2016

Statistical modeling of earthquake damage

Allison Nicole Waters
University of Northern Iowa

STATISTICAL MODELING OF EARTHQUAKE DAMAGE

A Thesis Submitted

in Partial Fulfillment

of the Requirements for the Designation

University Honors

Allison Nicole Waters

University of Northern Iowa

May 2016

Abstract

The purpose of this study was to build a statistical model of the economic damage that arises from earthquakes in order to better predict losses from future earthquakes. Though earthquakes are essentially a random event and cannot be fully anticipated, analyzing historical data and creating a statistical model can provide researchers with a more accurate estimate of future losses. The data set from which this model was built incorporated earthquakes occurring worldwide from 1915-2015 in which the total damage was recorded. The final model was a multiple linear regression model explaining total damage resulting from an earthquake through four independent variables: whether or not a tsunami occurred (*tsunami_dummy*), whether or not the earthquake occurred in a developed nation (*developed_dummy*), intensity (*intensity*), and number of injuries (*total_injuries*). Statisticians, specifically those at insurance companies, can use these results to provide rough estimates of potential losses after an earthquake occurs. This model is just a starting point for statisticians, however; more accurate and representative models can be created from insurance companies' historical losses in order to better estimate future losses.

This Study by: Allison Waters

Entitled: Statistical Modeling of Earthquake Damage

has been approved as meeting the thesis or project requirement for the Designation University Honors

__________ ____

Date Dr. Syed Kirmani, Honors Thesis Advisor, Mathematics

__________ ____

Date Dr. Jessica Moon, Director, University Honors Program

Statistical Modeling of Earthquake Damage

Catastrophes, though rare events, often hit unexpectedly, destroying buildings, leaving homes in ruins, and tearing communities apart. Many catastrophes, such as earthquakes, are unpredictable, but steps have been taken throughout history in order to prepare for these events. Buildings have been reinforced, weather monitoring has greatly improved, and past catastrophes have been meticulously studied in order to model, or predict, what the next catastrophe will bring. Statisticians create models of catastrophes—earthquakes, hurricanes, even terrorist attacks—in order to simulate the potential effects of such events. But with all the different variables of a catastrophic event, how can statisticians be confident in their models? How will the intensity of an earthquake impact the degree of economic damage in a city? How will the number of fatalities in an earthquake impact the total claim amount an insurance company may face? These kinds of questions are constantly scrutinized by modelers when developing and modifying catastrophe models.

The biggest problem with creating catastrophe models is the lack of credible data. The infrequency of catastrophic events leaves little data for statisticians to analyze. Also, catastrophes dating too far back often result in outdated data which is no longer relevant in the present day. However, does this mean catastrophe models cannot be created? Can a set of useful models to predict economic damage resulting from earthquakes be generated from a single data set? If so, what elements of an earthquake will prove to have the greatest impact on total economic damage? This study examined these questions and explored the field of earthquake catastrophe modeling further.

The purpose of this study was to build a statistical model of the economic damage that arises from earthquakes in order to better predict losses from future earthquakes. Though earthquakes are essentially a random event and cannot be fully anticipated, analyzing historical data and creating a statistical model can provide researchers with a more accurate estimate of future losses. Though many questions arose throughout the data analysis process, there were two initial research questions:

1. What impact do the number of deaths, injuries, or missing people resulting from an earthquake have on economic loss?
2. What impact does the intensity of an earthquake have on economic loss?

These questions shifted to reflect the data used in this study, but the original goal remained intact. This study was significant because it quantified the effect that variables related to earthquakes, such as intensity, number of deceased, and economic condition of the country in which the event occurs, have on the total economic loss from an earthquake.

Literature Review

Catastrophe modeling has been studied in great detail in order to better estimate future events and the potential losses associated with them. However, this is not an easy task; since catastrophes are rare events, there is limited data from which to build models of losses (Cristina & Alexandria, 2013). Catastrophe models are designed to estimate the potential frequency and severity of a catastrophic event, not to predict when an event of a particular severity will occur. This is especially the case in regards to earthquakes. Earthquakes are one of the most unpredictable natural phenomena because there are few warning signs of a potential seismic event, unlike those related to hurricanes or floods (Vere-Jones, 1995). Another caveat of catastrophe modeling is the constantly-changing landscape of insured properties, or exposures. Property values may fluctuate along with building structures and designs (Grace, Klein, Kleindorfer, & Murray, 2003). Advanced technology and a thorough understanding of the geophysics behind seismic events have allowed engineers to analyze the movement of a building in the event of an earthquake. These movements are then accounted for in the design of buildings in areas with high seismic-risk (Bolt, 1993). Because of the ever-changing nature of the insured landscape, data from the past may no longer be relevant.

In order to start forming these catastrophe models, researchers must fully understand the variables of the catastrophic events. Earthquake loss models often include measures of the magnitude and intensity (Cristina & Alexandria, 2013). Magnitude, historically measured using the Richter scale, is defined as "the logarithm to base ten of the maximum seismic-wave amplitude (in thousandths of a millimeter) recorded on a standard seismograph at a distance of 100 kilometers from the earthquake epicenter" (Bolt, 1993, p. 118). Various other measures of magnitude arose from Richter's original scale, most notably the moment magnitude scale (Bolt, 1993). A moment is the product of the size of a force and the distance between that force and the force opposite to it. In terms of earthquakes, the moment is the measure of a rupturing fault and the rebounding effect along that fault (Bolt, 1993). Because these moment values are often hard concepts to grasp mathematically, they are correlated with magnitude and measured on the moment magnitude scale. This scale is often used as a superior measurement because of its consistency across all sizes of earthquakes, unlike the original Richter scale.

Intensity is another measure of earthquake severity which is commonly measured by assessing the degree of damage a seismic event causes. This includes damage to structures, ground disturbances, and animal reactions to the earthquake (Bolt, 1993). The Modified Mercalli Intensity scale (MMI) is commonly used to measure earthquake intensity. This is a Roman numeral scale that provides a description of the degree to which an earthquake is felt or the damage caused by an earthquake. For example, an earthquake with an MMI of IX indicates an earthquake with violent shaking and resulting in "damage considerable in specially designed structures; well-designed frame structures thrown out of plumb. Damage great in substantial buildings, with partial collapse. Buildings shifted off foundations" (United States Geological Survey, 2015). Classification on the MMI is often done through questionnaires distributed to residents of the affected region, which provides a more qualitative measurement of earthquake severity as opposed to magnitude measurements, which are quantitative in nature.

Once researchers have a full grasp of the measurements involved in seismic analysis, they can begin to build models to estimate damages or loss from these catastrophic events. This damage arises from homes and buildings destroyed in an earthquake, lives lost, injuries sustained, and other physical effects resulting from the event. Insurance companies are often those most interested in these catastrophic models in order to estimate the potential claims they may need to pay out in the event of a catastrophe. These models are often developed by statistical modeling teams using historical losses within the company as well as industry data. These models are usually proprietary in order to protect any classified information from competitors or the public. Though the specific models are not published, researchers have discussed what types of statistical analyses are used to create these models.

AIR Worldwide, a catastrophe modeling firm, developed catastrophe modeling technology in the late 1980s that broke catastrophe modeling into three components: hazard, engineering, and financial (Grace et al., 2003). First, the hazard component simulates a catastrophic event and its intensity. Next, the engineering component estimates the amount of damage resulting from that simulated event. Lastly, the financial component assesses the economic value of the damage from the event. Figure 1 provides a flow chart of the AIR catastrophe modeling process. The output of this model includes probability distributions of losses over a certain time period or in terms of cumulative distribution functions (Grace et al., 2003).

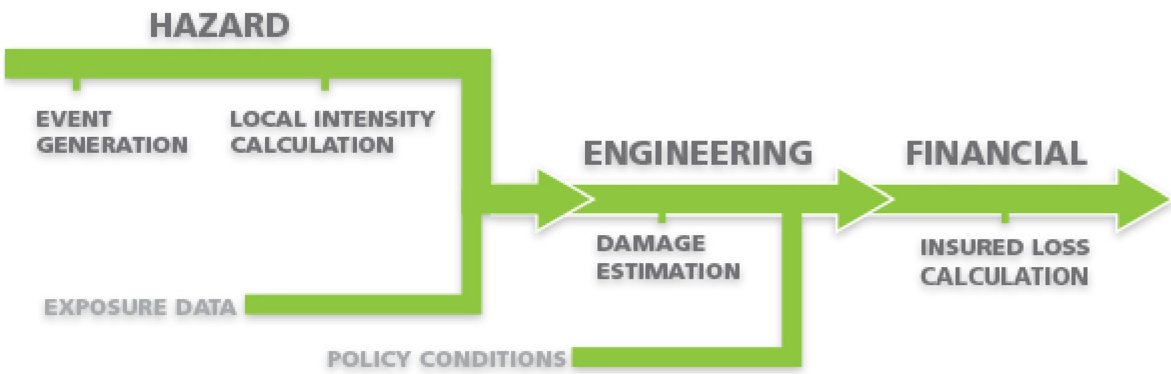

Figure 1. Catastrophe modeling framework. This figure provides a flow chart of the various steps AIR takes when creating a catastrophe model. Adapted from "About Catastrophe Modeling," by AIR Worldwide, n.d. (http://www.air-worldwide.com/Models/About-Catastrophe-Modeling/)

Statisticians also use a combination of two models to simulate a catastrophic event: one to estimate frequency and one to estimate severity. Frequency is a measure of when and how often an event occurs. Yilmaz, Erisoglu, and Celik (2004) suggested using Weibull distribution to model the time between two successive earthquakes. Weibull distribution is often used to measure "time-to-failure" or in the context of this study, the time until the next seismic event. Poisson models have also been used to model earthquake frequency (Weimer, n.d.). However, the Poisson distribution is more commonly used to model the number of events occurring in a specific area during a specific time period. Severity, on the other hand, is a measure of the potential damage or losses arising from a catastrophic event. One of the most common methods of modeling the severity of an earthquake is using a tapered Pareto distribution called modified Gutenberg-Richter law (Kagan & Schoenberg, 2001). This law models the relationship between earthquake magnitude and number of seismic events of at least that magnitude in a specified area during a given time period (Kagan & Schoenberg, 2001). The tapered Pareto distribution is designed to assign lower probabilities to earthquakes of extremely high magnitudes than the typical Pareto distribution. All in all, these models have their advantages and disadvantages, and they are often used in conjunction with each other for catastrophe modeling purposes.

Methodology

Data Selection

The first step in this study was to find a data set from which the model would be built. The National Oceanic and Atmospheric Administration (NOAA) has a group called the National Centers for Environmental Information which maintains a Global Significant Earthquake Database dating back to 2150 BCE (National Geophysical Data Center, 2016). The database allows a user to specify a date range, region, and country from which they want the records of the significant earthquakes meeting those specifications. All the earthquakes occurring worldwide between 1915 and 2015 were examined. Earthquakes without the damage in millions of dollars recorded were excluded from the data set. The results of that database search were exported to Excel for further modifications to better fit the design of this study.

The database provided estimates for damage in millions of dollars, and it also included an estimate for total damage in millions of dollars, which included damage from subsequent events such as a tsunami or another earthquake. Since the total economic effect of a seismic event was of interest in this study, only the total damage in millions of dollars was examined. After the earthquakes without total damage recorded were deleted, 360 entries remained. These earthquakes made up the final data set that was used in the modeling process.

Variable Selection

Next, the different variables were examined to determine whether or not they would be included as potential regressors. Some variables were excluded because they were irrelevant (e.g. latitude and longitude, region code, and state name), and others were not included because there were too many missing values to provide substantial analysis (e.g. houses destroyed and damaged, missing people, and various measures of magnitude). Other variables were transformed into more useful regressors, specifically economic development status and total damage scaled.

The original data set provided the country where the event occurred, but this was not relevant to the questions around which this study was focused. Instead, the country variable was transformed into a measure of economic development status, i.e. developed, economies in transition, or developing. This was done using a country classification table produced by the Development Policy and Analysis Division of the United Nations (2014). Thus, each of the earthquakes was classified as occurring in a location that was either developed, an economy in transition, or developing depending on the classification provided by the United Nations. This variable provided a way to compare the effect on total economic damage of various economic states, therefore enriching the analysis.

Another variable that required transformation was the total damage in millions of dollars. These dollar amounts were recorded in the value at the time of the event, so they needed to be scaled to current dollars, or 2015 dollars, for consistency. This was done using the Consumer Price Index (CPI) provided by the U.S. Department of Labor Bureau of Labor Statistics (US Inflation Calculator, n.d.). The average CPI of 2015 was divided by the CPI of years dating back to 1915 in order to produce a scaling factor for the non-current dollars. Then, the nominal dollar amounts for each of seismic events were multiplied by the corresponding scaling factor; the resulting dollar amount was coined as the variable total damage scaled.

The final data set included eight potential regressors: tsunami, economic development status, focal depth, eq primary, intensity, total deaths, total injuries, and total damage scaled.

Tsunami. The *tsunami_dummy* variable is a dummy variable that takes the value 1 if an earthquake resulted in a tsunami and 0 otherwise.

Economic development status. This variable, as mentioned before, is a transformation of the country in which a seismic event occurred. The variable is actually treated as two separate dummy variables: *developed_dummy* and *transition_dummy*. The first takes the value 1 if the country is developed and 0 otherwise, while the second takes the value 1 if the country is an economy in transition and 0 otherwise. There is no need for a third dummy variable for developing countries because this is accounted for if both *developed_dummy* and *transition_dummy* take the value 0.

Focal depth. The variable *focal_depth* is a measure of the depth of an earthquake, and it is given in kilometers.

Eq primary. The variable *eq_primary* was the most consistently recorded measure of magnitude in the data set. Magnitude is a measure of seismic energy of an earthquake taking the value 0 to 10; the higher the number, the more seismic energy an earthquake produced. The type of magnitude measure used was not noted, so this value could be the surface-wave magnitude, moment magnitude, compressional body wave magnitude, or another measure. However, these values are all measures of the magnitude and are similar, so they can be compared.

Intensity. The variable *intensity* is measured on the Modified Mercalli Intensity scale. Typically, the scale is given in Roman Numerals, but these were converted into values from 1 to 12 in the database. The scale is defined based on physical effects of an earthquake, such as damage to the frames of buildings, with 1 being a minor earthquake and 12 being the most damaging.

Total deaths. The variable *total_deaths* is the total number of deaths resulting from the earthquake and any secondary effects that may have occurred, such as a tsunami.

Total injuries. The variable *total_injuries* is the total number of injuries resulting from the earthquake and any secondary effects that may have occurred, such as a tsunami.

Total damage scaled. The variable *total_damage_scaled*, as mentioned before, is a transformation of the total damage variable originally included in the data set. This value is recorded in millions of dollars, scaled to 2015 dollars.

Results

Exploratory Analysis

The first step in model creation is to perform exploratory data analysis, which includes examining summary statistics and possible relationships between the regressors. SAS was used as the computer program for this study. Summary statistics of the eight independent variables and the sole dependent variable were calculated and can be seen in Figure 2. There is, on average, $3.251 billion dollars in damage for each earthquake, with a standard deviation of $16.661 billion. The high value for standard deviation for *total_damage_scaled*, *total_deaths*, and *total_injuries* can be explained by the great variability in earthquake effects. Some earthquakes can barely be felt by humans, while others completely decimate towns. For the three dummy variables, the most meaningful statistic is the sum, which would be equivalent to the number of values equal to 1. Of the 360 earthquakes, 117 were in developed nations, 21 were in economies in transition, and the remaining 222 were in developing nations. Another important note is that 94 of the 360 earthquakes also had a tsunami associated with them.

Simple Statistics						
Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
TOTAL_DAMAGE_SCALED	360	3251	16611	1170387	0.11540	231903
DEVELOPED_DUMMY	360	0.32500	0.46903	117.00000	0	1.00000
TRANSITION_DUMMY	360	0.05833	0.23470	21.00000	0	1.00000
TSUNAMI_DUMMY	360	0.26111	0.43985	94.00000	0	1.00000
EQ_PRIMARY	359	6.56435	0.98259	2357	2.10000	9.50000
FOCAL_DEPTH	352	27.40909	25.57426	9648	0	215.00000
INTENSITY	219	8.10959	1.49214	1776	3.00000	12.00000
TOTAL_DEATHS	292	6935	32188	2025083	1.00000	316000
TOTAL_INJURIES	268	9894	58362	2651516	1.00000	799000

Figure 2. Simple statistics of full data set.

In addition to the simple statistics, the relationship between potential variables also revealed a lot about the data set. The relationship between regressors is measured by the Pearson correlation coefficient (rho), which measures the strength and direction of a relationship. These values range from -1 to +1, with -1 being a perfectly negative relationship and +1 being a perfectly positive relationship. The correlation coefficient matrix can be seen in Figure 3. The scatterplot matrix in Figure 4 (p. 12) is a visual interpretation of the relationships between variables. The most highly correlated regressors were *total_deaths* and *total_injuries* (rho = 0.79911). This strong positive relationship is intuitively sound; one would expect that the more deaths resulting in an earthquake, the more injuries there would be as well. Another relatively strong relationship existed between *tsunami_dummy* and *eq_primary* (magnitude), with a rho of 0.52983. This relationship can also be justified through logic; the higher the magnitude of an earthquake, the greater likelihood that a tsunami will develop. The variables *intensity* and *eq_primary* also had a moderately strong positive relationship (rho = 0.50181). One would expect intensity to move in the same direction.

	total_damage_scaled	developed_dummy	transition_dummy	tsunami_dummy	eq_primary	focal_depth	intensity	total_deaths	total_injuries
total_damage_scaled	1.0000	0.1195	-0.0105	0.1795	0.1921	-0.0404	0.2571	0.1308	0.2075
developed_dummy		1.0000	-0.1727	0.0331	-0.1042	-0.1569	-0.0416	-0.0798	-0.0914
transition_dummy			1.0000	-0.0940	-0.1047	-0.0646	-0.0336	0.0250	-0.0351
tsunami_dummy				1.0000	0.5298	0.0337	0.2898	0.1676	0.0647
eq_primary					1.0000	0.1394	0.5018	0.2330	0.1515
focal_depth						1.0000	-0.1960	-0.0385	-0.0399
intensity							1.0000	0.3149	0.2175
total_deaths								1.0000	0.7991
total_injuries									1.0000

Figure 3. Pearson correlation coefficient matrix with notable relationships highlighted.

The relationship between the dependent variable and the various regressors were also of importance. Six regressors had positive relationships with the dependent variable, *total_damage_scaled*: *developed_dummy*, *tsunami_dummy*, *eq_primary*, *intensity*, *total_deaths*, and *total_injuries*. On the other hand, *transition_dummy* and *focal_depth* were negatively correlated with the dependent variable. With this knowledge, signs of coefficients could be predicted for the regressors; positive correlations correspond with positive signs, while negative correlations correspond to negative signs.

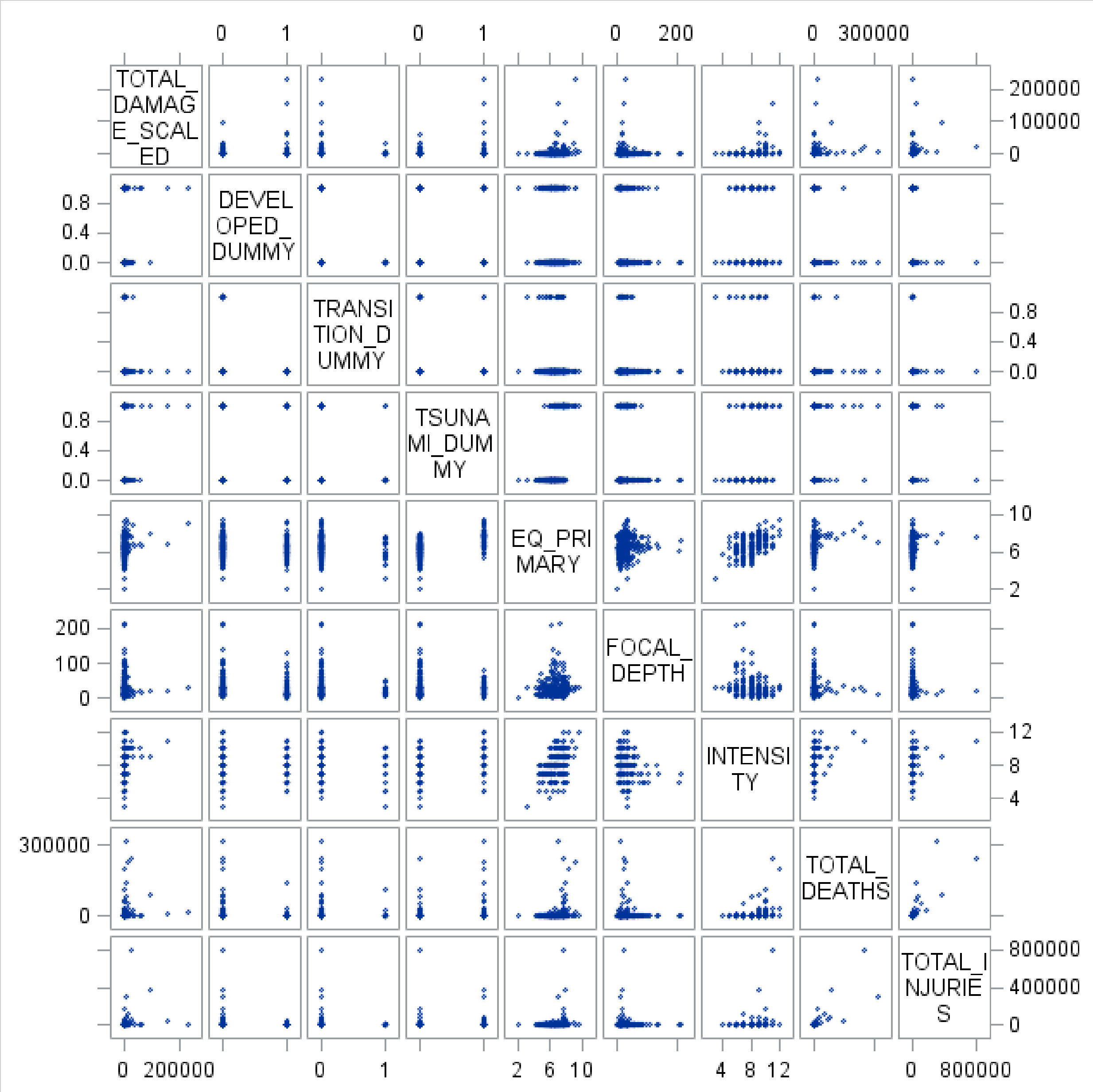

Figure 4. Scatterplot matrix output. This matrix displays the same information from the Pearson Correlation Coefficient matrix but in a visual manner.

Model Selection

The next step in creating this catastrophe model was experimenting with different variable combinations to design the model that best explained the data set. Three different selection processes were utilized in this project: forward, backward, and stepwise selection.

Forward selection begins with no variables in the model. During each step, SAS examines the F-statistic, a measure of significance of the model, for each of the independent variables. SAS then chooses the variable with the largest F-statistic, as long as the variable has a p-value lower than a set level (0.5), and enters it into the model. This process continues until no variables remain or no p-value is lower than the set level, and once a variable is added, it cannot be removed. Backward selection, on the other hand, begins with all of the variables in the model. The F-statistic for each of the variables is calculated, and the variables are deleted one at a time from the model if their p-values are greater than a specified level (0.1), beginning with the least significant. This process is repeated until all variables are significant at that level. Stepwise selection is a variation of the forward selection process. The main difference is that models may be removed from a model if their F-statistic is not significant at a specified level (0.1).

These three selection processes provided two distinct models; backward and stepwise produced the same model, while forward had an additional variable included. Backward and stepwise selection included *developed_dummy*, *tsunami_dumm*y, *intensity*, and *total_injuries*. The only difference with the forward selection was that *total_deaths* was also included as a regressor. However, due to the strong positive correlation between *total_deaths* and *total_injuries*, it seemed unnecessary to use both variables in the same model in fear of multicollinearity, or in other terms, redundancy. Thus, the model chosen to move forward with further modifications included only *total_injuries*.

A multiple regression was conducted using the aforementioned regressors to model the dependent variable, *total_damage_scaled*. All of the variables were significant at a 5% level except for *developed_dummy*, which had a p-value of 0.1033. Typically, variables included in a model are significant at least at the 10% level, so the variable was dropped, and the results were examined. Dropping this dummy variable lowered the r-squared of the model from almost 20% to 18%, so the researcher decided to include the variable even though it was seemingly insignificant. This was done because of the meaningful interpretations that could arise from the inclusion of the economic nature of a location of a seismic event in the modeling of resulting economic damage. The variables *total_injuries* and *total_deaths* were swapped to see what effect it had on the model. The results led to a lower r-squared and lower significance levels of other regressors, so *total_injuries* remained in the model. Thus, the final model included those variables selected in the stepwise and backward processes.

Assessing Influential Points

Once the model was chosen, potentially influential points and outliers were examined to determine whether or not they should be included in the data set. The four measures used to assess influence in this project were the student residual, DFITS, DFBETAS, and Cook's D. The student residual is a measure of the difference in an observation's actual value and its predicted value, scaled by its standard deviation. These values are considered potentially overly influential if they have a value greater than 2. DFITS measures the change in an observation's predicted value if that observation is deleted from the data set; large values indicate influence, especially those greater than 2. DFBETAS is the change in the coefficient of each regressor if an observation is deleted. A value of 2 is also used as a cutoff for this measure, and those exceeding that cutoff may have undue influence on the value of the coefficients. Lastly, Cook's D is a measure of difference in the predicted values of a model before and after the deletion of an observation. In practice, a value greater than 1 is evidence that an observation is overly influential.

In examining these four measures, a few observations stood out as they raised concerns in at least two of the four tests. Observation 83 failed each of the four tests, while observation 241 failed three of the four. Observation 323 was also noteworthy because it failed two of the four tests. In order to test whether or not the observation should be included in the model, Akaike's Information Criteria (AIC) was calculated. AIC is a measure of goodness-of-fit of a model, in which the model with the lowest AIC is considered the best model. The formula for AIC utilizes various measures included in the SAS output: $AIC = n \times \ln\left(\frac{SSE}{n}\right) + 2p$, where n is the number of observations, SSE is the sum of squared errors, and p is the number of parameters in the model (including the intercept). The AIC of the model with all observations included was 3040.51. When observation 83 was deleted, however, the AIC dropped to 2995.37; since the AIC without the observation was lower, that observation was permanently deleted from the data set. The AICs of the models without the other notable points were calculated, but they were either only slightly lower or higher than the model excluding observation 83. Because of this, only observation 83 was removed from the data set due to its extreme influence on the parameter estimates and predicted values.

Examining Potential Multicollinearity

With the data set and model finalized, the next step was to check for potential multicollinearity by examining the variance inflation factors and collinearity diagnostics of the model. Variance inflation factors (VIF) measure how much the variance of a parameter estimate is inflated due to collinearity. Values greater than 2 signify collinearity issues; however, all VIFs for the parameters were less than 2 and did not raise any concerns. Another measure of collinearity can be seen in the collinearity diagnostics produced by SAS (Figure 5). When the condition index of an eigenvalue is greater than 30 and the corresponding proportions of variation are also large, extreme collinearity is suspected. However, as seen in Figure 5, none of the condition indices are greater than 30, so no issues are suspected.

Collinearity Diagnostics							
		Condition	Proportion of Variation				
Number	Eigenvalue	Index	Intercept	DEVELOPED_DUMMY	TSUNAMI_DUMMY	INTENSITY	TOTAL_INJURIES
1	2.96126	1.00000	0.00261	0.03585	0.03810	0.00255	0.01130
2	1.00165	1.71941	0.00015867	0.10588	0.02020	0.00007230	0.70065
3	0.60115	2.21945	0.00019178	0.22391	0.66555	0.00008615	0.22154
4	0.42333	2.64485	0.01164	0.63365	0.25155	0.01021	0.05547
5	0.01261	15.32237	0.98540	0.00070132	0.02460	0.98708	0.01105

Figure 5. Collinearity diagnostics table.

Model Validation

Model validation was approached differently for this model than most linear regression models. Model validation is a process where the predictive power of a model is assessed. Typically, a data set is split into two parts: a training set and a test set. The training set is a subset of the data set used to build the model, while the test set is the remainder of the data set used to test the reliability of the model created with the training set. Fifty-fifty and 75-25 splits are common for the training and testing sets, respectively. However, since the sample size of this study was already relatively small when excluding observations with missing values, breaking the data set into two would limit the results too much.

Because of this, the model was built using the full data set. Thus, the predicted residual error sum of squares (PRESS) and sum of squares due to error (SSE) were compared to assess the predictive ability of this model. SSE measures the fit of a model by examining the difference in the observed and the predicted value for each observation. The SSE for this model was 28,447,762,129. The PRESS statistic is also a measure of fit, but it examines the difference between an observed value and the predicted value of that observation provided by a model created without that observation; thus, the model used to predict is not built with the observation it is predicting. The PRESS of this model was 32,710,880,547. These values are much higher than statisticians normally see, but this is due to the size and scale of the dependent variable, *total_damage_scaled*, which was in millions. If *total_damage_scaled* is expressed in billions instead, the PRESS and SSE are scaled down by 100,000 to 32,711 and 28,448, respectively. Since the large values are simply a result of scaling, they do not pose an issue. The PRESS/SSE ratio for this model was about 1.15, which is close to 1, verifying that this model has predictive value. Another test of model validation is to examine a graph of the predicted values against the observed values. Figure 6 is a plot of the predicted value of damage against the observed *total_damage_scaled*. The line in the figure represents when the predicted and observed values are the same. Since there is no clear pattern of points consistently above or below the line, there is no observable issue with the model's predictive ability.

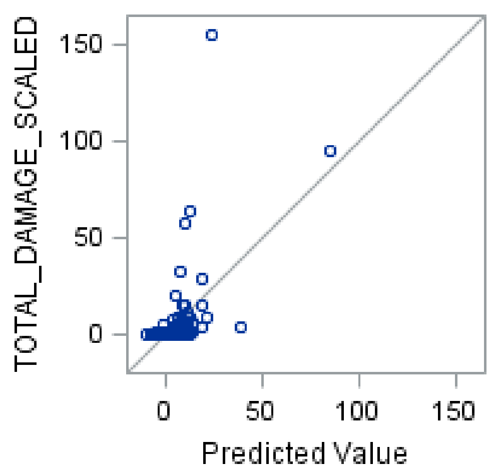

Figure 6. Plot of *total_damage_scaled* (Y) against the predicted value.

Interpretations of Final Model

The final model included *developed_dummy* (X_1), *tsunami_dummy* (X_2), *intensity* (X_3), and *total_injuries* (X_4). The r-squared of the model was 0.3167, which means that 31.67% of variability can be explained by the model. This is relatively high when considering the scientific nature of the data. The AIC was 2995.37, which was lower than previous models. The model itself and all but one of the dependent variables were significant at the 2.5% level, while *tsunami_dummy* was significant at a 10% level. Figure 7 provides the parameter estimates of the model.

Parameter Estimates							
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Tolerance	Variance Inflation
Intercept	1	-19971	6855.93315	-2.91	0.0041	.	0
DEVELOPED_DUMMY	1	5335.97889	2290.46656	2.33	0.0211	0.97542	1.02520
TSUNAMI_DUMMY	1	4203.50991	2418.49009	1.74	0.0842	0.93899	1.06497
INTENSITY	1	2333.30055	845.89335	2.76	0.0065	0.93538	1.06908
TOTAL_INJURIES	1	0.21428	0.03224	6.65	<.0001	0.94273	1.06075

Figure 7. Parameter estimates for final model.

In equation form, the model simply describes *total_damage_scaled* (Y) as a linear combination of the regressors:

$$Y = -19971 + 5335.98X_1 + 4203.51X_2 + 2333.30X_3 + 0.21428X_4$$

Note: coefficients are in millions of 2015 dollars

The dummy variables and continuous variables must be interpreted differently. Since the dummy variables only have two possible values (0 or 1), the additional damage resulting from an earthquake will be equal to the coefficient of that variable, or 0. For example, X_1 represents *developed_dummy*, so if an earthquake occurs in a developed nation, the damage of an earthquake will increase by $5335.98 million. The continuous variables are a bit more intuitive. For every one unit increase in intensity or total injuries, the damage resulting from an earthquake increases by $2333.30 million and $0.21428 million, respectively. In practice, the values of the variables would be inputted into the model and summed, and the resulting value would be the predicted total damage resulting from an earthquake with those specific characteristics.

Discussion

These results are meaningless unless they can be interpreted in the context of seismic activity. There are four characteristics that have a significant impact on the total damage resulting from an earthquake: whether or not that earthquake occurs in a developed nation, whether or not there is a tsunami associated with that earthquake, its intensity, and the total number of injuries from the earthquake. The variable that had the largest positive effect on damage was whether the nation was developed or not, while number of total injuries had the smallest effect. Statisticians, specifically those at insurance companies, can use these results to provide rough estimates of potential losses after an earthquake occurs. They can also run previous earthquake data through this model to obtain predicted losses, and these can be compared to their actual losses to get a gauge for their exposure. This knowledge can be used in pricing models, so insurance companies can minimize their losses by pricing insurance more accurately. This model is just a starting point for statisticians, however; more accurate and representative models can be created from insurance companies' historical losses in order to better estimate future losses.

This study had a few limitations in addition to its sole reliance on public, non-company-specific data. One limitation is that there were many missing data points in the database. Many variables, such as information about damaged houses, had to be excluded from the data set because so few observations had recorded values. This raises the concern of improper and incomplete data collection methods for seismic activity. One would expect values to be missing for some observations but certainly not the majority. This missing data and exclusion of variables could have resulted in leaving an influential variable out of the model.

Conclusion

Earthquakes can occur with little to no warning, and they can be detrimental to society by destroying buildings, killing and injuring citizens, and leaving towns in ruins. In order to combat and prepare for the damage caused by earthquakes, statisticians create catastrophe models to help predict the outcomes of these seismic events. The purpose of this study was to use catastrophe modeling as a basis to identify the key drivers of economic loss in an earthquake. Using multiple linear regression, total damage resulting from an earthquake was explained through four characteristics: whether or not a tsunami occurred, whether or not the earthquake occurred in a developed nation, intensity of the earthquake, and number of injuries resulting from the earthquake. Having these drivers identified provides a starting point for statisticians working in the insurance field. This study can be used as a framework to create a similar model using proprietary information and actual historical losses, further strengthening its usefulness as a loss estimation tool. A model explaining economic loss will give statisticians the information necessary to make informed decisions about expected losses from future earthquakes and provide insight into the source of those losses.

References

AIR Worldwide. (n.d.). Catastrophe modeling framework [Online image]. Retrieved from http://www.air-worldwide.com/Models/About-Catastrophe-Modeling/

Bolt, B. A. (1993). *Earthquakes.* New York, NY: W. H. Freeman and Company.

Cristina, C. & Alexandria, C. R. (2013). Catastrophic events modeling. *Annals of the University of Oradea, Economic Science Services*, 987-996.

Development Policy and Analysis Division. (2014). *World Economic Situation and Prospects country classification*. Retrieved from http://www.un.org/en/development/desa/ policy/wesp/wesp_current/2014wesp_country_classification.pdf

Grace, M. F., Klein, R. W., Kleindorfer, P. R., & Murray, M. R. (2003). *Catastrophe insurance: Consumer demand, markets and regulation.* Norwell, MA: Kluwer Academic Publishers.

Kagan, Y. Y. & Schoenberg, F. (2001). Estimation of the upper cutoff parameter for the tapered pareto distribution. *Journal of Applied Probability*, *38*(A), 901-918.

National Geophysical Data Center. (2016). *Significant earthquake database* [Data set]. doi:10.7289/V5TD9V7K

United States Geological Survey. (2015). The modified mercalli intensity scale. Retrieved from http://earthquake.usgs.gov/learn/topics/mercalli.php

US Inflation Calculator. (n.d.). Consumer price index data from 1913 to 2016. Retrieved from http://www.usinflationcalculator.com/inflation/consumer-price-index-and-annual-percent-changes-from-1913-to-2008/

Vere-Jones, D. (1995). Forecasting earthquakes and earthquake risk. *International Journal of Forecasting*, *11*, 503-538.

Weimer, S. (n.d.). Earthquake statistics and earthquake prediction research. Unpublished manuscript, Institute of Geophysics, Zurich, Switzerland.

Yilmaz, V., Erisoglu, M., & Celik, H. E. (2004). Probabilistic prediction of the next earthquake in the NAFZ, Turkey. *Dogus Universitesi Dergisi*, *5*(2), 243-250.

Appendix

I_D	TSUNAMI (1=yes, 0=no)	COUNTRY	ECONOMIC DEVELOPMENT STATUS	FOCAL_DEPTH	EQ_PRIMARY	INTENSITY (MMI)	TOTAL_DEATHS	TOTAL_INJURIES	TOTAL_DAMAGE_MILLIONS_DOLLARS (SCALED TO 2015)
3025	0	ITALY	Developed	10	7.5	11	29978		1408.0
3056	1	PANAMA	Developing	25	7	9			0.5
3124	1	USA TERRITORY	Developed	60	7.3	9	144		62.8
3165	1	CHINA	Developing	25	8.3	12	200000		296.3
3227	1	JAPAN	Developed	35	7.9		142807	47000	8316.4
9672	0	USA	Developed	25	6.7	8			2.0
3270	0	USA	Developed	10	6.8	9	13		108.4
3305	1	USA	Developed		5.8		1		40.9
3306	1	JAPAN	Developed	10	7.3		3022	3295	544.9
3450	1	NEW ZEALAND	Developed	35	7.7	9	256	400	389.8
3534	1	PANAMA	Developing	60	7.7			4	0.9
3550	1	PAKISTAN	Developing	33	7.5	10	60000		432.5
3561	0	USA	Developed		6.2	8	2		328.7
3563	0	USA	Developed		6	8	2		103.8
3632	1	CHILE	Developing	60	8.3	10	30000		15687.5
6510	1	TURKEY	Developing	35	7.7	12	32700		341.0
3662	0	USA	Developed	16	7.2	10	9		558.7
3699	0	ECUADOR	Developing	30	7.9	9	200		36.4
3784	0	USA	Developed	12	5.6	8			26.9
3811	1	PAKISTAN	Developing	15	8	10	4000		329.2
3869	1	PHILIPPINES	Developing	33	8.3	9	74		34.4
3884	0	JAPAN	Developed	20	7.3		5131	11000	9834.7
3891	0	TURKMENISTAN	Economies in transition	18	7.3	10	110000		245.9
3898	1	USA	Developed		7	8	8		249.0
3923	0	USA	Developed			5			89.6
3950	1	INDIA	Developing	33	8.6	11	1530		196.7
4002	0	USA	Developed	16	5.8	8	2		89.4
4014	0	TURKEY	Developing		7.5		1070		31.7

4020	1	PAPUA NEW GUINEA	Developing	35	7.4				0.1
4021	1	CHILE	Developing	60	7.6	10	9	26	4438.5
4029	1	GREECE	Developed		7.2	10	476	2412	887.7
4055	1	ALGERIA	Developing		6.7	11	1243	5000	52.9
4139	1	MEXICO	Developing	25	7.9	7	65		210.9
4199	1	USA	Developed	5	7.7	10	28		89.6
4216	1	MOROCCO	Developing	33	5.9	10	13100	25000	960.9
4227	1	CHILE	Developing	33	9.5	12	2223	3000	8007.3
4276	0	IRAN	Developing	27	7.2	9	12225	2776	235.4
4285	0	LIBYA	Developing	5	5.4		300	375	38.7
4297	0	MACEDONIA	Economies in transition	5	6	10	1070	3300	7745.7
4311	1	USA	Developed	33	9.2	10	139		3058.3
4322	1	JAPAN	Developed	40	7.5	5	26	450	611.7
4333	0	ALGERIA	Developing	10	5.5		4	350	15.0
4341	1	GREECE	Developed	18	6.3		38	253	60.2
4351	0	USA	Developed	59	6.6	8	7		210.7
4353	0	EL SALVADOR	Developing	23	6	6	125	1138	263.4
4357	0	GREECE	Developed	38	6.3			60	29.3
4362	0	UGANDA	Developing	36	7		140		11.0
4369	0	NEPAL	Developing	23	6.3	8	80	100	7.3
4377	0	GREECE	Developed	17	5.9	8		25	102.4
4381	1	PERU	Developing	38	8.1	9	110	3000	256.0
4383	1	CHILE	Developing	47	7.8	8	3	6	2.9
4410	1	VENEZUELA	Developing	10	6.5	8	300	1536	993.5
4423	0	NICARAGUA	Developing	5	4.6				13.6
4425	0	ITALY	Developed	33	6		216	563	2179.5
4428	1	GREECE	Developed	15	7.2	10	20	18	4.1
4434	0	IRAN	Developing	34	5.3	7	38	100	6.8
4448	1	PHILIPPINES	Developing	36	7.3		270	261	54.5

4456	0	IRAN	Developing	13	7.3	10	10488	17000	238.4
4483	0	ETHIOPIA	Developing	35	6.2	9	40	160	2.1
4495	0	SOUTH AFRICA	Developing	33	6.3		12		155.0
4502	0	USA	Developed	10	4.8	8	1		53.9
4523	0	TURKEY	Developing	20	7.4	10	1086	1174	339.6
4531	1	PERU	Developing	43	7.9	10	66794	50000	3237.6
4543	0	IRAN	Developing	16	6.7	8	176	483	48.9
4546	0	COLOMBIA	Developing	8	6.6			2	2.4
4550	1	PAPUA NEW GUINEA	Developing	8	7.3	8	18	20	10.7
4551	0	PERU	Developing	25	7.6	10	82	350	36.7
4556	0	ITALY	Developed	33	4.6	8	24	150	239.9
4558	0	USA	Developed	8	6.5	11	65	2000	2955.4
4572	1	CHILE	Developing	58	7.8	9	83	447	1383.5
4619	0	NICARAGUA	Developing	5	6.2		10000	20000	16829.3
4632	1	PHILIPPINES	Developing	33	7.5	9	15	64	2.4
4634	0	COSTA RICA	Developing	33	6.5		26	100	1.1
4637	0	USA	Developed	48	6.5	8		11	30.7
4638	1	JAPAN	Developed	48	7.7	8	27		26.7
4666	1	PERU	Developing	13	8.1	9	78	2414	48.1
4671	1	PAKISTAN	Developing	22	6.2	7	5300	17000	15.6
4696	0	USA	Developed	15	5.6	8		10	13.2
4704	1	USA	Developed	2	7.7	9	2		17.6
4711	1	GUATEMALA	Developing	5	7.5	9	23000	76000	8943.3
4720	0	AUSTRIA	Developed	9	6.5	10	978	1700	14995.8
4731	0	INDONESIA	Developing	40	6.5		573	4755	812.3
4735	0	CHINA	Developing	23	7.5	11	242769	799000	23326.8
4739	1	PHILIPPINES	Developing	33	8	6	8000	10000	558.2
4760	0	ROMANIA	Developed	89	7.5		1641	10500	7822.3
4764	0	PHILIPPINES	Developing	37	7.3	7	1	9	0.4

4784	0	ARGENTINA	Developing	13	7.4	9	70	300	312.9
4806	1	JAPAN	Developed	44	7.7	8	28	10000	3144.5
4808	1	GREECE	Developed	3	6.4	8	50		908.8
4815	0	IRAN	Developing	33	7.8		20000		181.8
4828	1	MEXICO	Developing	49	7.6		5	35	97.9
4830	1	SERBIA AND MONTENEGRO	Economies in transition	10	6.9	9	131	1001	8814.7
4834	0	INDONESIA	Developing	25	5.8		22		13.5
4838	0	PANAMA	Developing	28	6.5				6.5
4851	0	INDONESIA	Developing	62	6.1		30	200	52.2
4858	0	COLOMBIA	Developing	108	6.4		72	600	65.3
4863	1	COLOMBIA	Developing	33	7.7	9	600	20000	26.1
4868	1	AZORES (PORTUGAL)	Developed	10	6.7	8	69	600	14.4
4870	0	USA	Developed	11	5.9	7		50	33.1
4877	0	SERBIA AND MONTENEGRO	Economies in transition	9	5.8	8		50	14.4
4878	0	USA	Developed	7	6.1	7		7	5.8
4882	0	GREECE	Developed	14	6.4	9	1		14.4
4887	0	USA	Developed	8	5.1	7		2	2.9
4888	0	NEPAL	Developing	18	6.5		200		704.7
4891	0	ECUADOR	Developing	55	5.6		8	100	14.4
4894	0	JAPAN	Developed	73	6	6	2	73	2.9
4896	1	ALGERIA	Developing	10	7.7	10	5000	9000	14957.4
4899	0	MEXICO	Developing	72	6.4	9	300		14.4
4902	0	PERU	Developing	71	4.9		7		2.9
4903	0	ITALY	Developed	20	6.9	10	4689	7700	57528.4
4907	0	IRAN	Developing	33	5.8		26		14.4
4908	0	IRAN	Developing	41	5.2		3	139	2.9
4909	0	INDONESIA	Developing	33	6.7		305		13.0
4911	0	CHINA	Developing	33	6.8	8	150	489	13.0
4913	0	ITALY	Developed	10	4.6	7	12		1.3

4915	1	GREECE	Developed	33	6.7	8	22	400	2117.2
4919	0	USA	Developed	6	6	7			3.9
4921	0	IRAN	Developing	33	6.7		3000		13.0
4922	0	PERU	Developing	24	5.2		10		13.0
4923	0	IRAN	Developing	33	7.1		3000	1000	2607.4
4925	0	BOSNIA-HERZEGOVINA	Economies in transition	16	5.5	8		44	13.0
4929	0	PAKISTAN	Developing	33	5.9		220	2500	13.0
4931	0	COLOMBIA	Developing	54	5.4	7	15		13.0
4939	0	INDONESIA	Developing	40	5.5			17	8.6
4941	0	PERU	Developing	95	6.1	6	3		12.3
4945	0	EL SALVADOR	Developing	82	6.2	7	43		12.3
4953	0	MACEDONIA	Economies in transition	21	5.5	8	1	12	12.3
4954	0	YEMEN	Developing	5	6	8	2800	1500	4912.3
4957	0	AFGHANISTAN	Developing	36	6.6		500	3000	2.5
4959	1	INDONESIA	Developing	33	5.9		13	390	3.6
4963	0	MACEDONIA	Economies in transition	24	4.7	7	12		1.2
4967	0	IRAN	Developing	33	4.9		30	61	11.9
4968	0	COLOMBIA	Developing	22	4.9	8	350	1200	977.8
4971	0	TURKMENISTAN	Economies in transition	33	4.6	5			11.9
4972	0	COSTA RICA	Developing	37	7.3		6		2.4
4973	0	INDONESIA	Developing	79	6.6	7		100	2.4
4975	0	PERU	Developing	104	6.5				1.2
4977	0	USA	Developed	10	6.2	8		45	73.8
4978	1	JAPAN	Developed	24	7.7	8	104	324	1903.8
4980	0	COSTA RICA	Developing	33	6.2	9	2	60	11.9
4982	0	USA	Developed	37	6.1	6			2.4
4983	0	IRAN	Developing	41	5	6	3	41	11.9
4987	1	PHILIPPINES	Developing	29	6.5	8	16	471	5.3
4991	1	CHILE	Developing	15	7.3	7	5	12	2.4

5677	0	USA	Developed	14	7.3	9	2	2	29.7
4999	0	TURKEY	Developing	12	6.9		1342	1142	59.5
5002	0	CHINA	Developing	19	5.3	7	34	2200	11.9
5004	0	BELGIUM	Developed	10	5	7	2	30	119.0
5006	0	USA	Developed	12	6.7	8		6	15.5
5009	0	PAPUA NEW GUINEA	Developing	26	6.4	5	10		59.5
5011	0	GUINEA	Developing	11	6.2	9	443	1436	19.0
5018	0	AFGHANISTAN	Developing	215	7.2	7	26	483	7.1
5019	1	INDONESIA	Developing	33	6.6		2	89	1.1
5022	0	AFGHANISTAN	Developing	33	5.8	4	1	35	11.4
5027	0	PAKISTAN	Developing	208	6.1	6	4	13	11.4
5028	0	UZBEKISTAN	Economies in transition	15	7	9		100	11.4
5030	0	USA	Developed	8	6.1	8		27	18.2
5035	0	ITALY	Developed	12	5.3	8	3	200	57.0
5036	0	ITALY	Developed	10	5.8	8	3	100	11.4
5043	0	INDONESIA	Developing	33	5.2			123	2.3
5046	0	JAPAN	Developed	10	6.1	6	29		98.1
5054	0	ARGENTINA	Developing	5	6.9	7	6	238	11.0
5056	1	CHILE	Developing	40	8	8	180	2575	3304.1
5064	0	CHINA	Developing	5	5.8	8	23	300	2.2
5068	0	PAPUA NEW GUINEA	Developing	27	7.1	8	1		2.2
5069	1	PAPUA NEW GUINEA	Developing	46	7.2	7			2.2
5070	0	AFGHANISTAN	Developing	99	6.6	8	5	38	4.4
5072	0	CHINA	Developing	7	7.5	7	71	162	11.0
5076	1	MEXICO	Developing	28	8.1	9	9500	30000	8811.0
5082	0	TAJIKISTAN	Economies in transition	16	5.9	9	29	80	440.6
5092	0	PERU	Developing	51	4.6	6	16	170	47.6
5094	0	TURKEY	Developing	10	5.9		15	100	10.8
5099	0	VENEZUELA	Developing	19	6.2	7	2	45	2.2

7877	0	PAPUA NEW GUINEA	Developing	102	7.1	7			1.1
5100	0	USA	Developed	12	6	7		29	9.7
5102	0	USA	Developed	6	5.8	6			1.5
5105	0	USA	Developed	9	6.2	6			2.2
5107	0	ROMANIA	Developed	132	6.9	8	2	558	1578.7
5111	0	EL SALVADOR	Developing	7	5.4		1100	20000	3243.8
5118	0	BRAZIL	Developing	5	4.8		1		10.8
5119	0	BULGARIA	Developed	21	5.6	7	3	60	10.8
5124	0	ALGERIA	Developing	10	4.3		1	7	2.1
5126	1	PAPUA NEW GUINEA	Developing	55	7.6	7	3		5.5
5127	0	NEW ZEALAND	Developed	19	6.6	10	1	25	438.1
5130	0	ECUADOR	Developing	10	7.2		5000		3129.6
5134	0	INDONESIA	Developing	11	6.6		2	22	2.1
5138	0	CHILE	Developing	70	6.9	7	5	112	2.1
5143	0	USA	Developed	10	5.7	8	8	200	746.9
5148	0	USA	Developed	5	6.2	6	2		6.3
5150	1	INDONESIA	Developing	33	6.5		125	108	10.4
5167	0	NEPAL	Developing	57	6.6	8	1091		263.5
5176	0	CHINA	Developing	18	7.3	10	738	3900	538.9
5184	0	ARMENIA	Economies in transition	5	6.8	10	25000		32457.1
5191	0	MALAWI	Developing	30	6.1		9	100	53.5
5192	0	CHINA	Developing	13	6.2		11	42	313.5
5197	0	CHINA	Developing	33	5.6		1	91	103.2
5199	0	IRAN	Developing	31	5.8		114		9.6
5213	1	USA	Developed	19	6.9	9	62	3757	10704.0
5218	0	CHINA	Developing	33	4.7		4	161	9.6
5220	0	PHILIPPINES	Developing	24	7.5	6	2		1.9
5221	0	AUSTRALIA	Developed	10	5.4	8	12	100	1911.4
5222	1	PHILIPPINES	Developing	26	6.6	7		300	1.6

5223	0	USA	Developed	5	5.5	7		30	23.0
5224	0	PAKISTAN	Developing	10	6.1		11	40	1.8
5233	0	CHINA	Developing	8	6.9		126	2049	105.8
5240	0	PERU	Developing	24	6.5	6	200	8130	1.8
5241	0	ROMANIA	Developed	89	6.7	6	14	700	43.0
5248	0	IRAN	Developing	19	7.7	7	50000	105000	14507.5
5253	0	PHILIPPINES	Developing	25	7.8	9	2412	3000	670.2
5264	0	IRAN	Developing	11	6.7		22	100	421.3
5265	0	INDONESIA	Developing	48	6.8		1	32	3.8
5268	1	ITALY	Developed	11	5.3	7	19	200	906.7
5271	0	COSTA RICA	Developing	17	5.7	8	2	350	35.4
5274	0	AFGHANISTAN	Developing	142	6.4	7	848	200	62.6
5282	1	COSTA RICA	Developing	10	7.6	10	89		887.5
5286	0	GEORGIA	Economies in transition	17	7	9	270		2958.4
5295	0	USA	Developed	11	5.1	7	2	104	58.3
5296	0	INDONESIA	Developing	29	6.5		28	181	13.4
5307	0	INDIA	Developing	10	7	8	2000	1800	104.4
5317	0	TURKEY	Developing	27	6.9	8	653	2000	1267.0
5319	0	NETHERLANDS	Developed	21	5.2	8	1	45	168.9
5321	1	USA	Developed	15	7.1	8		98	126.7
5323	0	KYRGYZSTAN	Economies in transition	50	6.2	7	4		52.4
5328	0	USA	Developed	1	7.6	9	3	400	155.4
5332	0	KYRGYZSTAN	Economies in transition	27	7.5	9	75		219.6
5339	0	EGYPT	Developing	22	5.3		545	6512	2027.2
5343	1	INDONESIA	Developing	28	7.8		2500	2103	168.9
5344	0	JAPAN	Developed	102	7.6	6	2		587.2
5357	1	JAPAN	Developed	17	7.7	8	231	233	1979.8
5360	1	USA TERRITORY	Developed	59	7.8	9		48	410.1
5364	0	INDIA	Developing	7	6.2	8	11000	30000	492.1

5366	0	PAPUA NEW GUINEA	Developing	25	6.9		60	200	8.2
5372	1	USA	Developed	18	6.7	9	60	7000	63972.2
5377	0	UGANDA	Developing	14	6.2		7		112.0
5379	0	INDONESIA	Developing	23	6.9		207	2000	272.6
5380	0	IRAN	Developing	6	6.1		6		2.2
5387	0	COLOMBIA	Developing	12	6.8		295		3.8
5394	1	PHILIPPINES	Developing	32	7.1		81	225	5.9
5395	0	USA	Developed	23	5.5	7			3.4
5397	1	JAPAN	Developed	27	7.8	9	3	200	272.5
5399	1	JAPAN	Developed	22	6.9	11	5502	36896	155523.0
5404	0	COLOMBIA	Developing	74	6.4		45	400	77.8
5405	0	CYPRUS	Developed	10	5.9	7	2	5	6.7
5408	1	GREECE	Developed	14	6.6	8		25	699.9
5412	1	RUSSIA	Economies in transition	11	7.1	9	1989	750	466.6
5413	1	GREECE	Developed	14	6.5	7	26	60	1026.5
5414	0	CHINA	Developing	13	6.8		11	136	56.1
5416	1	CHILE	Developing	46	8	7	3		2.8
5419	0	TURKEY	Developing	33	6.4	8	95	348	320.1
5424	0	CHINA	Developing	10	6.2		81	800	124.4
5430	0	CHINA	Developing	11	6.6	9	322	17057	764.4
5433	1	INDONESIA	Developing	33	8.2		164	423	6.3
5436	0	CHINA	Developing	28	6.3		24	128	36.3
5447	0	IRAN	Developing	10	6.5	8	88	1948	44.3
5459	0	TRINIDAD AND TOBAGO	Developing	5	6.7			2	36.9
5461	0	IRAN	Developing	10	7.2	10	1728	2300	147.7
5465	0	INDIA	Developing	36	5.8	8	56	1000	211.2
5467	1	VENEZUELA	Developing	20	7	5	81	522	119.6
5473	0	ITALY	Developed	10	6	10	14	100	6682.1
5474	0	INDONESIA	Developing	33	5.9		20	300	1.6

5475	0	CHILE	Developing	58	7.1		8	300	70.9
5481	0	CHINA	Developing	30	5.7	8	70	11500	415.1
5495	0	AFGHANISTAN	Developing	33	6.6		4700		14.5
5497	0	TURKEY	Developing	33	6.3	8	145	1500	799.8
5498	0	AZORES (PORTUGAL)	Developed	10	6.2		10	100	104.7
5509	0	CHINA	Developing	33	5.6		5	1543	101.8
5510	1	INDONESIA	Developing	33	7.7	7	41	107	290.8
5511	0	CHINA	Developing	10	4.5			84	32.9
5512	0	COLOMBIA	Developing	17	6.2		1185	4750	2642.4
5514	0	RUSSIA	Economies in transition	33	3.2	3			1.3
5515	0	SPAIN	Developed	10	4.8	7		20	62.6
5522	0	AZERBAIJAN	Economies in transition	33	5.4	5	1	18	7.1
5523	0	MEXICO	Developing	70	7		20	200	322.7
5527	1	TURKEY	Developing	13	7.6	10	17118	50000	28453.4
5531	0	GREECE	Developed	10	6	9	143	1600	5975.2
5535	0	TAIWAN	Developing	33	7.7	10	2297	8700	19917.4
5537	0	MEXICO	Developing	61	7.5	8	35	215	234.5
5539	0	TAIWAN	Developing	33	5.9	6	1	254	1.1
5541	0	CHINA	Developing	10	5.3			4	62.6
5543	0	TURKEY	Developing	10	7.2	9	894	4948	1422.7
5548	0	PHILIPPINES	Developing	33	7.3	7	5	40	2.6
5549	0	INDONESIA	Developing	56	6.5		5	220	5.5
5550	0	ALGERIA	Developing	10	5.6	7	24	175	86.7
5553	0	CHINA	Developing	33	5.9		7	2528	101.2
5557	1	INDONESIA	Developing	26	7.6		46	264	41.3
5560	0	INDONESIA	Developing	33	7.9	6	103	2174	8.3
8209	0	ICELAND	Developed	10	6.5			1	27.5
5566	0	ICELAND	Developed	10	6.5				16.5
5574	0	RUSSIA	Economies in transition	10	6.8	6		8	1.3

5575	0	CHINA	Developing	33	4.2		1	406	59.2
5576	0	USA	Developed	10	5	7		41	68.8
5578	0	JAPAN	Developed	10	6.7	9		130	206.5
5587	1	EL SALVADOR	Developing	60	7.7	8	844	4723	1007.8
5589	0	INDIA	Developing	16	7.7	10	20005	166836	3510.4
5592	0	EL SALVADOR	Developing	10	6.6	6	315	3399	466.4
5595	0	USA	Developed	52	6.8	8	1	400	2676.6
5596	0	JAPAN	Developed	50	6.8	9	2	161	669.2
5598	0	CHINA	Developing	33	5.5		2	605	48.2
5625	1	PHILIPPINES	Developing	31	7.5		15	100	2.3
5639	0	IRAN	Developing	10	6.5	8	261	1300	395.2
5648	0	ITALY	Developed	5	6		2	20	658.7
5660	0	ITALY	Developed	10	5.7		29	135	1048.7
5663	0	USA	Developed	5	7.9	9		1	73.8
5694	1	ALGERIA	Developing	12	6.8	10	2266	10261	6440.7
5695	0	JAPAN	Developed	68	7			143	300.1
5707	0	CHINA	Developing	10	5.9		16	584	96.6
5708	0	JAPAN	Developed	10	5.5			569	529.4
5724	1	JAPAN	Developed	27	8.3			755	115.9
5725	0	RUSSIA	Economies in transition	16	7.3	10	3	5	13.7
5732	0	CHINA	Developing	10	5.8		9	43	51.5
5749	0	USA	Developed	8	6.6	8	2	40	386.4
5751	0	IRAN	Developing	10	6.6	9	31000	30000	42.1
5768	0	CHINA	Developing	19	5.4			100	92.8
5807	0	JAPAN	Developed	16	6.6		40	3183	35132.2
5817	0	INDONESIA	Developing	10	7.1	8	32	130	69.0
5823	1	INDONESIA	Developing	30	9.1		227899		12547.2
6778	0	PAKISTAN	Developing	26	7.6	8	80361	71574	6310.7
7241	0	RUSSIA	Economies in transition	22	7.6			40	64.7

7245	0	INDONESIA	Developing	13	6.3		5749	38568	3644.6
7368	1	USA	Developed	39	6.7	8			85.8
7525	0	CHINA	Developing	5	6.1		3	329	354.4
7521	1	JAPAN	Developed	10	6.6		9	1088	14291.9
7843	1	CHINA	Developing	19	7.9	9	87652	374171	94673.4
8207	0	CHINA	Developing	19	6.3			3	4.4
8264	0	ITALY	Developed	9	6.3		306	1500	2762.0
8409	1	INDONESIA	Developing	81	7.5		1117	1214	2430.5
8552	0	INDONESIA	Developing	10	6.6		3		11.0
8451	0	INDONESIA	Developing	18	6.6		2	200	2.7
8693	0	TAJIKISTAN	Economies in transition	47	5.1				1.6
8712	0	USA	Developed	29	6.5			30	23.7
8732	1	HAITI	Developing	13	7		316000	300000	8695.6
8872	1	CHILE	Developing	23	8.8	9	558	12000	32608.6
9152	0	TAIWAN	Developing	21	6.3			96	1087.0
8932	0	MEXICO	Developing	4	7.2		2	233	1250.0
8972	0	CHINA	Developing	17	6.9		2220	12135	543.5
9492	0	NEW ZEALAND	Developed	12	7	9		2	7065.2
9779	0	NEW ZEALAND	Developed	6	6.1		181	1500	15805.4
9799	1	JAPAN	Developed	30	9		18457	6152	231902.8
9801	0	MYANMAR (BURMA)	Developing	8	6.8		75	123	3.8
9858	0	NEW ZEALAND	Developed	6	6		1	45	3161.1
9842	0	INDIA	Developing	50	6.9		111	177	23.5
9845	0	TURKEY	Developing	16	7.1		604	2608	1580.5
9849	0	CHINA	Developing	28	5.6				11.1
9870	0	PHILIPPINES	Developing	11	6.7		51	112	15.5
9901	0	ITALY	Developed	10	5.9		17	350	16310.8
9924	0	CHINA	Developing	18	6.3			52	70.2
9935	1	PHILIPPINES	Developing	28	7.6		1	1	0.3

9937	0	CHINA	Developing	10	5.6		81	821	1032.3
9975	0	TAIWAN	Developing	19	6		1	86	1.1
10036	0	USA	Developed	0	2.1		14	200	101.7
9991	0	INDIA	Developing	15	5.7		3	90	19.8
10048	0	PHILIPPINES	Developing	20	7.1		186	583	21.0
10069	0	GREECE	Developed	12	6.1			6	178.2
10082	0	USA	Developed	5	5.1	6			10.8
10099	0	GERMANY	Developed	10	4.1				1.4
10096	0	GREECE	Developed	10	6.9		3	266	4505.3
10110	0	USA	Developed	11	6		1	172	700.8
10125	0	CHINA	Developing	16	5.2				14.7
10134	0	NEPAL	Developing	15	7.8		8200	17866	10000.0
10156	1	CHILE	Developing	22	8.3	9	15	34	600.0