Validation of Investment Models for Actuarial Applications

Mary Hardy
AFIR Symposium
September 2005
Acknowledgements

• Keith Freeland
• Matthew Till
• Institute for Quantitative Finance and Insurance
• CIBC
• NSERC
Outline

• Some history
• The models
• Does it matter?
• Traditional model selection
• Bootstrap evidence
• Abusing the bootstrap
History

- Single premium equity linked insurance in North America
  - Segregated Funds in Canada
  - Variable Annuities in USA
- Carry guarantees on death and maturity
- Guarantee may be fixed or increasing
History

• 25 years ago, UK faced the same issue

• MGWP published paper in 1980
  - Stochastic simulation of liabilities (and underlying assets)
  - Quantile (VaR) reserve.
  - Early application of early Wilkie Model
Canadian Method

• Stochastic simulation of liabilities
• CTE (Tail-VaR) reserve
• Not much hedging
  ➢ If hedged, simulate and reserve for unhedged risk
• Equity model: ‘freedom with calibration’
Canadian Calibration Method

• Use any model
• Check the left-tail accumulation factor probabilities, using standard data set
• Adjust parameters to meet calibration fatness requirement
• Table calculated using ‘Regime-Switching Lognormal –2’ model
Accumulation Factors

• Let $Y_t$ represent log return in $t^{th}$ month

• 1-year accumulation factors are
  $\exp(Y_t + Y_{t+1} + \ldots + Y_{t+11})$

• Similarly for 5-year and 10-year

• 40 years data $\Rightarrow$ 4 non-overlapping observations of 10-Year accumulation factor
## Canadian Calibration Table

<table>
<thead>
<tr>
<th>Accumulation Factor</th>
<th>2.5 %ile</th>
<th>5 %ile</th>
<th>10%ile</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-year</td>
<td>0.76</td>
<td>0.82</td>
<td>0.90</td>
</tr>
<tr>
<td>5-year</td>
<td>0.75</td>
<td>0.85</td>
<td>1.05</td>
</tr>
<tr>
<td>10-year</td>
<td>0.85</td>
<td>1.05</td>
<td>1.35</td>
</tr>
</tbody>
</table>
US approach

- C3P2
- Similar to Canadian approach
- Calibration Table applied to left and right tails
- US table derived from ‘Stochastic Log-Volatility’ model
Some outcomes…

• UK
  ➢ no more maturity guarantees

• Canada
  ➢ cut back on generous guarantees
  ➢ Plethora of equity models proposed
  ➢ Still little hedging

• USA
  ➢ Hedging common
Some equity models ....

• Regime Switching Log Normal (Hardy, 2001 and CIA Seg Fund Report, 2001)
• GARCH(1,1)
• MARCH (Chan and Wong, 2005)
• ‘Stochastic Log Volatility’ (AAA C3-Phase 2)
• Regime Switching Draw Down (Panneton, 2003)
S&P data

- not much auto-correlation
- but correlation is not always a good measure of independence
- notice bunching of poor returns (eg last 2 years)
- and association of high volatility with crashes
  ➢ ie high down more than high up
\[ Y_t \mid \rho_t = \mu_{\rho_t} + \sigma_{\rho_t} \varepsilon_t \]

**REGIME 1 \( \rho_1 \)**
- Low Volatility \( \sigma_1 \)
- High Mean \( \mu_1 \)

\[ Y_t = \mu_1 + \sigma_1 \varepsilon_t \]

**REGIME 2 \( \rho_2 \)**
- High Volatility \( \sigma_2 \)
- Low Mean \( \mu_2 \)

\[ Y_t = \mu_2 + \sigma_2 \varepsilon_t \]
The RSLN-2 Model

• The regime process is a hidden Markov process
• 2 Regimes are usually enough for monthly data.
• 2 Regime model has 6 parameters:

  $\Theta = \{\mu_1, \mu_2, \sigma_1, \sigma_2, p_{12}, p_{21}\}$

• Regime 1: Low Vol, High Mean, High Persistence (small $p_{12}$)
• Regime 2: High Vol, Low Mean, Low Persistence (large $p_{21}$)
GARCH(1,1)

\[ Y_t = \mu + \sqrt{h_t} \varepsilon_t \]

\[ h_t = \alpha_0 + \alpha_1 (Y_{t-1} - \mu)^2 + \beta h_{t-1} \]

- Where \( \varepsilon_t \sim N(0,1) \), iid
- Given \( F_{t-1} \), \( \varepsilon_t \) is the only stochastic element
- We generally require \( \alpha_1 + \beta < 1.0 \)
MARCH (2;0,0;2,0)

\[ Y_t \mid F_{t-1} \sim \begin{cases} Q_1 \quad \text{w.p. } \alpha_1 \\ Q_2 \quad \text{w.p. } (1 - \alpha_1) \end{cases} \]

\[ Q_1 \sim ARCH(2); \quad Q_2 \sim ARCH(0) \]

\[ h_{1,t} = \beta_{10} + \beta_{11}(Y_{t-1} - \phi_1)^2 + \beta_{12}(Y_{t-2} - \phi_1)^2 \]

\[ h_{2,t} = \beta_{20} \]
MARCH(2;0,0;2,0)

- MARCH(K; \(p_1\ldots,p_K; q_1,\ldots,q_K\)) is a mixture of K AR-ARCH models,
- \(p_j\) and \(q_j\) are the AR-order and ARCH-order of the \(j^{th}\) mixture RV
- According to Chan and Wong, provides superior fit to 3\(^{rd}\) and 4\(^{th}\) moments of monthly log-return disn cf RSLN
\[ v_t = \log \sigma_t = (1 - \varphi)v_{t-1} + \varphi \log \tau + \sigma_v Z_{v,t} \]

\[ \mu_t = A + B\sigma_t + C\sigma^2_t \]

\[ Y_t = \frac{\mu_t}{12} + \frac{\sigma_t}{\sqrt{12}} Z_{y,t} \]

- \(Z_{v,t}\) and \(Z_{y,t}\) are standard normal RVs, with correlation \(\rho\)
- The \(v_t\) process is constrained by upper and lower bounds

\textbf{SLV}
SLV

• According to C3P2, SLV

  ➢ “Captures the full benefits of stochastic volatility in an intuitive model suitable for real world projections”

  ➢ Stoch vol models are widely used in capital markets to price derivatives...

  ➢ Produces very “realistic” volatility paths
Regime Switching Draw Down (RSDD)

\[ Y_t \mid (\rho_t = s) = \kappa_s + \phi_s D_{t-1} + \sigma_s \varepsilon_t \]

\[ D_{t-1} = \min(0, D_{t-2} + Y_{t-1}) \]

\[ \varepsilon_t \sim N(0,1), \text{iid} \]

\( \rho_t \) is a Markov regime switching process
RSDD

• 2 Regimes proposed by Panneton
• $D_t$ is the draw-down factor
• RSLN-2 is recovered when $\phi_{\rho}=0$, for $\rho=1,2$
• Captures ‘tendency to recover’
Does it matter?

• 5 models, each being championed by someone.
• 2 RS, 2 conditional heteroscedastic, 1 ‘stochastic volatility’.
• Each fitted by MLE (-ish) to S&P500 data
• Does it make any difference to the results for Equity-Linked Capital Requirements?
Two methods for Equity Linked Life Insurance

- Actuarial Approach:
  - Simulate liabilities,
  - apply risk measure,
  - discount at r-f rate

- Determines the economic capital requirement to write the contract for a given solvency standard.
Two methods for Equity Linked Life Insurance

• Dynamic Hedging Approach
  - Simulate Hedge under real world measure
  - Estimate distribution of unhedged liability
  - Apply risk measure and discount at r-f rate
  - Add to cost of initial hedge
Actuarial Approach

- Estimate 95% CTE (= mean of worst 5% of simulated outcomes) for 20-year GMAB.
- Single Premium
- Guarantee payable on death or maturity
- Guarantee ‘ratchet’ at t=10
- Issue age 50, rf rate 5%, MER=3% p.y.
- Guarantee risk premium = 0.2% p.y.
- Deterministic mortality and lapses
## Risk Measure, % of P; AA

<table>
<thead>
<tr>
<th>Model</th>
<th>90% CTE</th>
<th>95% CTE</th>
</tr>
</thead>
<tbody>
<tr>
<td>RSDD</td>
<td>0.64 (0.09)</td>
<td>2.25 (0.16)</td>
</tr>
<tr>
<td>MARCH</td>
<td>2.85 (0.14)</td>
<td>5.22 (0.19)</td>
</tr>
<tr>
<td>SLV</td>
<td>3.12 (0.15)</td>
<td>5.47 (0.20)</td>
</tr>
<tr>
<td>GARCH</td>
<td>3.60 (0.19)</td>
<td>6.27 (0.22)</td>
</tr>
<tr>
<td>RSLN</td>
<td>6.50 (0.19)</td>
<td>9.53 (0.23)</td>
</tr>
<tr>
<td>RSGARCH</td>
<td>6.33 (0.17)</td>
<td>9.18 (0.23)</td>
</tr>
</tbody>
</table>
Does the model matter using the actuarial approach?

Oh Yes !!!
Using hedging?

- Straight Black-Scholes Hedge
- Simulate additional cost arising from
  - Discrete hedge
  - Model Error (ie P-measure is GARCH/RSDD etc)
  - Transactions costs
## Risk Measure, % of single premium

<table>
<thead>
<tr>
<th>Model</th>
<th>90% CTE</th>
<th>95% CTE</th>
</tr>
</thead>
<tbody>
<tr>
<td>RSDD</td>
<td>4.20 (0.08)</td>
<td>4.62 (0.10)</td>
</tr>
<tr>
<td>MARCH</td>
<td>3.67 (0.06)</td>
<td>4.00 (0.09)</td>
</tr>
<tr>
<td>SLV</td>
<td>3.39 (0.05)</td>
<td>3.67 (0.09)</td>
</tr>
<tr>
<td>GARCH</td>
<td>4.12 (0.08)</td>
<td>4.52 (0.11)</td>
</tr>
<tr>
<td>RSLN</td>
<td>4.06 (0.08)</td>
<td>4.45 (0.09)</td>
</tr>
<tr>
<td>RSGARCH</td>
<td>4.62 (0.12)</td>
<td>5.15 (0.17)</td>
</tr>
</tbody>
</table>
Does the model matter using the hedging approach?

Not so much….
But

- Many companies are not hedging
- Pressure to adopt models giving lower capital requirements
- Can we use traditional methods to eliminate any of the models?
## Likelihood Comparison

<table>
<thead>
<tr>
<th>Model</th>
<th># parameters</th>
<th>Max LL</th>
</tr>
</thead>
<tbody>
<tr>
<td>RSDD</td>
<td>8</td>
<td>1047.1</td>
</tr>
<tr>
<td>MARCH</td>
<td>7</td>
<td>1039.8</td>
</tr>
<tr>
<td>SLV</td>
<td>7</td>
<td>1032.9*</td>
</tr>
<tr>
<td>GARCH</td>
<td>4</td>
<td>1030.1</td>
</tr>
<tr>
<td>RSLN</td>
<td>6</td>
<td>1042.0</td>
</tr>
<tr>
<td>RSGARCH</td>
<td>8</td>
<td>1054.9</td>
</tr>
</tbody>
</table>
Residual analysis

- Residuals for RS models – weighted from individual regimes
- Residuals for SLV – using simulated volatility paths
GARCH Residuals q-q Plot
So far …

• Likelihood based selection doesn’t help much

• AIC is too simple, BIC depends on sample size, LRT has technical limitations

• Residuals can be useful, but are tricky in multifactor cases
so good…

• The regime switching models look good on likelihood and on residuals
• But there is a vast difference in application between rsdd and rsln or rsgarch
• What causes the big difference?
• Which rs model should we believe?
1-Year Accumulation Factors

- RSLN
- RSDD

1-Year AF

Probability Density Function
10-Year Accumulation Factors

Probability Density Function

10-Year AF

RSLN
RSDD
RSGARCH
Bootstrapping time series

• The traditional bootstrap is applied to independent observations.

• Dependent time series require different treatment.

• Order matters.
S&P 1-year Acc Factors

- If we take 1-year factors starting in January, empirical percentiles are (from 48 observations):
  - 2.5%ile – 0.84
  - 5%ile – 0.85
  - 10%ile – 0.94
S&P 1-year Acc Factors

- If we take 1-year factors starting in September, empirical percentiles are (47 observations):
  - 2.5%ile – 0.75
  - 5%ile – 0.87
  - 10%ile – 0.89

- Ranges are: 2.5%ile (0.74, 0.89)
  5%ile (0.83, 0.91)
  10%ile (0.89, 0.95)
1-year Acc factors

- Can’t use all 1-year factors because of dependence
- If we only use (eg) January series, we are ignoring information
- Bootstrap the percentiles using time series bootstrap.
Time series bootstrap

• Bootstrap from original observations in blocks of \( b \) consecutive values.

• If the blocks are too small, lose dependence factor \( \Rightarrow \) results too thin tailed (if +vely autocorrelated)

• If blocks are too large lose data, \( \Rightarrow \) results too thin tailed (extreme results averaged out)
Block size

- So choose block size to maximize tail thickness.
- Other ways of selecting block size.
- No general agreement – see references.
- Randomized block length suggested.
- block resampling reduces exposure of end points → cycle from end to start.
### Bootstrap Quantile Estimates

**1-Year Accumulation**

<table>
<thead>
<tr>
<th>Model</th>
<th>2.5%ile</th>
<th>5%ile</th>
<th>10%ile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bootstrap 90% CI</td>
<td>0.67 → 0.87</td>
<td>0.76 → 0.91</td>
<td>0.84 → 0.97</td>
</tr>
<tr>
<td>RSDD</td>
<td>0.768</td>
<td>0.831</td>
<td>0.901</td>
</tr>
<tr>
<td>RSLN</td>
<td>0.764</td>
<td>0.829</td>
<td>0.908</td>
</tr>
<tr>
<td>RSGARCH</td>
<td>0.792</td>
<td>0.847</td>
<td>0.910</td>
</tr>
</tbody>
</table>

This doesn’t help us much.
10-year accumulation factor

- We can do the same thing
- But the original data only has 4 non-overlapping observations
- Minimum 10-year observed AF is estimate of $1/5 = 20\%$ile
- So we bootstrap B samples of 4 observations
### Bootstrap Quantile Estimates

#### 10-Year Accumulation

<table>
<thead>
<tr>
<th>Model</th>
<th>20 %ile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bootstrap 90% CI</td>
<td>0.95 → 2.83</td>
</tr>
<tr>
<td></td>
<td>1.706</td>
</tr>
<tr>
<td>RSDD</td>
<td>1.953 (1.92, 1.97)</td>
</tr>
<tr>
<td>RSLN</td>
<td>1.773</td>
</tr>
<tr>
<td>RSGARCH</td>
<td>1.660 (1.63, 1.68)</td>
</tr>
</tbody>
</table>

And this doesn’t help us much either.
Oversampling

- Bootstrapping re-samples from original data
- \(\Rightarrow\) Four 10-year accumulation factors from 584 observations
- What happens if we break the rules and keep sampling?
Oversampling

- If data are independent then oversampling
  → thin tails
  - I.e. positive bias for low quantiles
  - Bias should be small for large original sample

- If data are very auto-correlated and block size is not large enough
  → thin tails
  - I.e. positive bias for low quantiles
Oversampling

• If data are very auto-correlated
  ➢ Oversampling with small block size will fatten tails
  ➢ Overall effect depends on correlation

• But we are estimating AFs so we also look at these correlations.
Back to the data

- No significant negative autocorrelations...
- So oversampling should over-estimate left tail quantiles (on average)
## Left tail, 10-Year AFs

<table>
<thead>
<tr>
<th>Model</th>
<th>2.5%</th>
<th>5%</th>
<th>10%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bootstrap</td>
<td>1.041</td>
<td>1.228</td>
<td>1.478</td>
</tr>
<tr>
<td>(sort of…)</td>
<td>(1.03, 1.06)</td>
<td>(1.20,1.25)</td>
<td>(1.47,1.49)</td>
</tr>
<tr>
<td>RSDD</td>
<td>1.277</td>
<td>1.439</td>
<td>1.653</td>
</tr>
<tr>
<td>SLV</td>
<td>1.082</td>
<td>1.254</td>
<td>1.468</td>
</tr>
<tr>
<td>RSGARCH</td>
<td>0.905</td>
<td>1.086</td>
<td>1.315</td>
</tr>
<tr>
<td>RSLN</td>
<td>0.914</td>
<td>1.105</td>
<td>1.378</td>
</tr>
</tbody>
</table>
Summing up

• We need to pay attention to model econometrics
• Huge financial implications – especially with traditional actuarial methods
• Abusing the bootstrap offers some info
• Multiple state models for equity returns.
References

• C3P2  http://www.actuary.org/pdf/life/c3_june05.pdf

• Chan A and Wong W-S *Mixture Gaussian Time Series Modelling of Long-Term Market Returns* (2005) NAAJ To Appear

• Panneton C-M (2003)  
  www.actuaries.ca/meetings/stochasticsymposium/Papers/Panneton.pdf