Appraising Life Office Valuations

Angus S. Macdonald
Department of Actuarial Mathematics and Statistics
Heriot-Watt University
Edinburgh
EH14 4AS
United Kingdom
Telephone : 031-451-3202
Fax : 031-451-3249
Email : angus@cara.ma.HW.ac.uk

Summary

This paper uses a simulation model of a participating life office to measure and compare the effectiveness of different methods of solvency valuation. It develops further the concept of adequacy based on simulated run-off tests allowing for interactions between assets and liabilities. The main aim is to compare some different traditional approaches to valuation - net premium, gross premium, active and passive bases - and to assess their suitability under modern conditions.

Appréciation des évaluations de solvabilité des compagnies d'assurance vie

Angus S. Macdonald
Département de mathématiques et de statistiques actuarielles
Heriot-Watt University
Edimbourg
EH14 4AS
Royaume-Uni
Téléphone : 031-451-3202
Télécopie : 031-451-3249
Email : angus@cara.ma.HW.ac.uk

Résumé

Nous utilisons dans cette étude un modèle de simulation d'une compagnie d'assurance vie participante pour mesurer et pour comparer l'efficacité des diverses méthodes d'évaluation de la solvabilité. Nous explicitons en détail le concept de l'adéquation sur la base de tests simulés permettant des interactions entre l'actif et le passif. L'objectif visé dans l'étude est de comparer diverses approches traditionnelles à l'évaluation - prime nette, prime brute, bases actives et passives - et d'évaluer leur validité dans les conditions actuelles de la conjoncture.

1. Introduction - adequacy and solvency

A previous paper [Macdonald 1993] defined the *adequacy* of a life office. "Adequacy" meant the ability to close to new business at a given time, and run off the in-force business without incurring a deficit.

The idea of projecting run-offs to test the fitness of a life office is almost as old as life offices themselves - indeed it underlies all prospective valuation methods - so there is nothing intrinsically new about adequacy. However, traditional approaches use very simple assumptions, such as constant interest rates and fixed bonus rules. These assumptions comprise a *model* of the future. Until recently, the question of the model's realism was hardly worth broaching; there was no practical alternative. Now, however, it is possible to make projections using different, possibly more realistic, models, so the traditional valuation model can be subjected to scrutiny.

We may ask, if we project run-offs using a realistic (we hope) model, how do the results compare with those obtained under the traditional valuation model? Put another way, does a traditional solvency valuation lead to the "correct" offices being closed down by the supervisors?

This paper uses simulation methods to test several traditional solvency valuations. Section 2 sketches the background to the paper; a fuller account can be found elsewhere. Section 3 describes the model's assumptions, effectively the model of the future against which the traditional valuation models are compared. Section 4 describes the valuation bases which are considered, and Sections 5 and 6 present some results.

2. Background

Often, the ability to run off an office's in-force business means the same as the office being "able to meet its guarantees", and is therefore described as "solvency". However, if the guarantees are not pre-determined, but are largely at the discretion of the office's managers, solvency on the basis of the guarantees is too weak; such an office should recognise that its effective liability might be in excess of its guaranteed liability.

In the U.K., with-profits business has evolved into such a state. The conventional wisdom during the past few decades has been to invest the funds backing with-profit policies largely - in some cases entirely - in equities and similar assets. Given the U.K.'s propensity to suffer bouts of inflation, this strategy may be justifiable, but it causes other difficulties. It is risky to distribute equity price "gains" in reversionary form, leading to increased reliance on terminal bonus. The policy's accumulated asset share is commonly, but not universally, taken as a measure of the "fair" payout at maturity, and is often used to set terminal bonus rates.

The Insurance Companies Act, recognising that bonus can form a large proportion of policyholders' benefits, requires offices to meet their policyholders' "reasonable expectations", (usually referred to as "PRE") thus setting a standard above strict solvency. Unfortunately, PRE is not defined. One view of PRE is given by Brindley et. al. [1990].

The fact that statutory minimum valuations are based upon guaranteed liabilities, which might be held back in the interests of investment freedom, makes it hard to *prescribe* safeguards of PRE. Policyholders (unknowingly) pay premiums with the implicit instruction "just do your best with this and give me something back at the end of the day". How managers use the great discretion which they have, or in other words the *strategies* which they pick, is then a crucial determinant of PRE.

Solvency valuations may be strengthened to make allowance for liabilities in excess of the guarantees, but even if such allowance is explicit, as in a bonus reserve valuation, it remains a test based upon a static model of the future in which managers display unwonted inertia. This matters because the outcome depends so very much upon what the managers decide to do.

Adequacy for modelling purposes might be defined in the following terms: "*Define an office to be "adequate" at a given time if, upon being closed at that time, and not thereafter departing too far from its intended strategies, it has surplus assets after the last policy has matured.*" The key element is the adherence - in spirit at least - to intended strategies. This is the embodiment of PRE. In practice it may be difficult to discover any intended strategies, but modelling may at least illuminate the impact of different strategies. In this paper, the focus is on different solvency tests rather than on different strategies; adequacy is used to measure the relative effectiveness of several valuation tests of solvency.

3. A description of the model

The office modelled writes 10-year with-profit endowment policies with a fixed annual premium. For simplicity, mortality and expenses are ignored. The key elements of the model are its strategies for asset allocation and bonus distribution, which are similar to those described in [Macdonald 1993], and will only be sketched here.

The asset allocation strategy is based on the sensitivity of the U.K. statutory minimum valuation basis to the current yields on gilts and equities. The maximum valuation interest rate is a weighted mean of these current yields, excluding any allowance for price gains in the case of equities, with the result that investment in equities tends to reduce the valuation interest rate and worsen the statutory solvency position. In addition, the office must be able to set up statutory minimum reserves after a fall of 25% in equity prices and a ±3% shift in gilt yields. (This resilience test, as it is known, has been modified recently but the work was based on the parameters quoted above.) Therefore an office in danger of failing the combined solvency/resilience test may be able to pass by switching assets from equities to gilts. This feature of the U.K. minimum solvency standard has been discussed by Ross [1989] and by Ross and McWhirter [1991]. The strategy used by the model is to invest 100% in equities as long as these tests can be passed, but to switch to gilts if required, to just the extent needed to pass the tests.

Reversionary bonus rates are set by projecting the asset shares at maturity of the in-force policies, at rates of interest determined as a weighted mean of 5-year (geometric) moving averages of the yields on the assets underlying the policies. Dividend growth is such that the resulting interest rate is in fact close to a 5-year moving average of the gilt yields. Bonus rates are declared which would allow the payment of a 25% terminal bonus, averaged over the in-force policies. In line with common practice, extreme changes in bonus rates are not permitted; the maximum proportionate increase and decrease are 25% and 20% respectively.

These strategies are modified after closure to represent the more cautious management which might be appropriate in a closed fund. Investment in equities is limited to 50%, the terminal bonus target is doubled to 50% and the maximum proportionate cut in reversionary bonuses is doubled to 40%.

Terminal bonuses are determined by paying out 100% of the smoothed value of the assets underlying the asset share at maturity. The asset values are smoothed by valuing future income using 5-year moving averages of the net gilt yield (for gilts) or the net dividend yield (for equities). Investigation shows that this asset smoothing is unbiased and unskewed. In addition, limits are placed on the changes in the maturity values from year to year, of +10% and -9.09%. These correspond to methods which it has been suggested are in use in some U.K. life offices.

To assume that the office writes nothing but 10-year policies is not particularly realistic. However, short term business is usually more vulnerable to adverse conditions than long term business, so such a model emphasises features which in practice would be harder to discern.

The in-force is built up by running the model for 40 years in stable (deterministic) conditions, so that at $t=40$ the office has reached a state of steady expansion. New business is assumed to grow at the rate of Retail Price Inflation, both before and after $t=40$. The stochastic projections after $t=40$ are made using the Wilkie asset model with the Reduced Standard parameter set [Wilkie 1984]. 1,000 investment scenarios were generated using this model, and these were used for each experiment, so that all can be compared sample path by sample path.

4. The valuation bases

This section describes the valuation bases which will be compared in Sections 5 and 6. The first pair of valuation bases are *static*; that is they are not varied as conditions change.

BASIS 1 is a gross premium basis, using an interest rate of 5% and a bonus loading of 2.5%. This is, in fact, identical to the premium basis, so it represents the most common practice in E.C. countries of valuing on the premium basis, adapted to the method of premium calculation common in the U.K..

BASIS 2 is a net premium basis, using an interest rate of 3%. This is broadly comparable in strength to Basis 1. No Zillmer is used, since there are no expenses.

The U.K. is unusual among E.C. countries in not requiring valuations to be carried out on the original premium basis, but allowing a more active approach. The statutory minimum basis itself is an example; it is a net premium basis in which the maximum interest rate is 92.5% of the net running yields on equities or the net redemption yields on gilts, weighted by market value. In addition, money to be invested more than 3 years in the future cannot be assumed to earn more than 7.2% gross. This is the third basis which we test.

BASIS 3 is close to the U.K. statutory minimum basis, described above. The 7.2% restriction on gross yields for investments made more than 3 years in the future is represented by a limit of 5.5% on net yields, since the office is assumed to pay tax.

The U.K. statutory minimum basis poses a problem for offices which invest heavily in equities (which means most with-profit offices). Because growth in dividends or prices is excluded from the calculation of the maximum interest rate, the basis can be harsh in some conditions. The fourth basis which we test is dynamic, like the statutory minimum basis, but it takes a different approach to the balance between caution and permissiveness.

BASIS 4 is a net premium valuation basis in which the interest rate is half the current net redemption yield on gilts. There is no further restriction on the yields assumed in respect of future investments.

The ratios of assets at market value to liabilities (A/L ratios) on each of these bases, at time t=40 just before the stochastic projections begin, are as follows :

BASIS 1	1.203
BASIS 2	1.161
BASIS 3	1.153
BASIS 4	1.167

The ratio of assets to asset shares (A/AS ratio) at time t=40 is 1.

5. Methods and some results

In this section the methods used are described, and some results are given.

(1) The model office is subjected to the 1,000 different futures generated by the Wilkie model. The projections are for 20 years. The office continues to transact new business, and its liabilities are calculated each year on several valuation bases. The object of this step is to find out which offices fail the various valuation tests and at what times.

(2) The second step is to take each of the 1,000 futures and to close the office in each of the next 20 years, running off the in-force business. Thus, 1,000 simulations are run in which the office is closed after 1 year, 1,000 simulations are run in which the office is closed after 2 years and so on. The object of this step is to find out which offices suffer a deficiency on closure during the 20 year time horizon.

(3) The third step is to compare the results of the previous two

steps. If we suppose that an office in (1) would be closed down upon failing a solvency valuation for the first time, (2) tells us whether the valuation correctly identified offices in difficulties. Moreover, by identifying in (2) all the offices which would ever have shown a deficit had they been closed down, we can see whether the valuation missed any inadequate offices.

The immediate effect of the asset switching algorithm is to switch part of the fund into gilts, as the office is unable to satisfy the resilience test at time t=40. Figure 1 shows the quartiles of the proportion of the fund invested in equities, assuming the office remains open to new business.

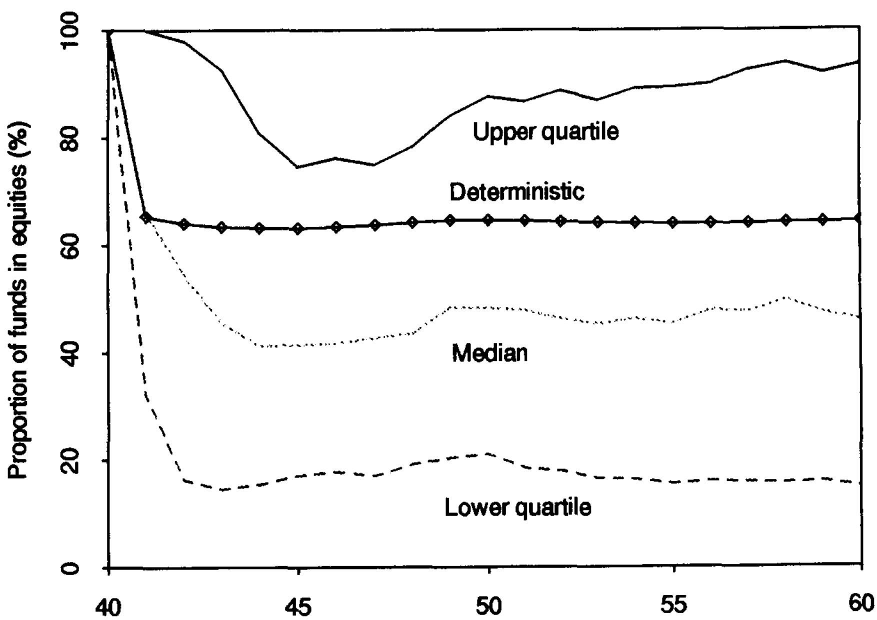

Figure 1

In 864 of the 1,000 projections (86.4%) the office was inadequate at some time during the 20 years t=41 to t=60. Compare this with the number of insolvencies under Bases 1 to 4.

Inadequacies	864
Insolvencies under Basis 1	724
Insolvencies under Basis 2	923
Insolvencies under Basis 3	468
Insolvencies under Basis 4	905

In each case, insolvency is taken to mean that the market value of the assets has fallen below the total policy values, or more concisely that the A/L ratio is less than 1. These overall figures do not show how accurate, or otherwise, the solvency valuations might be. Make the following definitions : a *correct* closure is an office which is insolvent and inadequate during the 20 years (not necessarily at the same time); a *wrong* closure is an office which is insolvent at some time but never inadequate; a *missed* closure is an office which is inadequate at some time but never insolvent. (The terminology is suggestive of closure following insolvency.) Then the four solvency valuations perform as follows :

TABLE 1

Comparison of numbers of inadequate and insolvent offices during a 20 year period

	Basis 1	Basis 2	Basis 3	Basis 4
Inadequacies	864	864	864	864
Total closures	724	923	468	905
Correct closures	658	816	455	805
Wrong closures	66	107	13	100
Missed closures	206	48	409	59

Even given a high proportion of inadequate offices, some solvency tests appear weak. They are, however, very sensitive to the addition of any solvency margins. A solvency margin expressed as a percentage of the liability is equivalent to the application of the solvency test using an A/L ratio other than 1. One way to show the sensitivity is to find the A/L ratio which, if used as the criterion of solvency, results in (nearly) equal numbers of wrong closures and missed closures. These and their effects are shown below.

TABLE 2

Solvency criteria at which errors of solvency valuations are (almost) equal

	Basis 1	Basis 2	Basis 3	Basis 4
A/L criterion	0.985	1.025	1.057	0.990
Inadequacies	864	864	864	864
Total closures	863	864	863	864
Correct closures	769	770	778	776
Wrong closures	94	94	85	88
Missed closures	95	94	86	88

Despite the differences when the solvency tests are applied using the same (A/L = 1) criterion, they show very similar results above. Notice the significant changes brought about by very small changes in the criterion of solvency. insolvencies. Under Basis 3, a change of 0.057 to the criterion, very little more than the E.C. solvency margin, almost doubles the number of insolvencies.

The previous figures relate to the whole projection period of 20 years. Calling a closure "correct", therefore, does not imply that it occurred at the same time as inadequacy, or even anywhere near it. Figure 2 shows the numbers of inadequate offices in each year t=41 to t=60, and the number insolvent under each of Bases 1 to 4 in each year.

At any time about half the offices are inadequate, as would be expected since (i) the aim is to pay out 100% of asset shares in the long run without additional assets to begin with, and (ii) the smoothing applied to asset prices is unbiased.

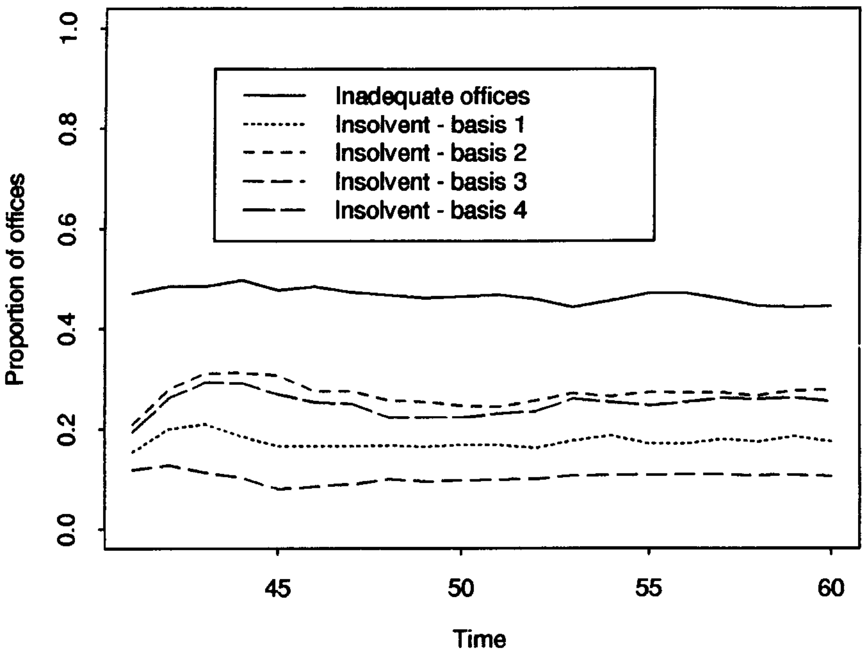

Figure 2

Although Figure 2 suggests lower levels of insolvency than of inadequacy, we know that about 80% of offices are insolvent at *some* time under three of the valuations. Inspection of the outputs (not shown) confirms that offices suffer temporary bouts of insolvency. Figures 3 to 6 show, for each solvency test, the numbers of offices each year which are adequate and solvent, adequate and insolvent, inadequate and solvent, and inadequate and insolvent.

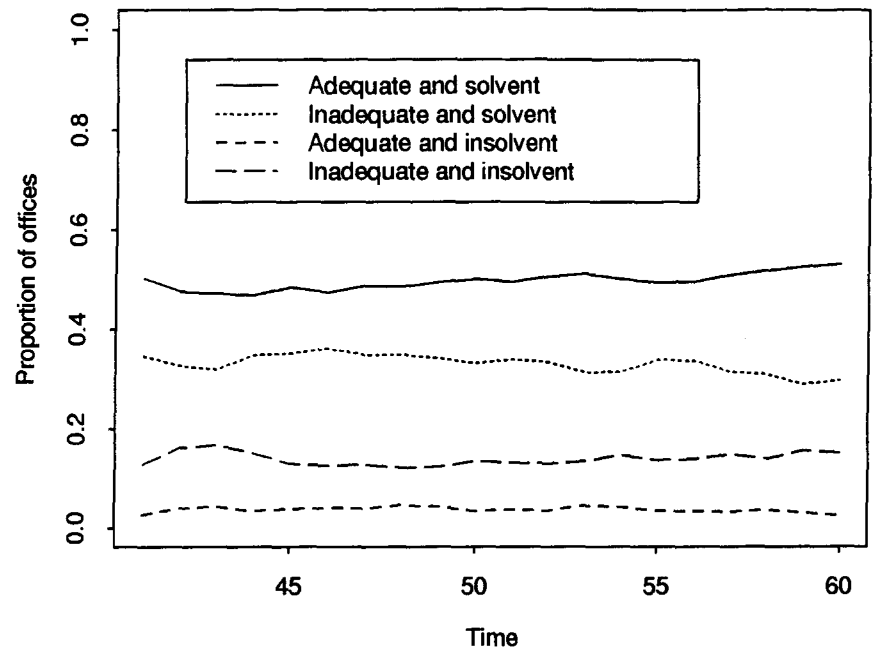

Figure 3 - Basis 1

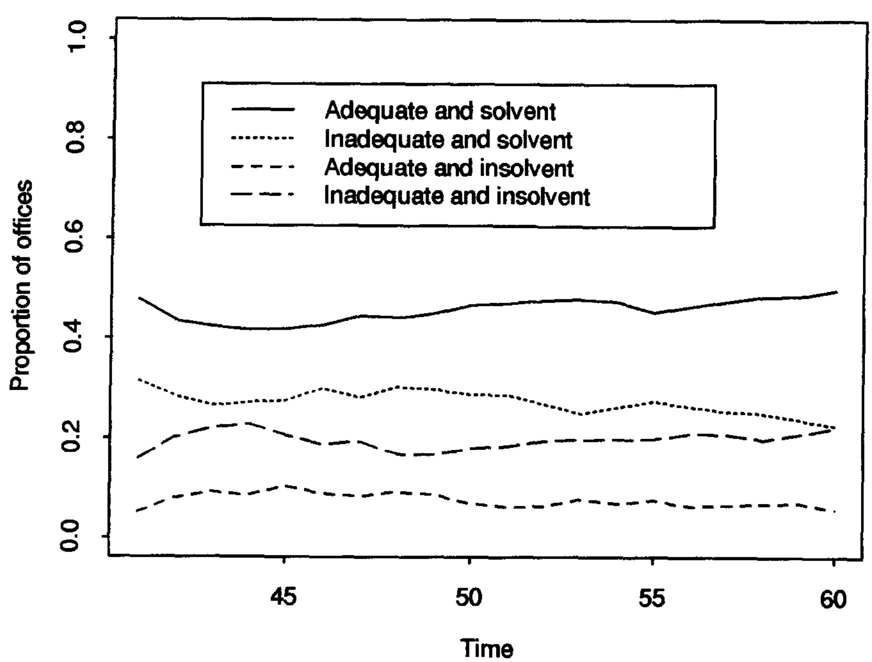

Figure 4 - Basis 2

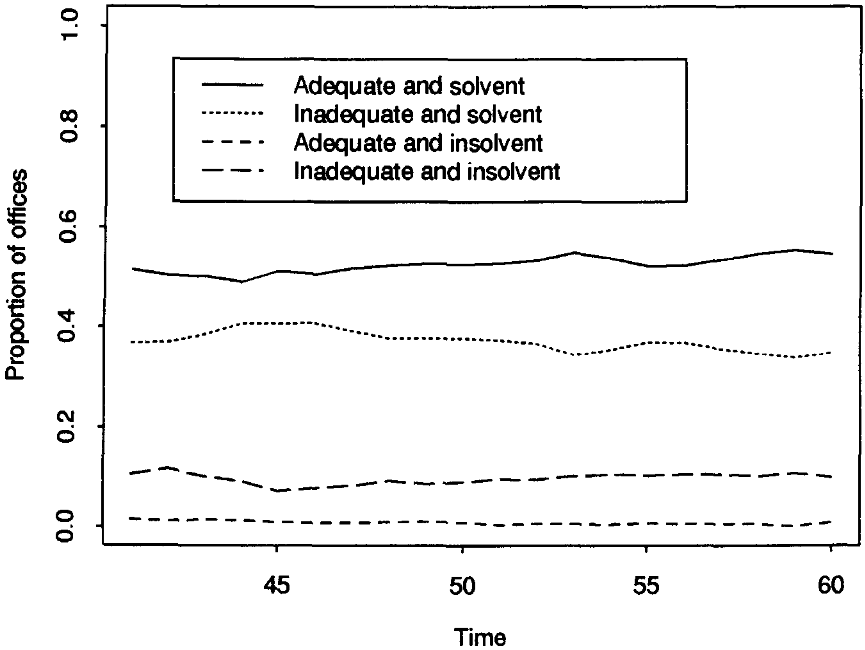

Figure 5 - Basis 3

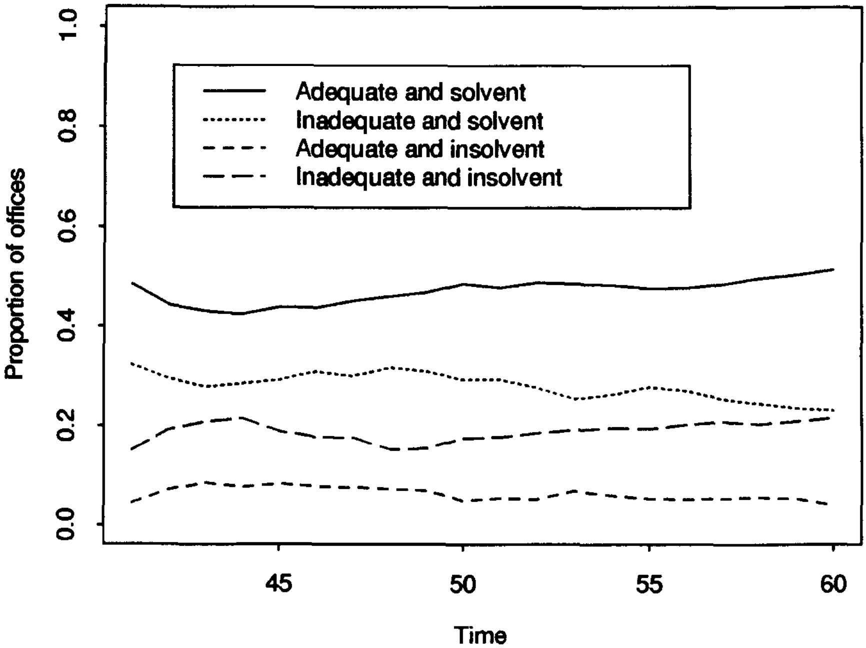

Figure 6 - Basis 4

These show some consistent features. At any given time the majority of offices are both adequate and solvent, and the smallest numbers are insolvent but adequate. The numbers both insolvent and inadequate are in between. A point in favour of the solvency valuations is that an office insolvent at any given time is more likely to be inadequate than adequate, but this might depend on the high prevalence of inadequacy.

The statutory minimum basis (Basis 3) shows the greatest disparity between adequacy and solvency, but it must be borne in mind that this office has (we hope) an unusually high rate of inadequacy because of its attempts to smooth maturity values without possessing capital with which to set up any kind of equalisation reserve.

6. The effect of an additional estate

The projections were repeated after adding an *additional estate* to the office at time t=40. An "additional estate" is a quantity of assets in excess of the asset shares of in-force policies. It is known in some quarters as "orphan surplus". In this case, assets equal to 10% of the total asset shares at time t=40 were added.

The effect upon solvency is perhaps more dramatic than the effect upon inadequacy. The number of inadequate offices during the 20-year time horizon falls from 864 to 405. The numbers of insolvencies (closures) under each valuation basis fall as follows :

Insolvencies under Basis 1	48
Insolvencies under Basis 2	76
Insolvencies under Basis 3	26
Insolvencies under Basis 4	75

One feature which is unchanged is the relative harshness of Bases 2 and 4. Both are net premium bases with an implicit margin in the interest rate, one fixed and one variable. It could be argued that if the margins were chosen more carefully, these bases would be no harsher than they needed to be, but this supposes a degree of foresight unattainable in practice.

For comparison with Figure 2, Figure 7 shows the proportions of inadequate and insolvent offices in each year of the projection period.

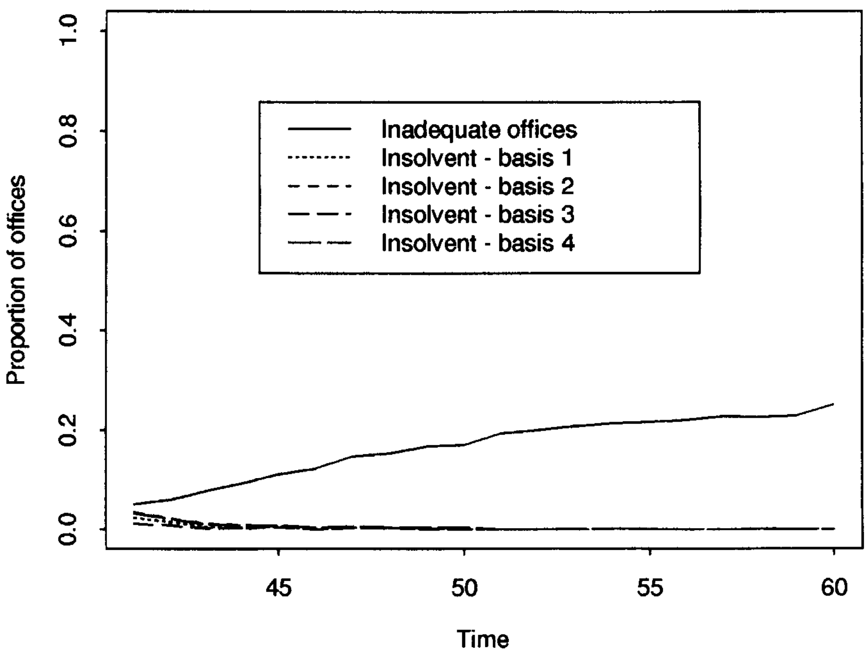

Figure 7

The pattern of rising inadequacies and falling insolvencies points once more to the rather arbitrary nature of the traditional solvency criterion. The pattern of insolvencies on the different bases is hardly discernible in this Figure, so Figure 8 shows, on a more detailed scale, the number of insolvent offices in each year.

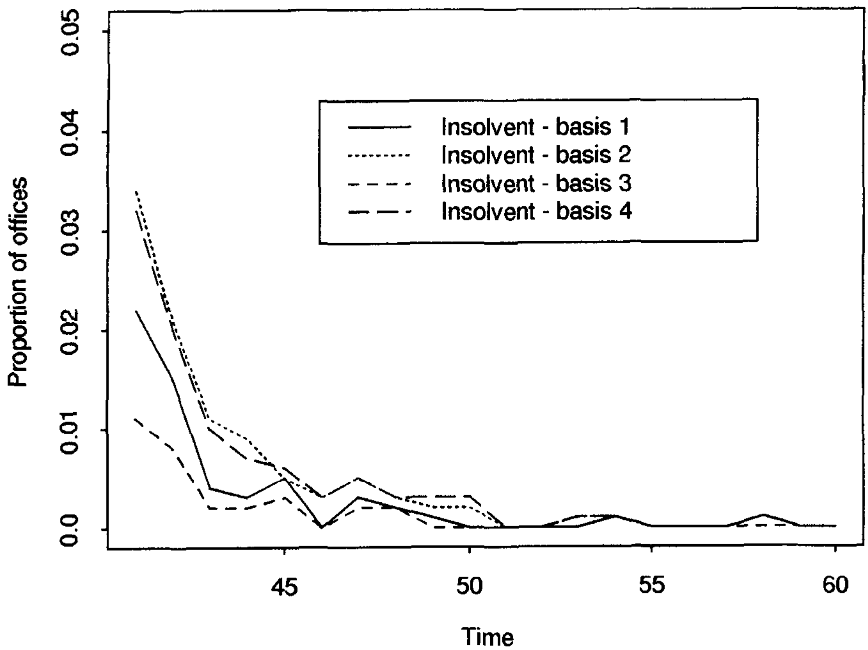

Figure 8

It is notable that nearly all of the insolvencies occur in the first 10 years. The most likely reasons for this are (i) switches into gilts just after time t=40, the effect of which on bonus algorithms is only gradual, and (ii) the rate of new business growth being less than the rate of return (on average) which causes the additional estate to increase as a proportion of the total asset shares.

Some other interesting features can be found in these results. If the solvency criterion - the A/L ratio which triggers closure - is chosen so that the numbers of errors - missed closures and wrong closures - are (nearly) equal, the results are as follows :

TABLE 3

Solvency criteria at which errors of solvency valuations are (almost) equal

	Basis 1	Basis 2	Basis 3	Basis 4
A/L criterion	1.220	1.174	1.213	1.184
Inadequacies	405	405	405	405
Total closures	406	404	405	406
Correct closures	211	210	213	212
Wrong closures	195	194	192	194
Missed closures	194	195	192	193

Compare this with Table 2. In both cases, the number of closures is practically the same as the number of inadequacies, and it is easy to see why this must be so. If we define

CC = No. of correct closures

IC = No. of incorrect closures

MC = No. of missed closures

TC = Total No. of closures

IA = Total number of inadequacies during the period

then $CC + IC = TC$ and $CC + MC = IA$ so if $IC = MC$, it follows that $TC = IA$. (The numbers are not exactly equal in Tables 2 and 3 because of rounding in the A/L ratio used as a criterion.)

What is not at all obvious is why the number of errors IC and MC should be practically the same for all four valuation bases. Given that the number of closures is equal to the number of inadequacies, why should the same number be correct closures? One suspects that some property of the sample paths may be involved, for example extreme falls in asset values which trigger insolvency under any "reasonable" basis. This is more likely as the same feature is present to a lesser degree in Table 2, in which the A/L thresholds are much lower. Further work is being done to investigate this.

It is also interesting to note that in no case did the ratio of assets to asset shares (the A/AS ratio) fall below 1.0 before the final contract had matured, regardless of the time of closure. This indicates that inadequacy caused by smoothing during the run-off period might not be detected using the A/AS ratio as a criterion.

7. Conclusions

Although these experiments are limited in scope, they have some interesting points worthy of wider investigation.

(1) If retrospective methods of surplus distribution are used, adequacy can be very sensitive to the level of assets in relation to the *asset shares*.

(2) On the other hand, solvency as measured by a traditional prospective valuation is very sensitive to the level of assets in relation to the *policy values*.

(3) None of the solvency valuations tested has any obvious direct link with the adequacy of the office. Moreover, a change in the circumstances of the office might itself be sufficient to alter the severity of any valuation test. It is difficult to say that any one approach to solvency based on prospective valuations (such as the U.K. statutory minimum basis) is clearly better or worse than any other similarly based approach.

References

B. Brindley, N. Dumbreck, C. Thomson, S. Thompson (1990). *Policyholders' Reasonable Expectations*. Report of Joint Working Party issued at seminar on Current Issues in Life Assurance, Birmingham, 10 July 1990.

A. S. Macdonald (1993). *What is the value of a valuation?* Proceedings of the 3rd International AFIR Colloquium, Rome 1993, **2** p.723.

M. D. Ross (1989). *Modelling a With-Profits Life Office*. Journal of the Institute of Actuaries **116** p.691.

M. D. Ross, M. R. McWhirter (1991). *The Impact on Solvency and Policy Results of the Valuation Regulations Restrictions on Equity Yields*. Unpublished paper.

A. D. Wilkie (1984). *A Stochastic Investment Model for Actuarial Use*. Transactions of the Faculty of Actuaries **39** p.341.