Transition Matrix Theory
And
Individual Claim Loss Development

John B. Mahon

Abstract

Motivation. Individual claim development is important for creating the average severity distributions that underlie most increased limits, and reinsurance pricing analyses, but most current methods do not adequately represent the true process.

Method. Transition Matrix Theory is applied to a large database of reinsurance data. The data is processed to isolate GL data, and the Transition Matrix process is described in detail.

Results. Individual claim size development is characterized as a distributional process. The effect of this distributional process on pricing parameters is contrasted with traditional methods.

Conclusions. Individual Claim Size development is a distributional process, and can be measured and introduced into procedures for calculating average severity distributions. A simple five parameter formula can model this process. The Transition Matrix process may overstate the distribution of the ultimate distribution, but this can be measured and corrected. Pricing parameters are affected by this process and its effect should be factored in when possible.

Keywords. Transition Matrix, Average Severity, Individual Claim Loss Development, Distributional Loss Development

1. INTRODUCTION

Loss development has long been considered to be an aggregate phenomenon, and not applicable to individual claims. Rating procedures require accurate estimates of individual claim size development in order to estimate the average severity distribution curves that underlie increased limits ratemaking, and reinsurance excess layer pricing. Current methods have limitations based on sparse data at high layers, or are based on assumptions that may introduce errors. This study applies the Transition Matrix Theory approach to a large collection of reinsurance individual large losses, and, characterizes individual claim size development as a distributional procedure. It was found that a simple five parameter distribution will model the process.

1.1 Research Context

Several approaches have been used to apply loss development to individual claims to adjust them for increased limits calculations and for pricing reinsurance excess layers. Transition matrix theory as applied to losses was introduced at the International Congress of Actuaries in 1980 by Charles Hachemeister [1]. A more recent presentation of this method can be found in Ole Hesselager's [2] 1994 paper where he presents a time continuous method for computing transition matrices. The present research uses a time discrete method of computing Markov transition matrices to represent the age to age loss development of a large body of reinsurance general liability claims.

A weakness of the transition matrix approach is that it generates a large number of parameters which make it unwieldy, and prone to parameter error.

1.2 Objective

This study uses a straightforward interpretation of the transition matrix theory and applies it to a large body of reinsurance individual large losses. This yields vectors which can be used to develop individual claims from an arbitrary size and evaluation to ultimate. The behavior of the loss development forecasts suggests a five parameter model that can be used to characterize the development of an open claim as a future distribution. This model is modified to reflect the fact that observed variation appears to be smaller than that provided by the Transition Matrix process.

1.3 Outline

The remainder of the paper proceeds as follows**.** Section 2.1 will provide a background for individual claim development. Section 2.2 describes the details of the Transition Matrix method as applied here. Section 2.3 describes the application of the Transition Matrix method to aollection of reinsurance data. Section 2.4 describes a comparison between transition matrix results and initial to final transitions. It discusses an adjustment to the Transition Matrix results to reduce excess variation introduced by the Transition Matrix method, and, proposes a model for distributional loss development. Section 2.5 describes an effect of the distributional loss development method on pricing parameters. Section 3 discusses the applicabililty and limitations of the method used here, and Section 4 collects the conclusions.

2. BACKGROUND AND METHODS

In this section, the background, method, data and application are described.

2.1 Background

Loss development has long been important to both reserving and pricing activities. Loss development for reserving has concentrated on the aggregate behavior of the losses.

Triangles of sums of losses or claim counts are subjected to procedures which measure the behavior of aggregate losses. The behavior of individual claims, for the most part, is not important. It only becomes important where large individual claims are near or at limits, and further development may distort the result. Treatment in this situation usually involves isolating these claims from the aggregate data, and handling them on an individual and ad hoc basis.

The rating discipline needs to address individual claim loss development at a more detailed level. Increased limits pricing for primary business, and excess layer pricing for reinsurance business require the correct estimate of large size losses. The issue of individual claim loss development becomes a critical factor in determining the correct probability of large losses used to determine pricing in these two business applications.

A variety of solutions been developed to deal with this problem, some, better than others. Elimination of the problem by using closed claim data has been successful to the extent that the data is available, and, not too stale. Fitting immature loss size data to severity distributions and measuring loss development by counts within empirical intervals of size, or changes in the parameters has been successful for creating increased limits factors for subline pricing for many years. It is limited by the fact that it requires large amounts of data and many man-hours to complete. This eliminates it from use in reinsurance pricing exercises.

This most common experience rating method used in reinsurance involves combining features of aggregate loss development that can be applied to individual losses. The losses are trended, then layered into the excess layer of rating interest, and then, the appropriate excess layer loss development factor is applied. This method suffers from two problems. One is that the excess loss development may be very different from the factors that are used, and the other is that there may be no losses in the higher layers after trending. Both of these can lead to significant errors.

Another method commonly used is to apply trend and average severity development factors to individual claims, then use the adjusted claims to fit a theoretical severity distribution. This severity distribution is then used to evaluate excess layers using exposure rating techniques. The first thing to say here is that it is incorrect to apply average severity loss development factors to individual losses, and call the result the ultimate value of that claim. This has to do with the nature of loss development of an individual claim. A claim can have a wide range of outcomes as it matures to ultimate. To simply say that when it matures to ultimate, it will have a value some "X" percent larger than current, misses the variability of the loss development process.

Exhibit 1 shows a typical adjustment for an individual loss to prepare it for fitting

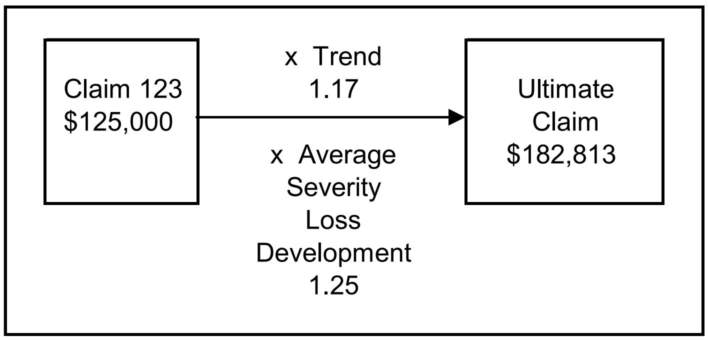

Exhibit 1 showing a typical trend and development adjustment to an indivdual claim

a severity curve. The reality of the situation is that an open claim has four possible outcomes at ultimate, it may stay the same size, it will grow in

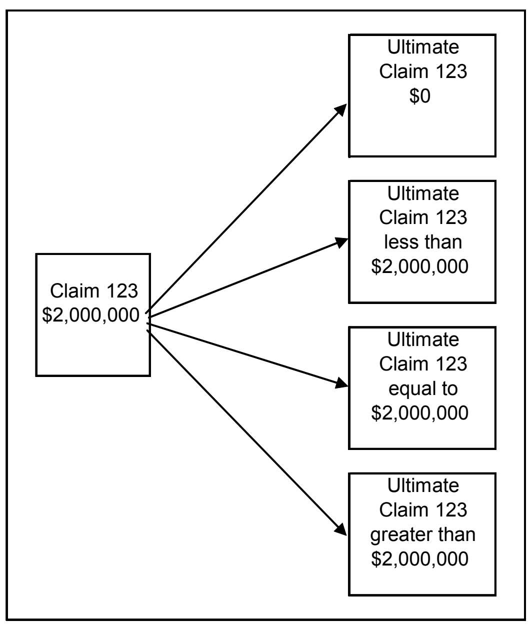

Exhibit 2. The four possible states for the ultimate settlement of a claim.

size, it will settle for a lesser amount, or it will close with no payment as shown in exhibit 2. Transition matrix theory accommodates the variation of possible outcomes.

2.2 The Transition Matrix Approach

A Markov Transition Matrix is a square matrix that contains the probabilities of moving from one state to another state [3]. For our purposes, the states will be the combination of open or closed, and size of loss. Exhibit 3 shows the complete list of states for our example. Note that the endpoints of the size intervals are determined exponentially. They increase by a constant factor, two, in this case. The interval end points can be arbitrary, but selecting exponential ones will provide additional insight into the results of this study.

Class	Open/ Closed	Interval bottom	Interval top	Count
0	Open	0	0	0
1	Open	0	200,000	0
2	Open	200,000	400,000	0
3	Open	400,000	800,000	0
4	Open	800,000	1,600,000	0
5	Open	1,600,000	3,200,000	0
6	Open	3,200,000	6,400,000	1
7	Open	6,400,000	12,800,000	0
8	Open	12,800,000	25,600,000	0
9	Open	25,600,000	51,200,000	0
10	Closed	0	0	0
11	Closed	0	200,000	0
12	Closed	200,000	400,000	0
13	Closed	400,000	800,000	0
14	Closed	800,000	1,600,000	0
15	Closed	1,600,000	3,200,000	0
16	Closed	3,200,000	6,400,000	0
17	Closed	6,400,000	12,800,000	0
18	Closed	12,800,000	25,600,000	0
19	Closed	25,600,000	51,200,000	0
As of 36 months				

Exhibit 4. This shows the state of the same claim shown in exhibit 3, but it is now $3,500,000 with a maturity of 36 months. It is now in class 6.

Class	Open/ Closed	Interval bottom	Interval top	Count
0	Open	0	0	0
1	Open	0	200,000	0
2	Open	200,000	400,000	0
3	Open	400,000	800,000	0
4	Open	800,000	1,600,000	0
5	Open	1,600,000	3,200,000	1
6	Open	3,200,000	6,400,000	0
7	Open	6,400,000	12,800,000	0
8	Open	12,800,000	25,600,000	0
9	Open	25,600,000	51,200,000	0
10	Closed	0	0	0
11	Closed	0	200,000	0
12	Closed	200,000	400,000	0
13	Closed	400,000	800,000	0
14	Closed	800,000	1,600,000	0
15	Closed	1,600,000	3,200,000	0
16	Closed	3,200,000	6,400,000	0
17	Closed	6,400,000	12,800,000	0
18	Closed	12,800,000	25,600,000	0
19	Closed	25,600,000	51,200,000	0
As of 24 months				

Exhibit 3. This shows our claim of $2,000,000 with a maturity of 24 months. All of the possible states of size and open or closed are shown. These states are labeled with a Class number which will be used to track them.

Also shown is an open claim of $2,000,000 as a count of one in class 5.

We now consider this claim as it matures to the 36 month evaluation, and it changes in value to $3,500,000. Exhibit 4 shows its state as a class 6.

Transition from 24 to 36 months																				
Final	Initial Class																			
Class	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
0																				
1																				
2																				
3																				
4																				
5																				
6						1														
7																				
8																				
9																				
10																				
11																				
12																				
13																				
14																				
15																				
16																				
17																				
18																				
19																				

Exhibit 5. The transition matrix for the sample claim that is a class 5 at 24 months and a class 6 at 36 months

We can now construct a transition matrix for the transition from 24 to 36 months for this loss as shown in exhibit 5. Consider if we have 755 class 5 losses at 24 months and they are entered into the transition matrix. This would result in the matrix shown in exhibit 6.

Transition from 24 to 36 months																				
Final	Initial Class																			
Class	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
0						13														
1						22														
2						35														
3						57														
4						91														
5						146														
6						91														
7						57														
8						35														
9						22														
10						4														
11						7														
12						11														
13						19														
14						30														
15						48														
16						30														
17						19														
18						11														
19						7														
Total						755														

Exhibit 6. The matrix showing all 24 month class 5 claims populated into their final class at 36 months.

Now we consider a complete collection of fictitious claims, in our example there are 4,259, of all sizes and open or closed status. These claims are mapped into this transition matrix and this results in the matrix shown in exhibit 7.

Transition from 24 to 36 months																					
Final	Initial Class																				
Class	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	Total
0	0	2	3	5	8	13	22	5	1	0	0	0	0	0	0	0	0	0	0	0	59
1	1	3	5	8	13	22	35	8	2	0	0	0	0	0	0	0	0	0	0	0	97
2	2	2	8	13	22	35	57	13	3	0	0	0	0	0	0	0	0	0	0	0	155
3	3	1	5	22	35	57	91	22	5	1	0	0	0	0	0	0	0	0	0	0	242
4	5	0	3	13	57	91	146	35	8	2	0	0	0	0	0	0	0	0	0	0	360
5	9	0	2	8	35	146	234	57	13	3	0	0	0	0	0	0	1	0	0	0	508
6	15	0	1	5	22	91	375	91	22	5	0	0	0	0	0	0	2	0	0	0	629
7	9	0	0	3	13	57	234	146	35	8	0	0	0	0	0	0	1	0	0	0	506
8	5	0	0	2	8	35	146	91	57	13	0	0	0	0	0	0	0	0	0	0	357
9	3	0	0	1	5	22	91	57	35	22	0	0	0	0	0	0	0	0	0	0	236
10	0	0	1	1	2	4	7	1	0	0	0	0	0	0	0	0	0	0	0	0	16
11	0	1	1	2	4	7	11	2	0	0	0	0	0	0	0	0	0	0	0	0	28
12	0	0	2	4	7	11	19	4	1	0	0	0	1	0	0	0	0	0	0	0	49
13	1	0	1	7	11	19	30	7	1	0	0	0	0	2	0	0	0	0	0	0	79
14	1	0	1	4	19	30	48	11	2	0	0	0	0	0	6	0	0	0	0	0	122
15	3	0	0	2	11	48	78	19	4	1	0	0	0	0	0	17	0	0	0	0	183
16	5	0	0	1	7	30	125	30	7	1	0	0	0	0	0	0	45	0	0	0	251
17	3	0	0	1	4	19	78	48	11	2	0	0	0	0	0	0	0	17	0	0	183
18	1	0	0	0	2	11	48	30	19	4	0	0	0	0	0	0	0	0	6	0	121
19	1	0	0	0	1	7	30	19	11	7	0	0	0	0	0	0	0	0	0	2	78
Total	67	9	33	102	286	755	1905	696	237	69	0	0	1	2	6	17	49	17	6	2	4259

Exhibit 7. This is a transition matrix populated with a complete inventory of losses starting at a 24 month maturity, and ending in a 36 month maturity. The value shown is claim count.

This matrix can be converted to a Markov transition matrix by dividing each column by the total at the bottom of the column. This normalizes each column so that it sums to one, and each value represents the probability that a selected initial class will make the transition to the selected final class. The complete Markov transition matrix is shown in exhibit 8.

Transition from 24 to 36 months																				
Final	Initial Class																			
Class	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
0	0	0.22	0.09	0.05	0.03	0.02	0.01	0.01	0	0	0	0	0	0	0	0	0	0	0	0
1	0.01	0.33	0.15	0.08	0.05	0.03	0.02	0.01	0.01	0	0	0	0	0	0	0	0	0	0	0
2	0.03	0.22	0.24	0.13	0.08	0.05	0.03	0.02	0.01	0	0	0	0	0	0	0	0	0	0	0
3	0.04	0.11	0.15	0.22	0.12	0.08	0.05	0.03	0.02	0.01	0	0	0	0	0	0	0	0	0	0
4	0.07	0	0.09	0.13	0.2	0.12	0.08	0.05	0.03	0.03	0	0	0	0	0	0	0	0	0	0
5	0.13	0	0.06	0.08	0.12	0.19	0.12	0.08	0.05	0.04	0	0	0	0	0	0	0.02	0	0	0
6	0.22	0	0.03	0.05	0.08	0.12	0.2	0.13	0.09	0.07	0	0	0	0	0	0	0.04	0	0	0
7	0.13	0	0	0.03	0.05	0.08	0.12	0.21	0.15	0.12	0	0	0	0	0	0	0.02	0	0	0
8	0.07	0	0	0.02	0.03	0.05	0.08	0.13	0.24	0.19	0	0	0	0	0	0	0	0	0	0
9	0.04	0	0	0.01	0.02	0.03	0.05	0.08	0.15	0.32	0	0	0	0	0	0	0	0	0	0
10	0	0	0.03	0.01	0.01	0.01	0	0	0	0	1	0	0	0	0	0	0	0	0	0
11	0	0.11	0.03	0.02	0.01	0.01	0.01	0	0	0	0	1	0	0	0	0	0	0	0	0
12	0	0	0.06	0.04	0.02	0.01	0.01	0.01	0	0	0	0	1	0	0	0	0	0	0	0
13	0.01	0	0.03	0.07	0.04	0.03	0.02	0.01	0	0	0	0	0	1	0	0	0	0	0	0
14	0.01	0	0.03	0.04	0.07	0.04	0.03	0.02	0.01	0	0	0	0	0	1	0	0	0	0	0
15	0.04	0	0	0.02	0.04	0.06	0.04	0.03	0.02	0.01	0	0	0	0	0	1	0	0	0	0
16	0.07	0	0	0.01	0.02	0.04	0.07	0.04	0.03	0.01	0	0	0	0	0	0	0.92	0	0	0
17	0.04	0	0	0.01	0.01	0.03	0.04	0.07	0.05	0.03	0	0	0	0	0	0	0	1	0	0
18	0.01	0	0	0	0.01	0.01	0.03	0.04	0.08	0.06	0	0	0	0	0	0	0	0	1	0
19	0.01	0	0	0	0	0.01	0.02	0.03	0.05	0.1	0	0	0	0	0	0	0	0	0	1
Total	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Exhibit 8. A Markov transition matrix for the transition from 24 to 36 months. Note that each column sums to one.

With this matrix populated, it is possible to observe structural details of loss development. To do this, we consider four different types of transitions, open to open, open to closed, closed to closed, and closed to open. This corresponds to the 4 quadrants of the transition matrix. The section of the matrix that shows open to open is shown in exhibit 9. Here we see that transitions with no size change (same initial and final class) has the highest probability. This forms a diagonal ridge across the matrix. Note that the columns do not sum to one because some of the probability is carried in the part of the matrix representing the open to closed transitions which is shown in Exhibit 10. This shows a similar diagonal ridge which represents claims that close in the same size range that they were open at the beginning of the transition.

Transition from 24 to 36 months										
Final Class	Initial Class									
	0	1	2	3	4	5	6	7	8	9
0	0	0.22	0.09	0.05	0.03	0.02	0.01	0.01	0	0
1	0.01	0.33	0.15	0.08	0.05	0.03	0.02	0.01	0.01	0
2	0.03	0.22	0.24	0.13	0.08	0.05	0.03	0.02	0.01	0
3	0.04	0.11	0.15	0.22	0.12	0.08	0.05	0.03	0.02	0.01
4	0.07	0	0.09	0.13	0.2	0.12	0.08	0.05	0.03	0.03
5	0.13	0	0.06	0.08	0.12	0.19	0.12	0.08	0.05	0.04
6	0.22	0	0.03	0.05	0.08	0.12	0.2	0.13	0.09	0.07
7	0.13	0	0	0.03	0.05	0.08	0.12	0.21	0.15	0.12
8	0.07	0	0	0.02	0.03	0.05	0.08	0.13	0.24	0.19
9	0.04	0	0	0.01	0.02	0.03	0.05	0.08	0.15	0.32

Exhibit 9 Transition Matrix for open to open losses.

Transition from 24 to 36 months										
Initial Class	Final Class									
	10	11	12	13	14	15	16	17	18	19
0	0	0	0	0.01	0.01	0.04	0.07	0.04	0.01	0.01
1	0	0.11	0	0	0	0	0	0	0	0
2	0.03	0.03	0.06	0.03	0.03	0	0	0	0	0
3	0.01	0.02	0.04	0.07	0.04	0.02	0.01	0.01	0	0
4	0.01	0.01	0.02	0.04	0.07	0.04	0.02	0.01	0.01	0
5	0.01	0.01	0.01	0.03	0.04	0.06	0.04	0.03	0.01	0.01
6	0	0.01	0.01	0.02	0.03	0.04	0.07	0.04	0.03	0.02
7	0	0	0.01	0.01	0.02	0.03	0.04	0.07	0.04	0.03
8	0	0	0	0	0.01	0.02	0.03	0.05	0.08	0.05
9	0	0	0	0	0	0.01	0.01	0.03	0.06	0.1

Exhibit 10 Transition Matrix for open to closed losses.

A third type of transition to be considered is the closed to closed transition as shown in exhibit 11. This looks as expected, where, the transitions with the same initial and final size form a 100 percent ridge forming a diagonal across the page. There is one exception in

Transition from 24 to 36 months										
Final Class	Initial Class									
	10	11	12	13	14	15	16	17	18	19
10	1	0	0	0	0	0	0	0	0	0
11	0	1	0	0	0	0	0	0	0	0
12	0	0	1	0	0	0	0	0	0	0
13	0	0	0	1	0	0	0	0	0	0
14	0	0	0	0	1	0	0	0	0	0
15	0	0	0	0	0	1	0	0	0	0
16	0	0	0	0	0	0	0.92	0	0	0
17	0	0	0	0	0	0	0	1	0	0
18	0	0	0	0	0	0	0	0	1	0
19	0	0	0	0	0	0	0	0	0	1

Exhibit 11 Transition Matrix for close to closed losses

our example, in the initial class of 16 where the probability is less then 100 percent. The rest of the probability is carried in the fourth type of transition the closed to open transitions shown in exhibit 12. Although this type of transition is rather rare, they are shown to illustrate that this portion of the matrix can contain real data and should not be ignored. This quadrant of the matrix must exist in order to accommodate the few claims that may fall into it. Otherwise, when coding to process data, errors can appear.

Transition from 24 to 36 months										
Final Class	Initial Class									
	10	11	12	13	14	15	16	17	18	19
0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0.02	0	0	0
6	0	0	0	0	0	0	0.04	0	0	0
7	0	0	0	0	0	0	0.02	0	0	0
8	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0

Exhibit 12 Transition Matrix for close to open losses

A second quantity that needs to be developed is a vector representing the probability of a claim being in a class state at an evaluation. This is performed as follows: One selects the evaluation of interest and assigns class values to all claims based on size and open status at the evaluation. Then, the count for each class is divided by the total number of claims in the evaluation. This will produce a vector of probabilities, an example of which, is shown in exhibit 13.

If we take the square Markov transition matrix for the transition from 24 to 36 months shown in exhibit 8 and multiply it by the 24 month initial vector shown in exhibit 13, the result will be a one dimensional vector that contains the final probabilities at 36 months. The Markov transition matrix chain is then

Initial Values 24 month evaluation		
Class	Claim Count	Prob.
0	67	0.016
1	9	0.002
2	33	0.008
3	102	0.024
4	286	0.067
5	755	0.177
6	1905	0.447
7	696	0.163
8	237	0.056
9	69	0.016
10	0	0.000
11	0	0.000
12	1	0.000
13	2	0.000
14	6	0.001
15	17	0.004
16	49	0.012
17	17	0.004
18	6	0.001
19	2	0.000
Total	4259	1.000

Initial Vector for Matrix multiplicatio
0.016
0.002
0.008
0.024
0.067
0.177
0.447
0.163
0.056
0.016
0.000
0.000
0.000
0.000
0.001
0.004
0.012
0.004
0.001
0.000

Exhibit 13. A vector of initial probabilities for 24 month evaluation.

established by using this final at 36 months vector, and using it as the initial vector at 36 months and multiplying the 36 to 48 month transition matrix by it resulting in the final at 48 month probability vector. This process is continued in maturity order until the oldest transition matrix is used. The last transition matrix may have to be judgmentally adjusted so that all claims are closed after it is used. This is accomplished by using a matrix where the first and last quadrants, exhibits 9 and 12 contain all zeros, and 100% of the probability is in the other two quadrants, exhibits 10 and 11.

We now have all the tools necessary to evaluate loss development by transition matrix theory. Let us consider the ultimate loss development of an individual claim. For an example, let us select a $2,000,000 claim that is open at 24 months. Then, we raise the question, what does the ultimate development for this claim look like? To arrive at the answer we simply place this claim in an initial vector, and we multiply this vector by all the transition matrices in maturity sequence forming a Markov transition chain.

Initial Values 24 month evaluation		
Class	Claim Count	Prob.
0	0	0
1	0	0
2	0	0
3	0	0
4	0	0
5	1	1
6	0	0
7	0	0
8	0	0
9	0	0
10	0	0
11	0	0
12	0	0
13	0	0
14	0	0
15	0	0
16	0	0
17	0	0
18	0	0
19	0	0
Total	1	1

Exhibit 14. Initial vector for a claim of $2,000,000 size and a maturity of 24 months.

The initial vector for this is a special case where all the probability is concentrated in one class, and our example is shown in exhibit 14. When this vector is multiplied by the transition matrix for the 24 to 36 month transition, the result is a final value vector which contains the contents of the initial class 5 column of the transition matrix, as shown in exhibit 15.

Final Value 36 Month Maturity	
Class	Prob.
0	0.017
1	0.029
2	0.046
3	0.075
4	0.121
5	0.193
6	0.121
7	0.075
8	0.046
9	0.029
10	0.005
11	0.009
12	0.015
13	0.025
14	0.040
15	0.064
16	0.040
17	0.025
18	0.015
19	0.009
Total	0.000

Exhibit 15. Final vector for a claim of $2,000,000 size and a maturity of 24 months at 36 months.

This final at 36 month vector serves as the initial vector for 36 months to multiply with the transition matrix for 36 to 48 months forming the next step in the Markov chain. This process is repeated until all of the transition matrices are used. The final value vector that results is the ultimate loss development for the claim in our example. In order to illustrate this example, artificial data in transition matrices is used. The results of these various transitions are shown in a graph in Exhibit 16. One obvious feature of this graph is the two distinct peaks. The first one is in the open claims range, (Classes 0 to 9) and the second one in the closed range, (Classes 10 to 19) This shows that as claims mature, the peak decreases on the left, and increases on the right, corresponding with a decrease in open claims and an increase in closed claims. The "Final" line, indicated by triangles, shows a peak, centered around class 15 (the same size as our starting size class 5) and classes 0 through 9 have zero probability signifying there are no open claims. It is interesting to note that there is about 3% probability in class 10 which is the closed, zero size class. This allows us to make a statement about the potential loss development of an open claim. We can say that its ultimate value will be distributed with the probabilities contained in the "Final" line shown on this graph. It has some probability of closing with no payment, and the rest of the probability is distributed with the indicated distribution of size.

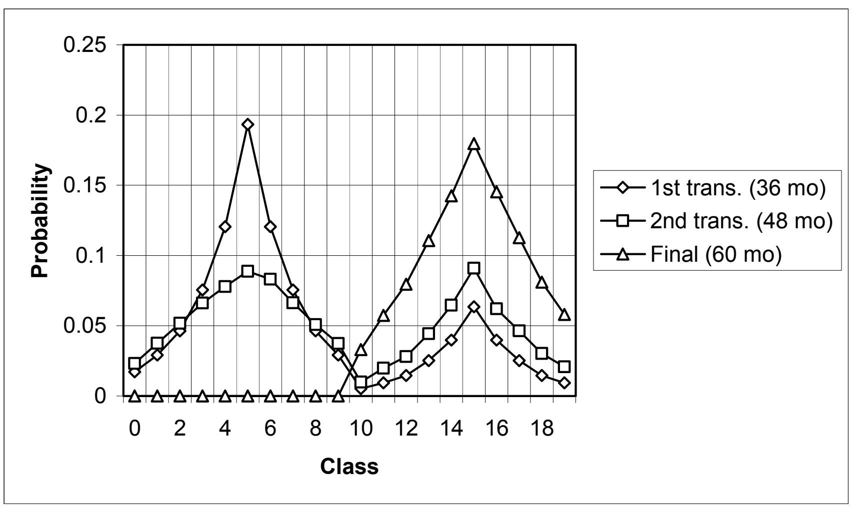

Exhibit 16. Graph showing intermediate and final results for a initial class 5 claim at 24 months maturity.

It is possible to select the points corresponding to the closed with non-zero size on the "Final" line and renormalize the probabilities and fit a distribution. This will be discussed further when the process is applied to real data.

2.3 Transition Matrix Applied to Real Data

A very large database was available for testing out this procedure. This data consists of all the claims that are submitted to a very large reinsurance intermediary for claims processing. The details of processing this data is contained in appendix A.

Class	Lower Limit	Upper Limit	Interval Average	Loss Average in Interval	Interval Count	Cumulative Count	Cumulative Percent	% Exceeding Upper Limit
001	-	5,423	2,712	1,461	2039	2,039	7.3%	92.7%
002	5,423	9,714	7,569	7,422	528	2,567	9.1%	90.9%
003	9,714	17,400	13,557	13,171	735	3,302	11.8%	88.2%
004	17,400	31,168	24,284	23,963	1009	4,311	15.4%	84.6%
005	31,168	55,828	43,498	42,707	1212	5,523	19.7%	80.3%
006	55,828	100,000	77,914	76,343	1850	7,373	26.3%	73.7%
007	100,000	179,121	139,561	137,018	2366	9,739	34.7%	65.3%
008	179,121	320,845	249,983	246,534	2916	12,655	45.1%	54.9%
009	320,845	574,702	447,774	438,115	3266	15,921	56.7%	43.3%
010	574,702	1,029,416	802,059	772,980	3667	19,588	69.8%	30.2%
011	1,029,416	1,843,905	1,436,661	1,364,345	3273	22,861	81.4%	18.6%
012	1,843,905	3,302,830	2,573,368	2,434,316	2071	24,932	88.8%	11.2%
013	3,302,830	5,916,079	4,609,455	4,382,069	1367	26,299	93.7%	6.3%
014	5,916,079	10,596,969	8,256,524	7,850,304	875	27,174	96.8%	3.2%
015	10,596,969	18,981,451	14,789,210	13,711,652	467	27,641	98.5%	1.5%
016	18,981,451	33,999,861	26,490,656	24,389,644	258	27,899	99.4%	0.6%
017	33,999,861	60,901,062	47,450,462	43,658,304	106	28,005	99.8%	0.2%
018	60,901,062	109,086,897	84,993,980	83,690,124	38	28,043	99.9%	0.1%
019	109,086,897	195,398,091	152,242,494	138,646,237	17	28,060	100.0%	0.0%
020	195,398,091	350,000,000	272,699,046	262,891,307	9	28,069	100.0%	0.0%
021	350,000,000	626,923,500	488,461,750	457,301,309	2	28,071	100.0%	0.0%

Exhibit 17. This is the size of loss profile of the claims used in this study after trend and at latest evaluation. Also shown are the interval end-points for defining the size classes.

2.3.1 Data Attributes

It is interesting to explore the size of loss distribution of the claims used in this study. Exhibit 17 contains the loss size distribution of the 28,000 claims in the study as of 2003. The size boundaries in this LSD, at first, appear to be unusual. They were selected to provide 14 intervals between $100,000 and $350,000,000. They were also selected to increase exponentially, and each is 1.79121 times the last one. Here we see about 3/4 of the claims are over $100,000, about 43% are over $574,702, 1/10 are over $3,302,830, and about 1% are over $18,981,451. It would appear that there is enough population in all the size bands to allow a valid study.

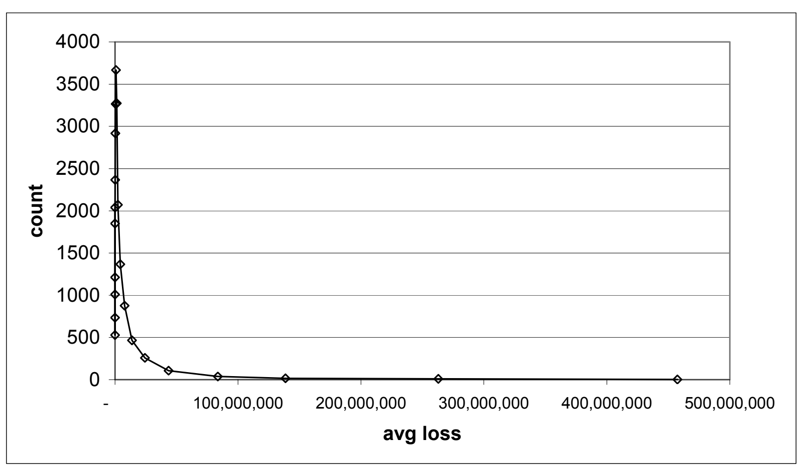

Exhibit 18. Graph of the distribution of the claim sizes used in the study.

2.3.2 Lognormal Behavior

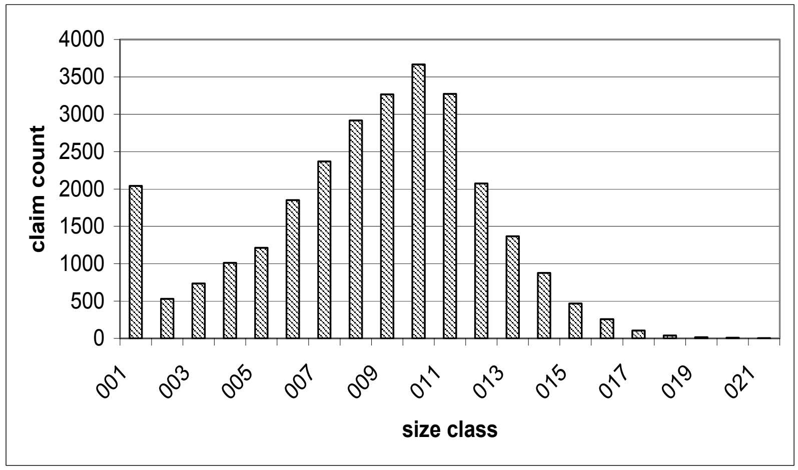

Exhibit 19. Histogram plot of claim counts by size class for losses in study.

Exhibit 18 shows a graph of the size of loss profile. This shows the claim count in asize interval verses the average size of the interval (the average of the upper and lower bound) This view shows a typical heavy tailed distribution with a significant skew to the right, but, one can get little other insight from it. Exhibit 19 shows the claim count by interval, plotted as a histogram of the intervals. This reveals much more about the distribution of the data. We see a bell shaped curve with little skewing left or right except for the elevated first interval. This occurs because the boundaries of the intervals increase exponentially. This behavior suggests that the losses are log normally distributed. Exhibit 20 shows the claim count plotted verses the interval average on a log scale. Again, we see the bell shaped curve with little skewing. The elevated first interval is probably due to the fact that it does not follow the exponential pattern of the other intervals. It contains all losses between 0 and $5,000 which would have been distributed over several intervals had they been defined with narrower (and exponential) boundaries. We find further support for the hypothesis that these losses are distributed log normally when we check the moments of this distribution. If we use the grouped data and take the log of the interval average as the distribution, we get a mean of 12.15, a standard deviation of 2.04, a skewness of -.4 and a kurtosis of 2.9. The last two are of particular interest as a distribution with a skewness between -0.5 and 0.5 is considered to be symmetrical and a normal distribution has a kurtosis of 3. This suggests that the lognormal distribution is consistent with this data.

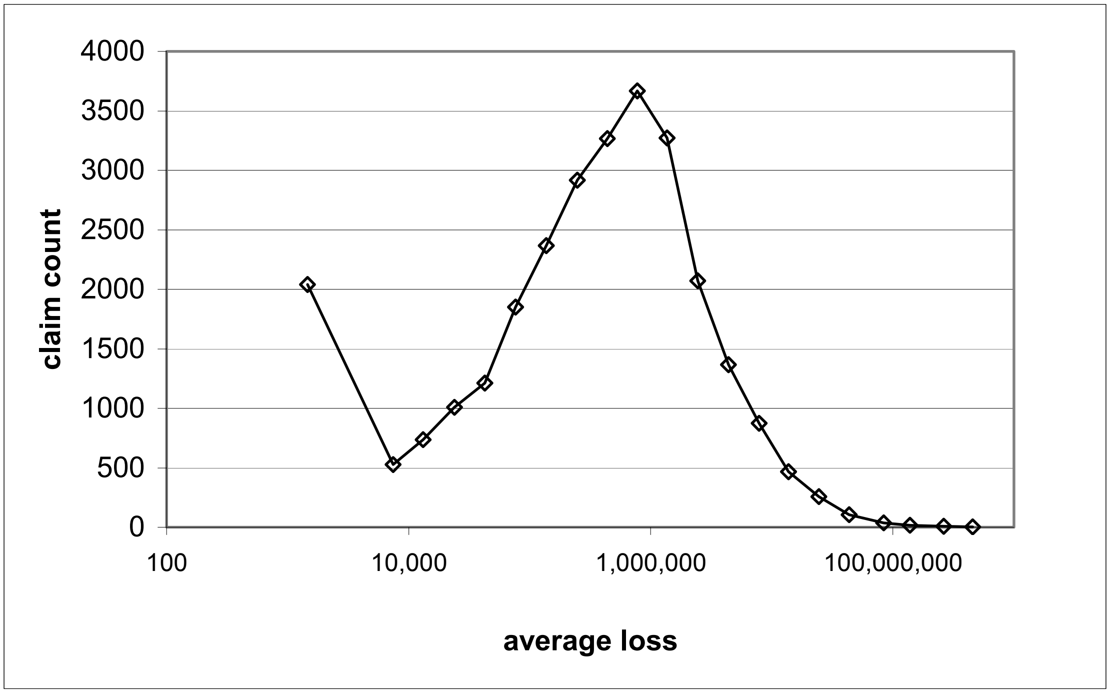

Exhibit 20. Log plot of claim counts of losses used in study.

Final	Initial size category ("O" for open, "C" for closed)																																										
Class	O000	O001	O002	O003	O004	O005	O006	O007	O008	O009	O010	O011	O012	O013	O014	O015	O016	O017	O018	O019	O020	O021	C000	C001	C002	C003	C004	C005	C006	C007	C008	C009	C010	C011	C012	C013	C014	C015	C016	C017	C018	C019	C020	C021
O000	0.00																																										
O001		0.35		0.01	0.01		0.00																																				
O002		0.01	0.34				0.01		0.00																																		
O003		0.00		0.32		0.01																																					
O004		0.00	0.03	0.04	0.40	0.01	0.01						0.01																														
O005		0.01		0.01	0.02	0.39	0.00																																				
O006		0.00		0.01	0.02	0.03	0.42	0.01				0.00																															
O007		0.00	0.02	0.03	0.02	0.05	0.03	0.43	0.01	0.01		0.00																															
O008		0.02		0.01	0.02	0.02	0.03	0.07	0.46	0.02	0.00	0.01																															
O009		0.01		0.02		0.02	0.01	0.03	0.06	0.48	0.04	0.01		0.03																													
O010		0.01		0.03		0.01	0.02	0.03	0.03	0.07	0.51	0.05																															
O011		0.03					0.00	0.02	0.03	0.03	0.05	0.48	0.08	0.03																													
O012		0.01					0.00	0.01	0.01	0.00	0.04	0.08	0.54	0.04	0.02																												
O013		0.00				0.01		0.00		0.00	0.01	0.02	0.05	0.61	0.10																												
O015		0.00						0.00			0.00	0.00	0.01		0.45	0.15																											
O016					0.01			0.00				0.00		0.01	0.02	0.41	0.18																										
O017										0.00				0.01		0.07	0.47																										
O018																	0.06	0.75																									
O019																			0.33																								
O020																				0.67	1.00																						
O021																						1.00																					
C000		0.41	0.39	0.29	0.31	0.16	0.14	0.12	0.12	0.11	0.09	0.11	0.03	0.07	0.17	0.04							1.00																				
C001		0.05	0.08	0.09	0.04	0.03	0.02	0.01	0.03	0.01	0.00	0.01	0.01										1.00																			
C002		0.01	0.02	0.01		0.01	0.01	0.02		0.00	0.00	0.00													1.00																		
C003		0.00	0.05	0.03	0.02	0.01	0.02	0.01		0.00																1.00																	
C004		0.00	0.02	0.01	0.02	0.02	0.00		0.01			0.00															1.00																
C005		0.00		0.01	0.03	0.09	0.03	0.01	0.01	0.00	0.00																1.00															
C006		0.00	0.02	0.02	0.02	0.07	0.10	0.03	0.02	0.01	0.00																	1.00														
C007		0.01	0.02	0.03	0.02	0.01	0.03	0.13	0.02	0.02	0.00	0.00																	1.00													
C008		0.01		0.03	0.02	0.01	0.04	0.03	0.12	0.05	0.02	0.00		0.01																1.00												
C009		0.01		0.01	0.01	0.01	0.01	0.03	0.04	0.11	0.04	0.02	0.01																		1.00											
C010		0.01	0.02	0.01	0.01		0.02	0.01	0.02	0.04	0.14	0.02	0.02																			1.00										
C011		0.00			0.01	0.01	0.00		0.01	0.02	0.03	0.15	0.05	0.01	0.02																		1.00									
C012			0.02						0.00	0.00	0.01	0.02	0.15	0.03	0.02																			1.00								
C013		0.00								0.00	0.00	0.01	0.04	0.11	0.02	0.04	0.06																		1.00							
C014							0.00					0.00	0.01	0.03	0.17	0.15																				1.00						
C015														0.01		0.07																					1.00					
C016		0.00					0.00									0.07	0.18		0.17																				1.00				
C017																	0.06	0.25																					1.00			
C018																			0.50																					1.00		
C019																																										1.00	
C020																				0.33																						1.00
C021																																												1.00

Exhibit 21 An example of a transition matrix for GL data for the transition of 24 to 36 months

Final Size Class	Initial Size Class																					
Class	000	001	002	003	004	005	006	007	008	009	010	011	012	013	014	015	016	017	018	019	020	021
000	0.00	0.80	0.54	0.45	0.42	0.42	0.41	0.36	0.36	0.35	0.33	0.29	0.29	0.27	0.30	0.19	0.10	0.14	0.17	0.06	0.23	
001		0.04	0.10	0.14	0.13	0.09	0.06	0.03	0.02	0.02	0.02	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
002		0.00	0.09	0.02	0.03	0.02	0.02	0.02	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
003		0.01	0.03	0.05	0.05	0.03	0.03	0.02	0.01	0.01	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
004		0.01	0.05	0.07	0.10	0.05	0.04	0.02	0.02	0.02	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
005		0.01	0.05	0.08	0.08	0.14	0.05	0.04	0.02	0.02	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
006		0.01	0.03	0.07	0.06	0.08	0.14	0.06	0.04	0.02	0.02	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
007		0.01	0.04	0.02	0.04	0.06	0.08	0.19	0.08	0.03	0.03	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
008		0.01	0.01	0.03	0.04	0.04	0.06	0.10	0.23	0.08	0.03	0.02	0.01	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00	
009		0.02	0.01	0.01	0.01	0.03	0.04	0.06	0.10	0.24	0.08	0.04	0.02	0.02	0.00	0.01	0.00	0.00	0.00	0.00	0.00	
010		0.02	0.01	0.02	0.02	0.01	0.03	0.05	0.07	0.12	0.28	0.11	0.04	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00	
011		0.03	0.02	0.02	0.02	0.02	0.02	0.03	0.03	0.05	0.13	0.32	0.12	0.05	0.02	0.03	0.00	0.00	0.00	0.00	0.01	
012		0.01	0.01	0.01	0.00	0.00	0.01	0.01	0.01	0.02	0.03	0.09	0.29	0.08	0.04	0.01	0.01	0.01	0.00	0.00	0.01	
013		0.01	0.00	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.04	0.11	0.34	0.11	0.05	0.01	0.02	0.00	0.01	0.05	
014		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.02	0.05	0.13	0.38	0.10	0.04	0.04	0.00	0.02	0.01	
015		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.02	0.05	0.09	0.45	0.15	0.06	0.01	0.08	0.01	
016		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.02	0.03	0.11	0.51	0.16	0.02	0.07	0.01	
017		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.05	0.15	0.55	0.21	0.72	0.04	
018		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.02	0.39	0.03	0.01	
019		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.19	0.00	0.04	
020																					0.57	0.67
021		0.00			0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.00	0.00					0.00	0.33

Exhibit 22. This is an ultimate matrix for the transition from 24 months maturity to ultimate. A claim open at 24 months will have an ultimate size that has a probability distribution described by the probabilities under its initial class size.

2.3.3 Transition Matrix Results

This data was processed through a Markov transition matrix analysis, that produced transition matrices which were multiplied together to yield ultimate matrices. Exhibit 21 shows an example of a transition matrix from this study. This one is for the transition from 24 to 36 months. Note that the classes indicating size and open status have been modified from the earlier example. Open claims are indicated with a class starting with an "O" and closed with a "C". Size ranges from 0 to 20 where 0 is a loss size of $0.00.

2.3.4 Distributional Development

An Ultimate Matrix is shown in exhibit 22. Note that the open and closed status has been collapsed into the closed status. This was accomplished by assuming that the final status of the last set of transitions was always closed. Since, at 23 years more than 95% of the transitions were closed to closed, this is not thought to be an unreasonable assumption. This Ultimate Matrix can be thought of a series of one dimensional vectors stacked next to each other. Each vector provides the prediction of ultimate size based on the initial size. This provides a critical insight. This suggests that it is possible to describe the loss development of an individual open claim. This vector of final possible outcomes provides some probability of closing with no payment, or an array of probabilities of closing at various sizes. We can study the conditional probability given that a claim closes with some payment by removing the probability of closing with no payment into a separate category. After the zero claims are removed, what is left is a probability distribution of final size given an initial size and initial maturity. Exhibit 23 shows the final distribution of a loss that had an initial size category 7 at a maturity of 24 months.

Size Category	Probability
000	0.362
001	0.029
002	0.019
003	0.020
004	0.024
005	0.036
006	0.059
007	0.189
008	0.101
009	0.064
010	0.047
011	0.030
012	0.009
013	0.007
014	0.002
015	0.001
016	0.000
017	0.000
018	0.000
019	0.000
020	0.000
021	0.000

Exhibit 23 - Ultimate size vector for size category 7 at 24 months.

Size Category	Lower Limit	Upper Limit	Average Loss size	Log of avg Loss Size	Probability (class) 007	Normalized Probability	x*prob	x^2*prob	x^3*prob	x^4*prob
000	0	0		0.00	0.3617					
001	0	5,423	2,712	7.91	0.0294	0.0460	0.36	2.88	22.74	179.73
002	5,423	9,714	7,569	8.93	0.0192	0.0301	0.27	2.40	21.48	191.86
003	9,714	17,400	13,557	9.51	0.0204	0.0320	0.30	2.90	27.59	262.51
004	17,400	31,168	24,284	10.10	0.0244	0.0383	0.39	3.90	39.39	397.78
005	31,168	55,828	43,498	10.68	0.0363	0.0568	0.61	6.48	69.21	739.18
006	55,828	100,000	77,914	11.26	0.0591	0.0927	1.04	11.75	132.39	1,491.19
007	100,000	179,121	139,561	11.85	0.1888	0.2958	3.50	41.51	491.77	5,825.66
008	179,121	320,845	249,983	12.43	0.1005	0.1575	1.96	24.33	302.40	3,758.64
009	320,845	574,702	447,774	13.01	0.0642	0.1005	1.31	17.02	221.52	2,882.39
010	574,702	1,029,416	802,059	13.59	0.0469	0.0735	1.00	13.59	184.77	2,512.00
011	1,029,416	1,843,905	1,436,661	14.18	0.0304	0.0476	0.67	9.57	135.68	1,923.60
012	1,843,905	3,302,830	2,573,368	14.76	0.0085	0.0134	0.20	2.91	42.94	633.78
013	3,302,830	5,916,079	4,609,455	15.34	0.0066	0.0103	0.16	2.42	37.08	568.94
014	5,916,079	10,596,969	8,256,524	15.93	0.0021	0.0033	0.05	0.83	13.28	211.51
015	10,596,969	18,981,451	14,789,210	16.51	0.0010	0.0015	0.02	0.41	6.70	110.69
016	18,981,451	33,999,861	26,490,656	17.09	0.0003	0.0005	0.01	0.14	2.42	41.30
017	33,999,861	60,901,062	47,450,462	17.68	0.0001	0.0001	0.00	0.04	0.72	12.74
018	60,901,062	109,086,897	84,993,980	18.26	0.0000	0.0000	0.00	0.01	0.22	4.02
019	109,086,897	195,398,091	152,242,494	18.84	0.0000	0.0000	0.00	0.00	0.03	0.58
020	195,398,091	350,000,000	272,699,046	19.42	0.0000	0.0000	0.00	0.00	0.00	0.00
021	350,000,000	626,923,500	488,461,750	20.01	0.0000	0.0000	0.00	0.01	0.23	4.59
Total					1.000					

mean (mu)	E(x^2)	E(x^3)	E(x^4)
11.86	143.12	1,752.57	21,752.68
	variance	3rd moment	4th moment
	2.39	-1.95	21.95
	std dev(sigma)	skewness	kurtosis
	1.55	-0.53	3.84

initial class	007
average size	139,561
log of avg size	11.85
mu	11.86
sigma	1.55
ratio mu/log(avg)	1.00
ratio sigma/mu	0.13
skewness	-0.53
kurtosis	3.84

Exhibit 24 - Moments of ultimate size vector for Category 7 size at 24 months.

We can apply the same log normal analysis, used previously, to this distribution, and we get the results shown on Exhibit 24. Here we see a mu of 11.86, a sigma of 1.55, a skewness of -0.53, and a kurtosis of 3.84. The skewness close to zero suggests that the ultimate values are lognormally distributed.

It is instructive to divide the final mu and sigma by the log of the initial value. Here we find that the final mean is very nearly equal to the initial value of the loss. Also, the standard deviation is a small fraction of the initial mean. This result suggests the remarkable conclusion. It seems that the loss development potential of a claim is that its ultimate value will be log normally distributed with a mu equal to the natural log of the initial value, and a sigma that can be predicted. Although this seems to be counter intuitive, we must remember that the formula for the mean of a lognormal has the following formula:

$$\text{mean} = e^{(mu+((sigma^2)/2))}$$

which leads to an increase in the mean as sigma grows. So, the resulting average loss at ultimate will be greater then the current evaluation.

2.3.5 Variation Over Initial Size

The next question to investigate is how mu and sigma of the ultimate distribution vary with initial size and maturity. Shown in exhibit 25 are the ultimate distributions for three initial claims sizes, Classes 6, 7, and 8, initially at 24 months maturity. This shows a peak of 25 to 30 percent in the Class 0 (closed with no payment) category. To the right, each curve has a distinctive bell shaped curve that peaks in the final size category that is the same as the initial size category. Each curve looks symmetrical, and, remarkably like each other.

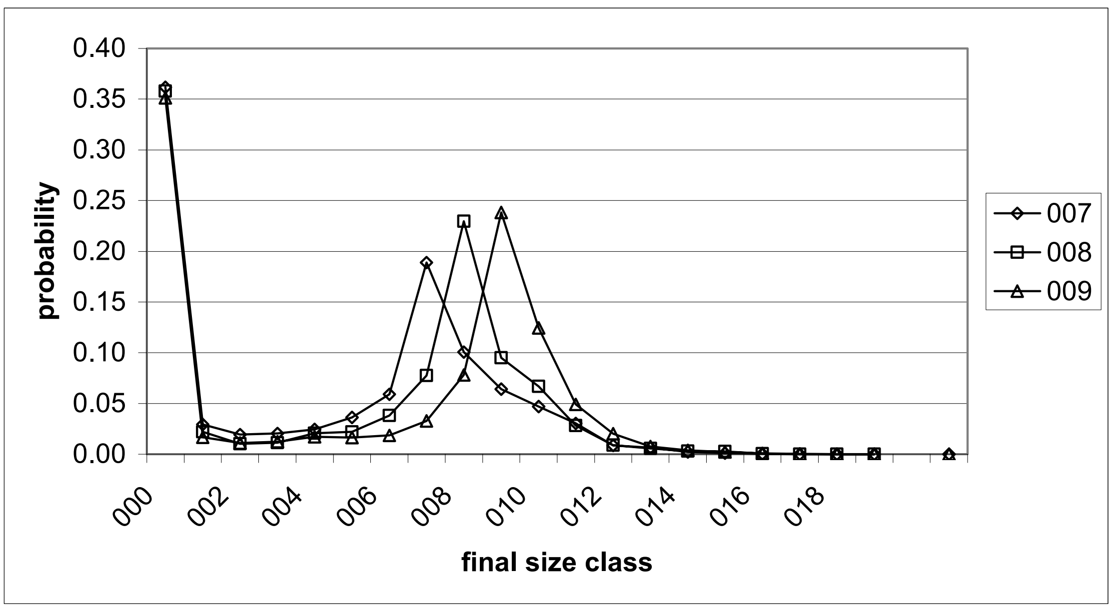

Exhibit 25 Graph of final distribution of 24 month initial open claims with intital sizes of Class 6, Class 7, and Class 8

The graphical appearance of this data alone suggests a possible behavior where the final size is related to the initial size, and the spread of the distribution does not vary with initial size.

Exhibit 26 is similar to the previous graph with a wider range of initial sizes. This covers size classes 7 to 14. The next exhibit is a graph of all the ultimate distributions for the 24month to ultimate transition. Each curve is for a different initial size class. They all demonstrate the bell shaped appearance seen previously, suggesting a lognormal distribution.

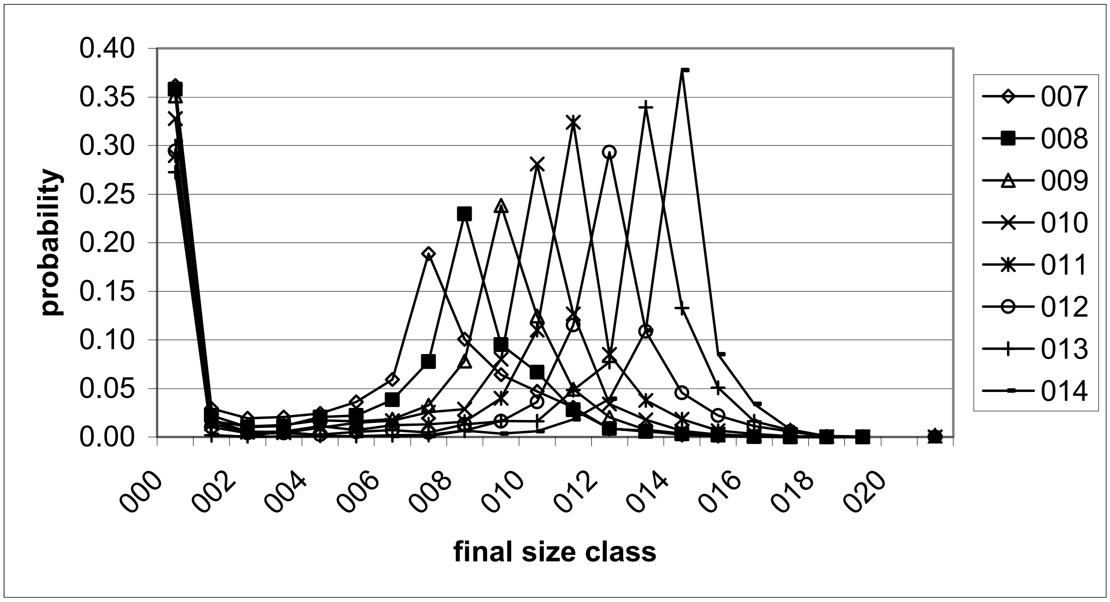

Exhibit 26 Graph of final distribution of 24 month initial open claims with intital sizes of Class 4 to Class 11.

The striking aspect of this graph is how much the bell shaped curves resemble each other in the range of size class 5 to 10. At first glance, these appear as identical curves which are offset from each other by a constant amount. Since the boundaries of the classes were defined with a multiplicative factor, the constant spacing occurs because of the logarithmic nature of the scale, and the fact that the mu of each distribution has a relation to the initial loss size. This relationship is that, the mu is nearly the same as the logarithm of the original loss size. The fact that the shape of the curves do not change as they progress from left to right suggest that the spread parameters are very similar for all the curves.

initial class	average class size	log of avg size	Parms of ult dists. mu	sigma	Skewness	ratio mu/log(avg)	ratio sigma/mu
001	2,712	7.91	11.93	2.57	-0.34	1.51	0.22
002	7,569	8.93	10.16	1.95	0.78	1.14	0.19
003	13,557	9.51	10.37	1.97	0.45	1.09	0.19
004	24,284	10.10	10.49	2.00	0.62	1.04	0.19
005	43,498	10.68	10.84	1.84	0.21	1.01	0.17
006	77,914	11.26	11.24	1.77	-0.04	1.00	0.16
007	139,561	11.85	11.86	1.55	-0.53	1.00	0.13
008	249,983	12.43	12.27	1.43	-0.90	0.99	0.12
009	447,774	13.01	12.75	1.45	-1.31	0.98	0.11
010	802,059	13.59	13.28	1.43	-1.76	0.98	0.11
011	1,436,661	14.18	13.81	1.46	-1.94	0.97	0.11
012	2,573,368	14.76	14.54	1.44	-2.16	0.98	0.10
013	4,609,455	15.34	15.25	1.14	-2.07	0.99	0.08
014	8,256,524	15.93	15.75	0.95	-3.14	0.99	0.06

Exhibit 27 - Parameters for ultimate distributions by size class for 24 month initial losses

Exhibit 27 shows the mu, sigma, skewness, and kurtosis of all the 24 month to ultimate distributions by initial size classification. Here we see that the mu's are all very close to the log of the initial, and the sigma values are all very similar. In exhibit 28 we graph the mu's and the ratio of mu to the natural log of the average loss in the interval. This shows a strong relationship between these two values.

The sigma values shown in exhibit 27, and graphed in exhibit 29, show a gradual decrease as the loss size increases. It may be possible to find a relationship between the loss size and sigma. It appears that a linear relationship between sigma and the natural log of the initial loss may describe the behavior of sigma.

The sknewness values shown in exhibit 27, and, graphed in exhibit 30, are positive for the small losses and negative for the large losses, and close to zero for the mid size losses. This is a relatively complex behavior but it can be understood based on the nature of the reinsurance claims that constitute the data.

Remember that these are only claims that are submitted for reinsurance recoveries. The full inventory of claims are not represented here. The positive skewing of the smaller claims can be understood as being caused by the submission of small claims that are expected to

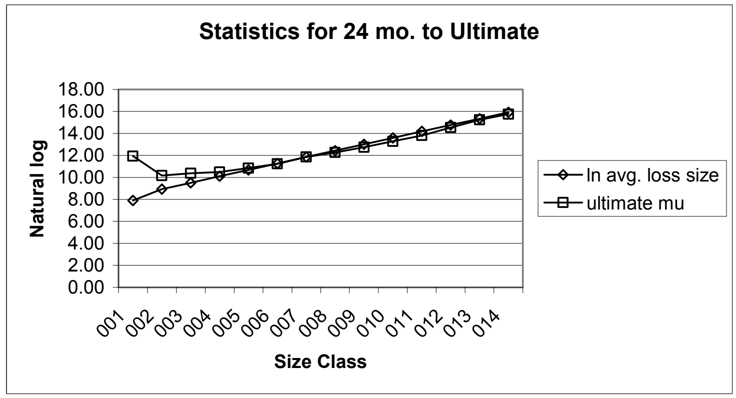

Exhibit 28 - This graph shows the natural log of the average loss size, and the mu of the lognormal distribution of the ultimate loss size distribution for claims with a current maturity of 24 months.

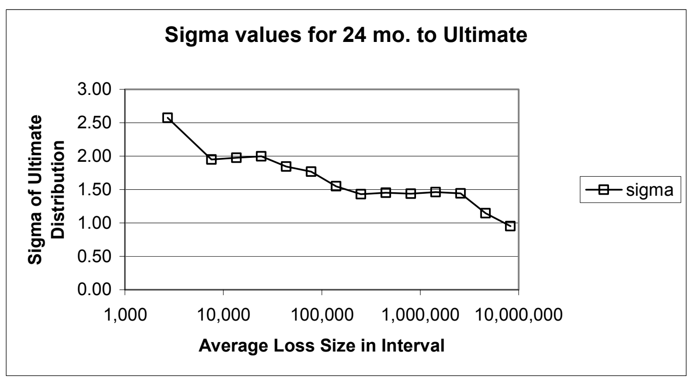

Exhibit 29 - Graph showing the sigma values for the lognormal distributions for the ultimate values for claims at 24 months.

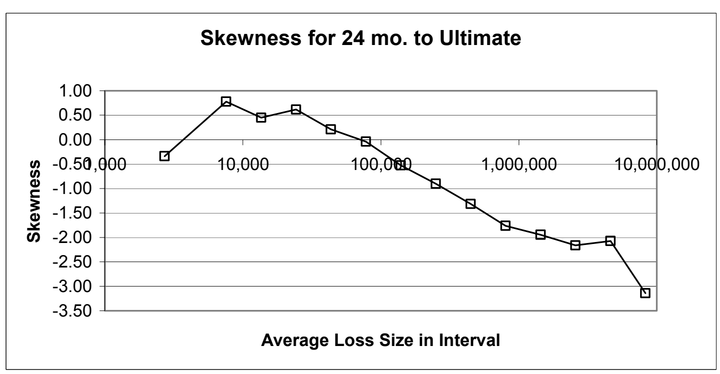

Exhibit 30 - Graph showing the skewness values for the lognormal distributions for the ultimate values for claims at 24 months.

settle are larger amounts and enjoy a reinsurance recovery.

The larger claims may skew negatively because they are large enough to feel the effect of policy limits. As large claims settle, they are always free to settle at smaller then currently reserved amounts, but, any tendency to settle at higher values may be limited when the policy limit is reached.

Thus, we can understand the appearance of the positive skewing of the small claims and the negative skewing of the large claims as being due to data reporting and policy limit effects and is not an essential element of the loss development. The use of a lognormal distribution to describe the ultimate development of an individual claim continues to be consistent with the observations.

At this point, there is enough evidence to postulate a model for ultimate loss development for open claims at 24 months maturity. A lognormal distribution with a mu equal to the natural log of the open claim size, and a sigma which is described by a linear relationship between sigma and the natural log of the open claim size is consistent with the current observations.

One must remember that this is a conditional distribution, based on the condition that the claim does not close with no payment. We must remember that the transition matrix process contains an ultimate size category 000 which contains a significant number of claims that close with no payment. One needs only to refer back to exhibit 23, and pick the value from the first line under the correct initial loss size, to get the probability that the claim closes with no payment.

Another consideration in the use of this model is that the primary policy limit distribution must be applied after the loss development is applied.

Size Category	Average Size in Interval	Ratio of Mu/natural log of average size							
		Maturity in Months							
		24	36	48	60	72	84	96	108
001	2,712	1.51	1.52	1.48	1.46	1.44	1.46	1.43	1.44
002	7,569	1.14	1.15	1.14	1.18	1.11	1.10	1.17	0.99
003	13,557	1.09	1.11	1.08	1.06	1.07	1.01	1.05	0.99
004	24,284	1.04	1.02	1.04	1.05	1.06	0.99	0.98	0.96
005	43,498	1.01	1.00	1.00	0.98	0.98	0.95	0.96	0.99
006	77,914	1.00	0.99	0.98	0.99	0.99	0.99	1.00	0.99
007	139,561	1.00	0.98	0.96	0.98	0.98	0.97	0.97	0.97
008	249,983	0.99	0.98	0.98	0.98	0.97	0.97	0.98	0.98
009	447,774	0.98	0.98	0.98	0.97	0.98	0.98	0.98	0.99
010	802,059	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98
011	1,436,661	0.97	0.98	0.98	0.98	0.98	0.98	0.98	0.98
012	2,573,368	0.98	0.99	0.99	0.98	0.99	0.99	0.99	0.99

Exhibit 31 - This shows the ratio of the mu of the ultimate distribution divided by the natural log of average size in the initial size category.

2.3.6 Exploration over Maturities

Thus far, we have explored loss development behavior at one selected initial maturity. In order to create a system that can accommodate a complete selection of real data, we need to be able to describe this process for losses over the entire range of maturities. In order to do this, we first look at the comparison of the modeled mu to the natural log of the average loss size in each initial size category. We are looking at the ability of the initial loss size to forecast the mu of the ultimate distribution. Shown in exhibit 31 is the ratio of the mu of the ultimate distribution divided by the natural log of the average loss within the size category. If the forecast of mu is perfect, then this ratio should be 1.00, according to the postulated model. What we find is a relative flat surface except for the turned edge, as shown in exhibit 32.

The turned edge may be caused by the origin of the data. Since it is a collection of claims

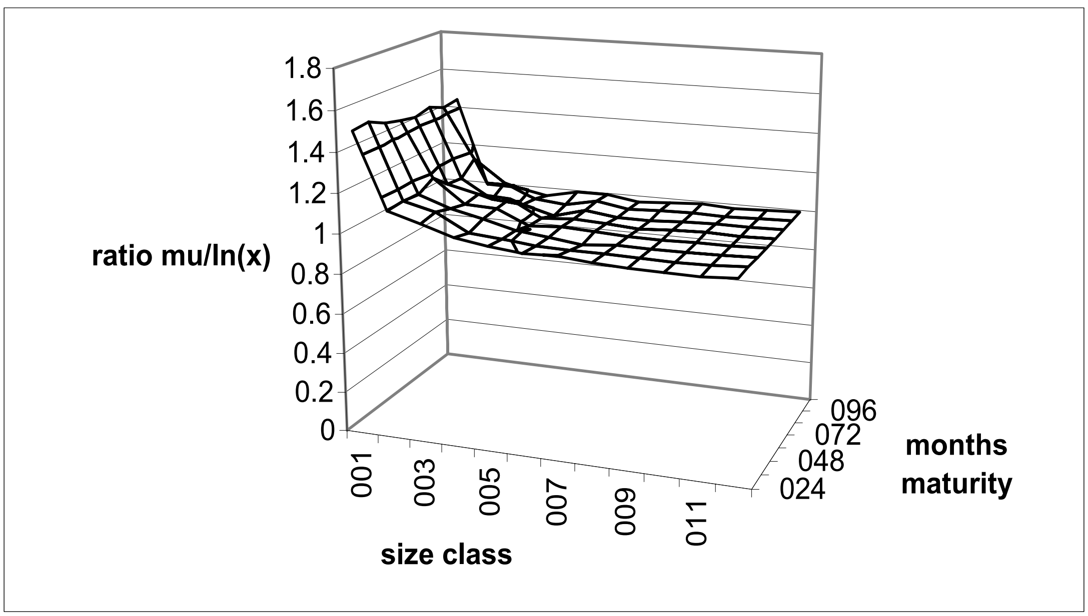

Exhibit 32 - Surface of ultimate mu / ln(avg loss size) over the range of size categories and months maturity of initial size observation.

that anticipate a reinsurance collection, it may be biased to develop larger. A more complete collection of claims may not have this bias. This suggests that the natural log of the current value of an open claim is a good predictor of the mu for the lognormal distribution describing the ultimate loss size.

If we conclude that this surface is a plane, and we ignore the first size category (001), all the remaining points have an average of 1.005 and a standard deviation of 0.05 Our model for estimating mu then becomes mu= 1.005 * ln(x), where x is the current open claim size. If one is not satisfied with the accuracy of applying one number over all size, maturity combinations, one could interpolate over the surface given in exhibit 31 and apply interpolated numbers to individual claims.

2.3.6 The Sigma Surface

We can also review the fitted sigma values as they vary by initial loss size and by initial maturity to see if a pattern emerges. These values are shown in exhibit 33, and graphed as a

Initial Size Class	Sigma at Ultimate Initial Maturity in Months									
	012	024	036	048	060	072	084	096	108	120
001	2.668	2.572	2.653	2.809	2.803	2.682	2.604	2.512	2.530	2.708
002	2.075	1.950	1.847	1.917	2.414	2.313	2.392	2.714	1.727	1.892
003	2.071	1.973	2.159	2.257	2.423	2.344	2.115	2.367	1.936	1.712
004	2.011	1.997	1.995	2.000	2.223	2.382	2.008	2.048	1.948	2.217
005	1.734	1.844	1.800	1.765	1.867	1.919	1.793	1.849	1.733	1.785
006	1.795	1.766	1.766	1.811	1.866	1.746	1.655	1.632	1.838	1.519
007	1.532	1.547	1.578	1.575	1.510	1.599	1.580	1.338	1.441	1.211
008	1.544	1.430	1.397	1.369	1.353	1.413	1.389	1.353	1.306	1.172
009	1.426	1.452	1.370	1.374	1.317	1.261	1.247	1.161	1.134	1.117
010	1.363	1.435	1.352	1.260	1.230	1.198	1.138	1.195	1.111	1.050
011	1.516	1.460	1.309	1.271	1.207	1.143	1.035	1.019	0.961	0.943
012	1.444	1.444	1.183	1.184	1.105	1.011	1.029	0.984	0.867	0.861
013	1.144	1.144	1.072	1.029	1.027	0.999	1.033	0.896	0.849	0.894
014	0.951	0.951	1.013	1.136	1.222	1.179	1.181	1.125	0.925	0.859
015	0.982	0.982	0.956	1.051	0.968	0.954	0.998	1.056	0.993	1.127

Exhibit 33 - These are the sigma's of the ultimate distributions
by initial size class and by initial maturity

surface in exhibit 34. This data appears to have some structure associated with it. There is a pronounced decrease in sigma as the size of the loss increases, and there is a modest decrease as the maturity of the claim increases. To further understand how sigma varies, it is instructive to graph it as one of the other variables change. First we look as the size changes.

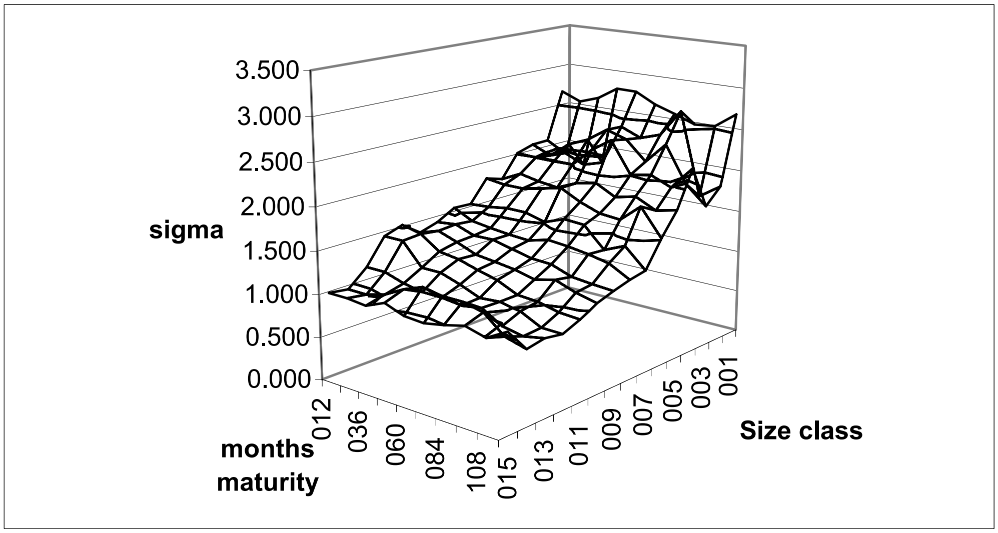

Exhibit 34 - 3D surface map of the sigma's of the ultimate distributions plotted verses maturity and size of current open claim.

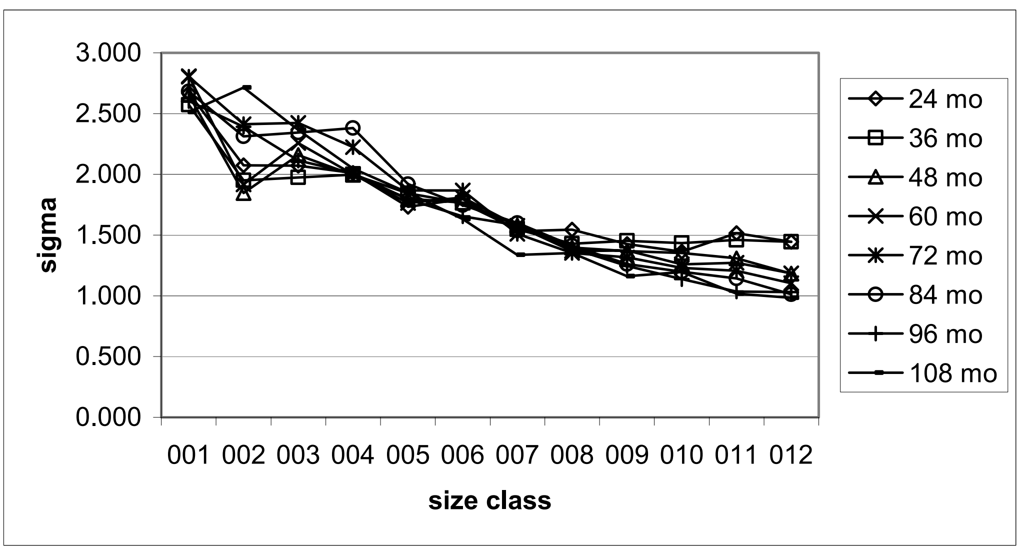

Exhibit 35 - Graph showing the variation of sigma with the size of the claim.

In exhibit 35, sigma is plotted as the size changes, and, each line represents a different maturity. This shows the decrease over time, which becomes more gradual as time progresses. If we assume that there is no structure in the maturity direction and all the variation is noise, we can average each size evaluation and plot the result. This is shown in exhibit 36. This shows more clearly the slowing of the decrease over time.

Exhibit 36 - Sigma's averaged over maturities and plotted verses size class.

It was found that the inverse of sigma behaves in a more orderly fashion. This variable appears to be linear when plotted against size class. The plot of the inverse of the average sigma is shown in exhibit 37. Here we see it is increasing at a constant rate. This variable has an additional benefit in that it behaves well at its extremes. At very large class sizes sigma becomes small, which is a believable result, and at very small class sizes it becomes very large, and then undefined. This is acceptable because there is no interest in modeling very, very small claims.

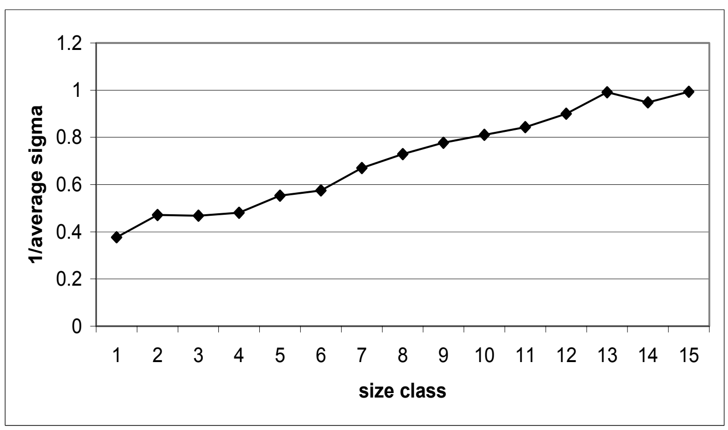

Exhibit 37 - Plot of 1/(avg. sigma) verses loss size.

The favorable behavior of this variable encourages us to explore the behavior of this transformed variable verses maturity. Exhibit 38 shows the plot of the inverse of sigma verses maturity, where, each line represents a size class. The overall impression is there is a bit of increase as maturity progresses. We average over the sizes and produce one value for each maturity and check to see if this varies as maturity. These values are plotted in exhibit 40. This shows a slight increase with maturity. There is nothing to suggest that this increase is more complicated than a first order linear effect.

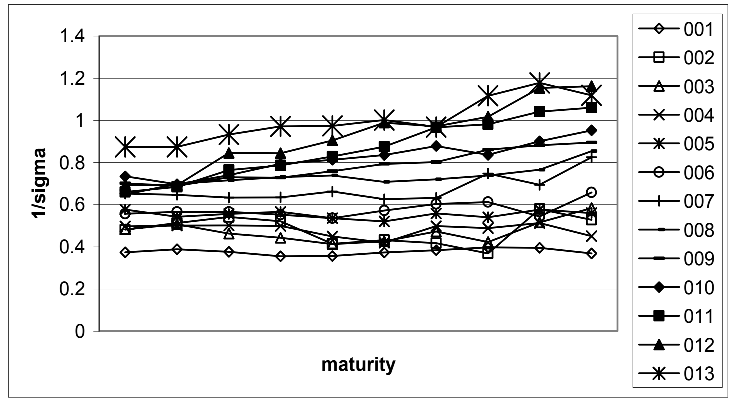

Exhibit 38 - Plot of 1/sigma verses maturity. Each line represents a different size class.

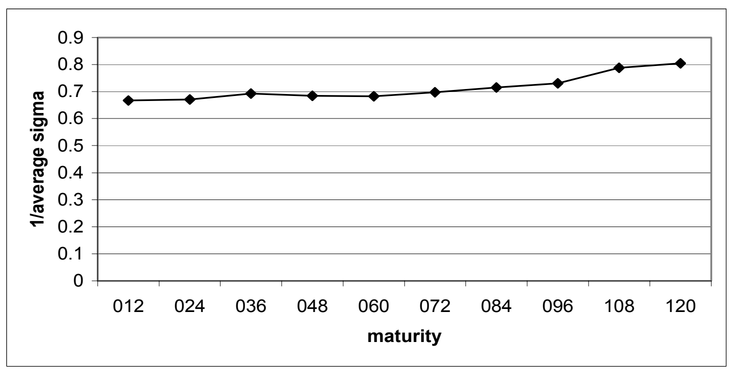

Exhibit 39 Plot of average of 1/sigma verses maturity.

2.3.7 Fitting the Sigma Surface

These two observations of linear behavior of 1/sigma verses maturity and verses loss size suggests that a linear fit to the surface will allow use

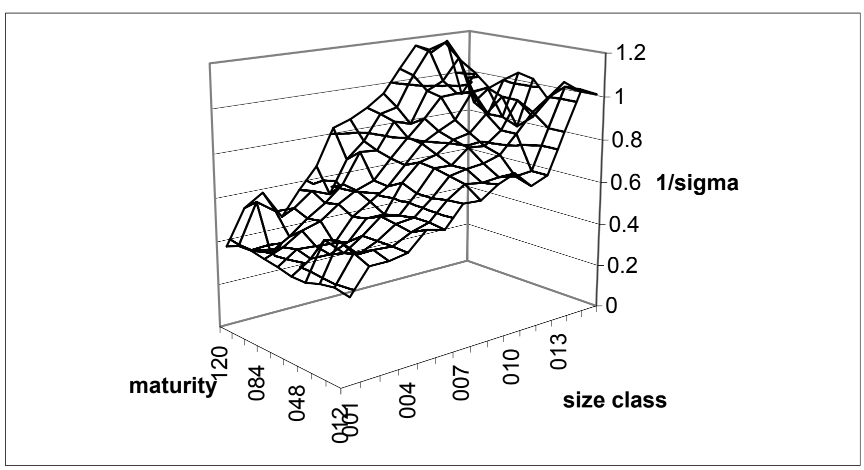

Exhibit 40 - Plot of 1/sigma verses maturity and size class.

to model sigma as a function of size and maturity. The sigma values in exhibit 33 are inverted and are graphed in exhibit 40. These values are analyzed by general linear regression against the dependent variables of, maturity in months, and, natural log of the average interval claims size. This resulted in a fitted regression of:

$$\text{Sigma} = 1/(\text{maturity} * 0.001205 + \ln(\text{loss size}) * 0.078874 - 0.34447)$$

A table of fitted sigma values are shown in exhibit 41, and these values are plotted in exhibit 42.

initial size class	initial maturity									
	012	024	036	048	060	072	084	096	108	120
001	3.41	3.25	3.10	2.97	2.85	2.73	2.63	2.53	2.44	2.36
002	2.67	2.57	2.48	2.39	2.31	2.24	2.17	2.10	2.04	1.98
003	2.38	2.30	2.23	2.16	2.09	2.03	1.97	1.92	1.87	1.82
004	2.14	2.08	2.02	1.96	1.91	1.86	1.81	1.76	1.72	1.68
005	1.95	1.90	1.85	1.80	1.75	1.71	1.67	1.63	1.59	1.56
006	1.79	1.75	1.70	1.66	1.62	1.59	1.55	1.52	1.48	1.45
007	1.65	1.62	1.58	1.54	1.51	1.48	1.45	1.42	1.39	1.36
008	1.54	1.50	1.47	1.44	1.41	1.38	1.36	1.33	1.31	1.28
009	1.44	1.41	1.38	1.35	1.33	1.30	1.28	1.25	1.23	1.21
010	1.35	1.32	1.30	1.27	1.25	1.23	1.21	1.19	1.17	1.15
011	1.27	1.25	1.22	1.20	1.18	1.16	1.14	1.12	1.11	1.09
012	1.18	1.18	1.16	1.14	1.12	1.10	1.09	1.07	1.05	1.04
013	1.12	1.12	1.10	1.08	1.07	1.05	1.03	1.02	1.00	0.99
014	1.06	1.06	1.05	1.03	1.02	1.00	0.99	0.97	0.96	0.95
015	1.01	1.01	1.00	0.98	0.97	0.96	0.94	0.93	0.92	0.91

Exhibit 41 - Table of fitted sigma's.

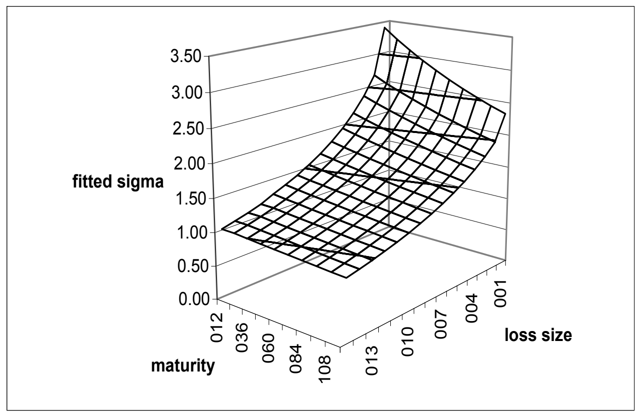

Exhibit 42 - Plot of fitted sigma's. Compare this with exhibit 34.

The error values for this regression are shown in exhibit 43. These values are the fitted sigma values minus the actual sigma values observed at each maturity and size class. The average of all these values is 0.00. These values are graphed in exhibit 44.

The only structure revealed in this graph is a sharp rise at early maturities and small size classes. This model tends to overestimate sigma in this region, but, since there is little interest in this region, this is an acceptable error. For those who need high accuracy in this area, it would be best to interpolate values directly from exhibit 33.

A review of the skewness as size and maturity is varied shows the same tendencies as noted earlier, positive skew for small losses and negative skew for large losses. The skewness values are shown in exhibit 45 and the surface is shown in exhibit 46. It may well be that the negative skewness and the decrease in sigma for large claims is caused by policy limit censoring. It is a long held view that small claims tend to develop larger, and large claims tend to develop smaller. This evidence certainly supports that view. One might be concerned that the proposed model will overdevelop large claims. One should reexamine exhibit 26 and observe that the graphed distributions for the large initial size classes

size class	Error values of sigma regression initial maturity 012	024	036	048	060	072	084	096	108	120
001	0.74	0.67	0.45	0.16	0.04	0.05	0.03	0.02	-0.09	-0.35
002	0.60	0.62	0.63	0.48	-0.10	-0.07	-0.22	-0.61	0.31	0.09
003	0.31	0.33	0.07	-0.10	-0.33	-0.31	-0.14	-0.45	-0.07	0.10
004	0.13	0.08	0.02	-0.04	-0.32	-0.53	-0.20	-0.29	-0.23	-0.54
005	0.22	0.05	0.05	0.03	-0.11	-0.21	-0.12	-0.22	-0.14	-0.23
006	0.00	-0.02	-0.06	-0.15	-0.24	-0.16	-0.11	-0.12	-0.35	-0.07
007	0.12	0.07	0.00	-0.03	0.00	-0.12	-0.13	0.08	-0.05	0.15
008	-0.01	0.07	0.08	0.07	0.06	-0.03	-0.03	-0.02	0.00	0.11
009	0.01	-0.05	0.01	-0.02	0.01	0.04	0.03	0.09	0.10	0.09
010	-0.02	-0.11	-0.06	0.01	0.02	0.03	0.07	-0.01	0.05	0.10
011	-0.25	-0.21	-0.08	-0.07	-0.02	0.02	0.11	0.11	0.15	0.15
012	-0.27	-0.27	-0.02	-0.04	0.02	0.09	0.06	0.09	0.19	0.18
013	-0.03	-0.03	0.03	0.05	0.04	0.05	0.00	0.12	0.15	0.10
014	0.11	0.11	0.03	-0.10	-0.21	-0.18	-0.19	-0.15	0.03	0.09
015	0.03	0.03	0.04	-0.07	0.00	0.00	-0.05	-0.12	-0.07	-0.22

Exhibit 43 - This table contains the error values of the regression for sigma. The values shown are (fitted sigma - actual sigma).

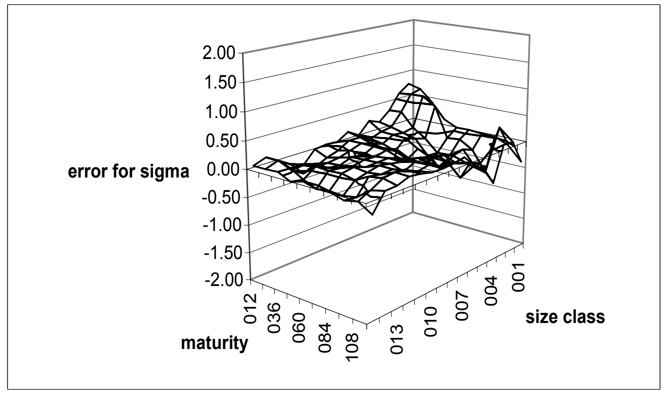

Exhibit 44 Plot of error values of regression for sigma. Vertical scale is : (modeled sigma - actual sigma)

Initial Size class	average size in interval	Skewness Initial Maturity in Months 024	036	048	060	072	084	096	108	120
001	2,712	-0.34	-0.34	-0.15	-0.09	-0.17	-0.30	-0.21	-0.15	-0.07
002	7,569	0.78	0.86	1.05	0.54	1.12	1.29	0.95	2.67	2.46
003	13,557	0.45	0.56	0.61	0.85	0.81	1.27	0.94	1.32	1.70
004	24,284	0.62	0.57	0.65	0.84	0.90	1.23	1.32	1.21	1.13
005	43,498	0.21	0.14	0.23	0.41	0.53	0.65	0.81	0.82	0.85
006	77,914	-0.04	0.03	0.15	0.19	0.27	0.44	0.59	0.76	0.21
007	139,561	-0.53	-0.59	-0.40	-0.32	-0.14	0.08	0.49	0.63	0.09
008	249,983	-0.90	-0.81	-0.74	-0.52	-0.36	-0.54	-0.67	-0.64	0.30
009	447,774	-1.31	-1.34	-1.21	-0.96	-0.99	-0.86	-0.87	-0.67	-0.93
010	802,059	-1.76	-1.68	-1.78	-1.88	-1.88	-1.93	-2.10	-2.17	-2.10
011	1,436,661	-1.94	-2.02	-1.97	-2.05	-2.14	-2.11	-2.26	-2.23	-2.36
012	2,573,368	-2.16	-2.34	-2.27	-2.24	-2.41	-2.60	-2.54	-2.66	-3.18
013	4,609,455	-2.07	-1.51	-2.49	-2.71	-2.92	-3.15	-2.62	-3.42	-3.73
014	8,256,524	-3.14	-3.20	-3.25	-3.57	-3.68	-3.15	-3.61	-2.79	-2.78
015	14,789,210	-2.81	-3.30	-3.02	-3.34	-3.63	-3.76	-4.21	-5.15	-5.11

Exhibit 45 Table of skewness values of ultimate distributions verses maturity and size.

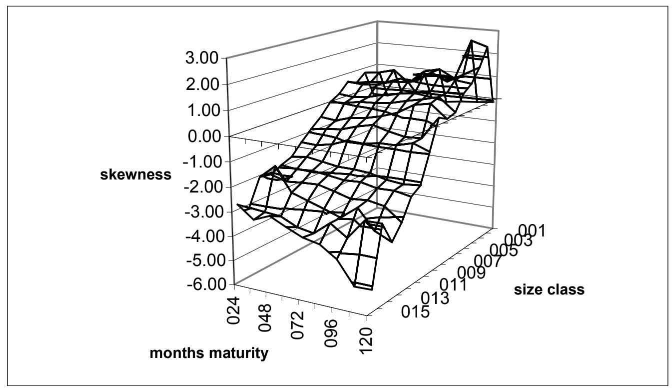

Exhibit 46 Graph of skewness values of ultimate distributions verses maturity and size.

demonstrate a high level of symmetry. It seems that the skewness is resulting from an extended negative tail of small values. The use of a model that does not pick up the negative skewness of large initial claims may only be missing a small probability of these small ultimates.

The concern about underdeveloping smaller claims may be unnecessary. It may be observed here because it is a characteristic of the reinsurance nature of the data. More often, this process is applied to " primary" data which will contain the complete inventory of small losses, not just the ones anticipating a reinsurance recovery. These claims should develop in a less skewed manner.

The Transition Matrix analysis of this data provides us with a method to model the future ultimate distribution of an individual open claim of a given size x, and maturity m. An open claim can be represented at ultimate as a lognormal distribution with:

mu= 1.005 * ln(loss size)

and,

Sigma = 1/(maturity *0.001205+ln(loss size)*0.078874-0.34447)

Where maturity is in months and loss size is in US dollars

2.3.8 Effect of Policy Limits

This study assumed that policy limits affected large losses and sought to avoid its effect. Due to this, the resulting distribution of ultimate losses, have no policy limit censoring. It is necessary to introduce it to arrive at the correct final ultimate value. The final distribution will be the lognormal distribution, given earlier, which has been censored by the policy limit for the claim. If this is not available, then a reasonable assumed policy limit must be used.

2.3.9 Effect of Zero Dollar Claims

The transition matrix process produces an estimate for claims that close with no payment at every maturity. The reader will remember in exhibit 24 all the statistical values are calculated using a conditional probability, after the probability of closing with no payment is removed. Any estimate of future development must reflect this. When taking an open non-zero claim to ultimate, the exhaustive range of outcomes must include zero value of probability P(x=0) and the proposed lognormal with a mu and sigma as previously discussed that has a probability of (1-(P(x=0)). A table of probabilities of closed with no payments is not included in this study because of the biased nature of the data. Since the claims used here are only those submitted for a possible reinsurance recovery, it is expected that the numbers of CWNP claims will differ from a general population of claims.

2.4 Comparison with Direct Transitions

2.4.1 Lack of Memory

The transition matrix process suffers the possibility of a troublesome error. This is due to the implied independence from transition to transition. It is as if each transition has no memory of earlier transitions. A claim, at some value x, does not care how it came to have this value. Its future transitions only depend on its current value. This is clearly not the real world situation. Each claim has a complete history from occurrence to settlement, and each succeeding value has some dependence on earlier values.

It is easy to imagine a scenario in the transition matrix approach that might be counter intuitive. Say, a $1,000,000 claim that goes through ten transitions where each transition happens to reduce the value by half. This results in the value of the claim reducing to $976. Remember that this transition is simply the movement from one class to the adjacent lower one. The transition matrix approach allows this possibility (abet with low probability), but, intuition tells us that this doesn't occur. One might guess that the dispersion caused by future development is exaggerated by the transition matrix approach.

Final	Initial Class																						Grand
Class	000	001	002	003	004	005	006	007	008	009	010	011	012	013	014	015	016	017	018	019	020	021	Total
000		2162	146	315	396	432	547	702	821	1038	919	722	443	263	134	44	18	8	3	2			9115
001		1478	33	68	84	95	83	59	43	35	31	16	8	3	2	1							2039
002		18	287	9	25	27	46	49	22	21	11	11	1	1									528
003		18	14	448	27	34	44	51	30	38	14	9	3	2	2	1							735
004		22	10	27	590	66	79	70	57	46	19	17	3	2	1								1009
005		18	10	23	41	784	87	84	59	49	33	17	4		2	1							1212
006		18	8	24	45	84	1302	117	99	73	50	19	8	1	2								1850
007		29	13	15	24	49	117	1701	194	117	65	24	13	2	1	1		1					2366
008		35	1	13	20	47	89	199	2058	275	112	41	12	9	3	2							2916
009		43	4	9	7	35	55	130	233	2352	265	91	26	12	2	2							3266
010		57	5	10	11	9	35	71	135	379	2588	293	55	14	3	1	1						3667
011		53	6	10	15	19	20	46	71	160	324	2302	192	37	13	4	1						3273
012		35	2	5	4	4	11	13	23	62	100	227	1434	123	24	1	2	1					2071
013		20		5	2	7	7	10	15	19	44	75	173	878	88	18	5	1					1367
014		10		1	3	4	6	3	7	14	5	27	47	98	588	52	8	1			1		875
015		5	1		4	4	5	3	3	5	9	22	19	28	58	282	16	2	1				467
016		2	1		1		2		2	5	2	6	9	8	19	35	157	5	3	1			258
017		1										1	4	4	7	16	12	55	4	2			106
018					2							1	1	1			2		30	1			38
019													1						3	13			17
020																				1	6	2	9
021														1								1	2
Grand Total		4024	541	982	1301	1700	2535	3308	3872	4688	4591	3921	2456	1487	949	461	222	74	44	20	7	3	37186

Exhibit 47 - Counts of claims for initial report to final report (current) classified by by initial and final size. All initial maturities are shown here.

2.4.2 An Alternative Method

In order to assess how much distortion might be caused by this, a study involving direct observation was conducted. The previous transition matrix approach involved observing each transition from year to year, and multiplying each transition until ultimate matrices were created. This study utilized direct observation of the transition from first report to current value which is taken as a proxy for ultimate. These observations were categorized as to initial size class, initial maturity, and final (current) size class at the latest evaluation, which allows these observations to be directly compared with the ultimate loss development matrices.

This study used the same data as previously and was adjusted by the same trend. To provide an overview of the outcome of this study, exhibit 47 shows the number of claims for all initial maturities when classified by initial and final (current) size.

When we view any of the initial size class columns, the bell shaped distribution becomes obvious. We can take the counts of initial claims in any column and divide it by the total counts in that column, and it represents the probability of a final size outcome given the initial size. This is graphed in exhibit 48. For clarity only the odd numbered size classes

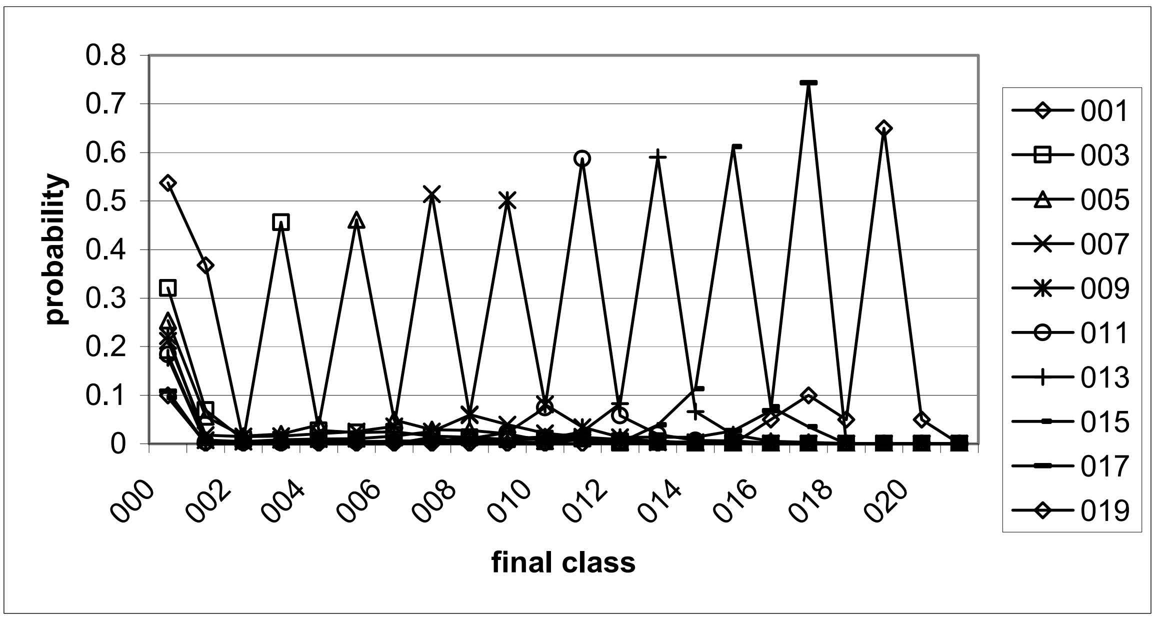

Exhibit 48 - Plot of probability for final size class for a selection of initial size classes for the initial to final transition.

are shown. The graph continues to show the strong "normal distribution like" behavior that we have seen previously. We will look at the distribution statistics for each "ultimate" distribution to see if this empirical ultimate transition is similar to the multiplied results from the transaction matrix method. The counts are further classified by initial maturity, and the statistics, mean, standard deviation, skewness, and kurtosis are calculated allowing us to compare the results directly with those of the transition matrix study.

2.4.3 Similar Results

The mu for these distributions are computed and then divided by the natural log of the average size within the interval similarly to what was done earlier for the transition matrix results. Exhibit 49 is a table of the mu to ln(x) ratio for the initial to "ultimate" transition data.

Size Category	Average Size in Interval	Ratio of Mu/natural log of average size Maturity in Months											
		012	024	036	048	060	072	084	096	108	120	132	144
001	2,712	1.60	1.55	1.57	1.53	1.60	1.45	1.59	1.49	1.42	1.40	1.69	1.43
002	7,569	1.08	1.06	1.10	1.04	1.06	1.10	1.02	1.00	1.07	1.10	1.08	1.16
003	13,557	1.07	1.05	1.04	1.06	1.06	1.09	1.00	1.08	1.04	1.06	1.08	1.03
004	24,284	1.04	1.05	1.05	1.02	1.03	1.03	1.03	1.04	1.00	1.01	1.03	1.06
005	43,498	1.03	1.03	1.03	1.02	1.01	1.00	1.01	1.00	0.99	1.02	1.01	1.02
006	77,914	1.01	1.02	1.01	1.01	1.00	1.00	1.01	1.00	0.99	0.98	0.98	1.00
007	139,561	1.00	1.01	1.00	1.00	1.01	0.99	0.99	0.99	1.00	1.00	1.00	1.00
008	249,983	1.00	1.00	0.99	1.00	0.99	1.00	0.99	1.00	0.98	0.99	0.99	1.00
009	447,774	0.99	1.00	0.99	0.99	0.99	1.00	1.00	1.00	0.99	0.99	1.00	1.00
010	802,059	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00	0.99	1.00	0.99	1.00
011	1,436,661	0.99	0.99	0.99	0.99	0.99	0.99	1.00	0.99	0.99	1.00	0.99	1.00
012	2,573,368	0.98	1.00	0.99	1.00	1.00	0.99	0.98	0.99	0.99	0.99	0.98	0.99
013	4,609,455	0.99	1.00	1.00	1.00	0.99	0.98	0.99	1.00	0.97	1.00	0.99	1.00
014	8,256,524	0.99	1.00	0.99	0.99	0.99	1.00	1.00	0.99	1.00	1.00	1.00	0.96
015	14,789,210	0.99	1.00	0.99	0.97	0.98	1.01	0.99	1.00	1.00	1.00	1.00	1.00

Exhibit 49 - Table of ratio of mu / ln(x) for distributions of initial to final transitions.

This appears very similar to its transition matrix counterpart shown in exhibit 31. The edge towards the small size classes is turned up as it is in the transition matrix study, though it doesn't seem to be any pattern to the differences. If one accepts this elevated mu's for small initial claims to be caused by the data collection process, then this data supports the assertion that the mu of the distribution can be estimated by the natural log of the loss size. This surface is graphed in

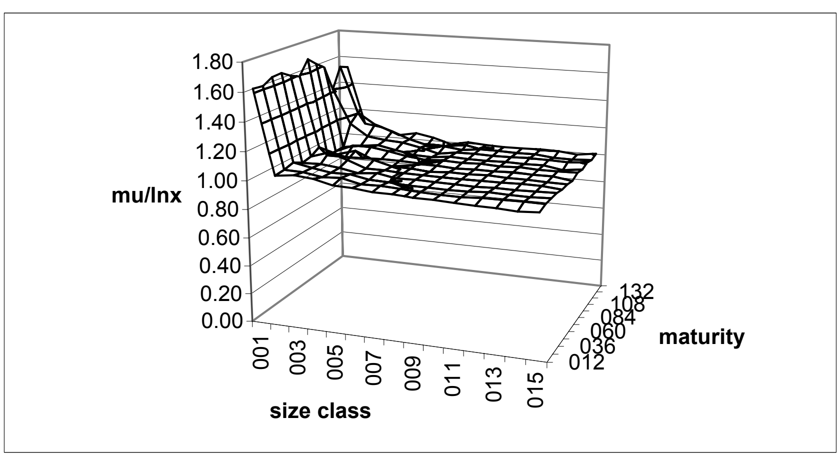

Exhibit 50 - Graph of ratio of mu / ln(x) for distributions of initial to final transitions.

Exhibit 50.

The other important statistical parameter to check is the standard deviation of the log of the "ultimate" loss size. Exhibit 51 shows the sigma's of the ultimate distributions of the

Initial Size	Sigma at Ultimate Initial Maturity in Months											
Class	012	024	036	048	060	072	084	096	108	120	132	144
001	2.40	2.20	2.29	2.15	1.88	2.09	1.99	1.74	1.54	1.09	1.83	2.56
002	1.15	1.25	1.51	0.93	1.18	2.01	0.48			2.16	2.86	1.71
003	1.43	1.23	1.32	1.34	1.98	1.01	0.36	1.55	0.86	1.48	2.12	1.40
004	1.42	1.22	1.18	1.38	1.45	2.24	0.45	1.80	1.57	1.50	1.88	1.59
005	1.18	1.45	1.28	1.14	1.14	1.30	1.08	1.43	1.00	1.25	0.90	1.47
006	1.22	1.24	1.17	1.27	1.02	0.89	1.07	0.89	0.81	1.29	1.46	1.45
007	1.01	1.18	1.17	0.99	1.16	1.01	0.93	1.00	0.86	0.70	0.82	0.85
008	1.10	1.08	1.08	0.92	0.94	0.86	0.89	0.99	0.58	0.72	0.69	1.20
009	1.13	1.07	1.07	1.03	0.96	0.88	1.04	0.96	0.89	0.68	0.59	0.85
010	1.03	1.06	1.01	0.97	0.94	0.92	0.80	0.66	0.79	0.90	0.42	0.48
011	1.16	1.22	1.00	1.01	0.75	0.76	0.95	0.68	0.34	0.52	0.69	0.17
012	1.23	1.21	0.86	0.89	0.77	0.77	0.55	0.48	0.77	0.72	0.40	0.32
013	1.11	1.63	0.94	0.99	0.55	0.47	0.82	0.50	0.37	0.21	0.53	0.34
014	1.08	0.57	1.26	0.63	0.96	0.79	0.47	0.53	1.12	0.62	1.18	0.48
015	1.25	0.54	0.75	0.71	1.30	0.96	1.19	0.30	1.06	0.17	0.42	0.17

Exhibit 51 - These are the sigma's of the ultimate distributions for the emperical initial to "ultimate" observations by initial size class and by initial maturity.

empirically observed initial to "ultimate" transitions, and, these are graphed in exhibit 52. Note that two values, initial size class 002, maturity 96 and 108 months, are missing due to sparse data.

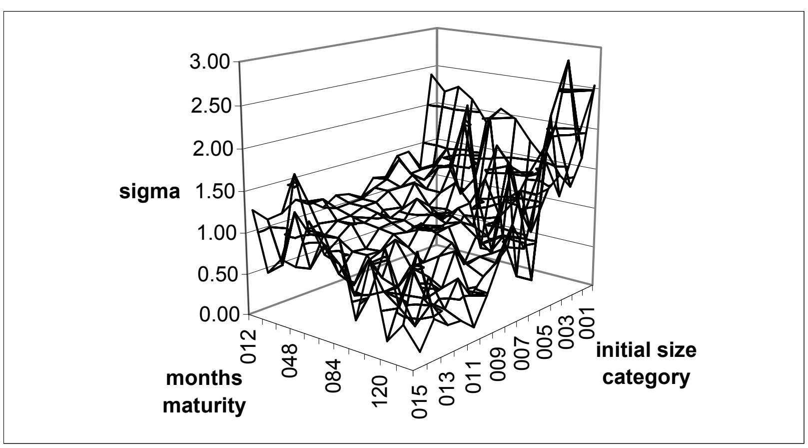

Exhibit 52 - Graph of the sigma's of the ultimate distributions for the emperical initial to "ultimate" observations by initial size class and by initial maturity.

When one compares them to the transition matrix values in exhibit 33 we see that the general shape of the surface is similar, but the values are somewhat higher. On average the transition matrix values are 1.4 times higher then the empirical values. When we look at the surface of the empirical sigma's we see a similar structure to that observed earlier for the sigma's of the transition matrix study. Exhibit 52 shows the plot of the values as a surface, and is comparable to the plot in exhibit 34. This shows higher values and a higher volatility in the small size classes.

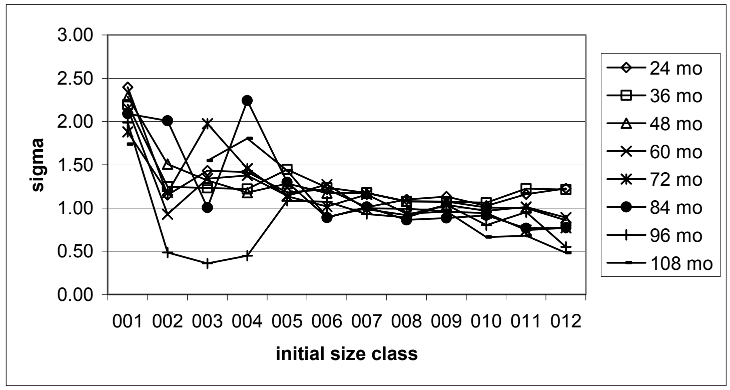

Exhibit 53 - Plot of emperical sigma's by initial size class. Each line represents a different initial maturity.

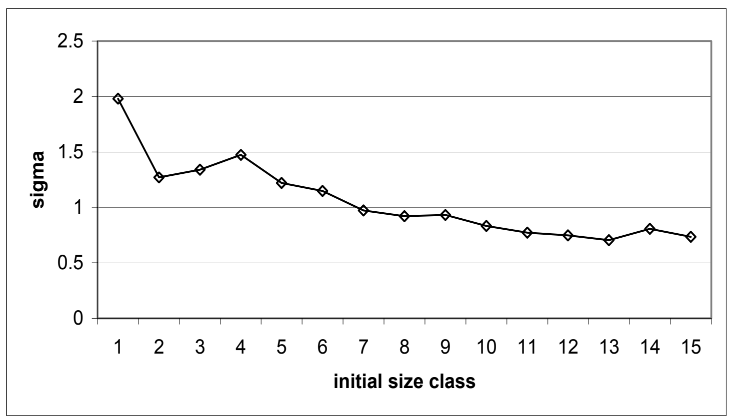

Exhibit 54 - Plot of emperical sigma's averaged over maturity and plotted by initial size class

The "sideways" view of this plot shown in exhibit 53 shows these higher values at small size classes, and then a leveling off as initial size class increases

If we take the average across the maturities and plot these averages verses the initial size class, we get a better sense of how sigma changes with initial size. A plot of this is shown in exhibit 54 and we see a gradual decrease with increasing size. If we take the inverse of this average sigma we see an increasing linear relation as shown in exhibit 55. Again, this is consistent with our earlier model for sigma.

Looking at the other dimension, change in maturity, we find that the sigma values show a gradual decrease with increasing maturity. These values are plotted in exhibit 56, which is comparable to what was seen in the transition matrix values. This is confirmed with a review of exhibit 39, which shows a gradual increase in 1/sigma.

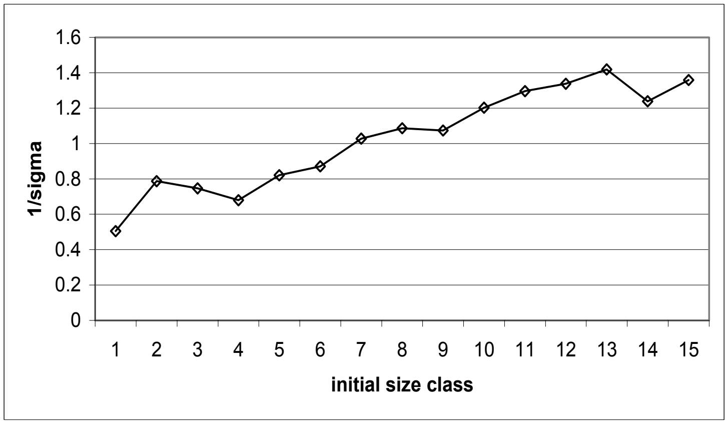

Exhibit 55 - Plot of inverse of average sigma's verses initial class size.

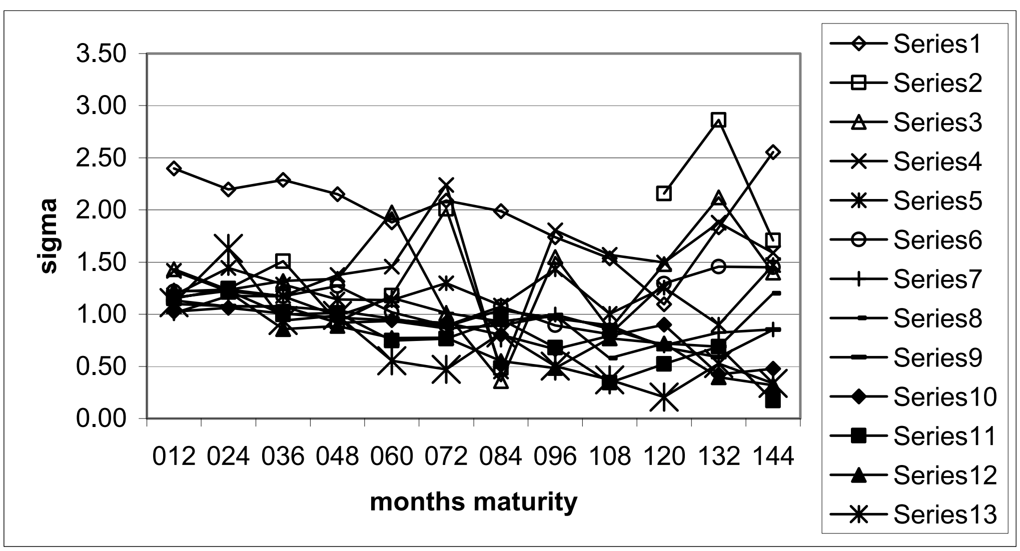

Exhibit 56 - Plot of emperical sigma's by months maturity. Each line represents a different initial size class.

The empirical sigma data is averaged over all sizes and plotted by maturity to show the decreasing trend with maturity as shown in exhibit 57. The inverse of the average

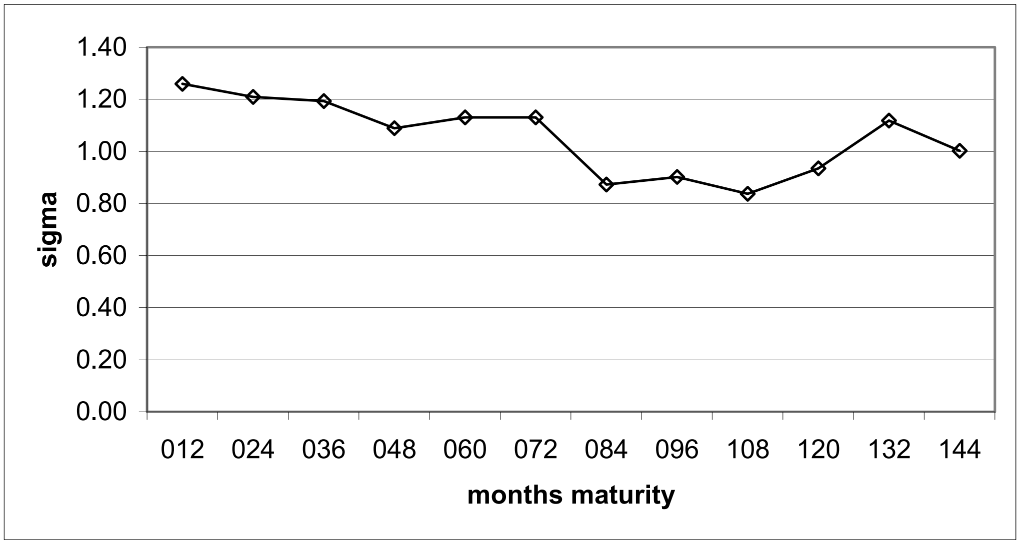

Exhibit 57 - Plot of emperical sigma's averaged over initial size class and ploted by maturity

sigma when plotted verses maturity shows the same increasing trend observed in the transition matrix data. This is plotted in exhibit 58 and can be compared with exhibit 39.

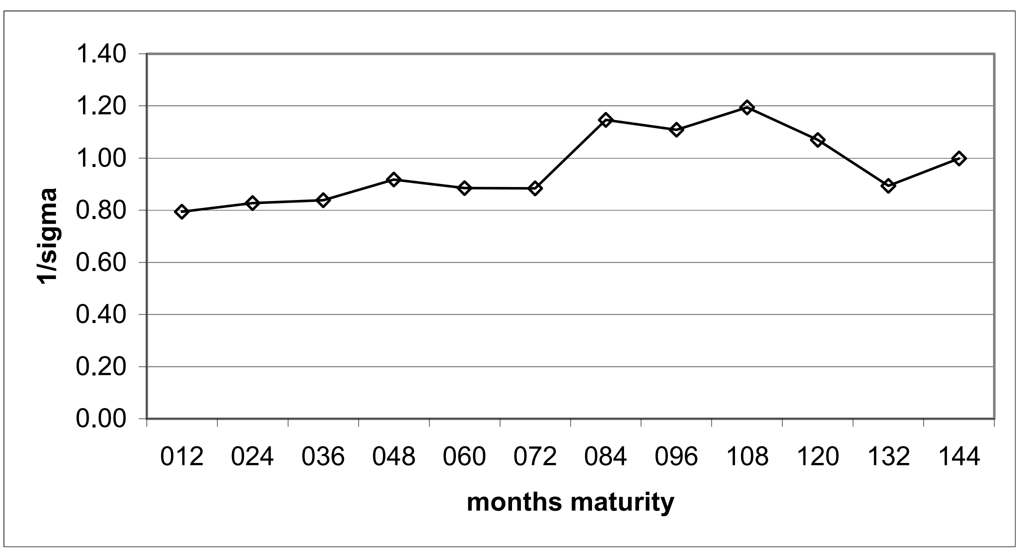

Exhibit 58 - Plot of inverse of average sigma's verses maturity.

A review of the skewness statistics from the empirical study shows similar behavior to the transition matrix study. The skewness is positive for small initial sizes and negative for larger sizes. It has no trend as maturity varies. The average for all observations is -0.33.

2.4.4 Sigma Differs

The overall impression provided by the empirical study is that the lognormal model of developed from the transition matrix study does a good job of describing the loss development, but it needs adjustment. The estimates of mu's from both are very similar, and the skewness follows the same pattern. The estimates of sigma follow a similar pattern but the values of the estimates differ. The transition matrix values are about 1.4 higher than the empirical estimates. If we accept the earlier argument that the transition matrix process may generate more variability then is present in reality, then it is necessary to find a way to reduce the variability. We can accept the observed sigma values in the empirical study, but this has limited application. Since the data has only one observation per claim it is limited and contains more noise. Since sigma behaved similarly in both studies, and differed only by scale, it is better to accept the aggregate level of the empirical sigmas and to try to adjust the fitted sigma surface from the transition matrix study. To do this we need to take a detailed look at the difference between the transition matrix and the empirical sigmas.

We can measure the difference between these two by dividing the empirical sigmas from exhibit 51, by the transition matrix sigmas from exhibit 41, which results in the table and

Initial Size	Sigma Ratio Surface Emperical Sigma's divided by TransitionMatrix Sigma's Initial Maturity in Months										
Class	012	024	036	048	060	072	084	096	108	120	avg.
001	0.90	0.85	0.86	0.77	0.67	0.78	0.76	0.69	0.61	0.40	0.73
002	0.56	0.64	0.82	0.48	0.49	0.87	0.20			1.14	0.65
003	0.69	0.62	0.61	0.59	0.82	0.43	0.17	0.65	0.44	0.86	0.59
004	0.70	0.61	0.59	0.69	0.65	0.94	0.22	0.88	0.81	0.68	0.68
005	0.68	0.78	0.71	0.65	0.61	0.68	0.60	0.77	0.58	0.70	0.68
006	0.68	0.70	0.66	0.70	0.55	0.51	0.65	0.55	0.44	0.85	0.63
007	0.66	0.76	0.74	0.63	0.77	0.63	0.59	0.74	0.60	0.58	0.67
008	0.71	0.75	0.77	0.67	0.69	0.61	0.64	0.73	0.44	0.61	0.66
009	0.79	0.74	0.78	0.75	0.73	0.70	0.83	0.83	0.79	0.61	0.76
010	0.75	0.74	0.74	0.77	0.77	0.77	0.70	0.56	0.71	0.86	0.74
011	0.76	0.84	0.76	0.80	0.62	0.67	0.92	0.67	0.36	0.55	0.69
012	0.85	0.84	0.72	0.75	0.70	0.77	0.53	0.49	0.88	0.84	0.74
013	0.97	1.42	0.88	0.96	0.54	0.47	0.79	0.56	0.44	0.23	0.73
014	1.14	0.60	1.25	0.55	0.79	0.67	0.39	0.47	1.21	0.72	0.78
015	1.27	0.55	0.79	0.67	1.34	1.00	1.19	0.28	1.06	0.16	0.83
avg.	0.81	0.76	0.78	0.70	0.71	0.70	0.61	0.63	0.67	0.65	

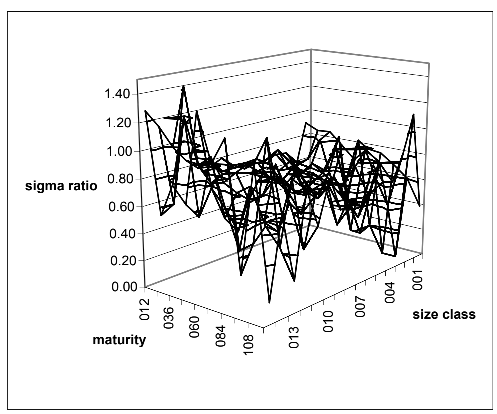

Exhibit 59 - This is the ratio of the emperical sigmas in exhibit 52 divided by the transition matrix sigmas in exhibit 34. A plot of this surface is shown at right. The average of this surface is 0.704

graph shown in exhibit 59. This is the surface of the ratio that is the correction factor to take the fitted transition matrix sigma value to the actual empirical sigma value. In a perfect world, every single value would be equal to each other. But, since there is noise in the data, variation is observed across the surface.

2.4.5 Correction Factor

We want to look for structure by taking the average across maturity and the average across size, which are displayed in the last column and the bottom row respectively. First, we consider changes with size, and we plot the individual maturity data, and then the averages as shown in exhibit 60. This reveals a slight upward trend with increasing initial loss size, however, the fluctuations in this line is well within the noise of the individual data points. Looking at the behavior as maturity varies in exhibit 61, we see a similar result, a slight decreasing trend as maturity increases which is much smaller than the noise of the original data. In interest of parsimony we select this surface to be a level plane with a value of its average, 0.704.

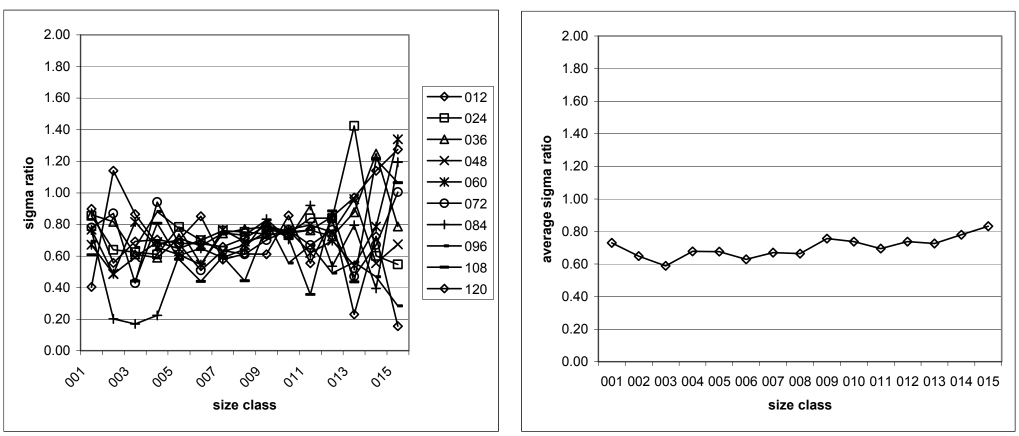

Exhibit 60 - Plot of sigma ratio surface showing how it changes with size. Individual maturities are shown at left, and values averaged over maturity is shown at right

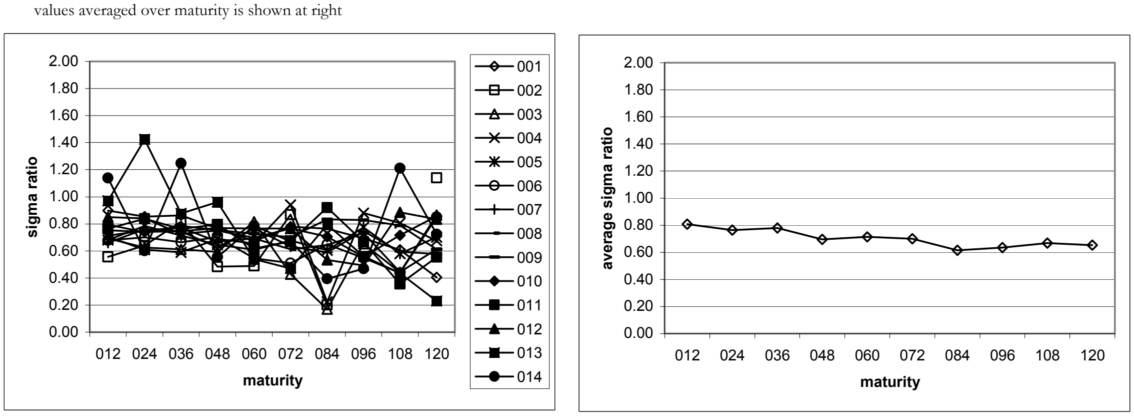

Exhibit 61 - Plot of sigma ratios as maturity varies. The left shows each size class as an individual line. The right shows the values averaged over maturity.

initial size class	fitted sigma values after factor adjustment initial maturity									
	012	024	036	048	060	072	084	096	108	120
001	2.40	2.29	2.18	2.09	2.00	1.92	1.85	1.78	1.72	1.66
002	1.88	1.81	1.74	1.68	1.63	1.58	1.53	1.48	1.44	1.39
003	1.67	1.62	1.57	1.52	1.47	1.43	1.39	1.35	1.31	1.28
004	1.51	1.46	1.42	1.38	1.34	1.31	1.27	1.24	1.21	1.18
005	1.37	1.34	1.30	1.27	1.23	1.20	1.17	1.15	1.12	1.10
006	1.26	1.23	1.20	1.17	1.14	1.12	1.09	1.07	1.04	1.02
007	1.16	1.14	1.11	1.09	1.06	1.04	1.02	1.00	0.98	0.96
008	1.08	1.06	1.04	1.01	0.99	0.97	0.95	0.94	0.92	0.90
009	1.01	0.99	0.97	0.95	0.93	0.92	0.90	0.88	0.87	0.85
010	0.95	0.93	0.91	0.90	0.88	0.86	0.85	0.83	0.82	0.81
011	0.89	0.88	0.86	0.85	0.83	0.82	0.80	0.79	0.78	0.77
012	0.83	0.83	0.82	0.80	0.79	0.78	0.76	0.75	0.74	0.73
013	0.79	0.79	0.77	0.76	0.75	0.74	0.73	0.72	0.71	0.70
014	0.75	0.75	0.74	0.73	0.72	0.70	0.69	0.69	0.68	0.67
015	0.71	0.71	0.70	0.69	0.68	0.67	0.66	0.66	0.65	0.64

Exhibit 62 - Adjusted fitted sigma surface (exhibit 42) after the application of the adjustment factor.

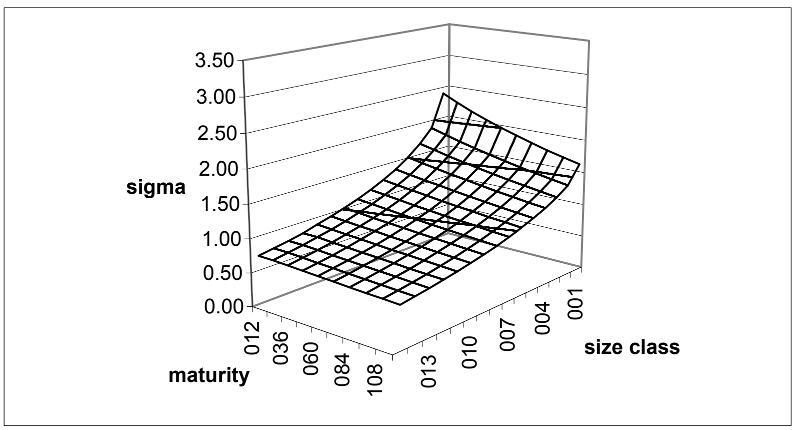

Exhibit 63 - Plot of fitted sigma values after adjusting to the level of the emperical sigmas. This is the data in exhibit 63 and can be compared to plot in exhibit 43.

With this surface approximated as a single value, we adjust the fitted sigma values in exhibit 41, and compare them to the empirical values. Exhibit 62 contains the revised fitted sigmas and this surface is plotted in exhibit 63.

In order to test how well we have approximated the empirical sigmas we can subtract the two surfaces, the adjusted fitted sigmas in exhibit 62 and the empirical sigmas in exhibit 51.

The resultant table of differences is shown in exhibit 64 and plotted in exhibit 65. Here we see the overall average of 0.01 of this surface is very close to zero indicating that the adjusted fitted sigmas are a good approximation to the average level of the empirical sigmas.

initial size class	Error Surface - Difference between adjusted, fitted TM sigmas and Emperical sigmas (TM - E) initial maturity										
	012	024	036	048	060	072	084	096	108	120	avg.
001	0.00	0.09	-0.10	-0.06	0.12	-0.17	-0.14	0.04	0.18	0.57	0.05
002	0.73	0.56	0.24	0.76	0.45	-0.43	1.04			-0.76	0.32
003	0.24	0.39	0.25	0.18	-0.50	0.42	1.03	-0.20	0.46	-0.20	0.21
004	0.09	0.25	0.24	0.00	-0.11	-0.93	0.82	-0.56	-0.36	-0.32	-0.09
005	0.19	-0.11	0.02	0.12	0.10	-0.09	0.09	-0.28	0.12	-0.16	0.00
006	0.04	-0.01	0.03	-0.10	0.12	0.23	0.02	0.17	0.24	-0.27	0.05
007	0.16	-0.04	-0.06	0.10	-0.09	0.03	0.09	0.00	0.12	0.26	0.06
008	-0.02	-0.02	-0.04	0.10	0.06	0.11	0.06	-0.05	0.34	0.18	0.07
009	-0.12	-0.08	-0.10	-0.08	-0.02	0.03	-0.14	-0.08	-0.03	0.17	-0.05
010	-0.08	-0.13	-0.09	-0.07	-0.06	-0.05	0.05	0.17	0.03	-0.09	-0.03
011	-0.27	-0.35	-0.14	-0.17	0.09	0.05	-0.15	0.11	0.44	0.24	-0.01
012	-0.40	-0.38	-0.04	-0.09	0.02	0.00	0.21	0.27	-0.03	0.01	-0.04
013	-0.32	-0.84	-0.17	-0.23	0.20	0.27	-0.09	0.21	0.34	0.49	-0.01
014	-0.34	0.18	-0.53	0.10	-0.24	-0.08	0.23	0.16	-0.44	0.04	-0.09
015	-0.54	0.18	-0.05	-0.02	-0.61	-0.28	-0.53	0.36	-0.41	0.46	-0.14
avg.	-0.04	-0.02	-0.04	0.04	-0.03	-0.06	0.17	0.02	0.07	0.04	
									overall average		0.01

Exhibit 64 - Error Surface of difference between adjusted fitted Transition Matrix sigma's and emperical sigmas.

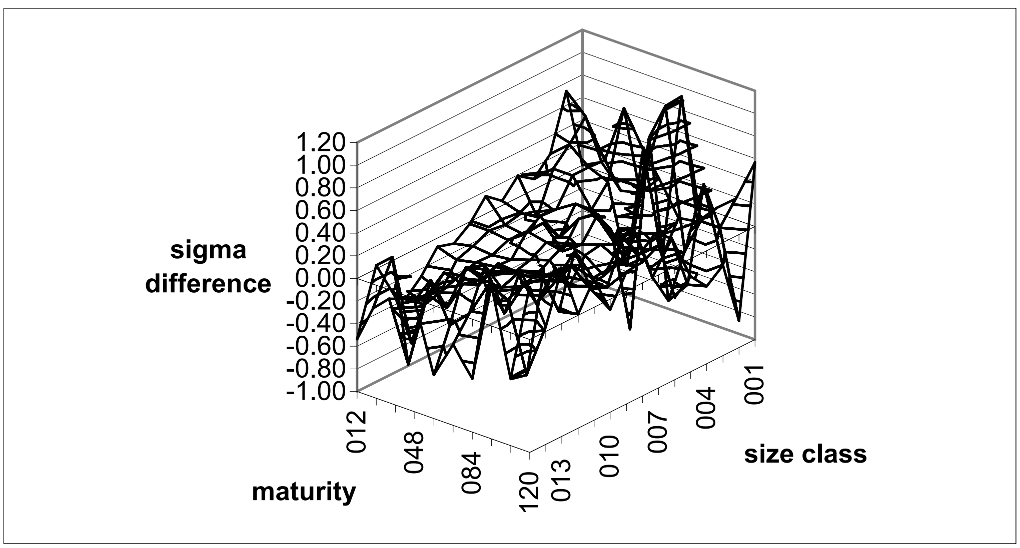

Exhibit 65 - Plot of Error surface of difference between adjusted, fitted Transition Matrix sigma's and emperical sigma's.

Observing the average column and row in exhibit 64 there may be some residual behavior, the average difference seems to go from positive to negative with increasing size, and from negative to positive with increasing maturity. These trends are small, and may be noise. We also must consider that the empirical study used the transition from initial to current report for a proxy for ultimate development. By adjusting the variation in transition matrix model to the level of the empirical data we have the shape of the transition matrix model, but the value levels of the empirically observed data.

2.4.6 The Distributional Loss Development Model

With the sigmas estimated, it is now possible to propose a model that describes the loss development of an open claim of a given size at a given maturity.

An open claim of a given loss size x and a maturity m its ultimate size can be expressed as a log normal distribution with:

Mu = ln(loss size x) * 1.005

and

Sigma = 0.701*(1/(maturity *0.001205+ln(loss size)*0.078874-0.34447))

Where maturity is in months and loss size is in US dollars

2.5 Ratemaking Considerations

2.5.1 Synthetic Data

It is instructive to explore the effect of distributional loss development on estimation of limited expected values and increased limits. In order to do this, a collection of claim values were simulated using a lognormal distribution with a mu of 13 and a sigma of 1.0. Ten thousand values were simulated and 50 were selected using stratified sampling. This was done by sorting them in order and selecting the first percentile, and then every second percentile thereafter, ending with the 99th percentile value. These values are shown in exhibit 66. When these values are graphed in order on a log scale the lognormal

42,151	201,660	357,243	595,631	1,059,857
68,964	215,841	374,978	628,044	1,143,719
86,054	231,946	393,343	657,576	1,236,435
104,213	245,987	414,947	689,340	1,353,120
119,440	260,120	435,752	730,062	1,503,952
134,623	273,892	457,429	772,238	1,670,692
148,901	289,921	480,840	818,901	1,916,141
161,543	305,308	505,404	869,830	2,341,089
175,838	321,387	535,753	925,229	2,897,278
189,464	339,692	567,292	993,142	4,564,144

Exhibit 66 - Simulated loss values, lognormal distribution mu = 13, sigma = 1.0

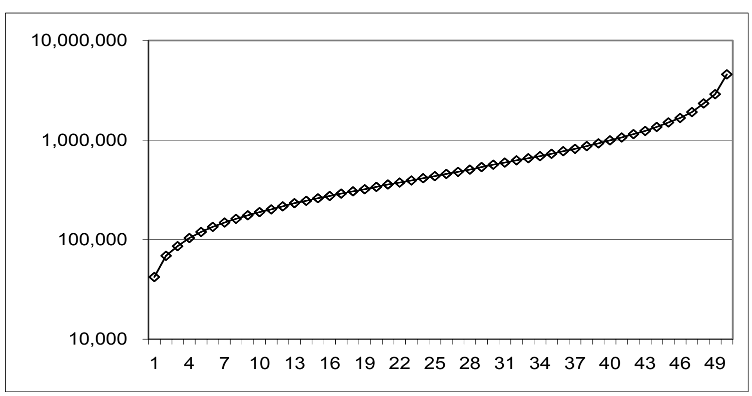

Exhibit 67 - Simulated values graphed in order on a log scale.

distribution becomes obvious as shown in exhibit 67.

2.5.2 Application of Loss Development

We select to apply distributional loss development to these claims using the lognormal model. This process is illustrated in exhibit 68 where eight of the fifty claims are shown for illustration. The claim value is shown in column 1 and its log is shown in column 2.

posted value (1)	dev mu (2) ln(1)	dev sigma (3) see below	ln(dev mean) (4) (2)+((3)^2)/2	dev mean (5) exp(4)	dev factor (6) (5)/(1)
42,151	10.65	1.37	11.59	108,440	2.57
161,543	11.99	1.14	12.64	308,745	1.91
260,120	12.47	1.07	13.04	462,449	1.78
374,978	12.83	1.03	13.36	635,643	1.70
535,753	13.19	0.99	13.68	871,713	1.63
772,238	13.56	0.95	14.01	1,210,565	1.57
1,236,435	14.03	0.90	14.44	1,858,590	1.50
4,564,144	15.33	0.80	15.65	6,270,935	1.37

formula for (3) =0.701*(1/(12*0.001205+LN(1)*0.078874-0.34447))

Exhibit 68 - Application of distributional loss development to eight of the 50 claim values. Note that in the formula for column (3), the log value is of the posted value in column (1).

assume that mu for the loss development model is 1.00 time the log of the loss size. Columns four and five are used to calculate the average loss size of the developed loss. The ratio of column 5 divided by column 1 is the implied loss development factor for the traditional loss development method. These will be used to create developed losses to compare with the distributional developed losses. In this way, we will compare losses whose averages are the same, and differ only in the change in the shape of the distribution. This process was applied to all fifty selected claims. When done, we have three lists of losses, the original shown in column 1, the traditionally developed losses in column 5 and the distributional developed losses represented by mu in column 2 and sigma in column 3.

2.5.3 Comparing Cumulative Density Functions

A good technique to compare the different distributions is to look at the cumulative density functions. Since we have a collection of losses, we can calculate an empirical cumulative density function as shown in exhibit 69. This method involves counting the number of

Original Claims					
Probabliity that the Limit exceeds the Loss					
	Limit				
Loss	25,000	100,000	500,000	1,000,000	5,000,000
42,151	0	1	1	1	1
161,543	0	0	1	1	1
260,120	0	0	1	1	1
374,978	0	0	1	1	1
535,753	0	0	0	1	1
772,238	0	0	0	1	1
1,236,435	0	0	0	0	1
4,564,144	0	0	0	0	1
Count	0	1	4	6	8
Probability	0.000	0.125	0.500	0.750	1.000

Exhibit 69 - Computation of cumulative probability for original claims.

claims that exceed a collection of arbitrary selected limits. The limits run across the top of the table, and the claims are in the first column. A count of one is placed in the field of the table for each intersection representing a claim exceeding a limit. The counts are totaled at the bottom and divided by the number of claim to yield the probability. Again, for display purposes we show eight claims, where fifty claims were used in the study. The final cumulative probability is the ordered pair of the limits running across the top of the table, and the probability running across the bottom of the table.

Claims with Traditional Loss Developemnt					
Probabliity that the Limit exceeds the Loss					
	Limit				
Loss	25,000	100,000	500,000	1,000,000	5,000,000
108,440	0	0	1	1	1
308,745	0	0	1	1	1
462,449	0	0	1	1	1
635,643	0	0	0	1	1
871,713	0	0	0	1	1
1,210,565	0	0	0	0	1
1,858,590	0	0	0	0	1
6,270,935	0	0	0	0	0
Count	0	0	3	5	7
Probability	0.000	0.000	0.375	0.625	0.875

Exhibit 70 - Computation of cumulative probability for claims with traditional loss development.

Exhibit 70 shows a similar treatment for the losses with traditional loss development. Even with the small sample of eight we see a shift in the distribution.

2.5.4 CDF for Distributional Development

The computation of the cumulative probability distribution for the losses with the distributional loss development applied is a bit more complicated. In this case, each developed loss is a distribution. But, this allows us to estimate a probability that a claim is less than a limit. Exhibit 71 shows the detail of this calculation. Across the top is the limit,

Probabliity that the loss is less than the Limit.							
			Limit, ln of limit				
			25,000	100,000	500,000	1,000,000	5,000,000
Loss	mu	sigma	10.13	11.51	13.12	13.82	15.42
42,151	10.65	1.37	0.35	0.74	0.96	0.99	1.00
161,543	11.99	1.14	0.05	0.34	0.84	0.95	1.00
260,120	12.47	1.07	0.01	0.19	0.73	0.90	1.00
374,978	12.83	1.03	0.00	0.10	0.61	0.83	0.99
535,753	13.19	0.99	0.00	0.04	0.47	0.74	0.99
772,238	13.56	0.95	0.00	0.02	0.32	0.61	0.98
1,236,435	14.03	0.90	0.00	0.00	0.16	0.41	0.94
4,564,144	15.33	0.80	0.00	0.00	0.00	0.03	0.55
Sum			0.42	1.42	4.10	5.44	7.44
Probability			0.05	0.18	0.51	0.68	0.93

Exhibit 71 - Illustration of method to calculate cumulative probability of claims with log normal distributional development applied.

and also shown is the log of the limit. We capitalize on the characteristic of the lognormal distribution that the log of the value is normally distributed. We can take the mu and sigma of the lognormal as the mean and standard deviation of a normal, respectively, and calculate the probability interval represented by the

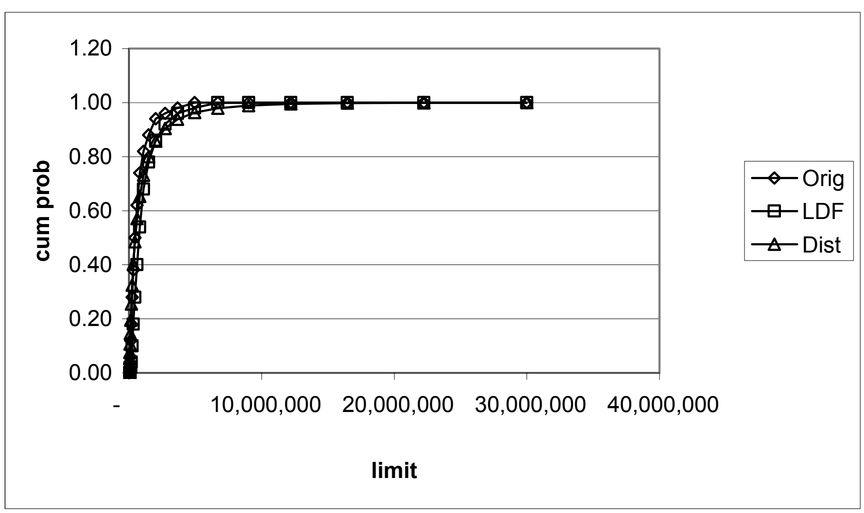

Exhibit 72 - Plot of cumulative probability curves for original losses, traditionally developed losses, and distributionally developed losses.

log of the limit using a normal distribution table. By choosing the correct tail of the normal, we will get the probability that the developed loss is less than the indicated limit. These probabilities are summed down the column and the total is divided by the total number of claims yielding the cumulative probability verses the limit.

Note that our example assumes that all the claims are open, and are subject to development. In a real life situation the collection of claims would be a mixture of open and closed claims. The open claims would be treated in the manner shown in exhibit 71 while the closed claims would be treated as in exhibit 70 where a zero or one is assigned to the probability in the table. The interesting fact is that, by simply adding the count and the "sum" values of exhibits 70 and 71 and then dividing by the total number of claims, one has the probability for the open and closed claims. This provides a method of creating the cumulative probability distribution of the mixed claims. With a cumulative probability available for the three types of losses they can be compared by plotting as shown in exhibit 72. We get a clearer picture of the different behaviors of the various loss development methods when we rearranging the horizontal scale to a log scale, as shown in exhibit 73.

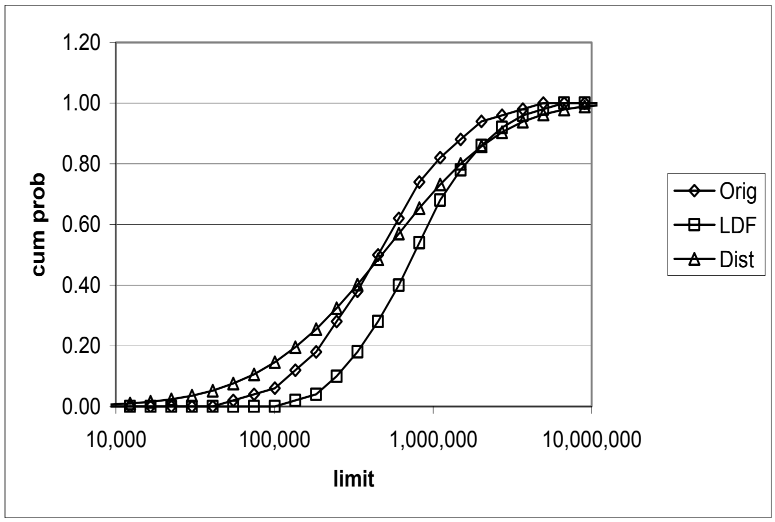

Exhibit 73 - Plot of cumulative probability curves for original losses, traditionally developed losses, and distributionally developed losses with a log scale for the limit.

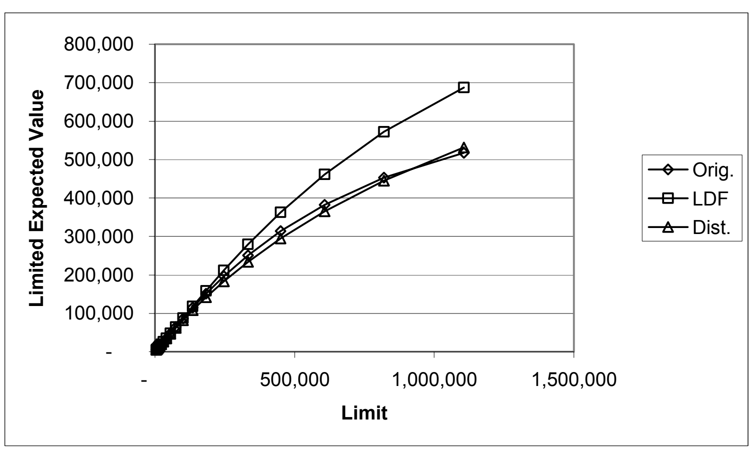

Exhibit 74 - Limited expected value in the range 0 to 1,000,000.

There are two comparisons in this graph. The first comparison is between the original claim data (diamonds) and the loss development factor data (boxes). Here, it would appear that the original line is shifted to the right, but maintains the same shape. The distributional development line (triangles) crosses the original line at the 50% range, but it shows more dispersion at small and large loss sizes.

2.5.5 Comparing Limited Expected Values

Limited expected values were calculated from the empirical cumulative density functions to see how they would behave relative to each other. Exhibit 74 shows the LEV for the range zero to one million in limit. Here we see similarity between the original and the distributional developed losses, while the LDF developed rises quickly. As we look at the range to $5,000,000 in limit we see the distributional rising up to meet the LDF adjusted data as shown in exhibit 75. Looking over the entire range up to $30,000,000 in limit, as shown in exhibit 76, we see that the distributional curve rises up and meets the LDF adjusted limited expected value. One can conclude that this is an expected conclusion since the values were formulated to have an equal mean. Remember, we selected the loss development factors so

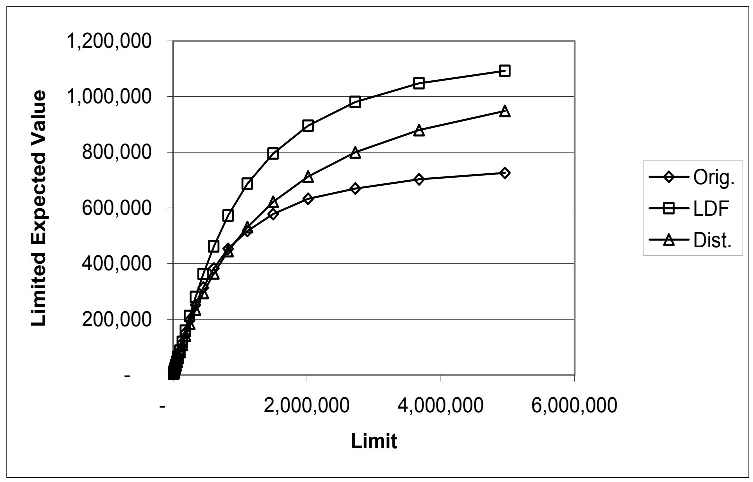

Exhibit 75 - Limited expected value in the range 0 to 5,000,000.

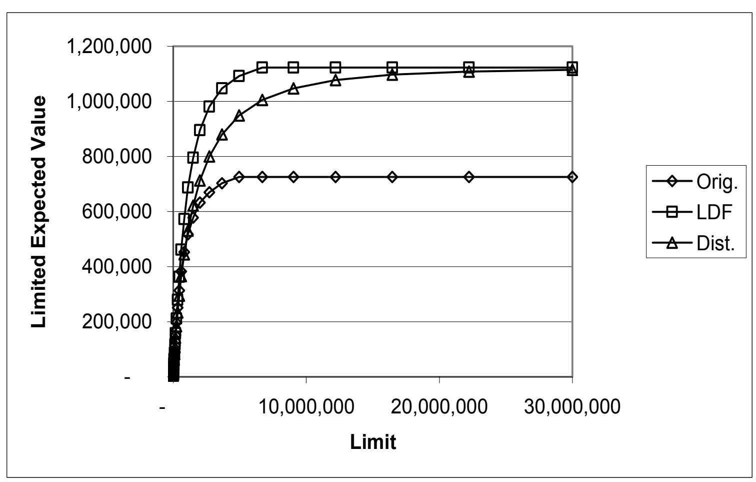

Exhibit 76 - Limited expected value in the range 0 to 30,000,000.

that the average for each individual loss would be the same for the LDF data and the distributional data. We only find this average converging at very high limits where they have little impact on the distribution.

2.5.6 Comparing Increased Limits Factors

With limited expected value curves available we can calculate pure loss increased limits factors. First we select a basic limit of $100,000 and compute the increased limits. In exhibit 77 we show the increased limits for the range of $100,000 to $1,000,000. This shows that the factor method produces ILFs higher by 20 to 30% as compared to the distributional method. The undeveloped and distributional adjusted ILF's are very similar. If we look over a wider range, up to $5,000,000, as shown in exhibit 78, we see a change in behavior. The distributional curve rises from the original curve and begins to approach the factor curve. This is exactly what is seen in the limited expected value curves, and, it is no surprise since increased limits are simply ratios of LEV's with the same denominator. The last range to explore is increased limits factors up to $30,000,000

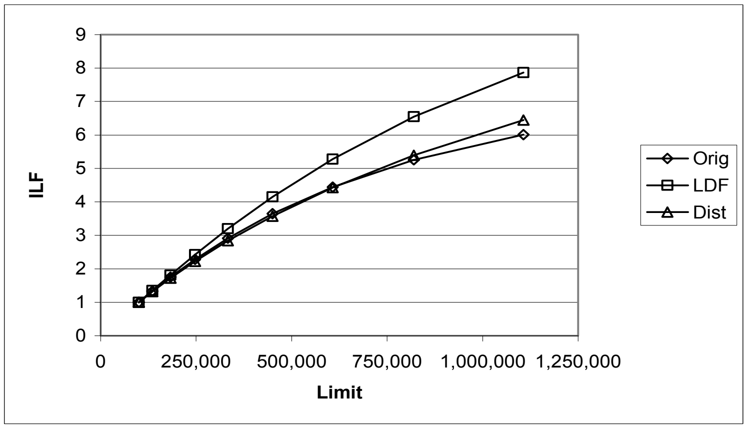

Exhibit 77 - Increased Limits Factors up to $1,000,000 where basic limit is $100,000.

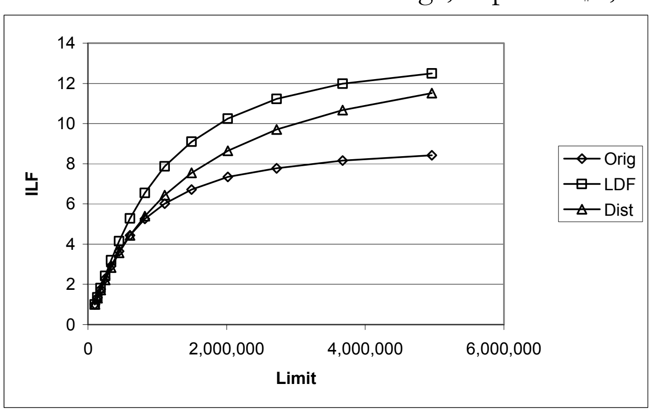

Exhibit 78 - Increased Limits Factors up to $5,000,000 where basic limit is $100,000.

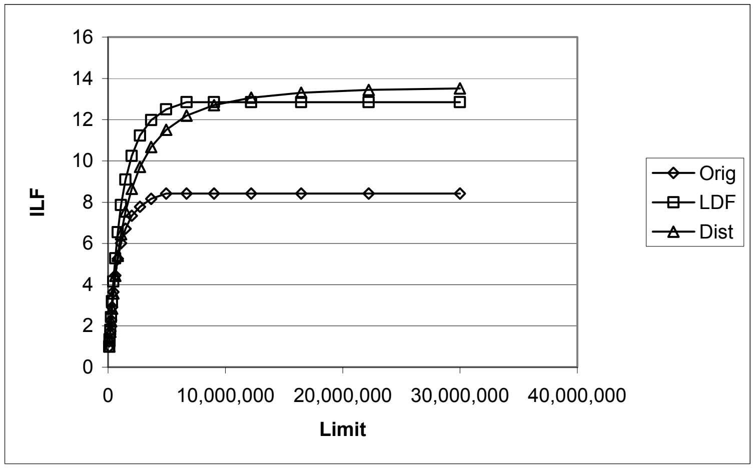

Exhibit 79 - Increased Limits Factors up to $30,000,000 where basic limit is $100,000.

as shown in exhibit 79. Here we see that the distributional line has risen up and exceeded the factor line. This is because the denominator for the distributional line is less at the $100,000 basic limit.

Varying the basic limit will change the behavior of the increased limits factors. Selecting $1,000,000 as basic limit and recalculating the increased limits factors produces the results

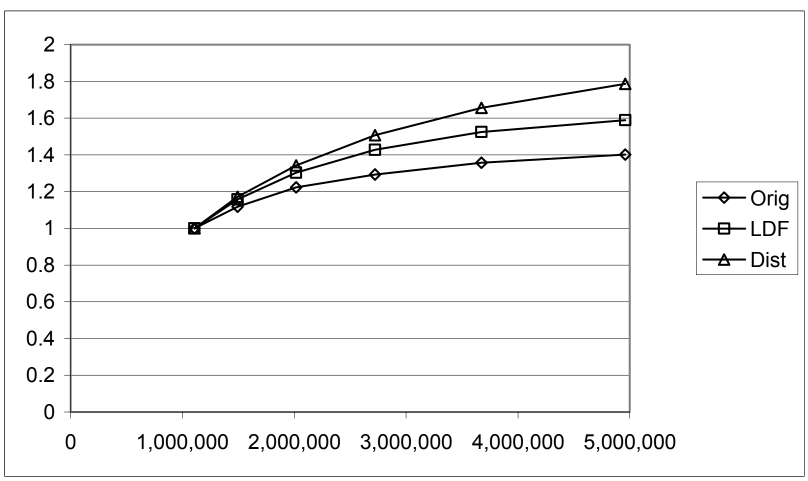

Exhibit 80 - Increased Limits Factors up to $5,000,000 where basic limit is $1,000,000.

shown in exhibit 80. Here we see that the distributional result is higher than the factor result, and this is consistent over this range and over the larger range, up to $30,000,000 as

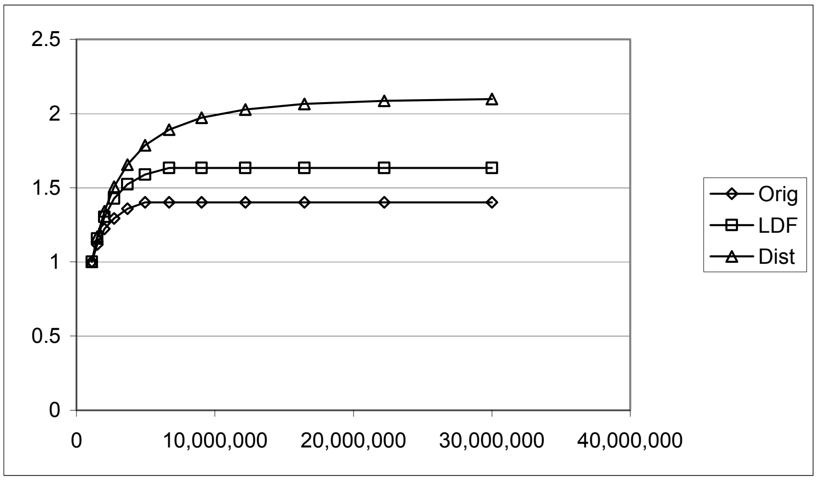

Exhibit 81 - Increased Limits Factors up to $30,000,000 where basic limit is $1,000,000.

shown in exhibit 81. Here we see a difference that is again in 30 percent range.

3. RESULTS AND DISCUSSION

It is well known that individual open claims will develop to an ultimate value that may be more, less, or the same as the current value. The exact nature of this distribution has never been clear. This study shows, with two different approaches, that it is a skewed distribution that can be modeled with well known severity distributions.

The lognormal distribution is particularly suited to modeling the severity distribution of the ultimate of an open claim. One difficulty of exploring loss development of individual claims is the excessive parameterization that occurs. The empirical transition matrix approach results in a very large number of parameters that vary by initial maturity, and claim size. Using them in a practical system to apply loss development to individual claims with the intention of arriving at an ultimate severity distribution would be cumbersome at the least. Applying the lognormal distribution to the transition matrix approach greatly reduces the parameters needed by characterizing the ultimate distribution of an individual loss as a lognormal with a mu which is a function of the initial size, and a sigma which a function of initial size and maturity. The math for combining a collection of individual lognormal distributions into an aggregate severity distribution is well known. This results in a practical method to apply loss development to individual claims that results in the ultimate severity distribution where the loss development recognizes the potential for claims to increase or decrease from the current size.

The lognormal individual claim loss development model appears to describe behavior in the absence of policy limits. It provides ample upward development of even very large claims. Policy limits must be applied after the loss development in order to correctly represent the potential loss of the developed severity distribution.

The observed dispersion in development appears to decrease and skewness becomes less positive as claim size increases. This is caused by an extended negative tail of small claims while the main peak remains symmetrical. The lognormal model does not capture this negative tail. Some practitioners may be uncomfortable in relying on policy limits to explain this and may want to adopt a more complex model that reflects this detail. A bi-modal lognormal treatment of the ultimate distributions may more accurately reflect this.

A comparison between Transition Matrix ultimate development and empirically observed first to last report transitions indicate that the Transition Matrix approach may result in more variation than actually present. This could be due to the independent nature of the Transition Matrix approach. It is reasonable to expect a certain amount of dependency as real claims progress to ultimate. It is interesting to note that the two models have the same shape and differ from each other only in the scale of one parameter, sigma. It is possible to adjust the lognormal development model resulting from the Transition Matrix approach so that it "balances" to the average values of the empirical loss development model. Another factor to consider is that the empirical loss development may very well underestimate sigma since it is missing data, but, does provides a minimum boundary. Further study is indicated to measure how much sigma is underestimated by the initial to final transitions, and, one may find that the truth may lay somewhere in between the empirical and the transition matrix approaches.

Both the Transition Matrix model and the empirically derived model exhibit a positive skewness for small claims and a negative skewness for larger claims. This conveniently fits with long held opinions in the casualty actuary community that small claims have a tendency to develop larger, and large claims have a tendency to develop smaller. The tendency for small claims to develop larger, in this case, may be a characteristic of the data because it is a subset that has been submitted for reinsurance recovery. It is not hard to imagine that the process of selecting claims for submission will exclude small simple claims that are not expected to develop. This results in a collection of claims that is biased towards larger development for the smaller claims.

The larger initial claims exhibit a symmetrical distribution (in the log transform space) about the most populated final size interval, with a small percentage of claims filling in the lower size intervals. Typically you see 2 to 4 percent of the claims in lower intervals

which are more that three size classes lower than the mode. Though not planned, the width of the size classes are about one standard deviation. So, the mode size class and the four adjacent classes account for about 95% of the claims. So, this 2% to 4% spread evenly in the lower size classes causes the negative skew values. If one ignores the extreme outliers, then the lognormal loss development model is a very good fit. In the future, it may do well to revisit the fit of these distributions with a bi-modal lognormal distribution. One mode will pick up the sharp peak around the initial loss size, and the other will be low probability with wide dispersion to pick up the claims that develop to much smaller values.

It is important to note that the study to measure ratemaking impacts was designed to stress the difference in the result between the two methods. It uses a claim set of all open claims. In most realistic situations, 60 to 75 percent of the claims would be closed and not experiencing any additional development. This would cut down this observed difference from the 30 percent range to less than 10 percent.

When considering the effect on pricing measures, it is important to point out that this study only compares the distributional loss development with single factor development, which is a method used in an ad-hoc manner to adjust small datasets for reinsurance rating. It does not imply a comparison with published industry standard increased limits factors, which are prepared with sophisticated methods that correctly reflect the distributional nature of individual loss development.

And last, the reader is guarded against a direct comparison of these increased limits factors with published increased limits factors since industry factors are prepared with a ballasting of a large collection of small losses, and that is clearly not the case in the losses used here.

4. CONCLUSIONS

Applying Transition Matrix theory to individual claims allows one to build up a picture of individual loss development, which has been seldom seen before. Using this method, it is possible characterize individual claim development as, distributional process where, the claim, at a known current and open amount, will, at ultimate, be a value which is forecast by a claim distribution. For general liability claims in the United States, the ultimate loss development of an individual claim can be represented as a heavy tailed skewed distribution, which closely resembles a log-normal distribution. It is possible to develop a simple functional relationship between the size and maturity of the open claim, and its ultimate lognormal distribution using four parameters.

The Transition Matrix approach may introduce excessive dispersion into the forecast of ultimate loss due to its independence assumption, but, it is possible to measure and adjust for it. This results in a model that allows one to take individual open claims, and adjust for development to ultimate, before fitting these claims to a severity curve. It can be shown that the distributional loss development process will change the shape of the ultimate size of loss distribution in a way that will affect loss cost estimates in a range of a few percent to 10 to 20 percent. It is important to reflect the distributional nature of loss development when evaluating individual loss data in order to avoid these errors.

Acknowledgement

The author wishes to acknowledge the assistance of, Jack Barnett who provided me his procedure for implementing the transition matrix method which is describe in this paper, Gary Venter who conceived the idea of applying the transition matrix method to our internal data, and whose comments have guided this paper, Spencer Gluck, and Jose Couret, for their review and comments which have influenced this paper.

Appendix A

The following describes the data used in this study,and, its treatment. This data was submitted as first dollar, 100% ground up insured loss. These claims were entered into the initial computer system as incremental transactions. The critical timing elements captured were date of loss, date the claim notice was submitted to the intermediary, and the date the claim was entered into the system. The various money types were maintained separately, such as indemnity, expense, subrogation, etc. A line of business code was entered for each claim, which was used for segregation into broad categories. A description of cause of loss field contained a detailed text description of the loss. This was used to isolate claims into sublines. It was found that simple terse descriptions were used repeatedly, and these were useful for identifying sublines. All the unique descriptions were isolated, and each was assigned to sublines. The goal was to create an Other Liability collection of data by identifying and removing Workers Comp, Medical Professional, Lawyers Professional, Pollution, and Auto. Also removed were claims arising from special events or circumstances such as the World Trade Center Disaster, toxic waste, environmental clean up, tobacco, cancer, etc. The remainder was deemed to be the Other Liability.

It was possible to isolate claims labeled as "other liability" using the Line of Business field. Inspection of this subset of the data revealed that it contained more than ordinary liability losses. A list of string fragments was assembled to eliminate claims based on the likelihood that the claim was another subline. For these claims, the description of cause of loss gave a good indication that the loss might be workers comp, legal liability, medical professional liability, auto liability, products etc. It also allows removal of special incidents, such as the World Trade Center disaster, tobacco losses, hazardous waste, environmental cleanup etc. The entire list of string fragments is shown in exhibit 17.

Strings for eliminating Claims				
FIRE	BREAST	IMPLANT	LEGAL	SURGERY
MOLD	CLASH	INS RAN	LIQUOR	SURGICAL
LAWYER	CLASS	INS REAR	M.V.	TABACCO
SURETY	D & O	INS VEH	M/	TOBBACL
ASBES	D & 0	INS. STRUCK	MED MAL	TRACTOR
AGG	D&O	INS.BACKED	MED. MAL	TRAILER
WASTE	D &O	INSD BACKED	MED PROF	TRUCK
ENV	D*O	COLLIDED	MED MAP	TYPHOON
CLEAN	D+O	CROSSED	MED NEG	VEH
SEX	D. & O.	INSD DRV	MED.MAL	VESSEL
CONTAM	D. AND O.	DRIVER	MED/MAL	W C
POL	D/O	INSD DV	MOLEST	W.C.
SITE	DIR &	INSD FAIL	MOTOR VEH	W/C
REMED	DIRECTOR	INSD HIT	MOTOR ACC	WC
WTC	E & O	INSD LOST	MOTORCYCLE	WORK
TOBAC	E&0	INSD R/E	MOTORIST	WORLD TRADE CENTER
CANCER	E&O	INSD RAN	MOTORVEHICLE	CAR CO
MEDICAL	CAR AC	INSD RE	MOTORYCLE	CAR CR
ACCOUNT	TRUCK	INSD REAR	MV	CAR FL
ACCT	VEH.	INSD ROLLED	NURSE	CAR HIT
AIDS	VAN	INSD RENTED	PEDEST	CAR IN
ATTNY	FEN PHEN	INSD SKID	PROD	CAR R
ATTORNEY	FIDELITY	INSD STRUCK	PROF	CAR S
ATTY	H.I.V.	INSD STUCK	REAR END	CAR T
AUDIT	HIV	INSD TRUCK	REAR-END	CAR/
AUTO	HOSPITAL	INSD TURN	REAREND	CAR\
Agg	HURRICANE	INTERSECTION	DRIVING	

Exhibit 17 This is a list of string fragments that were used to eliminate claims from the study.

Claims were eliminated if the string fragment was contained in the "Description of Cause of Loss" field.

The sum of the indemnity and ALAE was used as the loss in this study. The incremental transactions at irregular times were accumulated into year end evaluations for each claim. A claim was deemed to be closed if it's paid and incurred amounts were the same. The closure event was deemed to have occurred when the claim first arrived at this amount, and a flag was entered into the data, for each claim, to mark this.

The losses were trended using the Masterson trend factors published in Best's. The General Liability Bodily Injury trend indications were used. These were available from 1984 to present. 1980 to 1983 were adjusted by an additional annual trend of 8%.

Claims from accident years 1979 and earlier were excluded from this study in order to reduce the amount of computation. This left about 37,000 claims, of which, about 28,000 were non zero in 2003.

5. REFERENCES

Footnotes

[1] Hachemeister, Charles A., "A Stochastic Model for Loss Reserving," Trans. 21st Int. Congress of Actuaries, 1980, Vol. 1, pp. 185-194

[2] Hesselager, Ole, "A Markov Model for Loss Reserving,", ASTIN Bulletin, 1994, Vol 24-2, pp. 183-193

[3] http://economics.about.com/library/glossary/bldef-markov-transition-matrix.htm

Biography of the Author

John Mahon is a Vice President with Guy Carpenter Instrat, within Guy Carpenter & Co., Inc., where he has been providing actuarial services for 12 years. Prior to this, he was with American Re-Insurance Co., Inc. where he provided actuarial services for 7 years. Prior to this he was with ISO for 3 years in the increased limits and research areas. He was trained as a scientist, and holds a B.S. and a M.S. from Stevens Institute of Technology