AMERICAN ACADEMY *of* ACTUARIES

A RISK-BASED CAPITAL TREND TEST FOR P/C INSURERS:

REPORT TO THE NATIONAL ASSOCIATION OF INSURANCE COMMISSIONERS CAPITAL ADEQUACY TASK FORCE

American Academy of Actuaries
P/C Risk Based-Capital Committee

August 26, 2004

The American Academy of Actuaries is the public policy organization for actuaries practicing in all specialties within the United States. A major purpose of the Academy is to act as the public information organization for the profession. The Academy is non-partisan and assists the public policy process through the presentation of clear and objective actuarial analysis. The Academy regularly prepares testimony for Congress, provides information to federal elected officials, comments on proposed federal regulations, and works closely with state officials on issues related to insurance. The Academy also develops and upholds actuarial standards of conduct, qualification and practice, and the Code of Professional Conduct for all actuaries practicing in the United States.

1100 17th Street NW Seventh Floor Washington, DC 20036 Telephone 202 223 8196 Facsimile 202 872 1948 www.actuary.org

Summary

At the request of the National Association of Insurance Commissioners (NAIC), the American Academy of Actuaries' Property & Casualty Risk Based Capital Committee (P&C RBC) reviewed the suitability of using a life insurance type trend test in the P&C RBC calculation. We found that while a life insurance type trend test was not effective in differentiating companies that were more likely to fall below their Company Action Levels (CALs) in the subsequent year, a one-year Combined Ratio test did provide statistically significant differentiation.

Life Insurance Type Trend Test

The life insurance type trend test supplements the RBC calculation by looking at the trend in RBC results for a company. For companies that fall in the 200% to 250% of Authorized Control Level (ACL), the test subtracts from the current level the larger of the decrease in margin (the amount above 200%) for either the prior year or the average of the prior three years. If the result is a level of 190% or lower, the company is deemed to be in the CAL.

We applied the life insurance type trend test to the 143 P&C companies with RBC between 200% and 250%. Most companies in this band were flagged. This test did not differentiate between companies more and less likely to fall below the 200% threshold in the subsequent year.

Life Type Test 200%<RBC<250%	Flagged	Not Flagged	Total
Below 200% Next Year	13 (11.1%)	4 (15.4%)	17 (11.9%)
Above 200% Next Year	104 (88.9%)	22 (84.6%)	126 (88.1%)
Total	117 (100.0%)	26 (100.0%)	143 (100.0%)
Ratio of Companies Falling to <200% to Total	11.1%	15.4%	11.9%
Statistical significance	Not meaningful (tends in wrong direction)		

The test was also applied to 480 P&C companies that fell in the 200% to 350% ACL level. This population of companies had 39 that fell below the 200% level the next year and 441 that stayed above that level.

The results of the application of the life insurance type trend test are as follows:

Life Type Test 200%<RBC<350%	Flagged	Not Flagged	Total
Below 200% Next Year	24 (9.6%)	15 (6.5%)	39 (8.1%)
Above 200% Next Year	226 (90.4%)	215 (93.5%)	441 (91.9%)
Total	250 (100.0%)	230 (100.0%)	480 (100.0%)
Ratio of Companies Falling to <200% to Total	9.6%	6.5%	8.1%
Statistical significance	Not significant		

The results are not very good. Although 24 of the 39 companies that would fall below the 200% level the subsequent year were identified, the test incorrectly flagged more than half of the companies (226 of 441). There are more than 9 companies flagged that will remain above the CAL for each company that will fall below it. Applying a one tailed statistical test to the null hypothesis that the flagged sample has the same likelihood of subsequent failure as the overall sample, using the normal approximation to the binomial distribution, yields an outcome that is not significant, and thus the null hypothesis cannot be rejected.

Two-Tiered Combined Ratio Test

We found that a simple Combined Ratio test is the best predictor of the measures tested (see the Technical Appendix for additional detail) of future RBC declines to the CAL level:

RBC Ratio	Current Year Combined Ratio	Company Status
200%-300%	Greater than 120%	More Likely to Decline to CAL
	Less than 120%	Less Likely to Decline to CAL
300%-350%	Greater than 134%	More Likely to Decline to CAL
	Less than 134%	Less Likely to Decline to CAL
Above 350%	All	Less Likely to Decline to CAL

The Two-Tiered Combined Ratio test looks at the current year's combined ratio for companies that fall in the 200% to 350% level. For companies in the 200% to 300% level, if the combined ratio is 120% or more the company is flagged. For companies in the 300% to 350% level, the combined ratio criterion is 134%.

The results of the application of the Two-Tiered Combined Ratio Test were:

	Flagged	Not Flagged	Total
Below 200% Next Year	26 (17.2%)	13 (4.0%)	39 (8.1%)
Above 200% Next Year	125 (82.8%)	316 (96.0%)	441 (91.9%)
Total	151 (100.0%)	329 (100.0%)	480 (100.0%)
Ratio of Companies Falling to <200% to Total	17.2%	4.0%	8.1%
Statistical significance	> 99.5%		

For this test, there are 5 companies flagged that will remain above the CAL for each company that will fall below it, as compared to a 12:1 ratio of companies remaining above CAL in the entire sample of 480, and 25:1 of the 329 companies not flagged. Applying a one tailed statistical test to the null hypothesis that the flagged sample has the same likelihood of subsequent failure as the overall sample using the normal approximation to the binomial distribution yields an outcome that is highly significant, greater than the 99.5% level, so the null hypothesis can readily be rejected.

Discussion

There are no known criteria for how to measure the "goodness" of this type of test. There is benefit to improved identification of companies statistically more likely to experience financial difficulties. However, the costs of regulatory intervention should also be considered when the identification results in a high percentage of companies who may not need increased attention.

The table below compares the life insurance type trend test and the Two-Tiered Combined Ratio test to what would happen if the CAL were raised to 250% of RBC.

Comparison of Hypothetical Changes to the RBC Formula:

Hypothetical Formula Change	Additional Companies at CAL	True Alarms, Those Falling <200%	False Alarms	True Alarms to Total Flagged
Change the CAL to 250%	143	17	126	11.9%
Implement a life insurance type trend test for companies 200%<RBC<250%	117	13	104	11.1%
Implement a Combined Ratio test for companies 200%<RBC<250%	67	12	55	17.9%
Implement a Combined Ratio test for companies 200%<RBC<300%	123	22	101	17.9%
Implement a Combined Ratio test for companies 200%<RBC<350%	150	26	124	17.3%

Changing the CAL to 250% results in 17 additional companies correctly flagged, but at a cost of 126 false alarms. The life insurance type trend test does not improve the situation, reducing the true alarms by a greater percentage than the false alarms. If we compare the Combined Ratio test for companies 200%<RBC<300% to an increased CAL, we note that more companies that will subsequently fall to CAL are caught (22 vs. 17), but at a lesser cost in false alarms (101 vs. 126). A Combined Ratio test for companies with 200%<RBC<350% compared to an increase in CAL will also identify more companies that will subsequently fall to CAL (26 vs. 17) while the cost in false alarms is comparable (124 vs. 126).

Since the comparison shows a Combined Ratio test is a superior filter to other measures, but still flags many companies who remain above CAL, an alternative to modifying the RBC formula is to include a Combined Ratio test within the regulatory financial solvency framework, but not as part of the RBC formula. If CAL is not automatically triggered, the focus and degree of increased oversight might be more easily managed. Special circumstances that cause or contribute to a high Combined Ratio, such as reserve strengthening or a catastrophic event affecting results, could be considered. However, this approach does not provide the same authority for regulatory action.

A more precise test for this purpose may not exist. Additional analysis is being considered by our committee to more clearly understand the company/group relationships existing in the categories above. It is also possible that RBC ratios fall in the second subsequent year, which has not been tested. We are also interested in whether Combined Ratio trends over time may be significant, and whether capital arrangements made during the year may play a part.

The Combined Ratio test has statistical significance in differentiating companies whose current RBC ratio is between 200% and 350%. However, 82.8% of these companies did not fall below 200% in the subsequent year. We look forward to continuing discussions with the NAIC regarding efforts to identify potentially troubled companies as early as possible, and are available to provide additional research as needed.

The technical appendix that follows describes the methodology used to evaluate the life insurance type trend test, and to identify and evaluate other potential differentiators.

Technical Appendix

Background

In 2003, a working group of the American Academy of Actuaries was formed from the Committee on Property/Casualty Risk Based Capital to investigate the use of a trend test. This memorandum documents the analysis and recommendations of that working group.

A trend test has been used in the life insurance RBC formula to differentiate between companies that are above but near the threshold for company action level ("CAL") that may warrant regulatory attention from those that do not. The life insurance type trend test compares changes in RBC margin for the past year, and the average of the past three years, for companies with RBC ratios less than 250% but greater than the CAL of 200%. The greater of these "trends" is deducted from the current RBC margin, and if the result is below 190%, the company is deemed to fall into the CAL. No trend test is currently used in the Property/Casualty ("P&C") formula. Analysis was performed over the second half of 2003 and early 2004.

Approach

Before starting analysis work, a working group of the committee met to discuss the scope and approach of the work. Rather than stop at analyzing the results of the application of a life insurance type trend test on P&C companies, the working group felt the scope of the project included finding differentiators that worked well in determining which companies falling near the CAL were most likely to subsequently deteriorate into CAL, and which companies would most likely not deteriorate to CAL over the course of the subsequent year.

The approach taken was based on the statistical approach of "hypothesis testing". First, an informal review of past company failures or major RBC declines was conducted. Based on the review of the major causes of these past failures and declines, and the commonalities found, a number of working hypotheses were determined for testing. This process is referred to as analysis on the "micro" level, meaning individual company level. Publicly available information was used, such as data in AM Best Insurance Reports and press reports. The companies examined generally experienced the difficulties during the last five years. The hypotheses were stated in terms of a measure(s) that could be analyzed using data available from statutory annual statements. These hypotheses are stated in the "Results" section below.

Statistical Tests

The statistical tests of the hypotheses were performed in two ways. First, a series of regressions and correlations were run to determine the persistency and predictive power of the measure(s) underlying each hypothesis. This analysis was used to confirm and understand the results of the second phase of statistical analysis, which are discussed under "Retrospective Tests". The data used for this first phase of testing consisted of 2000 data predicting 2001, as well as 2001 data used to predict 2002. These statistical tests were performed before 2003 data became available.

Retrospective Tests

This second phase consisted of a retrospective test to confirm the predictive power of the posited causes of deterioration in RBC. In these tests, most of the hypotheses were brought forward from the first set of statistical tests, regardless of the outcome of the statistical tests.

Data was refreshed to include 2003. Since some of the hypotheses provided for predicting movements in RBC based on three historical years, this gave the ability to build two test years; first, to test the ability of data from 2000-2002 to predict movements in 2003 ("test year 2002, subsequent year 2003"), and similarly data from 1999-2001 to predict movements in 2002, ("test year 2001, subsequent year 2002"). These two sets of test data were combined in a database that initially consisted of 4795 observations, noting that one company could provide two observations, one each for test years 2001 and 2002. This data was then scrubbed to eliminate extreme and invalid data, reducing the number of observations to 3697. About 90% of observations eliminated were due to invalid data, rather than extreme values. The retrospective test is much less affected by extreme values as opposed to the statistical tests.

Retrospective tests were performed on two sets of data - first on the 314 observations with RBC ratios of 200% to 300%, and then on the broader set of 480 (which include the 314) observations with RBC between 200% and 350% (see page 3 of Exhibit 1.) These sets are referred to as the "border" companies. Of the 314 companies, 30 had RBC ratios deteriorate to the CAL in the subsequent year, and of the 480, 39 deteriorated to the CAL in the subsequent year. The metrics set up test the ability of various measures to correctly predict which companies deteriorate in the subsequent year. Based on the life insurance type trend test, the proportion of companies that fell to the CAL in the subsequent year based on RBC ratio in the current year was examined. Of companies with RBC between 200% and 300%, nearly 10% will deteriorate to the CAL in the subsequent year, and in the next layer of companies with RBC between 300% and 350%, 5.4% will deteriorate to CAL in the subsequent year. However, the proportion of companies deteriorating to CAL drops off sharply for higher RBC ratios as shown in Exhibit I, page 3. Therefore, we concluded the trend test should apply only to companies with RBC between 200% and 350%, and performed the metrics on the two bands discussed above (200%-300% and 300%-350%).

The retrospective test asked two questions for the border companies. First, did RBC decline to CAL in the subsequent year? Second, how effective was the hypothesis and its measure(s) in predicting that outcome? The second question was broken into three metrics:

1. What proportion of border companies that deteriorated into CAL in the subsequent year were correctly predicted by the measure? This is called "Ratio of failing companies flagged" in Exhibit I.
2. What proportion of border companies that did not deteriorate to CAL in the subsequent year were falsely predicted to deteriorate? This is called "false alarms" in Exhibit I, and has also been referred to as false positives.
3. What was the overall percentage of correct predictions for all border companies? This is called the "effectiveness ratio" in Exhibit I, and is the sum of true positives and true negatives.

Note that it is easy to get a successful result of 100% on Question 1 simply by predicting all companies would deteriorate to CAL. This is close to the result generated by the life insurance type trend test approach. That is why it is important to balance the other two metrics above.

There is a fourth, more subjective, criterion - simplicity. It is important that a proposed trend test is easy to apply and understand, and the reasoning behind its effectiveness is self-evident.

The approach for the retrospective test was to set a threshold on the hypothesized measure. For example, in testing the combined ratio, a threshold of 120% was set. If a company with RBC in the

border of 200%-300% also had a combined ratio that was over 120%, they "failed" and were flagged. If the combined ratio was better than 120%, the company "passed." In all cases we optimized the threshold to achieve the best possible differentiation for the measure(s) examined, using a trial and error approach.

These thresholds and the associated measures are shown in Exhibit I.

While applying these several metrics may seem complex, in the end the same measure gave the best answer for most of the metrics, so balancing the metrics didn't become a significant issue.

Statistical Data

Once the working hypotheses were determined, each hypothesis was tested using statistical data. These were the statistical and retrospective tests discussed above. This analysis is referred to as being conducted at the "macro" level. The data was obtained from the NAIC and was at an individual company level. A large number of fields were obtained for testing, generally using five years of history ending at 12/31/2003. The most critical of these fields were RBC ratio, capital, underwriting results, reserve runoff ratio, measures of leverage such as premium, reinsurance recoverable, and reserves, and measures of liquidity such as invested assets. Other fields related to capital changes, such as dividends, capital contributions, and realized and unrealized gains were also selected in order to gauge the impacts of capital changes on RBC.

The NAIC data was benchmarked against industry aggregates from AM Best in order to reconcile the totals and verify the quality of our data sample. For 2003, a small percentage of companies had not yet passed the NAIC edits, and were not included in our data.

The data was scrubbed for invalid and extreme values. The scrub for retrospective tests is described above. For the statistical tests, both the entire database, and by-test information required data scrubbing in order to have appropriate data available. Screening for extreme values was more stringent on the statistical tests, since extreme values have much greater impact on statistical tests than on the retrospective test.

Results

Hypotheses

Intuitive hypotheses regarding how companies become troubled involve underwriting results, leverage, reserve inadequacies, and liquidity. After examination of industry examples of rapid RBC ratio declines, the following hypotheses were posited for testing:

1. ***Life Insurance Type Trend Test***-The life insurance type trend test approach would be effective for P&C, that is, past changes in RBC ratio are effective predictors of future changes.
2. ***Test Year Combined Ratio***-The combined ratio in the test year is an effective predictor of changes in RBC ratio in the subsequent year.
3. ***Three Year Combined Ratio***-The three-year average combined ratio evaluated in the test year is an effective predictor of changes in RBC Ratio in the subsequent year.
4. ***Test Year Runoff Ratio***-The one-year runoff was defined for this test as the one-year reserve development divided by the prior year reserve. The posit is that the one-year reserve development in the test year is an effective predictor of changes in RBC Ratio in the subsequent year.

5. ***Three-Year Runoff Ratio***-The three-year average runoff ratio, defined similarly to the one-year ratio above, is an effective predictor of changes in RBC Ratio in the subsequent year.
6. ***Gross Leverage at end of Test Year***-The Gross Leverage was defined for this test as the ratio of the sum of (Gross WP + Gross Reserves + Reinsurance Recoverable on Paid Loss and LAE) to surplus. The posit is that the Gross Leverage is an effective predictor of changes in RBC Ratio in the subsequent year.
7. ***Net Leverage at end of Test Year***-The net leverage was defined for this test as the ratio of the sum of (Net WP + Net Reserves) to surplus. The posit is that this is an effective predictor of changes in RBC Ratio in the subsequent year.
8. ***Composite Combined Ratio and Reserve Runoff***-The posit is that a test constructed that combines thresholds of combined ratio and one-year reserve runoff in the test year is an effective predictor of changes in RBC ratio in the subsequent year.
9. ***Composite Combined Ratio, Reserve Runoff, and Leverage***- The posit is that a test constructed that combines thresholds of combined ratio, one-year reserve runoff, and leverage in the test year is an effective predictor of changes in RBC ratio in the subsequent year.
10. ***Two-Tiered Combined Ratio***- (Used for the broader set of border companies only). This is the same as the one-year combined ratio test, but allows for a differing threshold for companies with RBC between 200%-300% and those with RBC between 300%-350%. After testing, a threshold set at 170% of the former, for the latter group worked well was found to work well.
11. ***Liquidity***-The level of liquid assets compared to policyholder funds is an effective predictor of changes in RBC ratio in the subsequent year.
12. ***Fraud***-While there is no doubt fraud has played a part in some historical rapid declines in RBC, there did not seem to be a formulaic was to test this as part of an RBC trend test. Therefore, this hypothesis was not tested.
13. ***Asset Valuations***-Similar to fraud, the inflation of asset values and subsequent restatement may have played a role in past RBC declines. However, this hypothesis was not tested since we did not see a way to test the quality of assets in a formulaic way as part of an RBC trend test.

Results of Testing

Life Insurance Type Trend Test-This approach is to differentiate companies based on the past RBC changes. We ran several statistical tests checking correlations between movements in RBC ratios between years. In all tests, the correlation was low, or had a different sign than expected. We also did a simple sign test, to see if previous changes in RBC ratio were in the same direction as subsequent. We combined test years 2001 and 2002. The changes were in the same direction only 41% of the time. These tests are shown in models 1 and 2 of Exhibit II.

Our retrospective test was not any more promising. Using an approach similar to the life insurance type trend test, we found an effectiveness ratio of only 42% (see Exhibit I, page 1). This means the use of the life insurance type trend test was accurate only 42% of the time in correctly predicting if border companies would fall to the CAL in the subsequent year. Most other approaches produced much better metrics.

Simple trends in RBC are often driven by factors external to the P&C insurance industry. In the years we examined, changes in asset market valuations played a significant role in changes in RBC ratios

overall. We noted that exposure to volatile assets is already accounted for in the RBC formula, and trends in asset markets are unpredictable.

Based on the above considerations, we concluded a trend test based on past RBC change, similar to the life insurance type trend test, would not be effective for P&C insurance companies.

Combined Ratio-This predictor showed good promise of providing predictive capabilities. We ran time series correlations on the combined ratio. The statistical tests are shown in models 3 and 4 of Exhibit II. These show significant positive correlation between years, consistently in the 25%-35% range. We noted that the test years of 2000-2002 were also periods of rapid change in the insurance pricing environment, which may lead to less predictability in our test years than other years. This strong correlation resulted in tests of all companies, and on tests of only companies with low RBC ratios.

The retrospective tests were also promising. The one-year tests consistently produced better metrics than the three-year tests. The simple one-year test produced the best result in all three of our metrics when compared to the results of all other approaches. In the group of companies with RBC between 200%-300%, the approach was a correct predictor 65% of the time, and in the broader group of companies (200%<RBC<350%), that measure improved to 71% when the two-tiered test was applied (see Exhibit I, pages 1 and 2). This 71% success ratio was achieved while still flagging 67% of companies falling to CAL in the subsequent year, and raising false alarms in only 26% of companies. These metrics were better than any other approach we tested.

The simple two-tiered one-year combined ratio test also met our criteria of being simple, and easy to calculate and apply. The connection between the approach, the threshold, and the reasons regulators would be concerned was straightforward.

There is good reason to believe a company with poor underwriting results in one year, will be adversely affected in the subsequent year. In-force business will continue to run out and affect results. It takes time to change a renewal book, whether pricing, terms of coverage, or quality of underwriting is the issue, or all three. The market will usually constrain actions a company can take, and managements tend to be reluctant to leave troubled markets entirely.

It is natural to ask whether this information is already incorporated into the RBC formula due to the premium risk factors. The capital arising from premium risk is calculated based on a longer-term history of ratios, and the outliers in that history. Capital should be sufficient to cover the risk that those outliers will repeat, and the RBC formula reflects this. A trend test covers a different "risk", one that is more likely to occur. The combined ratio test applies when a company has a book of business that is generating losses currently (hence the "trend"), and has predictive power since it takes time to change that book. In a way, the one year combined ratio test provides for a risk that is much more likely to occur (risk of the trend of losses continuing) rather than a more unlikely risk embedded in the RBC formula (that of future underwriting results hitting an outlying value).

Due to the considerations above, we found the simple two-tiered one-year combined ratio test to be the best predictor of future RBC declines of any that we considered.

Reserve Runoff Tests-We tested the ratio of one-year reserve development to the original reserve from the statutory statement as both an auto-correlated predictor, and as a predictor of RBC ratio changes. This test also showed promise in the statistical tests. Models 5 and 6 of Exhibit II show high year-to-

year correlations of reserve runoff, generally in the 30%-40% range. This is true of all companies together, and when only companies with lower RBC are used.

The retrospective tests showed good results also, especially the test using three-year average runoff ratio to predict RBC declines. However, these results were not as effective as the combined ratio tests. This can be understood by considering that reserve runoff is already included in the combined ratio, since both are reflected in calendar year underwriting results. It makes sense that the three-year test is more predictive. If a company experiences a one time reserve adjustment, perhaps due to large claims or adverse court decisions, that would not be predictive of future RBC declines.

Based on these considerations, we concluded that while reserve runoff appears to show promise for use in a trend test, it is not as effective as a combined ratio test.

Leverage Tests-While the initial expectations of the working group was that leverage would have predictive potential, the statistical tests and retrospective tests showed this was not the case. Statistical models 7 and 8 show analysis of the impacts of leverage on subsequent RBC changes. The statistical tests show only the mildest association between gross and net leverage and subsequent changes in RBC, and often with a sign counter to expectations. The retrospective tests, as shown in Exhibit I, show an overall effectiveness ratio of the prediction near or less than 50%, implying random guessing would be as effective.

While several high profile insolvencies involved companies with high leverage, the analysis shows that statistically, high leverage is not strongly associated with RBC declines. The working group retains the opinion that high leverage is associated with high risk, that is, higher variation in results. However, high leverage can lead to higher levels of reward if reserves are correctly stated, and underwriting results are profitable. Similarly, high leverage can lead to rapid RBC decline if reserves or underwriting results are troubled, as a number of highly publicized past failures have shown. However, the statistical tests show that high variation is not the same as a tendency to decline. Therefore, we concluded that a leverage test would not be effective as a trend test for the P&C RBC formula.

Liquidity Tests-Tests on the predictability of liquidity on subsequent RBC declines are shown in Exhibit II, model 9. The results show only a mild association, with low statistical significance, and sometimes the wrong sign. In the discussions of the working group, we felt this indicated that liquidity problems were more of a lagging indicator for a P&C insurance company, rather than a leading indicator, since adverse financial results leading to liquidity problems should already be reflected in reserve accounts. Since liquidity is not an area addressed by the RBC formula, the basis for the hypothesis is not strong, and the statistical models show little potential, we did not carry this hypothesis through into the retrospective tests. We concluded there is little potential for liquidity measures to perform an effective trend test for P&C insurance RBC.

Composite Tests-The original expectation of the working group was that a composite measure combining several statistically effective measures would likely score well in the retrospective test metrics, and might become our recommendation. If this was the case, the trade-off between complexity and effectiveness would become key.

However, the retrospective test results shown in Exhibit I showed no increase in effectiveness by combining measures, in fact, adding other measures to the combined ratio test tended only to decrease the effectiveness. It didn't add ability to flag companies with subsequent declines, but it did add substantial numbers of "false alarms".

We set up retrospective tests in two ways, first combining underwriting and reserve measures, and then adding leverage measures. We optimized the thresholds at levels somewhat less restrictive than tests with measures used alone, as otherwise a large majority of companies would be flagged.

Since composite measures produced retrospective results that were less effective than using the combined ratio alone, we concluded that a composite measure would not be the most effective approach for a P&C trend test.

Exhibit I

Outcomes of the Metrics from the Retrospective Test

Predictors and thresholds for companies with RBC 200%<RBC<300%

Thresholds shaded

	Threshold	Ratio of failing cos flagged	False alarms/total	Effectiveness Ratio	Total # of Companies
Life Insurance Type Trend Test	trend	63%	55%	42%	314
Test Year CR Ratio	-20%	73%	32%	65%	314
Three Year Average CR Ratio	-17%	67%	38%	59%	314
Test Year Runoff Ratio	7%	67%	46%	51%	314
Three Year Average Runoff Ratio	5%	43%	41%	54%	314
Gross Leverage, End of Test Year	650%	67%	53%	44%	314
Net Leverage, End of Test Year	350%	60%	48%	48%	314
Composite CR and Runoff	Composite	63%	38%	59%	314
Composite test cr/runoff and leverage	Composite	70%	46%	51%	314

Composite Thresholds

	Both cr/runoff and leverage	either CR of Runoff
Combined Ratio	-25%	-26%
Runoff	45%	30%
Gross Leverage	1300%	xxx
Net Leverage	850%	xxx

Exhibit I

Outcomes of the Metrics from the Retrospective Test

Predictors and thresholds for companies with RBC 200%<RBC<350%

Thresholds shaded

	Threshold	Ratio of failing cos flagged	False alarms/total	Effectiveness Ratio	Total # of Companies
Life Insurance Type Trend Test	trend	62%	47%	50%	480
Test Year CR Ratio	-20%	67%	30%	67%	480
Three Year Average CR Ratio	-17%	64%	36%	61%	480
Test Year Runoff Ratio	7%	46%	43%	53%	480
Three Year Average Runoff Ratio	5%	46%	37%	59%	480
Gross Leverage, End of Test Year	650%	56%	48%	48%	480
Net Leverage, End of Test Year	350%	54%	46%	51%	480
Composite CR and Runoff	Composite	64%	34%	63%	480
Composite test cr/runoff and leverage	Composite	69%	42%	56%	480
Two Tiered Combined Ratio Test	-20%	67%	26%	71%	480

The two-tiered combined ratio test is based on a threshold for companies with 200%<RBC<300% which is the same as the one tiered test, but for companies with 300%<RBC<350%, the threshold is increased by the tier factor, shown below.

Tier Factor 1.7

Outcomes of the Metrics from the Retrospective Test

Analysis of companies which deteriorate to Company Action Level in the year subsequent to the test year.

How many companies with 200 <RBC < 300 in the test year ended up below 200 in the subsequent year?	30	out of	314	or	9.6%
How many companies with RBC > 300 in the test year ended up below 200 in the subsequent year?	25	out of	3268	or	0.8%
And of the 25:					
How many companies with 300 <RBC < 350 in the test year ended up below 200 in the subsequent year?	9	out of	166	or	5.4%
How many companies with 350 <RBC < 400 in the test year ended up below 200 in the subsequent year?	4	out of	205	or	2.0%
How many companies with 400 <RBC < 450 in the test year ended up below 200 in the subsequent year?	3	out of	176	or	1.7%
Of the rest……..	9	out of	2721	or	0.3%

Analysis of NAIC Company Data to Evaluate Effectiveness of Life Insurance Type Trend Testing
Summary of Models Run on the Underlying Data

Model number: 1

Hypothesis: Simple trend test based on change in RBC ratio works for all companies.
Approach: Correlation matrix for change in RBC, sign test, 2000-2002.
Results:

	Chg RBC 01-00	*Chg RBC 02-01*
Chg RBC 01-00	1	
Chg RBC 02-01	-23%	1

Count 1 if sign changes	1,118
Total Count	2,244
Proportion	49.8%

Conclusion: Cannot reject, sign of correlation is wrong, sign test not significant.

Model number: 2

Hypothesis: Same as model 1, but only on companies with RBC ratio<250
Approach: Correlation matrix for this group of companies, and sign test, cos with RBC<250.
Results:

	Chg RBC 01-00	*Chg RBC 02-01*
Chg RBC 01-00	1	
Chg RBC 02-01	1%	1

Count 1 if sign changes	90
Total Count	145
Proportion	62.1%

Output of sign test run on 2002 and 2001

	2002	2001	2001-2002
Declines successfully predicted	37	11	48
Number of Declines	54	48	102
Ratio	69%	23%	47%
Improvements successfully predicted	18	40	58
Number of Improvements	91	63	154
Ratio	20%	63%	38%
Overall success ratio	38%	46%	41%

Conclusion: Cannot reject, correlation not significant, and sign test tends wrong direction.

Analysis of NAIC Company Data to Evaluate Effectiveness of Life Insurance Type Trend Testing
Summary of Models Run on the Underlying Data

Model number: 3
Hypothesis: Correlation exists between year to year underwriting results
Approach: Correlation matrix for all companies
Results:

	2000	*2001*	*2002*
2000	100%		
2001	25%	100%	
2002	23%	34%	100%

Conclusion: Fairly strong and positive correlation year to year.

Model number: 4
Hypothesis: Model 3, but with only the companies with RBC<250
Approach: Correlation matrix for all companies with RBC<250
Results:

	CR 2001	*CR 2002*
CR 2001	100%	
CR 2002	75%	100%

	CR 2000	*CR 2001*
CR 2000	100%	
CR 2001	-15%	100%

Re-run of 2000, 2001 with elimination of outlier cos 2031

	CR 2000	*CR 2001*
CR 2000	100%	
CR 2001	33%	100%

Conclusion: Fairly strong positive correlation year to year.

Model number 5
Hypothesis: Correlation exists between year to year reserve runoff
Approach: Correlation matrix for all companies reserve runoff
Results:

	Runoff 2001	*Runoff 2002*
Runoff 2001	100%	
Runoff 2002	33%	100%

	Runoff 2000	*Runoff 2001*
Runoff 2000	100%	
Runoff 2001	37%	100%

Analysis of NAIC Company Data to Evaluate Effectiveness of Life Insurance Type Trend Testing
Summary of Models Run on the Underlying Data

Model number 6

Hypothesis: Correlation exists between year to year reserve runoff for RBC<250
Approach: Correlation matrix for all companies reserve runoff for companies<250
Results:

	Runoff 2000	*Runoff 2001*
Runoff 2000	100%	
Runoff 2001	35%	100%

	Runoff 2001	*Runoff 2002*
Runoff 2001	100%	
Runoff 2002	29%	100%

Conclusion: Fairly strong positive correlation year to year.

Model number 7

Hypothesis: Correlation exists between high gross leverage and adverse RBC change
Approach: Calculate gross leverage and correlate with RBC change
Definition of Gross Leverage: (Gross WP+gross Reserves+Recoverable on Paid)/surplus
Average Gross Leverage from Benchmark Total: 2001 347% 2000 287%
Results:

	Gr Lev 2001	*Chg RBC 2002*
Gr Lev 2001	100%	
Chg RBC 2002	1%	100%

	Gr Lev 2000	*Chg RBC 2001*
Gr Lev 2000	100%	
Chg RBC 2001	-1%	100%

Reperformed test including only RBC<250

	Gr Lev 2000	*Chg RBC 2001*
Gr Lev 2000	100%	
Chg RBC 2001	-5%	100%

	Gr Lev 2001	*Chg RBC 2002*
Gr Lev 2001	100%	
Chg RBC 2002	-3%	100%

Analysis of NAIC Company Data to Evaluate Effectiveness of Life Insurance Type Trend Testing
Summary of Models Run on the Underlying Data

Reperformed test for RBC<250 as a regression to get significance level for 2001
SUMMARY OUTPUT

Regression Statistics	
Multiple R	5%
R Square	0%
Adjusted R Square	-1%
Standard Error	1.28
Observations	107

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.39	0.39	0.24	0.62
Residual	105	171.48	1.63		
Total	106	171.87			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	0.543	0.158	3.44	0%	0.230	0.856	0.230	0.856
Gr Lev 2000	(0.004)	0.008	(0.49)	62%	(0.020)	0.012	(0.020)	0.012

Analysis of NAIC Company Data to Evaluate Effectiveness of Life Insurance Type Trend Testing
Summary of Models Run on the Underlying Data

Reperformed test for RBC<250 as a regression to get significance level for 2002
SUMMARY OUTPUT

Regression Statistics	
Multiple R	3%
R Square	0%
Adjusted R Square	-1%
Standard Error	1.73
Observations	124

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.36	0.36	0.12	0.73
Residual	122	365.82	3.00		
Total	123	366.18			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	0.597	0.214	2.794	1%	0.17	1.02	0.17	1.02
Gr Lev 2001	(0.004)	0.011	(0.345)	73%	(0.03)	0.02	(0.03)	0.02

Conclusion: In just about every test, there is only the mildest association between gross leverage and changes in RBC, but it is negative in some cases as expected.

Analysis of NAIC Company Data to Evaluate Effectiveness of Life Insurance Type Trend Testing
Summary of Models Run on the Underlying Data

Model number **8**

Hypothesis: Correlation exists between high net leverage and adverse RBC change
Approach: Calculate net leverage and correlate with RBC change
Definition of Net Leverage: (Net WP+Net Reserves)/surplus
Average Net Leverage from Benchmark Total: 2001 194% 2000 170%
Results using all data points after scrubbing

	Net Lev 2000	*Chg RBC 2001*
Net Lev 2000	100%	
Chg RBC 2001	3%	100%

	Net Lev 2001	*Chg RBC 2002*
Net Lev 2001	100%	
Chg RBC 2002	4%	100%

And for points with RBC<250

	Net Lev 2000	*Chg RBC 2001*
Net Lev 2000	100%	
Chg RBC 2001	16%	100%

	Net Lev 2001	*Chg RBC 2002*
Net Lev 2001	100%	
Chg RBC 2002	1%	100%

Using regression to test significance

SUMMARY OUTPUT-Leverage as a predictor of RBC change, test year 2000, RBC<250%

Regression Statistics	
Multiple R	16%
R Square	3%
Adjusted R Square	2%
Standard Error	2.90
Observations	156

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	34.64	34.64	4.11	4%
Residual	154	1,296.50	8.42		
Total	155	1,331.14			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	(1.70)	0.27	(6.37)	0%	(2.23)	(1.17)	(2.23)	(1.17)
Net Lev 2000	0.06	0.03	2.03	4%	0.00	0.13	0.00	0.13

Analysis of NAIC Company Data to Evaluate Effectiveness of Life Insurance Type Trend Testing
Summary of Models Run on the Underlying Data

SUMMARY OUTPUT-Leverage as a predictor of RBC change, test year 2001, RBC values below 250%.

Regression Statistics	
Multiple R	1%
R Square	0%
Adjusted R Square	-1%
Standard Error	1.73
Observations	125

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.06	0.06	0.02	88%
Residual	123	369.82	3.01		
Total	124	369.89			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	0.55	0.19	2.86	0%	0.17	0.93	0.17	0.93
Net Lev 2001	0.00	0.02	0.15	88%	(0.03)	0.04	(0.03)	0.04

Modelers note on tests 7 and 8. A more rigorous data scrub was used in model 8, eliminating the companies with 0 or negative leverage in model 8, but not applying this standard in model 7. Impact on conclusions is negligible.

Analysis of NAIC Company Data to Evaluate Effectiveness of Life Insurance Type Trend Testing
Summary of Models Run on the Underlying Data

Model number **9**

Hypothesis: Correlation exists between overall liquidity and subsequent RBC change.

Approach: Correlation matrix between RBC change and overall liquidity, used 2001 test year, 2002 subsequent year.

Results:

All Values

	RBC Change	*2001 Current Liquidity*
RBC Change	1	
2001 Current Liquidity Inv Asse	-4%	1

Values below 300% RBC

	RBC Change	*'001 Current Liquidity*
RBC Change	1	
2001 Current Liquidity Inv	2%	1

Approach: Perform regression to obtain "t" test

SUMMARY OUTPUT-Regression statistics for regression of liquidity on RBC change for entire sample

Regression Statistics	
Multiple R	4%
R Square	0%
Adjusted R Square	0%
Standard Error	2.57
Observations	1670

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	15.60	15.60	2.36	12%
Residual	1668	11,037.72	6.62		
Total	1669	11,053.32			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	(0.45)	7.2%	(6.27)	0%	(0.59)	(0.31)	(0.59)	(0.31)
2001 Current Liqui	(0.02)	1.3%	(1.54)	12%	(0.04)	0.01	(0.04)	0.01

Analysis of NAIC Company Data to Evaluate Effectiveness of Life Insurance Type Trend Testing
Summary of Models Run on the Underlying Data

SUMMARY OUTPUT-Regression statistics for regression of liquidity on RBC change for RBC ratios below 300%

Regression Statistics	
Multiple R	2%
R Square	0%
Adjusted R Square	0%
Standard Error	1.65
Observations	208

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	0.15	0.15	0.06	0.81
Residual	206	560.51	2.72		
Total	207	560.66			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	0.38	0.16	2.41	0.02	0.07	0.68	0.07	0.68
2001 Current Liqui	0.02	0.07	0.24	0.81	(0.13)	0.16	(0.13)	0.16

Conclusion on Liquidity: Not a significant predictor. Note that the "all values" F stat appears significant, but the sign is in thw opposite direction from what we would expect. For only companies with RBC below 300%, the sign is correct, but the significance is very low, with the CI for the coefficient centered around 0.