

October 16, 2014

Supplemental Report

Proposed Transfer of Runoff Business from OneBeacon Insurance Group to Armour Group Holdings Limited

Comments on:

“RRC Review of Towers Watson Response to July 23, 2014 Hearing Record Regarding One Beacon Form A Filing” dated October 2, 2014

by

Allan Kaufman
FCAS, MAAA, FIA (Hon), CPCU

Allan Kaufman

I. INTRODUCTION

1. This is a supplement to my Analysis of Financial Capacity, dated October 16, 2014 (the “Kaufman Report”), to respond to Risk & Regulatory Consulting’s (“RRC”) “Review of Towers Watson Response to July 23, 2014 Hearing Record Regarding One Beacon Form A Filing,” dated October 2, 2014 (the “RRC Letter”). The three-page RRC Letter provided the following additional information:
 - The magnitude of the most adverse National Indemnity Company (“NICO”) simulation¹ from the Towers Watson (“TW”) modeling (“worst simulation”);
 - RRC’s analysis of that worst simulation; and,
 - More details on the extent of RRC’s and TW’s use of sensitivity testing for certain issues.
2. With the benefit of that additional information I have prepared this supplement regarding the variability in the TW modeling and the use of sensitivity testing.

II. DOCUMENTS RELIED UPON

3. In addition to the documents I relied upon to form my opinions and conclusions in the Kaufman Report,² I have also relied upon the following documents to form my responses discussed throughout this supplemental report:
 - The RRC Letter (Cumulative Log, Document 111)
 - The Pennsylvania Insurance Department (“PID”) Letter, dated October 6, 2014 (Cumulative Log, Document 110)

III. VARIABILITY OF THE TW OUTCOME DISTRIBUTION

4. In RRC Points 1-3, RRC is addressing, at least in part, detailed issues I raised regarding the extent to which the TW distribution of outcomes, particularly the NICO outcomes, has sufficient variability, i.e., does the distribution of outcomes represent the chance of extreme simulations adequately enough to be relied on for assessing the chance of Payment Failure,³ or assessing capital adequacy generally for the Runoff Companies.⁴

¹ Refers to losses ceded under the NICO reinsurance agreement, as defined in: Cumulative Log, Document 082 (RRC Reserve Report, p. 6).

² Kaufman Report, Exhibit 2.

³ When the insurer runs out of assets. This is the point when there is “no money left in the bank.” (Kaufman Report, p. 20.)

⁴ Runoff Companies is defined in the Kaufman Report, p. 1.

5. Based on my review of the TW Response to Public Comments, dated August 11, 2014 (the "TW Reply"),⁵ I understand that TW believes that their outcome distribution has sufficient variability. However, the TW Reply is not sufficiently detailed to allow me, even with my level of expertise, but with access to only the public information, to adequately assess the TW work.
6. I understand that the RRC work is intended to evaluate TW's assumptions, methods, and results with access to confidential information unavailable to me. However, RRC does not report why RRC believes the TW assumptions and methods that determine the variability in the outcome distribution are appropriate. Rather, RRC reports on the frequency of TW-modeled outcomes at various probability levels. Therefore, I understand that the RRC conclusions are based heavily, if not totally, on the observed distribution of TW-modeled outcomes, rather than on TW's underlying assumptions and methods.
7. RRC's observations on the frequency of TW's calculated adverse outcomes are shown in columns A and B of Table 1 on the following page.

⁵ Cumulative Log, Document 108 (TW Reply).

Table 1
Distribution of NICO Outcomes

	(A)	(B)	(C)	(D)	(E)
	Percentile Probabilities	NICO Gross per TW (millions)	Ratio to Mean	Equitas Analysis (millions)	Ratio to Mean
(1)	Mean	\$ 843	1.00	\$ 7,800	1.00
(2)	90 th	\$1,097	1.30	\$11,600	1.49
(3)	95 th	\$1,189	1.41	\$13,300	1.71
(4)	99 th	\$1,405	1.67	\$18,100	2.32
(5)	99.9 th	N/A	N/A	\$27,800	3.56
(6)	99.99 th	\$1,686	2.00	N/A	N/A

Sources:

B1 = Estimated, as TW and RRC do not provide that value.⁶

B2, B3, B4 = Cumulative Log, Document 083 (RRC Stochastic Report, p. 10).

B6 = Cumulative Log, Document 111 (RRC Letter, p. 2).

Col. C = Col. B / B1

Col. D = Equitas Independent Expert Report, p. 136, available at <https://www.equitas-partvii.co.uk/Equitas-IndependentExpertsReport.pdf>.

Col. E = Col. D / D1

8. Column C shows the relative size of outcomes at various percentile probability levels by expressing each outcome to the ratio to mean value. For example, the value 1.67 in row 4 shows that the TW modeling of NICO lines implies that outcomes 67% or more above the mean value arise in 1-in-100 simulations, i.e., at the 99th percentile, or, equivalently, in 100 of the 10,000 TW simulations.
9. The value 2.00 in column C, row 6 is based on the following statement in the RRC Letter:

[T]he worst scenario in the stochastic model (with respect to NICO losses) shows gross NICO losses more than twice^[7,8] the level of the mean scenario. We note that such outliers show significant stressing by Towers.⁹

⁶ Calculated as:

- \$668 million reserves ceded to NICO (Potomac 2013 Annual Statement Note 23F); plus,
- \$230 million = \$198.3 million, the remaining limit within the NICO cover (Potomac 2013 Annual Statement, Note 33; Cumulative Log, Document 080 (TW Stochastic Report, p. 6).), plus \$31.7 million, or the reserve increase indicated by the TW central estimate, i.e., the \$10 million increase in TW's central estimate, plus \$21.7 million, or the amount that TW's central estimate exceeds the NICO limit of coverage (Cumulative Log, Document 080 (TW Stochastic Report, p. 6) and TW Reserve Report as of September 30, 2012, December 31, 2012 and March 31, 2013, p. 12.) less,
- \$55 million of payments or other reserve reductions on NICO, estimated as the NICO share of the \$85 million reserve reduction on NICO (64.4%) and General Reinsurance Corporation (35.6%) contracts combined, between December 31, 2013 and June 30, 2014. (Potomac 2013 Annual Statement, Note 23F; Potomac June 30, 2014 Quarterly Statement, p. 3, line 2501, difference between columns 2 and 1.)

⁷ On p. 1 of the RRC Letter, RRC states "nearly twice" rather than "more than twice." My comments apply in either case. (Cumulative Log, Document 111 (RRC Letter, p. 1).)

⁸ Obviously, to determine the specific probability of an "at or above twice the mean" event in the TW model would require more scenarios. However, I believe the information about this one observation is important because (a) RRC

10. Given that TW has modeled 10,000 simulations, the “worst scenario” would be a 1-in-10,000 event, or equivalently, an event at the 99.99th percentile.
11. In accepting the TW result, RRC is accepting that an outcome twice the size of the current TW central estimate has a probability of 1-in-10,000, and that the probability of outcomes at three times the mean or higher are essentially zero. For example, RRC is accepting that the probability of the event “claims at least twice the current TW central estimate (the mean)” is lower than the probability of a Triple-A bond default over two or more years.¹⁰ The chance of a bond initially rated Triple-A defaulting over 15 years is approximate 1%, or 1-in-100. The default probabilities for lower rated bonds are higher.
12. While the effect of an outcome twice the TW central estimate might be a “significant[ly] stressing” event, as RRC says, RRC does not explain what benchmarks or other methods it has applied to assess whether the 1-in-10,000 probability implied by the TW model is more reasonable than a 1-in-100, 1-in-1,000, or 1-in-1,000,000 probability. Moreover, RRC has not documented whether it requested any relevant benchmarks from TW, as TW has likely done stochastic analyses for other clients in the US and UK.
13. To provide a relevant benchmark, columns D and E in Table 1 show the results of modeling similar business, i.e., the Equitas transaction, as outlined in the Kaufman Report. These values are the result of modeling done by actuaries for Equitas (Equitas in the capacity analogous to OBIG), reviewed in detail by me as the Independent Expert (a role analogous, but with much broader scope, to the role of RRC) (the “Equitas Model”). The results of the Equitas Model were further reviewed by PricewaterhouseCoopers (consulting for Equitas) who compared the results to a summary of the confidence levels associated with asbestos and pollution modeling used in the analyses of business transfers¹¹ (analogous to the Ancillary Transactions and the Proposed Transaction¹²) by their actuaries in the London market.
14. It is important to note that the values in column C are lower than the values in column E. This shows that the distribution of outcomes from the TW model is not as variable as the

uses it; (b) the RRC discussion of the value illustrates the manner in which RRC has evaluated (on the public record) the TW modeling; (c) knowing the value of the highest simulation tells us that there are zero simulations with high simulated claim amounts; and (d) the observed event “zero claims at or above 2.0 times the mean” from the TW modeling can be compared to the event “zero claims at or above 2.0 times the mean” implied by the Equitas Model. The Equitas Model (Table 1, columns D and E, lines 3 and 4) implies 500 claims (95th percentile out of 10,000 simulations) for the event “claims at least 1.71 times the mean” and 100 claims (99th percentile out of 10,000) for the event “claims at least 2.32 times the mean.” By calculating the average weighted towards the 99th percentile, 150 or 200 claims out of 10,000 trigger the event “claims at least twice the mean.” That is, the Equitas Model expects 150 or 200 claims for the range of values where the TW model shows zero claims.

⁹ Cumulative Log, Document 111 (RRC Letter, p. 2, item 3).

¹⁰ Standard and Poor’s Rating Services, Default, Transition, and Recovery: 2013 Annual Global Corporate Default Study and Rating Transitions, March 19, 2014, Table 24, p. 56.

¹¹ “PwC has compared the results of the Equitas model to results that PwC prepared for other clients, relating to liabilities similar to those of Equitas, which were used for the purposes of Part VII of FSMA transfers and other transactions.” “PwC concluded that the Equitas model, with the base assumptions, produced slightly higher values for the statistics in Table 6-4 than those of their benchmarks.” (Equitas Independent Expert Report 2009, p. 136, 6.4.29 and 6.4.30, available at <https://www.equitas-partvii.co.uk/Equitas-IndependentExpertsReport.pdf>.)

¹² Both terms are defined in the Kaufman Report, p. 1.

distribution of outcomes from the Equitas Model. For the outcome with “claims at least twice the mean,” the expected probability from the Equitas Model is 1-in-100, not 1-in-10,000. More generally, across all probability levels, Table 1 demonstrates that if the Equitas Model had been applied to the Runoff Companies’ data, there would be more failures than predicted by the TW modeling.

15. Thus, there appears to be one or more assumptions in the TW model causing the results to be less variable than the Equitas Model. There is no information in the public record to indicate that RRC has investigated why the TW assumptions are more optimistic than the Equitas Model or what research RRC has done in comparing the TW assumptions to any standard. There is not enough detail in either the TW or RRC public reports to allow me to assess what that might be.
16. RRC further states, “We also note that the stress scenarios with the highest assumed loss development would cause solvency problems for the ‘as is’ runoff as well as the ‘Form A’ proposed structure.”¹³ However:
 - Even assuming that statement is true,¹⁴ it has no relevance as to whether the TW probability for that outcome is appropriate.
 - RRC does not define “as is” runoff, but I assume RRC means that to be OBIC at December 31, 2013. In that case:
 - The RRC statement regarding “solvency problems” arising from claims at twice the mean appears to be false. OBIC at December 31, 2013 has surplus of \$866 million¹⁵ and the additional liability would not be paid for many years.¹⁶
 - Thus, according to RRC, the NICO liabilities have a 0% chance of Payment Failure as of December 31, 2013, but contribute substantially to the nearly 12%¹⁷ Payment Failure probability in the event the Proposed Transaction were approved.
 - Thus, also according to RRC, policyholders of the Runoff Companies would be disadvantaged by the Proposed Transaction.

¹³ Cumulative Log, Document 111 (RRC Letter, p. 2, item 3).

¹⁴ TW and RRC have not reported that they have modeled an “as is” case.

¹⁵ OBIC 2013 Annual Statement, p. 3.

¹⁶ An increase in liabilities of \$875 million, or the mean of \$843 million (B1 in Table 1), plus the \$31.7 million increase to the TW central estimate (Refer to Footnote 6), would not cause Payment Failure. With reserve discounting, as assumed by TW for the Runoff Companies, the increase also would not cause technical insolvency, or the time when assets would not be sufficient to pay for estimated liabilities. I have not examined this in enough detail to assess whether there would be an indicated capital inadequacy for certain regulatory purposes.

¹⁷ The complement of the 88.28% “Successes after 30 years” rate. (Cumulative Log, Document 080 (TW Stochastic Report, Table 1, p. 5).)

IV. SENSITIVITY TESTING

17. TW, RRC and FTI seem to have a communication gap on the subject of sensitivity testing. In the July 18th FTI letter,¹⁸ I proposed that more sensitivity testing was appropriate to evaluate the extent to which certain parameters affected the Payment Failure probability. I did not propose that sensitivity testing replace the stochastic modeling, but rather that it complement it.
18. TW and RRC also seem to agree that sensitivity testing is useful in complementing the stochastic modeling process. For example, the TW Stochastic Report¹⁹ shows the effect three different investment strategies have on Payment Failure probabilities. I consider that work to constitute a sensitivity test.
19. Further, as now disclosed in the RRC Letter, RRC asked TW:
 - To do sensitivity testing regarding the “asset yields slightly lower than their baseline scenario.”²⁰ It appears TW provided the test, although RRC does not report the results.
 - “[I]f they [TW] had done any sensitivity testing with regard to the release of the SODs [Securities on Deposit]. They [TW] replied that they had . . .” and that the effect “did not significantly affect the model’s results.”²¹ The quantitative results are not reported by TW or RRC.
20. It appears that TW and RRC do not object to sensitivity tests as a general matter, but to the criticism that the TW Stochastic Report was inadequate because it did not appear to reflect any sensitivity testing. I understand that TW has explained that they believe the scope of their assignment did not require sensitivity testing.
21. Regardless of the details of the scope of TW or RRC work, the ultimate decision to be made by the PID is whether to approve, and if so, on what terms or conditions, the Ancillary Transactions and Proposed Transaction.
22. Based on my experience in business transfers in the UK, including my work as the Independent Expert for the Equitas Transfer, and in working with the UK insurance regulator in that context, I understand that it is important for the decision-maker (the PID in this case) to have information on the extent to which variations in the professional judgments drive the assessment of impact of the transaction on its stakeholders.
23. Sensitivity testing is one way to provide that information. For the decision maker, it is one situation if variations in professional judgments do not change the Payment Failure

¹⁸ Cumulative Log, Document 089 (AK Public Comment Letter, Exhibit 1).

¹⁹ Cumulative Log, Document 080 (TW Stochastic Report).

²⁰ Cumulative Log, Document 111 (RRC Letter, p. 2, item 1).

²¹ Cumulative Log, Document 111 (RRC Letter, pp. 2-3, item 4).

probability by more than 100 basis points, i.e., from 12%²² to between 11% and 13%. It is a different matter if alternative professional judgments might change the 12% Payment Failure to between 5% and 25%.

24. I believe the following sensitivity tests are relevant to the PID's evaluation of the Proposed Transaction, and should be reported to the PID and the public:

Three sensitivity tests that I believe have been done, or are straightforward:

- The Payment Failure probability indicated by the change in "asset yields slightly lower than their baseline scenario,"²³ work already completed by TW.
- The Payment Failure probability indicated by other sensitivity tests, if any, done by TW, but not yet disclosed.
- The Payment Failure probability indicated if the investment strategy needed to match the criteria given to Excalibur Reinsurance Corporation, i.e., to invest in "highly rated securities" and no equities.²⁴

Two sensitivity tests that are likely key drivers for the Payment Failure probabilities:

- The Payment Failure probability if the central estimate for NICO lines were 10% higher or 10% lower than currently estimated by TW.²⁵
- The Payment Failure probability if there were an increase in a key parameter or parameters that would increase the variability (e.g., the standard deviation) of the distribution of NICO lines outcomes by 25%.²⁶

One sensitivity test based on the judgments of TW and RRC:

- As TW and RRC are closest to the details of the modeling, the PID should request TW and RRC to identify the small number of other variables and professional judgments, if any, that would have more effect on the Payment Failure probability than the two tests immediately above.

25. These sensitivity tests should be applied to the technical insolvency analysis requested by the PID in its October 6, 2014 letter,²⁷ as well as to the Payment Failure analysis.

²² Ignoring the fact that I believe the 12% Payment Failure probability implies that the Proposed Transaction is not appropriate, regardless of other issues.

²³ Cumulative Log, Document 111 (RRC Letter, p. 2, item 1).

²⁴ Kaufman Report, p. 30, ¶ 100.

²⁵ This "10%" sensitivity test is an alternative to a request for a test of the Payment Failure probability indicated by the high or low estimate for NICO lines. If I understand TW's reply to a number of my questions, TW does not have separate assumptions underlying the low, central and high NICO estimates. Instead, TW may have selected the low and high NICO reserve estimates considering the range of outcomes given the central estimate.

²⁶ The limited TW description of its assumptions and method does not allow me to be more specific about which parameter(s) it should adjust.

²⁷ Cumulative Log, Document 110 (PID Letter, p. 3, item 13).

V. CONCLUSION

26. The information provided in the RRC Letter provides further evidence that:
 - The Payment Failure probability is higher than the 12% indicated by TW.
 - Runoff Companies policyholders had more security at December 31, 2009, or the status quo in the Kaufman Report, and at December 31, 2013, the “as is” runoff period from the RRC Letter, than would be the situation in the event that the Proposed Transaction is approved.
27. The PID and the public will benefit from sensitivity tests proposed in Section IV.