Two Economic Scenario Generators for Insurance ERM Calibrated to Australian Rates

Gary Venter

Abstract Two simplified economic scenario generators are introduced for application to insurance risk management and are calibrated to Australian zero-coupon rates. A validation methodology for testing their strengths and weakness is explored, and steps towards a more sophisticated model are discussed.

1 Introduction

Interest rate stochastic generators intended for insurance ERM have some different emphases than those used in finance. The latter are usually trading focused and include options pricing. For ERM, however, typically the risk to asset values over a few year horizon is quantified for selected portfolio strategies. When pricing is paramount, risk-neutral rates are targeted. For these, expected values have a risk margin built in, so no additional risk loading is needed. But for risk analysis of insurer assets and liabilities, the risk to real-world rates is the key.

Even for risk analysis, arbitrage-free generators are important. Although arbitrage is available in financial markets, it is not usually available for very long, so having arbitrage opportunities appear in a risk model could create distorted probabilities. For instance, a portfolio optimization study would say to put everything into taking the arbitrage.

We discuss here the three-factor CIR model and the correlated two-factor Vasicek model for the evolution of yield curves. These models are easy to use to simulate yield-curve scenarios, and not too difficult to calibrate, for instance through MCMC, which we illustrate. In our sample Australian data we have 37 zero-coupon maturities τ going from 1 to 10 years in quarterly increments, over 95 calendar months u from January 2009 through November 2016.

2 Affine Models

The CIR and Vasicek models are examples of affine models, which are almost but not quite linear. They are row factor times column factor models like actuaries use for reserving and mortality, but since row

Gary Venter
University of New South Wales, e-mail: gary.venter@gmail.com

parameters are multiplied by column parameters, the models are not strictly linear. The modeled yield $y_{u,\tau}$ for row u = time observed and column τ = bond maturity starts as a product of a row parameter r_u with a column parameter B_τ, but there is a constant term A_τ for each maturity as well, and there may be more than one B, r pairs of parameters. The two-factor version is:

$$y_{u,\tau} = A_\tau + r_{1,u}B_{1,\tau} + r_{2,u}B_{2,\tau} \tag{1}$$

A_τ can be thought of as the base yield curve by maturity that does not change by period u, while the Bs are tuning curves that get weighted by the values of the rs for each period. Figures 1 and 2 show these curves and weights for the CIR model discussed below.

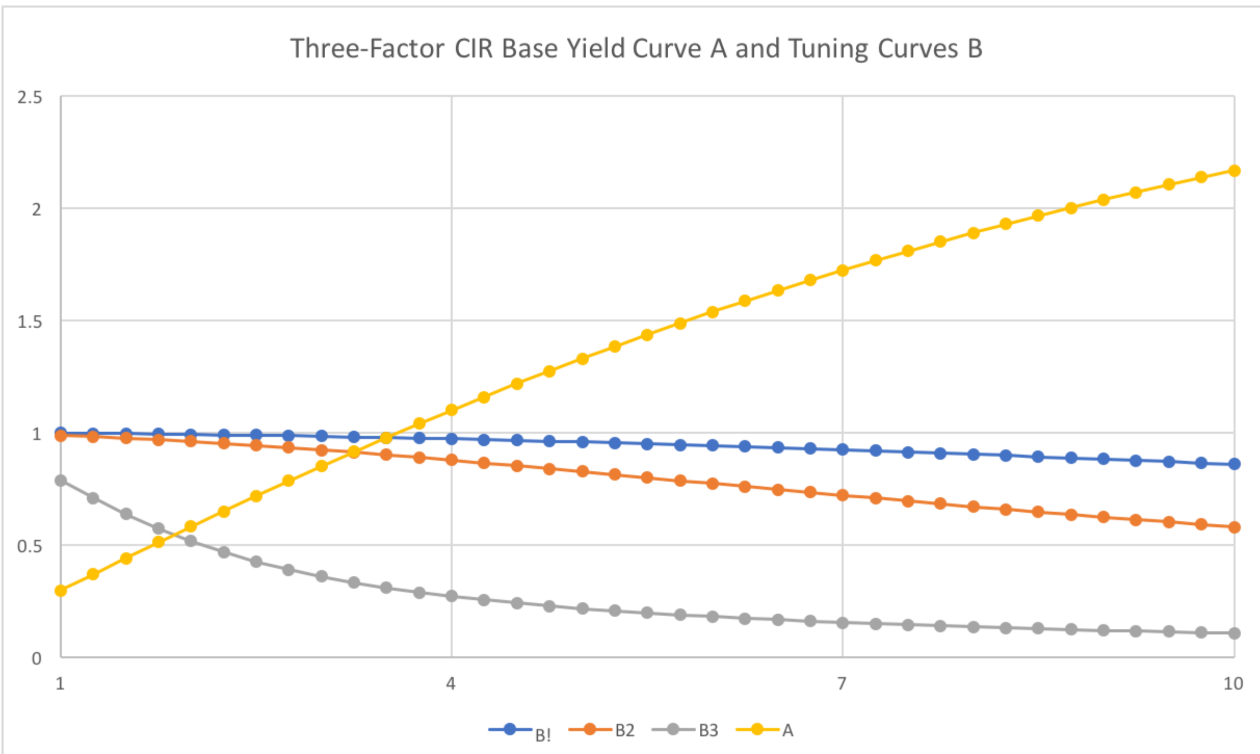

Fig. 1 Three-Factor CIR Base and Tuning Yield Curves Fit to Australian Zero Coupon Rates

The weights r are referred to as short rates. At $\tau = 0$, $A_\tau = 0$ and $B_\tau = 1$, so the rate is the sum of the rs. The B curves show the sensitivity to the rates at each maturity to a change in the short rates. When they are downward sloping, the longer rates do not vary as much as the shorter rates do, leading to a declining term structure of volatility and some flattening of the yield curve when the rates increase.

The short rates are constrained in how much they can change from one period to the next. The degree of constraint is a parameter for each r process. Here r_2 is the most constrained, and r_3 the least constrained. The fitted yield curve at any a point in time is the base curve plus the tuning curves multiplied by a vertical slice of the three r processes.

For the CIR and Vasicek processes, each r has three parameters – a long-term mean level, the speed of mean reversion, and the degree of variability allowed between periods. A value of each r process is fit

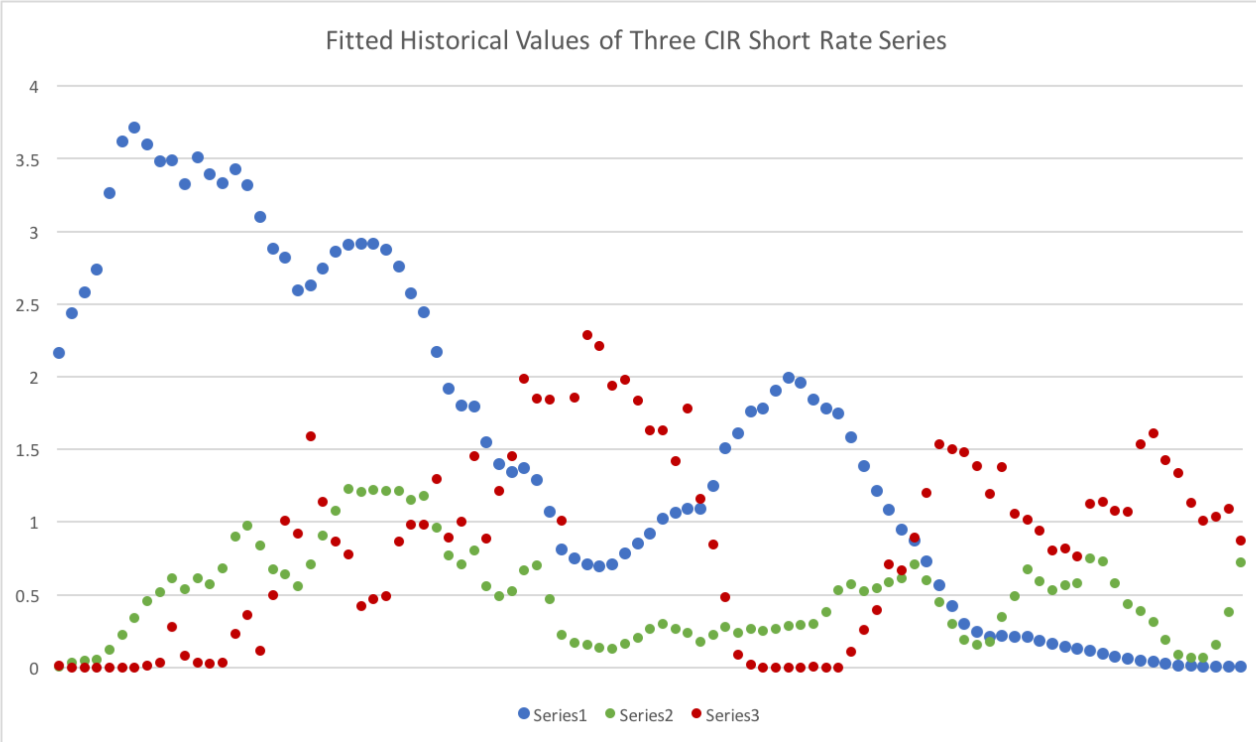

Fig. 2 Three-Factor CIR Tuning-Curve Weights r by Time Period

for each period in the dataset, but due to the restricted variability, fewer degrees of freedom are used by this fitting. That is, the number of effective parameters for a process is less than the number of periods, and can be quite a bit less for the highly constrained processes like r_2 here.

The A and B curves do not use any additional parameters. They are determined uniquely by the parameters of the short-rate processes.

3 The Vasicek Model

Bonds are traded continuously, so the yield curve needs to be calculable at any point in time u. To accommodate this, r_u is modeled using a continuous process called Brownian motion. This is defined as a normal distribution that starts at zero at time zero and at time t has mean zero and variance t. So from u to $u+d$ the change in the process has mean zero and variance d.

Each Vasicek factor r_j has three parameters – a long-term mean θ, a speed of mean reversion κ, and a volatility σ. Conditional on its value r_u at time u, its value at time $u+d$ is normally distributed with variance $\sigma^2 d$ and mean:

$$(1-\kappa d)r_u + \kappa\theta d \tag{2}$$

Thus it is gradually pulled towards its long-term mean by its speed of mean reversion.

Using no-arbitrage arguments, the interest rate $y_{u,\tau}$ at time u for a bond maturing at $u+\tau$, so with maturity τ, has been found to be:

$$y_{u,\tau} = A_\tau + B_\tau r_u \tag{3}$$
$$B_\tau = \frac{1-e^{-\kappa\tau}}{\kappa\tau} \tag{4}$$
$$A_\tau = \frac{1}{2}\frac{\sigma^2}{\kappa^2}\left(\frac{\kappa\tau}{2}B_\tau^2 + B_\tau - 1\right) + \theta\,(1-B_\tau) \tag{5}$$

Note that A and B do not change over time, just over maturity, and r changes over time but is constant by maturity.

3.1 Two-Factor Correlated Vasicek

In the two-factor versions, two r processes are used. In the independent version, these two processes are independent, and then the interest rates at each maturity are just the sums of the two partial rates at that maturity. This uses two $B_{j,\tau}$ rates and a new A_τ which is the sum of the two individual As.

In the correlated two-factor model, the processes evolve according to a bivariate normal. If the correlation is ρ, A_τ is adjusted by adding the correlation correction:

$$\frac{\rho\sigma_1\sigma_2}{\kappa_1\kappa_2}\left(\frac{1-e^{-\tau(\kappa_1+\kappa_2)}}{\tau(\kappa_1+\kappa_2)} - B_{1,\tau} - B_{2,\tau} + 1\right) \tag{6}$$

This also varies by maturity but not by time.

To simulate the bivariate standard normal with correlation ρ, take two independent standard normal draws v and x. Let $z = \rho x + (1-\rho^2)^{\frac{1}{2}} v$. Then x and z are standard normal variates with correlation ρ.

The advantage of the correlated version comes when ρ is negative. This allows more complex yield-curve shapes, comparable perhaps to those from three independent processes.

The framework for this model was presented in [Brigo and Mercurio(2006)], and the equations were put in the form used here by [Enev(2011)].

4 Calibrating to Australian Zero-Coupon Rates

The Vasicek model assumes continuously compounding interest rates with no interest payments until maturity – so applies to zero-coupon bonds. These are not always traded so empirical rates are rare, but the Reserve Bank of Australia posts constructed zero-coupon rates at daily intervals on its website. We used those (month-ending rates only) to calibrate the Vasicek model. There were 95 months in the data set at the time: January, 2009 to November, 2016.

The yield curves they present often are V-shaped, with a minimum rate around the one-year maturity. This is a difficult shape for many models to match, so we used maturities of one to ten years, which are available at three-month maturity increments – 37 maturities in all, and are the most relevant maturities for risk management purposes. It is not obvious how this V shape arose, but sometimes bond interest payments are considered to be short-duration zero-coupon bonds themselves. If these came from older, higher yielding bonds, the short duration rates could look higher.

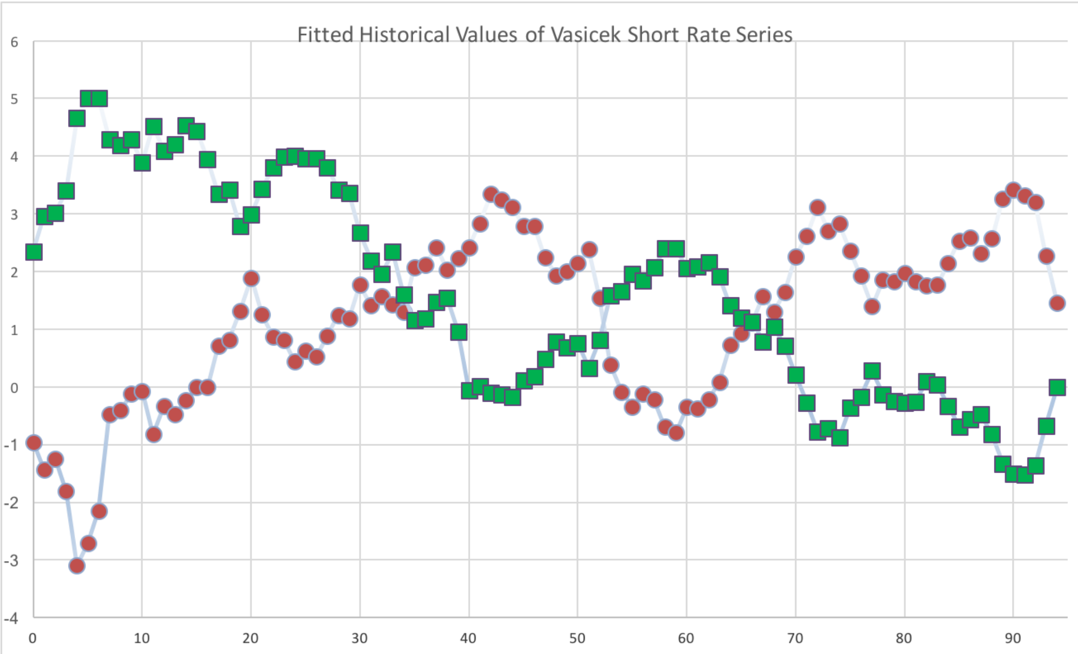

Fig. 3 Vasicek Fitted r_1 and r_2 Processes

Figure 3 shows the historical fitted r_1 and r_2 processes. The negative correlation is apparent in the mirroring of the plots. Some of the flexibility of the Vasicek yield curves comes from the rs being able to take negative values, which is not possible for the CIR process, as well as from the negative correlation. The underlying curves by maturity in Figure 4 are not that different than the CIR curves.

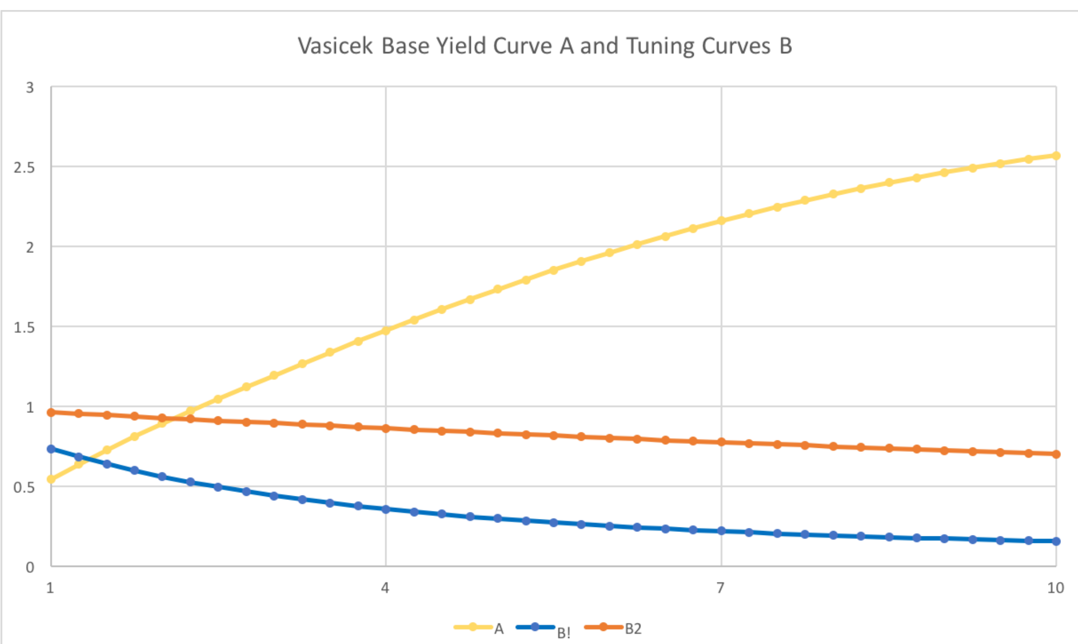

Fig. 4 Vasicek Base and Tuning Yield Curves

We estimated the parameters using the Stan MCMC package in R. MCMC is a form of Bayesian estimation that requires specification of prior distributions for each parameter, but does not require any specification as to what the posterior distributions of the parameters given the data might be. The posterior distributions are simulated numerically based on the model, the priors, and the data. The Stan package interface consists mainly of a programming language for model specification.

Here we assumed normally-distributed residuals with a common standard deviation. The two-factor model has 8 parameters – 3 for each factor, the correlation, and the residual standard deviation. However the model also needs estimates for the two r process values at each of the 95 months, so 190 more of what we will call quasi parameters, since they are quite constrained. These enter into the estimation of the 7 Vasicek process parameters, but are not used in model projections.

The prior distributions for $r_{1,u}$ and $r_{2,u}$ were those for the Vasicek process conditional on $r_{1,u-1/12}$ and $r_{2,u-1/12}$. For $r_{1,u}$, this is just the normal distribution with mean $(1-\kappa_1/12)r_{1,u-1/12}+\kappa_1\theta_1/12$ and standard deviation $\sigma_1/\sqrt{12}$. But for r_2, the correlation has to be taken into account. This makes $r_{2,u}$ conditional on both $r_{2,u-1/12}$ and $r_{1,u}$. Suppressing the conditioning on the $u-1/12$ values, the conditional distribution given $r_{1,u}$ is normal with mean and standard deviation:

$$E(r_{2,u}|r_{1,u}) = Er_{2,u}+\rho\frac{\sigma_1}{\sigma_2}(r_{1,u}-Er_{1,u}) \tag{7}$$

$$StdDev(r_{2,u}|r_{1,u}) = \sqrt{1-\rho^2}\sigma_2 \tag{8}$$

The priors and posterior means for the other parameters are shown below. All of these are uniform distributions, except for the residual standard deviation σ_y, which has prior proportional to $1/x$.

Parameter	κ_1	κ_2	σ_1	σ_2	θ_1	θ_2	ρ	σ_y	$r_{1,1}$	$r_{2,1}$
Prior	(0,3]	(0,3]	(0,10]	(0,10]	[0.1,0.6]	[12,18]	[-1,1]	(0,∞)	[-5,15]	[-5,15]
Posterior	0.650	0.076	1.40	0.521	0.279	17.4	-0.369	0.0540	-0.955	2.53

The priors for the reverting means, θs, look quite specific. Initially they were given wider ranges, but with θ_2 always forced to be greater than θ_1 for identifiability. The narrower priors here came about after observing the fits for various wider priors. Even now, some differences in θ_1 across chains can be seen to lead to noticeably different r values but with very comparable fitted interest rates. Just forcing this parameter to be a constant would seem to work fine.

5 CIR Model

Named after Cox, Ingersoll and Ross, this model is similar to the Vasicek model. For small d, the r process at time $u+d$, conditional on its value r_u at time u, is normally distributed with variance $r_u\sigma^2 d$ and mean:

$$(1-\kappa d)r_u+\kappa\theta d \tag{9}$$

The only difference from the Vasicek process is that the variance is proportional to r_u. This means that the process must stay non-negative. It does that since if it ever got to zero, its next value would be $\kappa\theta d$.

The A and B functions are a bit more intricate. For each process:

$$y_{u,\tau} = A_\tau + B_\tau r_u \tag{10}$$
$$h = \sqrt{\kappa^2 + 2\sigma^2} \tag{11}$$
$$C_\tau = 2h + (\kappa + h)\left(e^{h\tau} - 1\right) \tag{12}$$
$$B_\tau = 2\frac{e^{h\tau} - 1}{\tau C_\tau} \tag{13}$$
$$A_\tau = \frac{\kappa\theta}{\sigma^2}\left(\frac{2}{\tau}log\frac{C_\tau}{2h} - \kappa - h\right) \tag{14}$$

The normal distribution only holds for very small values of d. The continuously changing variance makes cumulative values not normal. In fact for larger d, r follows a non-central chi-squared distribution, which is somewhat gamma-like, with mean and variance:

$$E(r_{u+d}|r_u) = r_u e^{-\kappa d} + \theta\left(1 - e^{-\kappa d}\right) \tag{15}$$
$$Var(r_{u+d}|r_u) = \frac{\sigma^2}{2\kappa}\left(1 - e^{-\kappa d}\right)\left(2r_u e^{-\kappa d} + \theta\left(1 - e^{-\kappa d}\right)\right) \tag{16}$$

For parameter estimation we used a gamma distribution with these moments as the prior for $r_{u+1/12}$. Hopefully that approximation does not miss the optimum parameters for the r processes by much. The estimated parameters are:

Parameter	κ	θ	σ
r_1	0.630	0.460	0.106
r_2	0.022	16.6	0.100
r_3	0.004	23.5	0.945

The parameters are to some degree unusual. The 2nd process has a long-term mean interest rate of 16.6%, but a time to mean reversion of about 45 years. This creates a quite gradual upward drift in rates. The third process has a mean of 23.5% with time to reversion of 250 years. It also has a high standard deviation. That is why it moves more over time than the other processes as seen in Figure 2. It is almost a random walk. The first process is more usual. It has about the same volatility as the 2nd one but moves more, probably because of the faster mean reversion.

6 Testing the Models

6.1 Fit to Historical Yield Curves

The Vasicek residual standard deviation of 0.054% is much better than that of 0.121% for the CIR, but both are fairly small compared to a typical yield of 3%, indicating reasonably good fits to the historical data. Figure 5 compares actual and fitted yield curves for three months that show a variety of levels and curve shapes. Tight curvatures and multiple inflection points appear to be beyond the capacity of either model to capture, but the Vasicek fits do much better at this than the CIR does. The CIR curves are all pretty similar in shape.

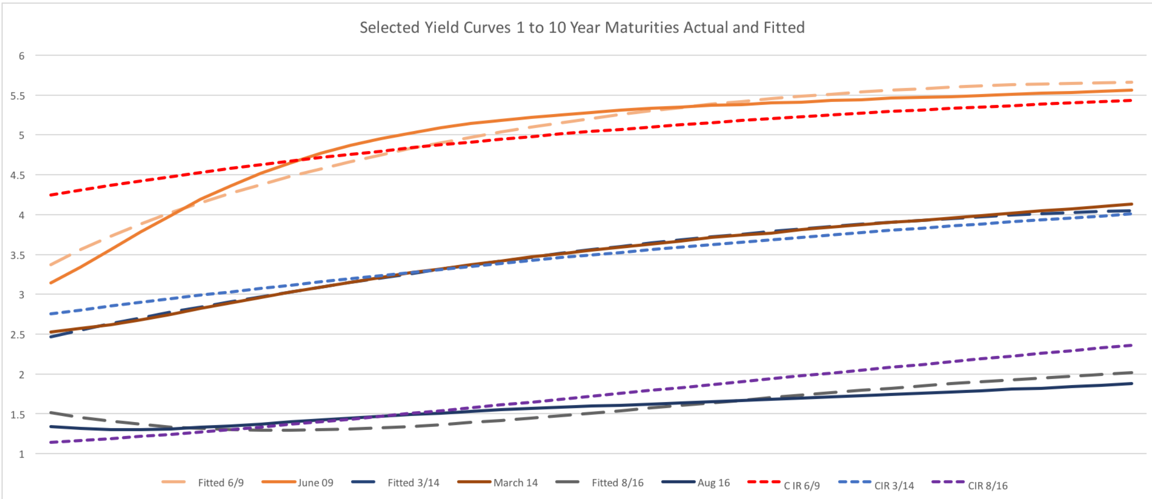

Fig. 5 Actual and Fitted Values for June 2009, September 2010, March 2014 and August 2016

6.2 Bayesian Fit Measures

Typically model fits are compared using adjustments to the loglikelihood to penalize it for number of parameters used relative to sample size. One of these, called the HQIC (Hannan Quinn information criterion) subtracts a penalty of $p = log(log(n))k$, where n is the sample size with k parameters. The small sample AIC, or AICc, has penalty $p = \frac{nk+1}{n-k}$.

MCMC estimation comes with its own adjustment to the loglikelihood to compensate for it being measured on the same sample the parameters were estimated from, which is what the traditional penalties are also attempting to do. It does this by a numerical integration for the estimated loglikelihood on out of sample points using a methodology called leave one out, or loo.

The loo-penalized loglikelihood for the two factor correlated Vasicek model for this data set is 5118.3, which includes a penalty of 274. This is only useful for comparison to other models fit to the data. It is not obvious what the HQIC or AICc penalties would be in this case, as the 190 quasi parameters $p_{1,u}$ and $p_{2,u}$ are constrained to not change too much from one month to the next by using the Vasicek process as a prior.

Since log(log(3515)) = 2.1, a penalty of 274 would arise from 130.5 parameters by HQIC. Backing into k from AICc would imply 254 parameters. Apparently loo is giving more of a penalty than AICc would, as there are only 198 parameters and quasi parameters in total. It seems more comparable to HQIC, but quite possibly with a little less penalty, as 130.5 effective parameters might be a bit low, given the moderate degree of constraint provided by the Vasicek priors.

The CIR fit had a loo measure of 2385, which is much worse (higher is better, as the penalty actually applies to the NLL). That includes a penalty of 103.5, or 49 equivalent parameters by the HQIC comparison. The much more constrained r curves apparently used many fewer parameters than Vasicek did, but also gave a poorer fit.

6.3 *Moments of the Yields*

These fits include the estimates of the underlying risk-neutral trends, so the moments of the fitted yields would be expected to match up well. A way to test the adequacy of the fitted parameters themselves is to look at simulated future series and compare various properties of the historical series to the simulated. This would be how the model would be used in risk analysis, so is a directly relevant test.

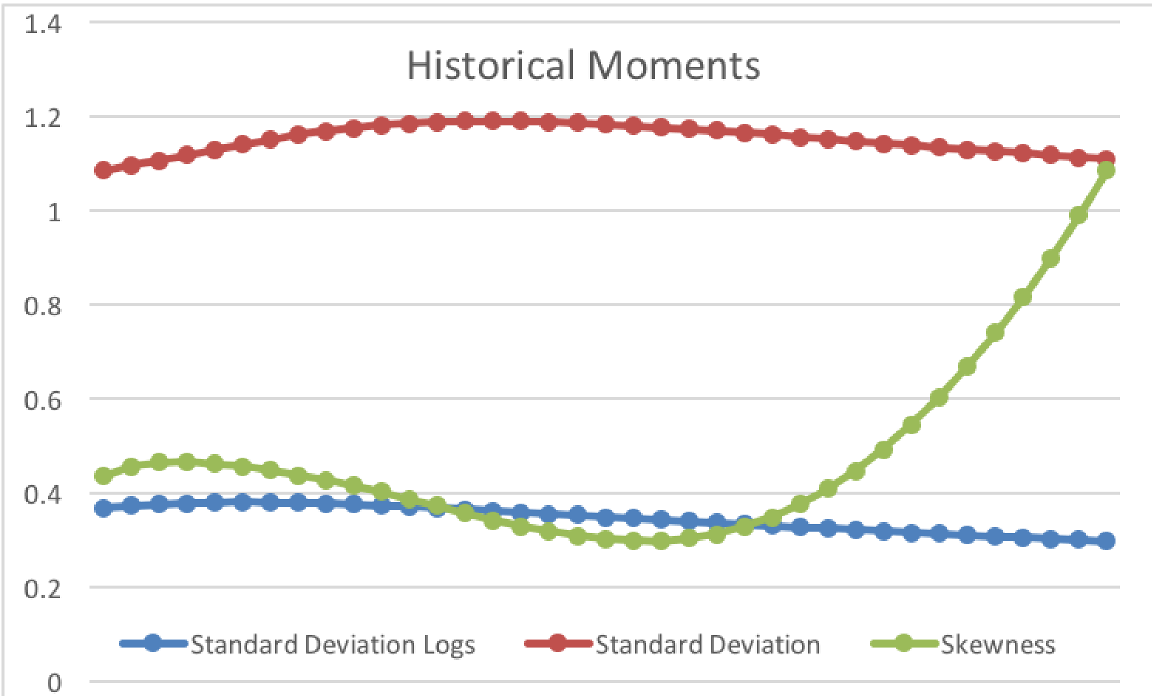

Fig. 6 Moments of Rates for 1 to 10 Year Maturities 2009 – 2016

As Figure 5 shows, standard deviations were fairly flat across all maturities, averaging about 1.15. The longer maturities show skewness around 1.0, with those for shorter maturities as low as 0.25. These are all fairly moderate skewnesses. The standard deviation of the log of the rates shows the variability relative to the rate. These go from about 40% at the short end down to 30% for the longest rates.

Quantifying the distribution of yield-curve shapes is difficult, as there are so many maturities. One aspect of shapes seen in many economies is that when the short rate rises, the longer rates tend to rise less, leading to a compression of the curve. To quantify this we did a regression of the 3-year to 10-year spread on the 1-year rate for the historical data. The slope of this would generally be negative because of this compression. For the 2009 – 2016 period the slope was -12%, and the standard deviation of rates around the fitted line was 0.24. This slope and standard deviation provide one picture of the distribution of curve shapes.

In the US the historical slope was more like -40%, with a standard deviation around the line about 0.1. The slope was not meaningfully measurable in the last several years with very low short rates, but the spread around the original slope has been consistent with history. The Australian short rates never got so low but the slope may be flatter with a wider spread recently due to the unusual rate performance.

To test moments and shape distributions with simulation, we simulated 10,000 scenarios of three years for maturities 1, 3, 5, 7, and 10 years for each model over a three-year horizon. Because some of the short rates have high reverting means with a slow rate of mean reversion, the rates will gradually increase rates over time, although more so for the Vasicek model. That one is probably useful for projections of a few years only, and both should be recalibrated from time to time.

It turns out that both models have some limitations that show up in these tests. Since some of these are in opposite directions, we also look at the combined set of 20,000 scenarios from the two models as a potential view of future possible scenarios.

In any case, the first three moments for the yield rates for each year are shown below, first for the Vasicek model, then for the CIR, and then for both.

yr, term	meanV	std devV	skewV	meanC	std devC	skewC	meanB	std devB	skewB
1, 1	2.77	0.76	0.02	1.88	0.49	2.41	2.33	0.78	0.79
1, 3	3.14	0.52	0.04	22.17	0.47	2..43	2.66	0.70	0.50
1, 5	3.49	0.43	0.04	2.46	0.46	2.44	2.97	0.68	0.31
1, 7	3.77	0.38	0.04	2.73	0.44	2.45	3.25	0.66	0.24
1, 10	4.00	0.33	0.04	3.09	0.40	2.45	3.54	0.59	0.28
2, 1	3.80	0.90	0.00	2.24	0.76	3.03	3.02	1.14	0.70
2, 3	4.12	0.65	-0.01	2.55	0.74	3.04	3.34	1.05	0.41
2, 5	4.41	0.56	-0.01	2.84	0.72	3.05	3.62	1.02	0.31
2, 7	4.63	0.50	-0.01	3.10	0.69	3.06	3.86	0.98	0.26
2, 10	4.79	0.45	-0.01	3.42	0.63	3.06	4.10	0.88	0.29
3, 1	4.77	0.98	0.04	2.63	0.99	3.25	3.70	1.46	0.51
3, 3	5.04	0.75	0.03	2.94	0.97	3.26	3.99	1.37	0.35
3, 5	5.27	0.65	0.02	3.21	0.94	3.26	4.24	1.31	0.29
3, 7	5.44	0.60	0.01	3.46	0.90	3.27	4.45	1.25	0.26
3, 10	5.52	0.54	0.00	3.75	0.83	3.27	4.64	1.13	0.28

The mean rates are increasing a fair amount over time for the Vasicek simulation, and less so for the others. This can be controlled a bit by the starting values for the simulated rs and should be made consistent with overall macro expectations. The standard deviations of the rates are lower than historical and decline across the maturities for Vasicek, are low but fairly flat for CIR and are like history but decreasing at longer maturities for both combined, depending to some degree on how far ahead the simulation is. The skewnesses from this normal distribution model are near zero and are quite high for CIR. The combined model is more realistic than either.

In the combined model, the standard deviation of the log of the rates declines from around 0.35 at the short end to about 0.25 at the long end, which is in reasonable agreement with history.

For the Vasicek model, the slopes of the 10 –3 spread on 1 year rate regressions were about-35%, with standard deviations around 0.1. These show a stronger compression effect than seen in the last few years, but are similar to historical US experience. For the CIR model, the slopes were -14%, which are like recent Australian data, but with standard deviations only about 0.02. This means that the yield curves were all about the same shape by this measure. The combined model had slopes around -20% and standard deviations about 0.13. This seems the best of them.

Still, in the combined model the two sets of yield curves maintain separate identities, and the distribution of shapes is not what the regression numbers would seem to imply. Figure 7 illustrates this.

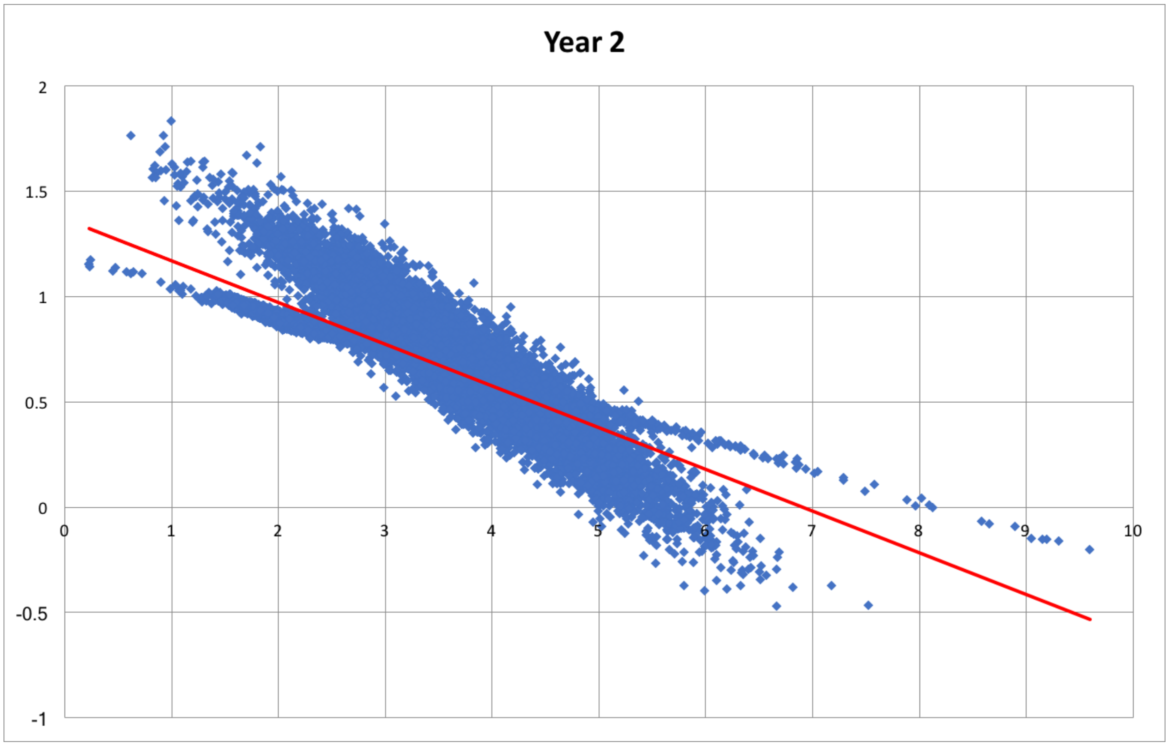

Fig. 7 3 year to 10 year spread as a function of 1 year rate, 2nd year of simulation

Overall the moment analysis suggests that the combined model output is fairly similar to historical data by moment measures.

7 Conclusions

The two-factor correlated Vasicek model is easy to simulate and gives a good fit to recent Australian zero-coupon bond yields, with a negative correlation between factors. Estimation is straightforward with MCMC. It has some degree of flexibility in yield-curve shapes, but these are not as complex as some in the data. Simulated future results appear reasonable for yield levels and curve shapes, but probably understate the tail risk for longer bonds. The three-factor CIR model does not have nearly as much variability in curve shapes so does not fit the data as well. Combining the models seems to take advantage of some of their separate somewhat offsetting limitations.

Some possibly better models are discussed in the Appendix.

References

[Brigo and Mercurio(2006)] Brigo, D., Mercurio, F., 2006. Interest rate models - theory and practice: With smile, inflation and credit. Springer Finance 2nd Edition.

[Enev(2011)] Enev, E., 2011. Gaussian affine term structure models within the duffie-kan framework and their use for forecasting. http://scriptiesonline.uba.uva.nl/document/338587.
[Feldhütter(2016)] Feldhütter, P., 2016. Can affine models match the moments in bond yields? http://feldhutter.com/RiskPremiumPaperFinal.pdf.

Appendix – Modeling Alternatives

A.1 BDFS

The next step in complexity is to make the reverting mean θ and volatility σ themselves stochastic. An example is the BDFS model that has a Vasicek process for the rate r and for the reverting mean θ, but a CIR process for the volatility v. Here we omit the subscript t:

$$dr = \kappa(\theta - r)dt + \sqrt{v}dW_1 \tag{17}$$
$$d\theta = \alpha(\beta - \theta)dt + \eta dW_2 \tag{18}$$
$$dV = a(b - v)dt + \phi\sqrt{v}dW_3 \tag{19}$$
$$dW_1 dW_3 = \rho dt \tag{20}$$

The last equation puts some correlation into the process.

This model has an almost closed form solution for the longer rates. The distinction seems to be that a closed form function can be computed on a phone app, log for example, where an almost closed form calculation requires a computer app. In any case, the rates can be programmed but require calling a function for the numerical solution of differential equations.

A.2 $A_2(3)$ *Model*

A more general system has been developed for categorizing multi-factor models with affine term structure. One of the common models from that is the $A_2(3)$ model. It is a 3-factor model where 2 of the 3 factors directly go into the stochastic component of the short rate. It is similar to the BDFS model in that it has stochastic mean reversion and stochastic volatility, but instead of having separate factors for these, they are both linear combinations of 2 of the 3 factors – the third being the short rate itself. It is defined by the following equations:

$$dz_1 = \kappa_{11}(\theta_1 - z_1)dt + \kappa_{12}(\theta_2 - z_2)dt + \sqrt{z_1}dW_1 \tag{21}$$
$$dz_2 = \kappa_{21}(\theta_1 - z_1)dt + \kappa_{22}(\theta_2 - z_2)dt + \sqrt{\beta z_2}dW_2 \tag{22}$$
$$\theta = a_\theta + b_\theta z_1 + c_\theta z_2 \tag{23}$$
$$v = a_\theta + b_v z_1 + c_v z_2 \tag{24}$$
$$dr = \kappa(\theta - r)dt + p\sqrt{z_1}dW_1 + q\sqrt{\beta z_2}dW_2 + \sqrt{v}dW_3 \tag{25}$$

Other series could also be postulated to be linear combinations of z_1 and z_2 – various credit spreads for risky bonds, for example.

An increasingly popular convention is to make the last process $r - z_1 - z_2$. Then the interest rate is the sum of the three process. Here there is no scaling parameter for the volatility of the first process, so that is taken up by the b_θ and b_v parameters. Another way to do that is to leave out the θ and v equations, and make the interest rate a linear combination of the three processes. These gives the standard canonical form for affine models.

A.3 General Affine Model Form

The general form uses matrix notation for an N-factor model. We will take $N = 3$, which is the most common form – the simplest that seems adequate. There are 3 process in a 3-vector X_t at time t. The interest rate r_t is a linear function of those processes. With a scalar δ_0 and a 3-vector δ_x, the interest rate is:

$$r_t = \delta_0 + \delta_x' X_t \tag{26}$$

Under the risk-neutral measure, the 3 X processes follow a diffusion given by:

$$dX_t = (K_0 - K_1 X_t)dt + \Sigma\sqrt{S_t}d\widetilde{W}_t \tag{27}$$

Here K_0 is a 3-vector, and K_1, Σ, S_t are all 3x3 matrices. S_t is a diagonal matrix, with the i^{th} diagonal element a linear combination of the X processes, i.e., = $\alpha_i + \beta_i' X_t$, with the α's constants and the β's 3-vectors.

At time t, the yield on a bond with term to maturity τ is then $-A_\tau/\tau + B_\tau' X_t/\tau$, where for each τ, A_τ and B_τ are the scalar and 3-vector that solve the system of ODEs

$$dA_\tau/d\tau = -K_0' B_\tau + \frac{1}{2}\sum_{i=1}^{3}(\Sigma' B_\tau)_i^2 \alpha_i - \delta_0 \tag{28}$$

$$dB_\tau/d\tau = -K_1' B_\tau - \frac{1}{2}\sum_{i=1}^{3}(\Sigma' B_\tau)_i^2 \beta_i + \delta_x \tag{29}$$

In this notation, the $A_2(3)$ model is defined by taking $K_0(3) = K_1(1,3) = K_1(2,3) = \alpha_1 = \alpha_2 = \beta_1(2) = \beta_1(3) = \beta_2(1) = \beta_2(3) = \beta_3(3) = 0$ and $\alpha_3 = \beta_1(1) = \beta_2(2) = 1$, with constraints on the non-zero parameters of: $K_0 > 0$, $\beta_3 > 0$, $[K_1^{-1}K_0] > 0$, $\delta_x(3) > 0$, $K_1(2,1) \leq 0$, $K_1(1,2) \leq 0$. See [Feldhütter(2016)] who provides the estimates on US data from that time of:

$$\delta_0;\ \delta_x(*);\ \beta_3(*);\ K_0(*) = 0.0215;\ 0.0052,\ 0.0012,\ 0.0023;\ 0.7379,\ 2.3153,\ 0.0;\ 0.3151,\ 0.1265,\ 0.0 \tag{30}$$

$$K_1(*);\ K_2(*);\ K_3(*) = 1.4202,\ -0.2215,\ 0.0;\ -0.1222,\ 0.0271,\ 0.0;\ 3.7514,\ -0.6462,\ 0.7034 \tag{31}$$

Thus there are 15 non-zero parameters for the expected yields in this model, compared to 7 for the correlated two-factor Vasicek.