indirectly, duration) to reflect pinching of the idealized hysteresis loop. Structural response is parameterized as the spectral displacement (S_d, in inches) and spectral acceleration (S_a, in g's) of the idealized oscillator. It is estimated as the

point—called the performance point—where the pushover curve intersects the idealized demand spectrum with the same effective damping ratio B_{eff}.

The idealized demand spectrum has two parts: a constant-acceleration portion and a constant-velocity portion, both adjusted to account for site soil amplification (F_a in the constant-acceleration region and F_v in the constant-velocity region) and effective damping (using R_A in the constant-acceleration region and R_V in the constant-velocity region). Figure 5 summarizes the methodology. Equations (1) through (7) give the equations for the demand spectrum. In Equation (6), *Area* is the area of one idealized full hysteresis loop whose upper right hand corner is the performance point, and B_E is the elastic damping ratio. See Porter (2009a) for the pushover curve, especially for the ellipse portion between yield and ultimate.
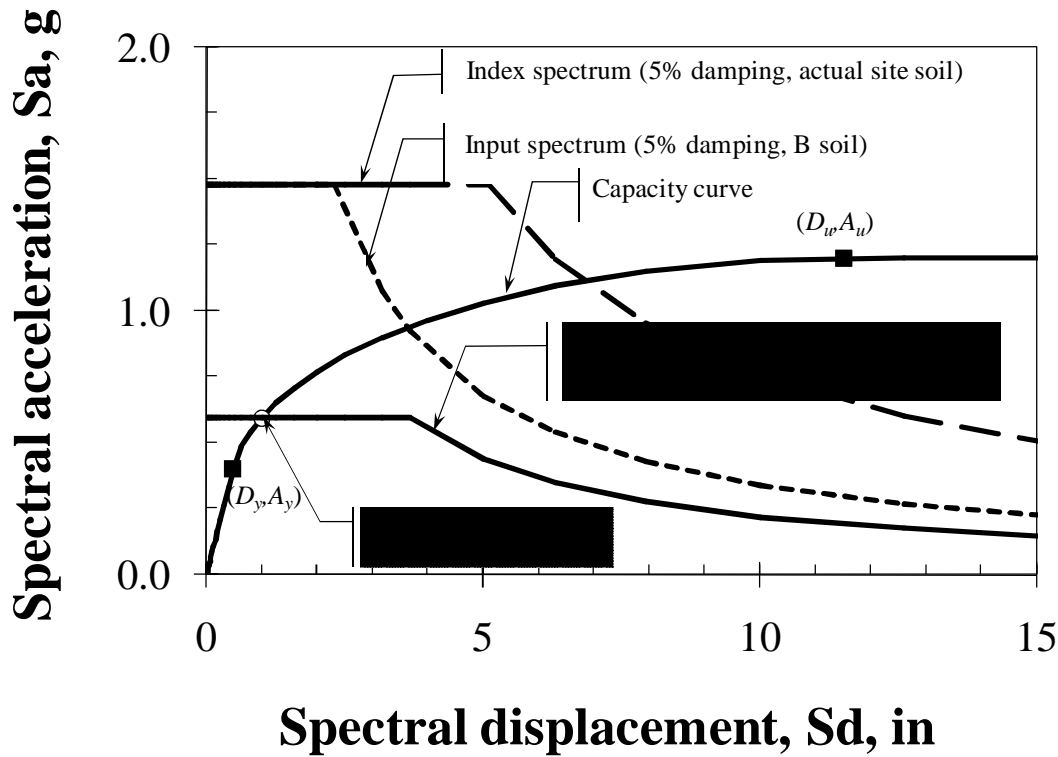

Figure 5. Capacity spectrum method of structural analysis as intended for HAZUS

$$S_a = S_S F_a / R_A \qquad 0 < T \leq T_{AVD} \tag{1}$$

$$S_a = S_1 F_v / (R_V T) \qquad T_{AVD} \leq T \tag{2}$$

$$T = 0.32 \sqrt{S_d / S_a} \tag{3}$$

$$R_A = 2.12 / \left(3.21 - 0.68\ln\left[100 B_{eff}\right]\right) \tag{4}$$

$$R_V = 1.65 / \left(2.31 - 0.41\ln\left[100 B_{eff}\right]\right) \tag{5}$$

$$B_{eff} = B_E + \kappa\left(Area / \left(2\pi S_d S_a\right)\right) \tag{6}$$

$$Area \approx 4 S_a\left(S_d - S_a / \left(A_y / D_y\right)\right) \tag{7}$$

Damage is estimated by inputting structural response (S_d, S_a) into lognormal fragility functions for three generalized buildings components (structural, nonstructural drift-sensitive, and nonstructural acceleration-sensitive) and each of four qualitative damage states (slight, moderate, extensive and complete), as shown in Equation 8 for the uncertain structural damage state D_s. In the equation, Φ is the cumulative standard normal distribution function, θ_i and β_i are parameters of the distribution for damage state i, and d is a particular value of D_s. Nonstructural drift-sensitive damage D_{nd} and nonstructural acceleration-sensitive damage D_{na} are estimated similarly. Repair cost is then estimated as the sum of component damage state probabilities times the mean repair cost given component damage, as shown in Equation (9). In the equation, $L_{s,d}$, $L_{nd,d}$, and $L_{na,d}$ represent the mean damage factor (repair cost as a fraction of total replacement cost V) for the structural, nonstructural drift-sensitive and nonstructural acceleration-sensitive components, respectively, given that they are in damage state d.

$$\begin{aligned}
P\left[D_s = d \middle| S_d = x\right] &= \Phi\left(\frac{\ln\left[x/\theta_d\right]}{\beta_d}\right) - \Phi\left(\frac{\ln\left[x/\theta_{d+1}\right]}{\beta_{d+1}}\right) \quad 1 \le d \le 3 \\
&= \Phi\left(\frac{\ln\left[x/\theta_4\right]}{\beta_4}\right) \qquad\qquad d = 4
\end{aligned} \tag{8}$$

$$E[L] = V\left(\sum_{d=1}^{4} P\left[D_s = d \middle| S_d = x\right] L_{s,d} + \sum_{d=1}^{4} P\left[D_{nd} = d \middle| S_d = x\right] L_{nd,d} + \sum_{d=1}^{4} P\left[D_{na} = d \middle| S_a = y\right] L_{na,d}\right) \tag{9}$$

A main difference between the HAZUS-MH approach and the cracking-an-open-safe approach is that the latter begins by selecting an S_d value for the performance point. One can then work backwards to calculate S_a, T, B_{eff}, R_A, R_V, F_a, F_v, S_S and S_1, which requires no iteration, and forwards to $E[L]$. By repeating the process at many values of S_d, one can create a lookup table relating $S_S F_a$ and $S_1 F_v$, which are the 5%-damped values of spectral acceleration response accounting for site soil, to $E[L]$, for a given combination of magnitude, distance, etc. The HAZUS-MH approach, by contrast, starts with S_S and S_1, calculating forward to the performance point (S_d, S_a), which *does* tend to require iteration, and thence to $E[L]$. The computational

demands of the iteration can be significant for a large portfolio or a probabilistic risk assessment. Having a vulnerability function lookup table greatly reduces the computational demands. Informal experiments by the USGS's PAGER project team (PAGER stands for Prompt Assessment of Global Earthquakes for Response; Porter et al. 2008) show that a scenario loss estimate for a large portfolio—thousands or tens of thousands of properties—can take seconds rather than hours typically required by HAZUS-MH.

METHOD FOR CALCULATING COV

With this background in mind, the calculation of COV is straightforward. At the end of Equation (8), we have the probabilistic damage state of the building, accounting for uncertainties in structural response given intensity and damage given structural response. We will need to account for uncertainty in loss given damage state, which is nowhere stated in the HAZUS-MH documentation. NIBS and FEMA (2003) give only the mean loss given damage state, denoted in Equation (9) by $L_{s,d}$, $L_{ns,d}$, and $L_{na,d}$.

It is assumed here that the loss conditioned on damage state is uniformly distributed between lower and upper bounds for each component and damage state. This is an important assumption, but it seems justified: we do not know the true distribution, and can only estimate its bounds. Information theory says that the uniform distribution is the maximum-entropy (minimum-information) distribution under those conditions. That is, to assume any other distribution implies more knowledge—less uncertainty—than seems available, and would produce a lower COV. Now, proceeding under this assumption, the second moment of loss given the performance point (S_d, S_a) is given by Equation (10):

$$
\begin{aligned}
E\left[L^2\right] = &\sum_{d=1}^{4} P\left[D_s = d \middle| S_d = x\right] \cdot \left(\frac{\left(b_{s,d} - a_{s,d}\right)^2}{12} + L_{s,d}^2\right) \\
&+ \sum_{d=1}^{4} P\left[D_{nd} = d \middle| S_d = x\right] \cdot \left(\frac{\left(b_{nd,d} - a_{nd,d}\right)^2}{12} + L_{nd,d}^2\right) \\
&+ \sum_{d=1}^{4} P\left[D_{na} = d \middle| S_a = y\right] \cdot \left(\frac{\left(b_{na,d} - a_{na,d}\right)^2}{12} + L_{na,d}^2\right)
\end{aligned}
\tag{10}
$$

In Equation (10), $b_{s,d}$ and $a_{s,d}$ refer to the upper and lower bounds of damage factor for the structural component given that it is in damage state d. Similarly the terms $b_{nd,d}$, $a_{nd,d}$ refer to the same bounds for the nonstructural drift-sensitive component, and $b_{na,d}$, $a_{na,d}$ refer to the bounds

for the nonstructural acceleration-sensitive component. In each sum, the first term gives the probability of the component being in a given damage state, the second gives the second moment of loss about its central value, and the third term gives the second moment of the central value about zero.

Finally, one can calculate the coefficient of variation of loss by recalling that variance is given by Equation (11) and COV by Equation (12):

$$Var[L] = E[L^2] - E^2[L] \tag{11}$$

$$COV[L] = \frac{\sqrt{Var[L]}}{E[L]} = \frac{\sqrt{E[L^2] - E^2[L]}}{E[L]} = \sqrt{\frac{E[L^2]}{E^2[L]} - 1} \tag{12}$$

Now the only missing pieces of information are the a and b values of Equation (10). First consider the values of $L_{s,d}$, $L_{nd,d}$, and $L_{na,d}$ proposed in NIBS and FEMA (2003). In general, Tables 15.2, 15.3 and 15.4 of the HAZUS-MH earthquake technical manual assume that repair of slight structural damage costs approximately 0.5% of building replacement cost, moderate generally costs 3%, extensive is approximately 10%, and complete is approximately 20%. For nonstructural drift-sensitive, slight, moderate, extensive and complete are roughly 0.5%, 4%, 11%, and 40%. For nonstructural acceleration-sensitive, the figures are roughly 1%, 6%, 20%, and 40%. It seems reasonable to infer the associated ranges as shown in Table 9. The table also shows the value of the second term in each summand of Equation (10), i.e.,

$$s^2 = \frac{(b-a)^2}{12} \tag{13}$$

Table 9. Approximate values of a and b in Equation (10), as inferred from NIBS and FEMA (2003) Tables 15.2, 15.3, and 15.4

d	Structural		Nonstructural drift-sensitive		Nonstructural acceleration-sensitive	
	a–b	s	a–b	s	a–b	s
Slight	0 – 0.01	0.0029	0 – 0.01	0.0029	0 – 0.02	0.0058
Moderate	0.01 – 0.05	0.0115	0.01 – 0.07	0.0173	0.02 – 0.10	0.0231
Extensive	0.05 – 0.15	0.0289	0.07 – 0.15	0.0231	0.10 – 0.30	0.0577
Complete	0.15 – 0.25	0.0289	0.15 – 0.65	0.1443	0.30 – 0.50	0.0577

Equation (12) was evaluated for every combination of HAZUS-MH's occupancy type, model building type, design level, magnitude (5, 6, 7, or 8), distance (10, 20, 40, or 80 km), seismic

domain (plate boundary or continental interior) and NEHRP site soil category (A, B, C, D or E), appended to the seismic vulnerability function tables detailed in Porter (2009b), and posted for free download at the home page of the Alliance for Global Open Risk Analysis (AGORA 2009; www.risk-agora.org; free registration is required).

The coefficient of variation generally decreases with increasing mean damage factor, as has been observed using other analytical vulnerability methodologies (e.g., Porter et al. 2006), as illustrated in Figure 3. The standard deviation of damage factor (mean times COV) modestly increases with MDF. In the figures, COV has been artificially limited to 0.5 or greater since lower values seem unrealistic, and only samples with MDF ≥ 0.01 are shown. The values of COV seem generally reasonable, at least on an order-of-magnitude basis, given the fairly broad definitions of the HAZUS-MH model buildings types and the various sources of uncertainty involved. The equations in the figures refer to the trendlines shown overlain on the sample data, and the value R^2 refers to the fraction of the total data variance accounted for by the trendline.
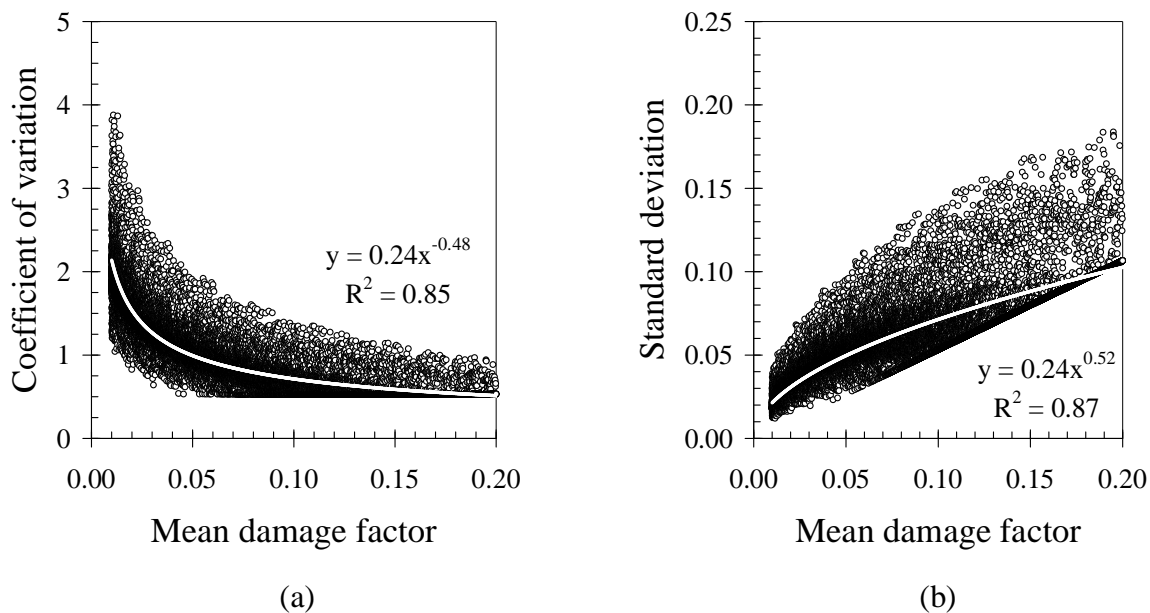

(a)   (b)

Figure 6. Trends in (a) COV versus MDF and (b) standard deviation versus MDF

CONCLUSIONS

In previous work, the "open safe" of the HAZUS-MH earthquake vulnerability methodology was cracked to produce tabular seismic vulnerability functions: mean damage factor versus either of two structure-independent intensity measures: 5%-damped spectral acceleration at 0.3-sec or 1.0-sec periods. The seismic vulnerability functions were tabulated for every combination of