Geographic Rating of Individual Risk Transfer Costs Without Territorial Boundaries

by Randall E. Brubaker, FCAS

Notice: A patent application is pending with regard to the procedure described in this paper. The rights to use this procedure may be subject to restriction.

"Geographic Rating of Individual Risk Transfer Costs Without Territorial Boundaries"

Abstract

This paper describes a geographic ratemaking procedure that does not require territories or territory boundaries. The procedure develops a unique rate for every point on the map. The result can be visualized as a smooth surface over a map, with the height of the surface at any point representing the rate for that point. Abrupt changes in rates such as those which occur at territory boundaries are eliminated, though "natural boundaries" can be provided on an exception basis.

The procedure described uses massive computing power (as available on personal computers), and geo-coded loss data. Policy rating requires use of a personal computer, and geographic software to determine latitude/longitude for a risk.

The paper discusses credibility and data-weighting concepts, and determination of the effect of a rate change. It also compares the traditional territorial ratemaking model to the proposed method in the context of a generalized model for determining rates for individual locations.

An example of the described method is provided based on an adaptation of zip code data.

Introduction

The third principle of ratemaking of the Casualty Actuarial Society is "a rate provides for the costs associated with an individual risk transfer." Risk location is known to be an important determinant of the cost of an individual risk transfer for certain types of insurance. This paper discusses the estimation of cost of an individual risk transfer, as it would be done one risk at a time, taking into consideration the unique geographic location of each risk. A ratemaking procedure is proposed that effectively repeats this process for all locations. The proposed ratemaking procedure does not use rating territories.

A ratemaking procedure that evaluates costs for every location will likely result in rates that respond gradually to changes in geographic location. Two risks not far apart will in most cases largely share similar influences of geographic location on expected losses. Thus expected losses evaluated separately for two close-by risks should not be much different and rates should change gradually in response to location. This will eliminate the discontinuities that occur at territorial boundaries. There may be situations in which "natural boundaries" such as a river separate risks that are close-by in distance, and the rate development procedure may need to recognize these particular situations. As will be described below, it is possible in the calculation of rates for individual geographic points to select the data that is most appropriate for the calculation of a rate at each point.

The proposed procedure for geographic pricing implies massive data processing requirements. It is necessary to perform separate calculations of indicated rates for many different geographic locations. It isn't necessary (and would be impossible) to calculate rates for all points, however. Rates may be calculated for a finite number of evenly-spaced points close enough together that an interpolation procedure is reasonable to determine rates for points in between. By this approach only a finite number of experience-based rate calculations are necessary.

The development of massive computing power (even in a desk-top PC) is one of the developments of the last decade or so that makes this proposed approach to geographic pricing practical.

An additional capability necessary to determine premiums based on individual risk location is the means to rate policies without the need for the policy-rater to manually determine latitude/longitude for each risk. The recent development of geographic software offers a practical solution to this requirement. Software is available that will determine latitude/longitude based on street address, and/or based upon indication by the policy rater of risk location on a map shown on a computer screen. These capabilities make it possible to determine an individual risk location's premium based on knowledge of street address, and/or indication of map location.

Ratemaking for Geographic Points

The orientation towards insurance rate-making offered in this paper is most different from current methods in that the objective is to develop rates for specific geographic points, as opposed to territories. It also is necessary to develop a ratemaking algorithm that the computer will use to calculate rates for a large number of points. As described above, the end product of the ratemaking process will be rates for a large number of evenly spaced pre-determined points, and an interpolation procedure that allows determination of a rate for any point that lies among the predetermined points.

Before proceeding further, the definition of a term will be useful. The predetermined points for which rates are calculated by the ratemaking algorithm will be called "grid" points.

The first task for an actuary in use of this procedure will be to establish the grid points. Since rates for points in between grid points will be based on interpolation, the consideration that must be made is how far apart the grid points may be for interpolation among the grid points to still be reasonable. This will depend upon the degree of variation in expected loss based upon location. This judgment may vary among different regions of a state. In metropolitan areas where expected loss varies over relatively short distances, grid points may need to be only a mile or less apart. In rural areas five or ten miles may be acceptable.

The number of grid points will affect the computing resources necessary. For a state 200 miles-square, grid points one mile apart will mean that the ratemaking algorithm must be repeated 201 x 201 times, or 40,401 iterations.

The next step in the ratemaking process is to provide the computer instructions on what loss and exposure data to use for calculation of the rate for each grid point. The procedure described here assumes that available data consists of historic geo-coded individual risk loss and exposure data. (The term "geo-coded" means that latitude and longitude for each risk location are part of the statistical loss and exposure records). Selection of data for calculation of a rate for each grid point should be based on criteria that indicate expected similarity to the grid point in terms of expected loss per unit of exposure. The most obvious criteria for similarity is geographic proximity. This can be translated into a rule for the ratemaking algorithm such as "for each grid point, calculate a rate using all loss and exposure data within a ten mile radius of the grid point." More sophisticated criteria are possible. If a radius criteria such as the above example is used, the radius could be different for different grid points. The geographic shape of the data set used for each grid point also does not need to be a circle. The actuary is free to use judgment in deciding what data to use for each grid point. There may be data from regions of a state relatively far away that may be considered useful for ratemaking for a particular grid point, or set of grid points. There also may be data close-by that is considered not useful.

Data for calculation of a rate for a grid point may also be selected from other locations based on criteria other than geographic. For example, it may be decided that population density is a useful criteria for similar expected loss. Data for calculation for a grid point could be selected on the combined criteria of geographic proximity and similar population density.

"Natural boundaries" may be dealt with in selection of data for a grid point. The computer could be instructed that for grid points in a certain region, data from an inappropriate region shall not be used.

Another way in which the ratemaking algorithm can be refined is by varying the weight assigned to data records selected for each grid point based upon a relative similarity criteria. For example, if data within a certain radius of a grid point is used it may be desired that data closer to the grid point within that radius receive a greater weight than data from further away. Carrying the example further, the actuary might choose a ten-mile radius, and also decide that data nearest the grid point should have three times the weight of data ten miles away. Data at points in between could have weights with appropriately varying proportions. A formula weighting scheme of this type will be described in the next section of this paper. Varying weights can also be based on criteria other than distance, e.g. similarity of population density.

A Data Weighting Example

One form of weight for individual loss and exposure records used in the calculation of the rate for a grid point would be a fraction with distance from the grid point in the denominator. Such a weight would decrease as distance from the point being rated increases. A general formula for such a weight is

$$W=\left(\frac{1}{D+1}\right)^{P}$$

W in the expression above is the weight assigned to a particular loss and exposure record. D is the distance from the location of the loss and exposure record to the grid point being rated. P is an exponent that varies the sensitivity of W to D. If P is near zero, all W's will be close to 1. As P increases, the sensitivity of W to D increases. The actuary will want to choose a P value that varies weight based on distance in a reasonable manner. Further discussion of this is provided in Appendix A.

The quantity +1 in the denominator above is present to prevent division by zero in case of a data point having the same location as a grid point. This quantity also prevents an inordinately large weight being given to a data point that is only a very small distance from the grid point.

The quantity D, or distance from the data point to the grid point can be calculated in the following manner based on latitude/longitude.

$$D = 3958 \text{ miles} \times \text{Arc Cos}\,[(\sin a \sin b) + (\cos a \cos b \cos g)]$$

where a = latitude of the grid point

b = latitude of the data point

g = degrees of longitude between the grid point and the data point.

The distance 3,958 miles is approximately one radian on the earth's surface. The formula above is adapted from a formula on page 35 of Elements of Cartography, a cartography textbook by Arthur Howard Robinson (Wiley, New York, 1969).

After W's are determined for all individual data records, each W would be divided by the sum of all W's.

The example above assumes that there is exactly one exposure unit associated with each data record. If the number of exposure units varies then the formula above would be used to assign a weight to each exposure unit.

Credibility

Credibility considerations and procedures should apply to ratemaking for points in the same manner as for ratemaking for territories. "Full credibility" criteria generally are based on the volume of historic data necessary for an estimated rate to have a specified probability of falling within a desired percentage of the true expected loss per unit of exposure. Credibility criteria for rates for points should be the same as those rates for territories, unless an uneven weighting of data based on relative similarity is used. An uneven weighting will increase the expected variance of an estimated rate, for a given amount of data. It follows that uneven weighting will require an increase in data required for full credibility. An example of the sensitivity of a credibility standard to weighting is discussed in Appendix A. In this example, a credibility standard is shown to increase by a factor of 1.1 to 2.05 depending on the degree of unevenness of weighting.

A credibility procedure requires an alternate rate indicator against which to apply the complement of credibility. In ratemaking for points, there are a number of candidates for an alternate indicator, including the following:

- The prior rate for the grid point, trended to current cost level (this is possible only if this ratemaking procedure has been used before).

- A prior rate for a grid point based upon a territorial rate structure, trended to current cost level (see the section below - "Use of Summarized Data").

- An indicated rate based on a relationship of expected losses to a variable other than location, e.g. population density.

Other alternate indicators may exist.

Supplemental Smoothing

After development of rates for grid points, and application of credibility procedures, it may be desirable to view a three-dimensional representation and/or contour charts of a "surface" made of the rates for the grid points. This will illustrate if there is "bumpiness" in the surface due to apparent randomness of underlying claim experience. If this is the case then supplemental smoothing of the grid point rates may be desired. This could be done by capping the influence of large claims in the underlying data and recalculating the rates. It would also be possible to smooth a surface of the grid points by running the grid point rates themselves through the rate calculation algorithm. It is anticipated that additional smoothing procedures will be developed as experience is gained using the general approach proposed in this paper.

Use of Summarized Data

If a company does not have geo-coded individual risk data, the procedure described in this paper can be applied to loss and exposure data that is summarized by geographic areas such as zip codes or rating territories. Generally, this can be done by establishing evenly - spaced simulated data locations over the area to be rated, and apportionment of the summarized data to the simulated locations. The simulated data could be based upon territorial data such as losses or territorial rates, or it could be zip code data. An example of this procedure is included in this paper.

Loss Development, Trend, Class Distribution

Loss development and trend must also be part of the ratemaking process, and may be taken care of either by adjusting the individual risk data appropriately, or by uniformly adjusting the rates for all grid points at the end of the process to balance to an aggregate indication that incorporates development and trend.

As in territorial ratemaking, there may be an inter-relationship between location of risks and non-geographic rating factors such as driver class for automobile or construction class for property. If this is true then adjustment of loss experience to offset these factors before calculation of the rates for the grid points would be appropriate. The procedures that are used to address this issue for territorial ratemaking should also apply to the calculation of rates for grid points.

Calculation of the Effect of Rate Changes

With the proposed rating procedure it will still be desired to know the overall effect of a rate change. Also, it may be desired to determine premium at present rates for the purpose of determining a rate indication. Even though each risk has its own unique rate based on this procedure, there is a way to calculate the effect of rate changes or premium at present rates other than re-rating all historic data. The procedure to do this relies on the recording of exposure units allocated to each grid point. While every risk in the rating process gets its own rate based on interpolation among the rates for grid-points, the rating process still generates exposure data assignable to grid points. This is because by the interpolation process every risk's rate is a weighted combination of rates from grid points. The identity of the grid points used for each risk, the exposures of the risk, and the weights assigned to each grid point can be retained as statistical data. The sum of weighted exposures so recorded for each grid point for all risks rated will be the exposures allocated to each grid point to be used in developing the effect of a rate change, or in calculating premium at present rates. With exposures allocated to grid points, the effect of a rate change can be calculated as the exposure-weighted average of the changes at the grid points. Similarly premium at present rates can be calculated by extending historic exposures recorded for the grid points by the current rates for the grid points.

Policy Rating - An Interpolation Example

Suppose a rate is desired for a location with latitude/longitude (x,y), and (x,y) lies among four grid points with latitude/longitude as shown below:

.a,b		.a,c
	* (x,y)	
.d,b		.d,c

Let $R_{(x,y)}$ represent the rate for the point (x,y), and assume that rates for the four grid points have been determined based on loss and exposure data.

A formula that can be used to calculate a rate for the point (x,y) is

$$R_{(x,y)} = R_{(a,b)}\left(\frac{d-x}{d-a}\right)\left(\frac{c-y}{c-b}\right) + R_{(a,c)}\left(\frac{d-x}{d-a}\right)\left(\frac{b-y}{b-c}\right) + R_{(d,b)}\left(\frac{a-x}{a-d}\right)\left(\frac{c-y}{c-b}\right) + R_{(d,c)}\left(\frac{a-x}{a-d}\right)\left(\frac{b-y}{b-c}\right)$$

The above formula can be thought of as a simultaneous two-way linear interpolation. It gives reasonable answers but there may also be other ways of interpolating. It seems likely that any two reasonable interpolation methods will give similar answers, considering that the rates at the four grid points should not be significantly different.

A Generalized Model and Territorial Rates as a Specific Case

A general model for determination of rates for specific points based on historic claim, exposure, and location data will now be described. It will also be shown that territorial rating as a means to develop rates for specific points is a special case of this model.

Let the coordinates (x,y) specify the latitude and longitude for the point to be rated.

Let it be assumed that loss and exposure data are available from N previously recorded insurance contracts.
For each contract the following data is available:

L_i denoting losses incurred for contract i

$(x,y)_i$ denoting the coordinates for the geographic location of the earned exposure for contract i.

It is assumed that there is one unit of exposure for each historic contract.

The general model for determination of a rate $R_{(x,y)}$ for a point (x,y) is

$$R_{(x,y)}=\sum_{i=1}^{N} W_{(x,y)_i} L_i$$

with

$$\sum_{i=1}^{N} W_{(x,y)_i}=1$$

where $W_{(x,y)i}$ is the weight assigned to data from contract i in determination of the rate for the point (x,y). Note that the $W_{(x,y)i}$ may be different among the various contracts. Also, a different set of weights may be assignable among the recorded contracts for every different (x,y) location being rated.

The Traditional Rating Territory Model

In the traditional model, rates are developed based on geographic location by grouping data into mutually exclusive territories. Assuming full credibility, the rate for all points in a territory is based on the total historic losses divided by total historic exposures for all contracts provided in the territory.

By this procedure

$$R_{(x,y)}=\frac{\sum_{i=1}^{N_j} L_i}{N_j}$$

where (x,y) is in a territory which will be designated as territory J, and the first N_j historic contracts are those located in territory J.

The expression above can also be written as

$$R_{(x,y)}=\sum_{i=1}^{N_j}\left(\frac{1}{N_j}\right)L_i$$

which is equivalent in the general model to defining

$$W_{(x,y)_i}=\frac{1}{N_j}$$

for contracts that were in Territory J, and

$$W_{(x,y)_i}=0$$

for contracts in other territories.

For contracts located in the territory that (x,y) happens to be in the weights are equal, and for contracts in other territories the weights are zero. There is no use of data outside of territory J, and no distinction among data points within territory J.

If a territorial rate is the credibility weighted combination of a territory indication and a statewide indication, the calculation for a rate at a point in territory J is

$$R_{(x,y)} = Z\frac{\sum_{i=1}^{N_j} L_i}{N_j} + (1-Z)\frac{\sum_{i=1}^{N} L_i}{N}$$

which can be written

$$R_{(x,y)} = Z\sum_{i=1}^{N_j}\frac{1}{N_j}L_i + (1-Z)\sum_{i=1}^{N}\frac{1}{N}L_i$$

By rearrangement and combination of terms of the summations the above expression is equivalent to the general model with $W_{(x,y)i}$ equal to a uniform value for contracts within territory J, and another uniform value for contracts outside territory J.

The above discussion provides a framework for comparison of the territorial model to other types of models in considering which might be best for developing rates that accurately estimate expected losses for specific locations. It is apparent that the territorial model is appropriate in situations where there are two uniform degrees of relevancy of data to the point being rated. Also, these two "degrees of relevancy" are independent of the location of a point being rated within a defined territory. This type of rating could be appropriate in a situation where whether or not a point (x,y) belongs within a particular defined territory is the only significant influence on expected losses. This might be the case if a location in a particular political subdivision is the only relevant geographic consideration in determination of expected losses. If the influence of geography on expected losses is more complex than a "two-weight" model allows for, then the generalized model opens the door to other alternatives.

Expense/Profit and Catastrophe Loads

Rates for grid points may be first calculated as expected loss rates. Then expenses and profit may be added on a flat dollar basis and/or a percentage load. Coverages that include catastrophe perils will need an element for expected catastrophe losses. This could be developed from a catastrophe simulation model that develops expected loss rates at individual geographic points, or by using the catastrophe element that is built into an existing territorial rate structure.

Rate Manual Format

It is anticipated that rating of policies using this procedure will be done using a Personal Computer. However a rate manual specifying the rate structure of the company will probably be necessary at least for rate filing purposes. The geographic base rates in a rate manual would consist of a listing of the grid points and their rates. An interpolation rule for rating of policies located in between the grid points would also be part of the rate manual. The rest of the rate manual need not be different from the current format. Only the determination of geographic base rates would change.

Communication Tools

A company using this pricing procedure will need to develop new ways of communicating with regard to rate levels.

Here are some tools that should be useful.

Contour maps of rate levels.
Contour maps of rate change percentages.
Contour maps of premium comparisons.

Average rates and average rate changes over geographic areas of interest.

Maps of grid points and rates at grid points.

Rates and rate changes at individual grid points.

Any of the above maps should be producible as overlays over street maps. This will facilitate use of the maps for underwriting and marketing management.

An Example Using Zip Code Data

If geo-coded individual risk data is not available to a company, an alternative for use of the procedure described in this paper is to use zip code data of exposures and losses. Zip Code areas generally are a finer geographic breakdown than are most territorial rating structures, and much of the refinement in geographic pricing developed by this procedure can be achieved with zip code data.

An example of use of zip code data is provided here, using Private Passenger Automobile Bodily Injury zip code data for 1991 for the industry obtained from the California Insurance Department. The geographic area for which the example is developed is the San Francisco Bay Area. The example to be developed includes use of the data to determine pure premiums for a company's rating territories, so that the results of territory ratemaking can be compared to the procedure described in this paper.

Exhibit 1 attached to this paper is a zip code map of the Bay Area, with the industry pure premium (losses divided by exposure) by zip code shown on each of the zip code areas. Exhibit 2 is a map of average pure premiums for the rating territories of a well known auto insurer, based on this industry data. This company defines their territories using zip codes. (The territory definitions were obtained from documents available to the public at the California Department of Insurance). The pure premium shown on Exhibit 2 for each of the territories is an exposure-weighted average of the zip code pure premiums shown on Exhibit 1, for the zip codes within each territory.

Exhibit 3 attached illustrates the first step in the procedure described in this paper, which is to determine the "grid points" for which rates will be determined. For this example, the intervals between grid points are four-tenths of a mile each.

Exhibit 3 shows the grid points over zip code boundaries. For this example, each grid point is assumed to be a location of risk data as well as a point for which we will determine a rate. For each assumed data location, we will assign an observed pure premium equal to the pure premium (loss/exposure) of the zip code area that the grid point is within.. To each assumed data location we also assign a number of exposures equal to the total exposures for the zip code that the data point is within, divided by the number of data points in the zip code.

The next step in this example is to determine a radius around each grid point from within which data will be used to calculate a rate (pure premium) for each grid point. For this example we used a radius of one and one-half miles. We also used the weighting formula based on distance described earlier in this paper, with a P value (exponent) of 1.0.

Exhibit 4 attached is an illustration of one grid point, (at the longitude, latitude shown on the exhibit) and the simulated data points around it that are used to determine a rate (i.e. fitted pure premium) for that grid point. A similar picture would apply for all other grid points in this example.

Exhibit 5 details the calculations of the rate for the grid point illustrated on Exhibit 4. All assumed data locations that are used to calculate the rate for the grid point are listed, with their distances from the grid point and weights based on distance and exposure. The total in Column (8) is the rate for the grid point. Rates for all other grid points are similarly calculated.

Exhibit 6 is a contour chart of the rate surface that results from rates calculated for every grid point as illustrated on Exhibit 5. The rates for points in between the grid points are assumed to be based on interpolation. Exhibit 6 also shows the territorial rates shown on Exhibit 2. The territorial rates may be compared to the surface contours, with regard to which is higher or lower, and with regard to how the rates vary by location.

It may be noted that the one and one-half mile radius used for this example was chosen in recognition that the example area is densely populated, and that substantial variation of cost by location may occur. As to credibility, use of more than one year of data would be appropriate for an actual application. Also, credibility procedures can be used as discussed above. To illustrate the volume of data used for a grid point in this example, it is estimated that 946 claims were within the circle shown around the grid point in Exhibit 4. For the zip codes that are only partially within this circle, this estimate includes a proportion of the total claims in each zip code equal to the proportion of data points of the zip code within the circle.

This example includes use of a "natural boundary" in creation of the pure premium surface. Alameda Island, which lies near Oakland in the East Bay, is isolated from the nearby mainland in a manner that could be expected to develop a distinct difference in expected losses. This island has its own zip code, and the data used for the grid points on this island was restricted to be from that zip code only. Also, data from this zip code was excluded from use for any grid points on the mainland. An area known as "Bay Farm Island" lies just below Alameda Island. For the purpose of this example we used the same pure premium as for Alameda Island, instead of the zero shown on Exhibit I. We eliminated rate contours over the area of the Oakland airport.

Summary and Conclusion

The geographic insurance pricing procedure described in this paper offers a new approach to insurance pricing based on geographic location. It eliminates the need to determine territorial boundaries, and also eliminates the discontinuities in territorial rates that occur at boundaries. It should be a more accurate procedure of evaluating insurance costs in relation to location, if such costs vary in a gradual manner. The procedure is made possible by the development in the last few years of massive computing power, and geographic software.

Weighting of Data
Based on Relative Similarity

This appendix discusses the choice of P in the weighting formula

$$W=\left(\frac{1}{D+1}\right)^{P}$$

Generally, the larger P is the more weight will be assigned to nearby data points in calculation of the rate for a grid point. This will increase the sensitivity of rates for the grid points to nearby experience, but it will also decrease credibility and increase random fluctuations in estimated rates.

To illustrate how P effects W, the following table shows how W is affected for data 0 to 9 miles from a grid point by varying P. The W values below have not been adjusted so that their sum is 1. This would be done in an actual application.

	W Values				
Miles from Grid Point	P=0.01	P=0.2	P=0.6	P=1.0	P=2.0
0	1.000	1.000	1.000	1.000	1.000
1	1.000	.871	0.660	0.500	0.250
2	.999	.803	0.517	0.333	0.111
3	.999	.758	0.435	0.250	0.063
4	.999	.725	0.381	0.200	0.040
5	.998	.699	0.341	0.167	0.028
6	.998	.678	0.311	0.143	0.020
7	.998	.660	0.287	0.125	0.016
8	.998	.644	0.268	0.111	0.012
9	.998	.631	0.251	0.100	0.010

A table such as above should aid an actuary in determining a reasonable P value. For example, more sensitivity of weighting to distance is probably desirable than results from P=0.01. Less sensitivity of weighting is probably desirable than results from P=2.0. For P=2.0, data from only two miles away gets only 11.1% as much weight as data very close by. Based on examination of the table above, an actuary might choose a P value in the area of 0.5 to 0.7. The choice of a P value must ultimately be based on a judgment of the relative value of data at varying distances from a point being rated. Such a judgment could also vary between rural and urban points.

The higher the P value, the more uneven is the resulting weighting. This will decrease credibility because it will increase the variance of an estimated rate. Under traditional territorial ratemaking, a rate indication is developed from the sum of recorded losses divided by the number of exposure units. If N is the number of exposure units and σ is the standard deviation of loss for a unit of exposure, the standard deviation of the sum of recorded losses divided by N is $(1/\sqrt{N})\,\sigma$. This assumes even weighting of the observed loss and exposure records, and independence of the loss and exposure records.

A more general formula for the standard deviation of the weighted sum of observed losses is

$$\left(\sum_{i=1}^{n} (W_i)^2 \sigma^2\right)^{1/2}$$

$$\textit{with} \qquad \sum_{i=1}^{n} w_i = 1$$

The above expression is minimized by having all Wi's equal 1/N. If the Wi's are uneven, the standard deviation is larger than the minimum.

If the Wi's are uneven and the distribution of exposures by distance from a point being rated is known, it is possible to determine the increase in data required for full credibility. As an example, if exposure units are approximately uniformly distributed over a circle of radius nine miles, and a P value of 0.6 is used in the weighting formula above, it takes approximately 10% more exposure units for the standard deviation of the unevenly weighted indication to be as low as for an evenly weighted indication.

The increase in exposure units needed goes up sharply as P goes up, as shown in the table below:

P Value	Increase in Exposure Units
0.6	+10%
0.8	+25%
1.0	+45%
1.2	+85%
1.4	+105%

The spreadsheet on the following page shows how the increase in exposure units above was determined for P=0.6. The same spreadsheet was used for the other P values. While uneven weighting can be expected to always increase the amount of data needed to meet credibility standards, the percentage increase needed can be expected to vary depending upon the weighting formula used, and the geographic distribution of data.

EFFECT OF UNEVEN WEIGHTING ON CREDIBILITY STANDARD

Miles from Point Grid	# of Risks	Weights Based on Distance (1/((1)+1))^P	Weighted # of Risks (2) x (3)	Normalized Weights (3) / (4 Tot)	Sum of the Variances (2)x(5)^2	Increased # of Risks (2)x(13)	Weighted # of Risks (3) x (7)	Normalized Weights (3) / (8 Tot)	Sum of the Variances (7)x(9)^2
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
0	10	1.000	10.0	0.002988	0.000089	11	11.000000	0.002716	0.000081
1	30	0.660	19.8	0.001971	0.000117	33	21.771881	0.001792	0.000106
2	50	0.517	25.9	0.001546	0.000119	55	28.450502	0.001405	0.000109
3	70	0.435	30.5	0.001301	0.000118	77	33.516197	0.001182	0.000108
4	90	0.381	34.3	0.001138	0.000116	99	37.692348	0.001034	0.000106
5	110	0.341	37.5	0.001020	0.000114	121	41.294729	0.000927	0.000104
6	130	0.311	40.4	0.000930	0.000112	143	44.491517	0.000845	0.000102
7	150	0.287	43.1	0.000858	0.000110	165	47.383807	0.000780	0.000100
8	170	0.268	45.5	0.000800	0.000109	187	50.037557	0.000727	0.000099
9	190	0.251	47.7	0.000751	0.000107	209	52.498426	0.000682	0.000097
TOTAL	1,000		334.7		0.001113	1100	368.136964		0.001012

Std. Dev. of Est. Rate	Evenly Weighted (11a)	Unevenly Weighted (11b)	Adjusted Weighted (11c)
	0.031623	0.03336325	0.031811

INPUTS:

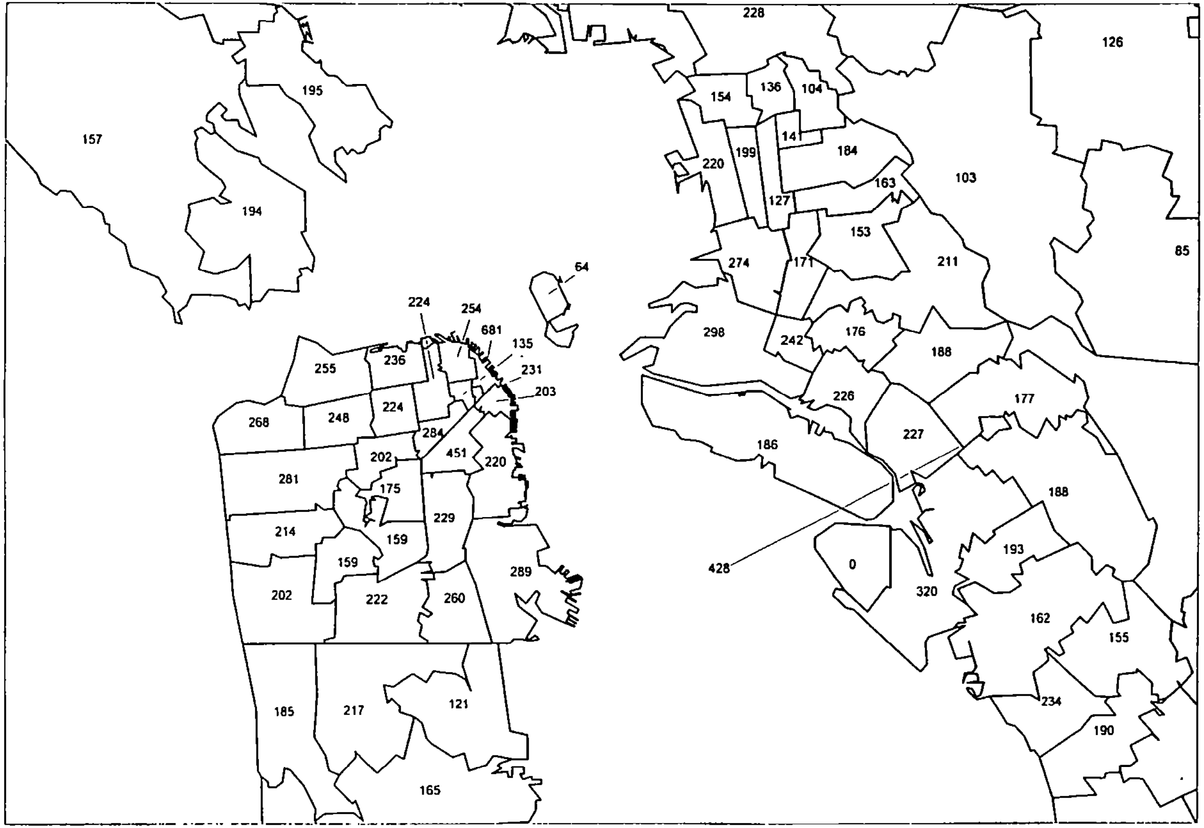
P-Factor (12)	# Risks Increase (13)
0.6	1.1

Notes:

(a) This exhibit assumes that the variance of loss experience for an individual risk equals 1.
(b) Item (11a) = 1 / 1000^.5
(c) Item (11b) = ((6) Total)^.5
(d) Item (11c) = ((10) Total)^.5
(e) Item (13) (# Risks Increase) is determined by trial and error such that Item (11c) is approximately equal to Item 11(a).

SAN FRANCISCO AREA INDUSTRY PURE PREMIUMS
BY ZIP CODE
BODILY INJURY 1991

Exhibit 1



——— 1994 ZIP CODE BOUNDARIES

Exhibit 2

SAN FRANCISCO AREA INDUSTRY PURE PREMIUMS BY RATING TERRITORY BODILY INJURY

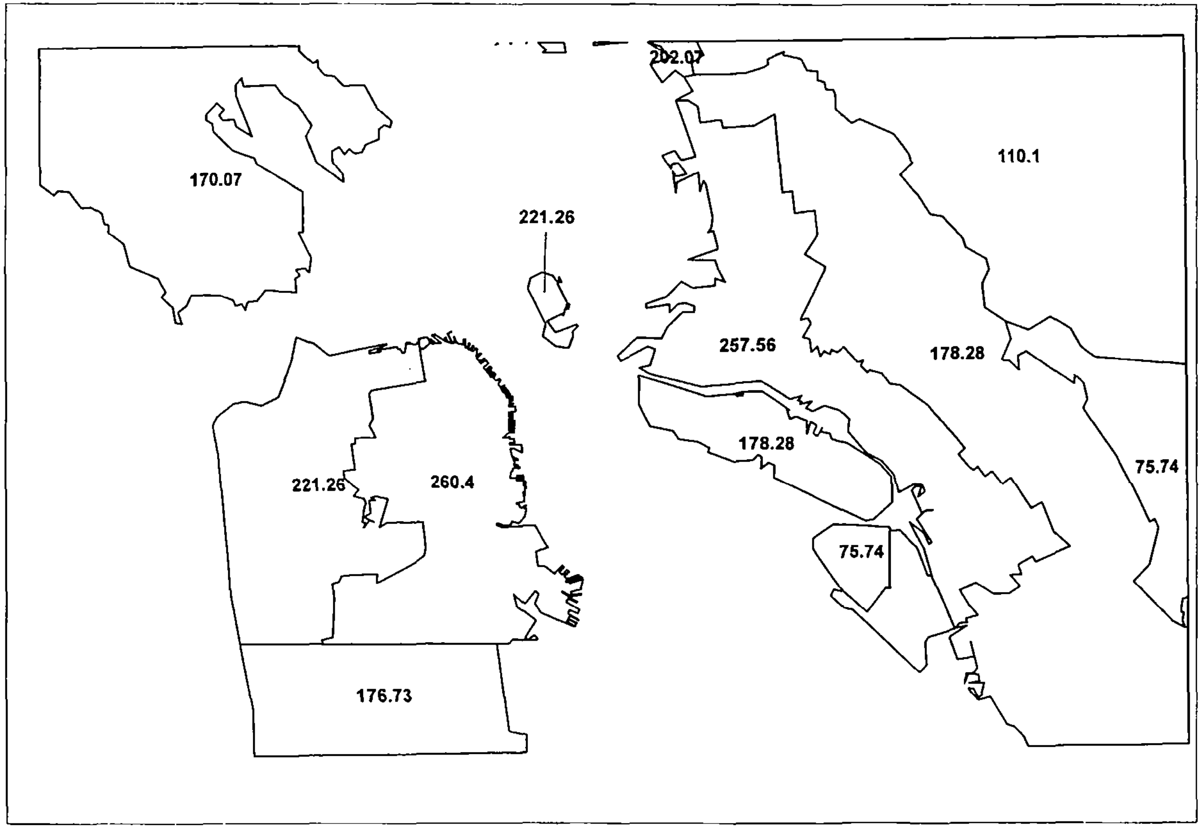
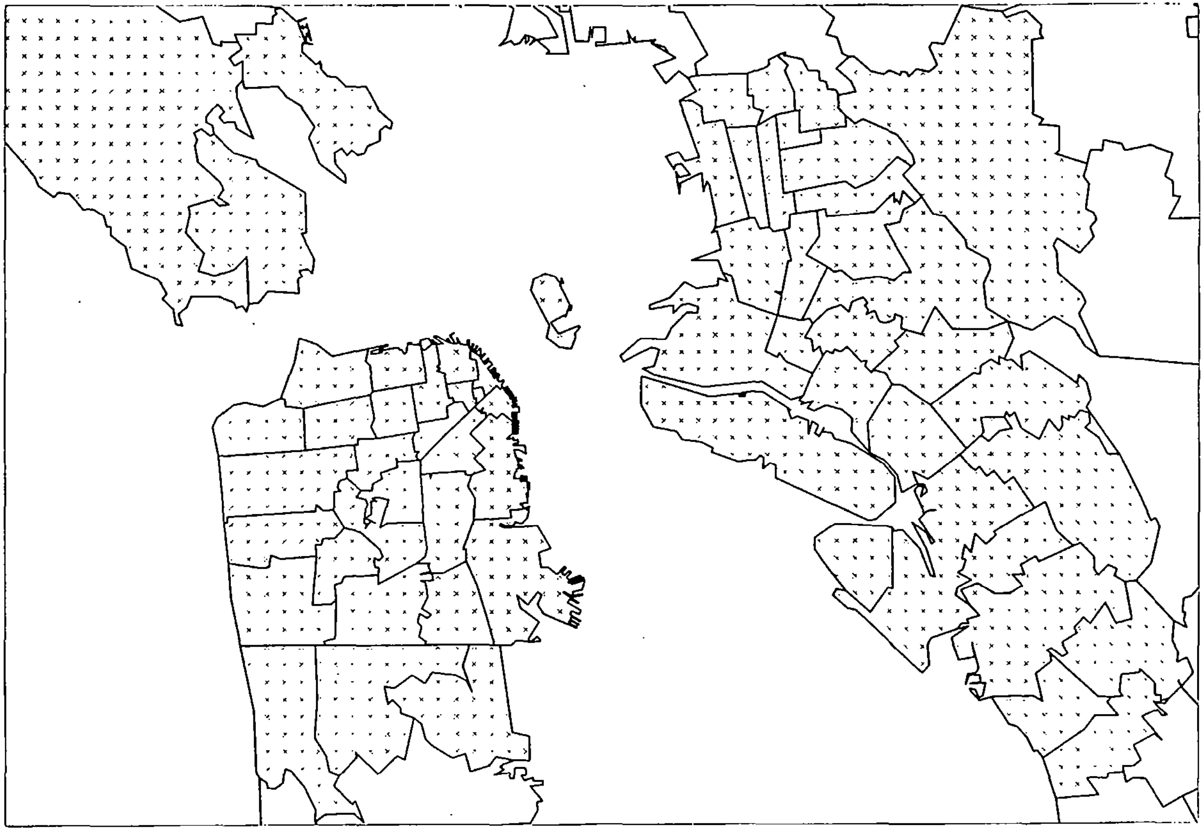


Exhibit 3

SAN FRANCISCO AREA GRID POINTS
0.4 MILES APART

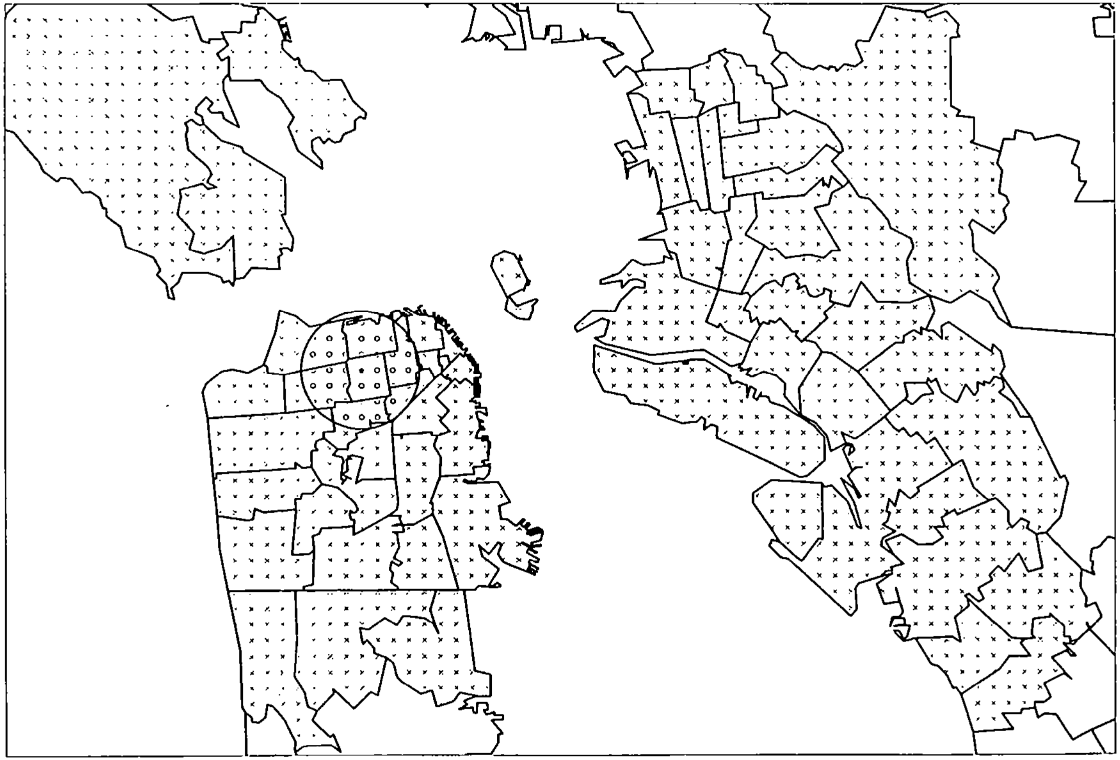


1994 ZIP CODE BOUNDARIES

Exhibit 4

CALCULATION OF BODILY INJURY PURE PREMIUM SURFACE FOR A GRID POINT

LONGITUDE = -122.439362, LATITUDE = 37.788797, RADIUS = 1.5 MILES



• GRID POINT   ∘ DATA POINT (INCLUDES GRID POINT)   ——— 1994 ZIP CODE BOUNDARIES

CALCULATION OF BODILY INJURY PURE PREMIUM SURFACE
FOR A GRID POINT
LONGITUDE = -122.439362, LATITUDE = 37.788797, RADIUS = 1.5 MILES, P=1.0

	Longitude	Latitude	Distance from Grid Point	Distance Weight	Exposures at Grid Point	Total Weight	Total Weight Normalized	Unweighted Zip Code Pure Premium	Weighted Smoothed Pure Premium
	(1)		(2)	(3)	(4)	(5)	(6)	(7)	(8)
GRID POINT	-122.439362	37.788797	0.000	1.0000	1958.00	1958.000	6.803%	223.66	231.48
	-122.461387	37.800391	1.443	0.4093	46.73	19.127	0.066%	254.86	
	-122.461378	37.794594	1.265	0.4415	46.73	20.631	0.072%	254.86	
	-122.461369	37.788797	1.200	0.4545	1353.09	614.979	2.137%	248.31	
	-122.461360	37.783000	1.264	0.4417	1353.09	597.660	2.077%	248.31	
	-122.461350	37.777203	1.441	0.4097	1353.09	554.361	1.926%	248.31	
	-122.454059	37.806188	1.443	0.4093	46.73	19.127	0.066%	254.86	
	-122.454051	37.800391	1.132	0.4690	46.73	21.916	0.076%	254.86	
	-122.454042	37.794594	0.895	0.5277	46.73	24.659	0.086%	254.86	
	-122.454033	37.788797	0.800	0.5556	1353.09	751.777	2.612%	248.31	
	-122.454025	37.783000	0.894	0.5280	1353.09	714.432	2.482%	248.31	
	-122.454016	37.777203	1.131	0.4693	1353.09	635.005	2.206%	248.31	
	-122.454007	37.771406	1.441	0.4097	799.81	327.682	1.138%	280.85	
	-122.446722	37.806188	1.265	0.4415	1505.71	664.771	2.310%	236.44	
	-122.446714	37.800391	0.895	0.5277	46.73	24.659	0.086%	254.86	
	-122.446706	37.794594	0.566	0.6386	46.73	29.842	0.104%	254.86	
	-122.446698	37.788797	0.400	0.7143	1353.09	966.512	3.358%	248.31	
	-122.446690	37.783000	0.565	0.6390	1353.09	864.625	3.004%	248.31	
	-122.446682	37.777203	0.894	0.5280	1353.09	714.432	2.482%	248.31	
	-122.446674	37.771406	1.265	0.4415	1319.13	582.396	2.023%	202.04	
	-122.439385	37.806188	1.200	0.4545	1505.71	684.345	2.378%	236.44	
	-122.439377	37.800391	0.800	0.5556	1505.71	836.572	2.907%	236.44	
	-122.439370	37.794594	0.400	0.7143	1505.71	1075.529	3.737%	236.44	
	-122.439355	37.783000	0.400	0.7143	1958.00	1398.599	4.859%	223.66	
	-122.439347	37.777203	0.800	0.5556	1319.13	732.909	2.546%	202.04	
	-122.439340	37.771406	1.200	0.4545	1319.13	599.545	2.083%	202.04	
	-122.432040	37.800391	0.894	0.5280	1505.71	795.015	2.762%	236.44	
	-122.432034	37.794594	0.565	0.6390	1958.00	1251.162	4.347%	223.66	

Exhibit 5
Page 2

CALCULATION OF BODILY INJURY PURE PREMIUM SURFACE
FOR A GRID POINT
LONGITUDE = -122.439362, LATITUDE = 37.788797, RADIUS = 1.5 MILES, P=1.0

Longitude	Latitude	Distance from Grid Point	Distance Weight	Exposures at Grid Point	Total Weight	Total Weight Normalized	Unweighted Zip Code Pure Premium	Weighted Smoothed Pure Premium
(1)		(2)	(3)	(4)	(5)	(6)	(7)	(8)
-122.432027	37.788797	0.400	0.7143	1958.00	1398.599	4.859%	223.66	
-122.432020	37.783000	0.566	0.6386	1958.00	1250.379	4.344%	223.66	
-122.432013	37.777203	0.895	0.5277	1319.13	696.105	2.419%	202.04	
-122.432006	37.771406	1.265	0.4415	1319.13	582.396	2.023%	202.04	
-122.424710	37.806188	1.442	0.4095	1505.71	616.588	2.142%	236.44	
-122.424704	37.800391	1.131	0.4693	1505.71	706.630	2.455%	236.44	
-122.424697	37.794594	0.894	0.5280	1825.67	963.954	3.349%	223.65	
-122.424691	37.788797	0.800	0.5556	1825.67	1014.342	3.524%	223.65	
-122.424685	37.783000	0.895	0.5277	1825.67	963.406	3.347%	223.65	
-122.424678	37.777203	1.132	0.4690	419.83	196.900	0.684%	284.05	
-122.424672	37.771406	1.443	0.4093	419.83	171.836	0.597%	284.05	
-122.417367	37.800391	1.442	0.4095	1825.67	747.612	2.597%	223.65	
-122.417361	37.794594	1.265	0.4415	1825.67	806.033	2.800%	223.65	
-122.417356	37.788797	1.200	0.4545	1825.67	829.767	2.883%	223.65	
-122.417350	37.783000	1.265	0.4415	419.83	185.355	0.644%	284.05	
-122.417344	37.777203	1.443	0.4093	419.83	171.836	0.597%	284.05	

Notes (by column)

(1) Location of each data point
(2) See formula in text of article
(3) $[1/[(2)+1]]^P$, where P = 1.0
(5) (3) x (4)
(6) (5)/[Sum of (5)]
(7) 1991 California Department of Insurance BI Industry Pure Premiums
(8) Sum of [(6) x (7)]

Exhibit 6

SAN FRANCISCO AREA
BODILY INJURY PURE PREMIUM SURFACE
Radius = 1.5 miles, P = 1.0