BAYESIAN ESTIMATION OF STATE SPACE RESERVING MODELS

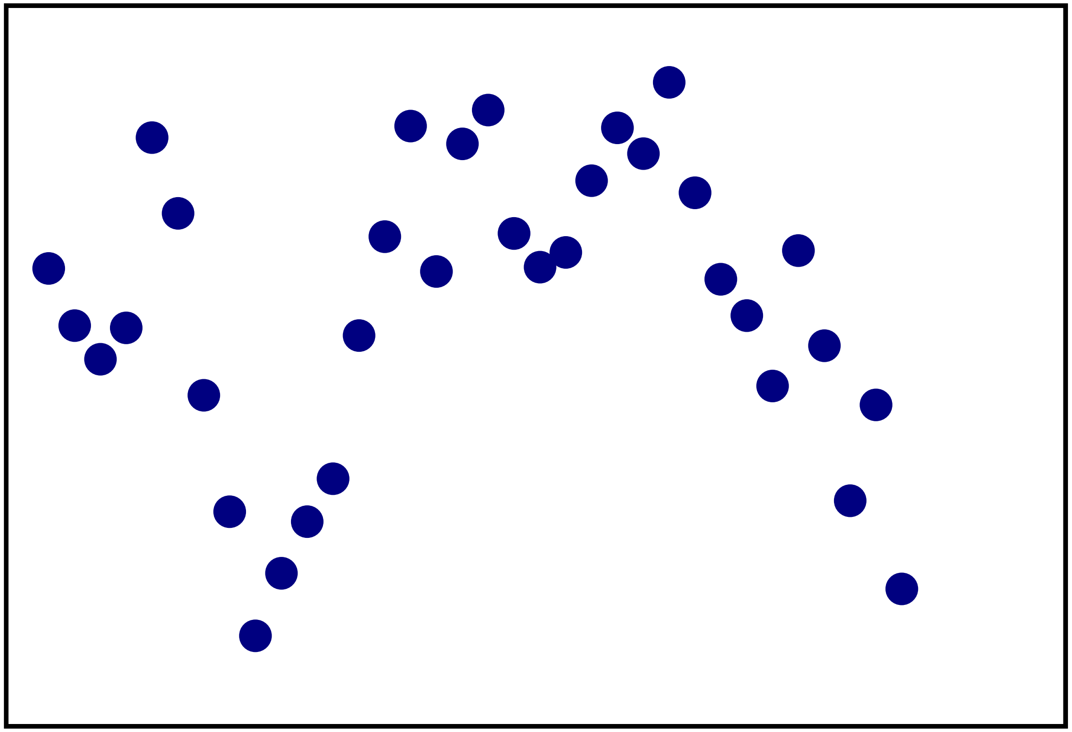
Casualty Loss Reserve Seminar
September 11-12, 2006

Frank Schmid, Director and Senior Economist, National Council on Compensation Insurance, Inc. 2006 Hicks-Tinbergen co-Medalist.

Jon Evans, Actuary, National Council on Compensation Insurance, Inc.

Why Use a State Space Model ?

In many situations, particularly time series, an estimate is needed for each observed data point and a traditional statistical model would either suggest too many parameters or be too rigid to fit the data.



What Is a State Space Model ?

Traditional Statistical Model

The parameter vector θ determines the distribution of the observed vector X.

A sample $x_1, \ldots, x_n$ is used to estimate θ.

$$X \sim F(x \mid \theta)$$

$$\hat{\theta} = S(x_1, \ldots, x_n)$$

State Space Model

The hyperparameter vector θ determines the distribution of the unobserved state vector W.

The state vector W and the hyperparameter vector θ determine the distribution of the observed vector X.

A sample $x_1, \ldots, x_n$ is used to estimate θ.

The estimated value of θ and the sample $x_1, \ldots, x_n$ are used to estimate W.

$$\hat{\theta} = S(x_1, \ldots, x_n)$$

$$W \sim g(w \mid \theta)$$

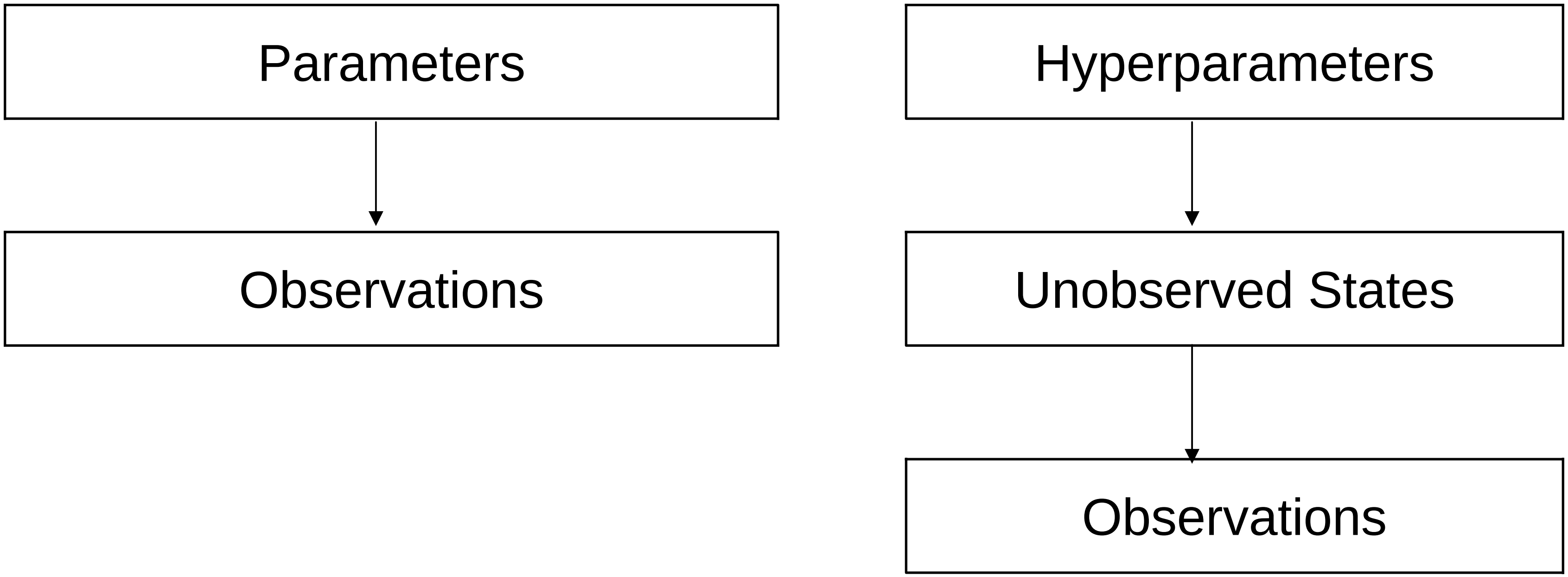
$$X_i \sim F(x_i \mid W_i, \theta)$$

$$\hat{\theta} = S(x_1, \ldots, x_n)$$

$$\hat{W}_i = T(x_1, \ldots, x_n, \hat{\theta})$$

Traditional Statistical Model

State Space Model



Random Walk with Stochastic Drift— A Basic Time Series State Space Model

Observation	$y_t = \mu_t + \varepsilon_t$	$\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$
Level State	$\mu_t = \mu_{t-1} + \beta_t + \eta_t$	$\eta_t \sim N(0, \sigma_\eta^2)$
Slope State	$\beta_t = \beta_{t-1} + \xi_t$	$\xi_t \sim N(0, \sigma_\xi^2)$

σ_ε^2, σ_η^2, and σ_ζ^2 are the hyperparameters.

1960 - The Kalman Filter Makes State Space Estimation of Time Series Practical

Kalman, R. E. "A New Approach to Linear Filtering and Prediction Problems“, *Transactions of the ASME - Journal of Basic Engineering,* Vol. 82: pp. 35-45 (1960).

- Extensively used up to the present day for signal processing, particularly in aerospace applications.
- A few hyperparameters specify the variance of measurement noise and the variance of movement of unobserved states, respectively.
- A sequential piecewise least squares estimate, the “filter” estimate requiring a trivial amount of calculation, is made of the unobserved states.
- The filter estimated states imply values for the observation noise, which can be used to calculate the likelihood function.
- The hyperparameters can be set based on prior knowledge of the system, MLE, or with modern computers even Bayesian estimation.

Likelihood Function for a Kalman Filtered Random Walk with Drift

$$L(y_1,\ldots,y_n,\hat{\mu}_1,\ldots,\hat{\mu}_n,\sigma_\varepsilon^2,\sigma_\mu^2,\sigma_\beta^2)=\prod_{i=1}^{n} f_{NormalDensity}(\hat{\mu}_i - y_i;0,\sigma_\varepsilon^2)$$

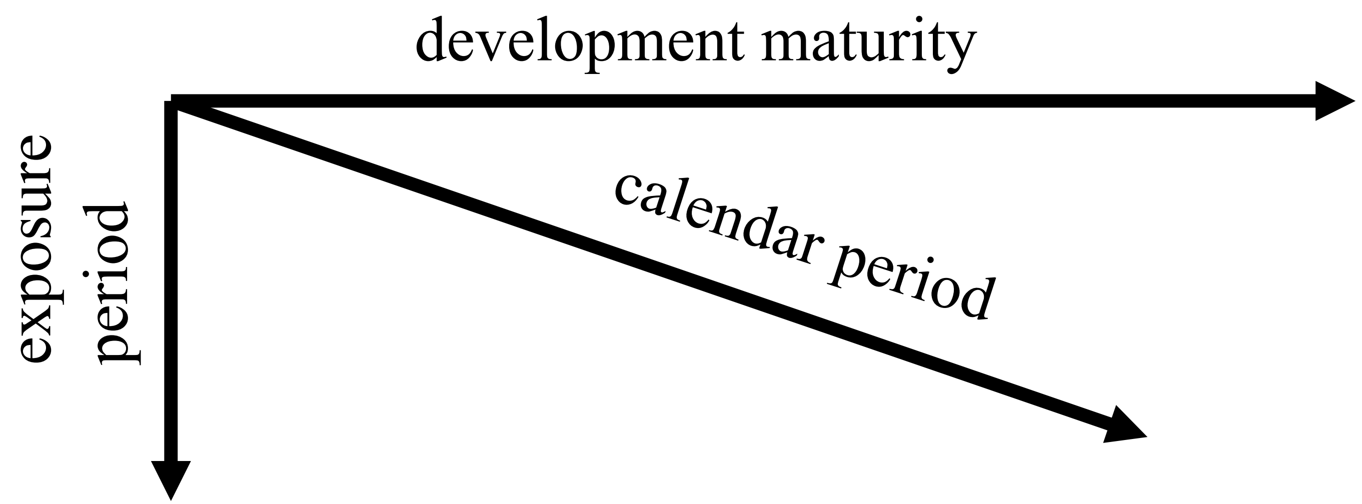
$$\hat{\mu}_i = KalmanFilter(y_1,\ldots,y_i,\sigma_\varepsilon^2,\sigma_\mu^2,\sigma_\beta^2)$$

Using the likelihood function σ_ε^2, σ_η^2, and σ_ζ^2 (and if necessary μ_0) can be estimated by either:

1. MLE, or
2. A Bayesian posterior assuming a diffuse prior distribution.

NCCI

The Challenge of Squaring The Liability Triangle



- only n(n+1)/2 observations for n exposure periods
- at least 3 directions of “trend” effects
- “trend” effects do not behave like rigid parameters

1983 - Zehnwirth and De Jong Use the Kalman Filter to Square Liability Triangles

de Jong, P.; Zehnwirth, Benjamin, "Claims Reserving State Space Models and Kalman Filter", *Journal of the Institute of Actuaries*, 1983. introduces a Kalman filtered states space model for squaring liability triangles:

- States are weights applied to a basis set of fixed payout patterns.
- Allows for inclusion of collateral information about calendar year inflation and volume by exposure year.
- Payment patten weights evolve over the time dimension of the exposure year.
- Kalman Filter is used allowing intial weights and the covariance parameters solved by MLE.

1980s to Present - Bayesian Estimation of Statistical Models Exploded in Popularity as Computer Power Made it Practical

- Gibbs sampling allows simulation of a random sample from a density that is known up to proportionality. (The normalization constant may be unknown.)
- This is exactly the case for the Bayesian posterior since the density is the product of the likelihood function and prior distributions.
- Additionally, computers can sometimes be used for numerical quadrature of the Bayesian formula or more rarely even closed form solution with computer algebra.
- Bayesian posteriors are advantageous in providing “complete” parameter uncertainty information in addition to process uncertainty information.

Complete Bayesian Approach to Hyperparameter Estimation Without the Kalman Filter

Rather than use the Kalman Filter to estimate the state values, diffuse prior distributions can be assumed for state values and the state values can be integrated out of both the numerator and denominator of Bayesian posterior formula.

$$P(\sigma_\varepsilon^2,\sigma_\mu^2,\sigma_\beta^2 \mid y_1,\dots,y_n) =$$

$$\frac{\displaystyle\int_{\hat{\mu}_1,\dots,\hat{\mu}_n} d\hat{\mu}_1 \dots d\hat{\mu}_n \quad L(y_1,\dots,y_n,\hat{\mu}_1,\dots,\hat{\mu}_n,\sigma_\varepsilon^2,\sigma_\mu^2,\sigma_\beta^2) P(\sigma_\varepsilon^2,\sigma_\mu^2,\sigma_\beta^2,\hat{\mu}_1,\dots,\hat{\mu}_n)}{\displaystyle\int_{\sigma_\varepsilon^2,\sigma_\mu^2,\sigma_\beta^2} d\sigma_\varepsilon^2\, d\sigma_\mu^2\, d\sigma_\beta^2 \int_{\hat{\mu}_1,\dots,\hat{\mu}_n} d\hat{\mu}_1 \dots d\hat{\mu}_n \quad L(y_1,\dots,y_n,\hat{\mu}_1,\dots,\hat{\mu}_n,\sigma_\varepsilon^2,\sigma_\mu^2,\sigma_\beta^2) P(\sigma_\varepsilon^2,\sigma_\mu^2,\sigma_\beta^2,\hat{\mu}_1,\dots,\hat{\mu}_n)}$$

Late 1990s To Present - Scollnik and Others Use Bayesian Methods without the Kalman Filter to Square Triangles

As a recent example, Scollnik, David P.M. "Implementation of Four Models for Outstanding Liabilities in WinBUGS: A Discussion of a Paper by Ntzoufras and Dellaportas (2002)", *North American Actuarial Journal*, Vol 6 No 1, 2002. uses WinBUGS software to simulate the Bayesian posterior of several state space models for squaring liability triangles:

- Complete Bayesian approach without the Kalman Filter.
- Includes stochastic exposure year, calendar year, and development lag effects in states
- Allow for inclusion of collateral information about counts of claims settled.
- Pure Bayesian approach with posterior distribution of state values simulated also.

WinBUGS – Free Bayesian Software that Estimates almost any Statistical Model

- Most traditional statistical models, such as GLM, correspond to cases where a tractable calculation exists for MLEs of the parameters.
- Modern computer power allows for tractable Bayesian estimation of a very general range of much more compex statistical models.
- WinBUGS software allows users to define random variables and relations between them (i.e., the parameters of the distribution of one variable are a function of another) in relatively arbitrary graphs that can have many layers. Observation data is input and those random variables that are not observed are treated as parameters whose Bayesian Posterior is estimated.
- WinBUGS is widely used by academic researchers and is free on the internet.
- It even has a GUI so you can create the model by drawing graphs (called "doodles").

A Recent Version of Schmid's Triangular Liability Model under Development at NCCI

- empirical measurement of log of incremental → $y_{i,j} \sim \mathrm{N}(b_{i,j}, \sigma_y^2)$, $i = 1,\dots,N$, $j = 1,\dots,N+1-i$
- Cusum constraint → $\hat{\mathbf{z}}_i{}' - \mathbf{z}_i{}' \sim \mathrm{N}(\mathbf{0}, \Sigma)$, $i = 1,\dots,N$
- summing incrementals → $\hat{z}_{i,j} = \log\left(\sum_{k=1}^{j} \exp\left(\hat{y}_{i,k}\right)\right)$, $i, j = 1,\dots,N$
- extra noise on incremental of log effect → $\hat{y}_{i,j} \sim \mathrm{N}(b_{i,j}, \sigma_y^2)$, $i, j = 1,\dots,N$
- Initial upper left corner state → $b_{1,1} \sim \mathrm{N}(y_{1,1}, \sigma_{11}^2)$
- "random walking down" the lag 1 column → $b_{i,1} = b_{i-1,1} + \Lambda_{i,1}$, $i = 2,\dots,N$; $\Lambda_{i,1} = \eta_{i,1} + \kappa_{i+1}$, $i = 2,\dots,N$
- random walking exposure year effect → $\eta_{i,1} \sim \mathrm{N}(\eta_{i-1,1}, \sigma_\eta^2)$, $i = 3,\dots,N$
- Use of informed prior (rate of inflation) for calendar year effect → $\kappa_{i+j} \sim \mathrm{N}(\pi_{i+j}, \sigma_\kappa^2)$, $i + j = 4,\dots,N+1$
- "random walking down" across columns of development maturity → $b_{i,j} = b_{i,j-1} + \Lambda_{i,j}$, $i = 1,\dots,N$, $j = 2,\dots,N$; $\Lambda_{i,j} = \delta_j + \kappa_{i+j}$, $i = 1,\dots,N$, $j = 2,\dots,N$
- random walking development maturity effect → $\delta_j \sim \mathrm{N}(\delta_{j-1}, \sigma_\delta^2)$, $j = 3,\dots,N$

The row vectors $\hat{\mathbf{z}}_i$ and $\mathbf{z}_i$ comprise, respectively, the fitted and actual logarithmic cumulative payments of exposure year i.

Summary

- Actuaries face increasing demands to estimate reserve uncertainty, requiring true statistical models of triangular liability.
- State space models for triangular liability have been around for over 20 years and are the most promising approach to the challenge.
- The Kalman Filter estimate states based on the hyperparameters and observed data.
- The Kalman Filter allows for MLE of the hyperparameters given the the observed data.
- Bayesian estimation of hyperparameters with or without the Kalman Filter has also become practical with increasing computer power.
- State space models of triangular liability can be formulated and estimated with Kalman Filter/MLE, Kalman Filter/Bayesian, or complete Bayesian approaches.
- WinBUGS is a free statistical software package that allows for simulation estimation of the Bayesian posterior for extremely general models.

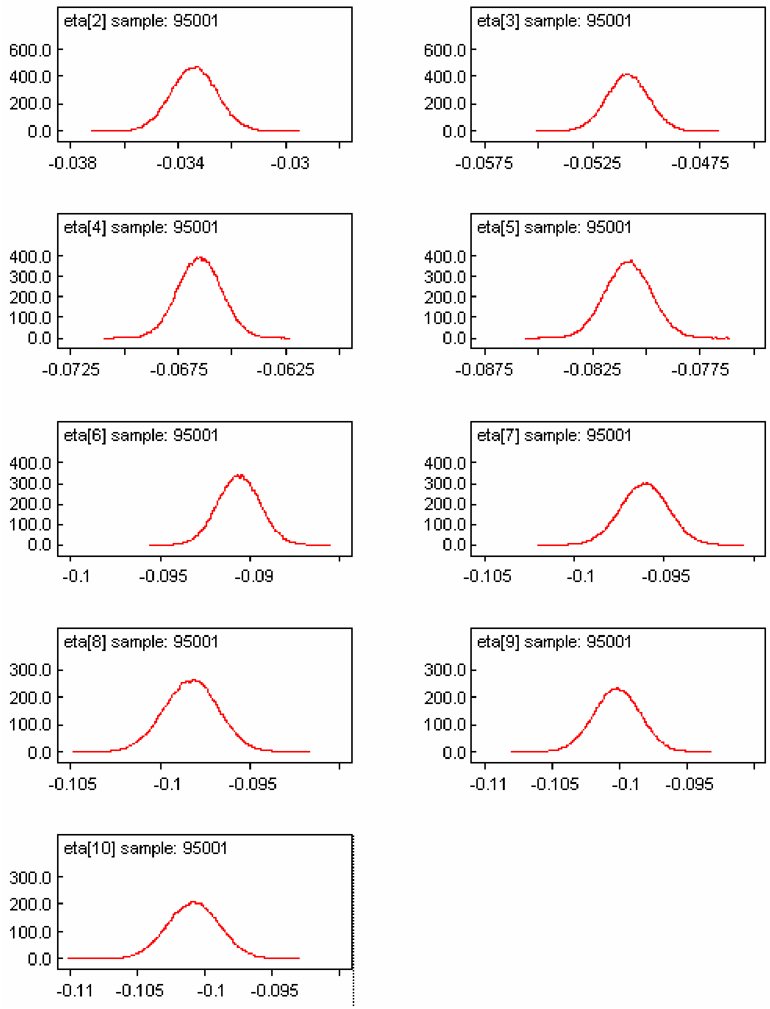
Example Output Of Schmid's Model For A Reinsurance Data Set*

*Automatic Facultative Reinsurance on General Liability (excl. asbestos and environmental), *Historical Loss Development Study*, 1991 edition, Reinsurance Association of America. Cited after Mack (1995).

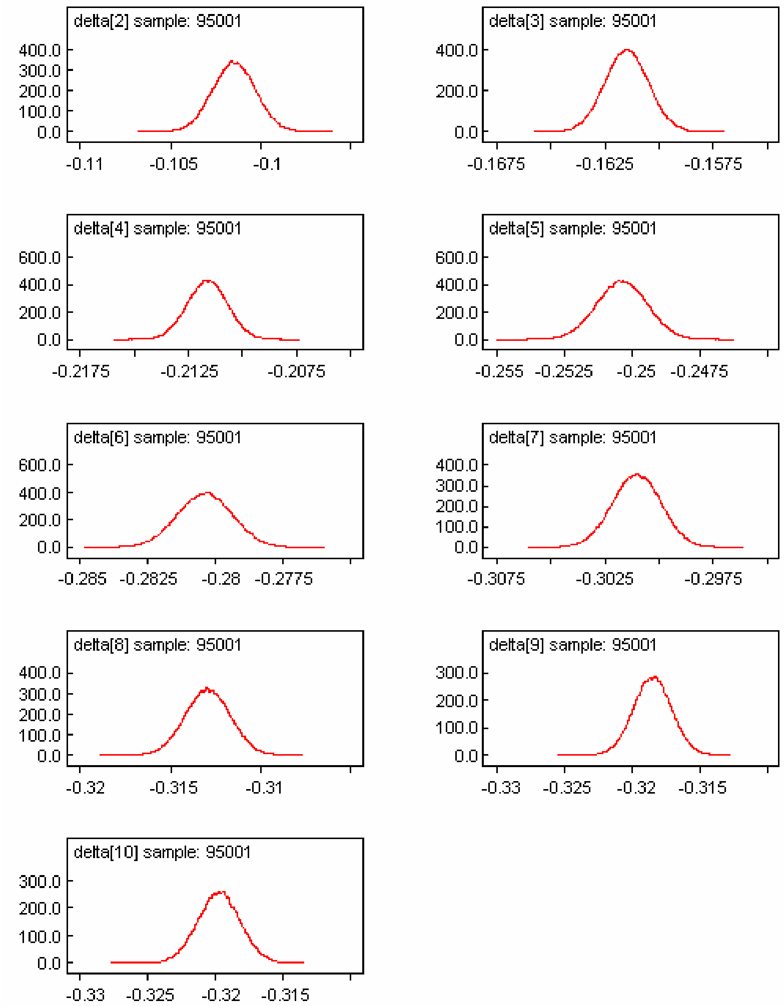
AFG Data Set, Nodes, Statistics

node	mean	sd	MC error	10.0%	median	90.0%	start	sample
delta[2]	-0.101467	0.00118289	9.1801E-6	-0.102991	-0.101466	-0.0999526	5000	95001
delta[3]	-0.161491	0.00100368	8.24836E-6	-0.162775	-0.161492	-0.160207	5000	95001
delta[4]	-0.211628	9.26326E-4	8.45663E-6	-0.212817	-0.211625	-0.210445	5000	95001
delta[5]	-0.25037	9.34228E-4	1.02354E-5	-0.25157	-0.250368	-0.249175	5000	95001
delta[6]	-0.280395	0.00100947	1.26785E-5	-0.28169	-0.280387	-0.279099	5000	95001
delta[7]	-0.301028	0.00111734	1.55165E-5	-0.302459	-0.301023	-0.299604	5000	95001
delta[8]	-0.312929	0.00125036	1.8375E-5	-0.314531	-0.312929	-0.311324	5000	95001
delta[9]	-0.318473	0.00140123	2.06877E-5	-0.320261	-0.318471	-0.316676	5000	95001
delta[10]	-0.319743	0.0015596	2.17991E-5	-0.321749	-0.31974	-0.317745	5000	95001
eta[2]	-0.0334044	8.47765E-4	4.689E-6	-0.0344922	-0.033402	-0.0323194	5000	95001
eta[3]	-0.0508612	9.66245E-4	6.67209E-6	-0.0520967	-0.0508633	-0.0496235	5000	95001
eta[4]	-0.0664967	0.00101118	8.41603E-6	-0.067792	-0.0664955	-0.0651992	5000	95001
eta[5]	-0.0808292	0.00107057	1.05204E-5	-0.0821995	-0.0808335	-0.0794575	5000	95001
eta[6]	-0.0906137	0.00117565	1.39413E-5	-0.0921132	-0.0906153	-0.0891115	5000	95001
eta[7]	-0.0960968	0.00132867	1.82239E-5	-0.0978015	-0.0960903	-0.0944045	5000	95001
eta[8]	-0.0982696	0.00151669	2.26636E-5	-0.100217	-0.0982529	-0.0963377	5000	95001
eta[9]	-0.100164	0.00171727	2.56105E-5	-0.102371	-0.100159	-0.0979697	5000	95001
eta[10]	-0.10086	0.00192312	2.67303E-5	-0.103326	-0.100851	-0.0984036	5000	95001
kappa[3]	0.0359923	0.00106799	4.48936E-6	0.0346225	0.0359974	0.0373605	5000	95001
kappa[4]	0.0459253	0.00103993	4.08892E-6	0.0445903	0.0459272	0.0472548	5000	95001
kappa[5]	0.0186086	0.00101017	4.04802E-6	0.0173156	0.0186043	0.0199046	5000	95001
kappa[6]	0.0339335	0.00103146	4.12277E-6	0.0326148	0.0339309	0.0352493	5000	95001
kappa[7]	0.0276944	0.00102075	4.14079E-6	0.0263816	0.027694	0.0290008	5000	95001
kappa[8]	0.0100997	0.00100528	4.1972E-6	0.00881511	0.0100986	0.0113947	5000	95001
kappa[9]	0.0274553	0.00102374	3.7581E-6	0.0261397	0.0274533	0.0287709	5000	95001
kappa[10]	0.032554	0.00102385	3.68503E-6	0.0312467	0.0325556	0.0338647	5000	95001
kappa[11]	0.0414044	0.00103348	3.41325E-6	0.0400782	0.0414046	0.042725	5000	95001

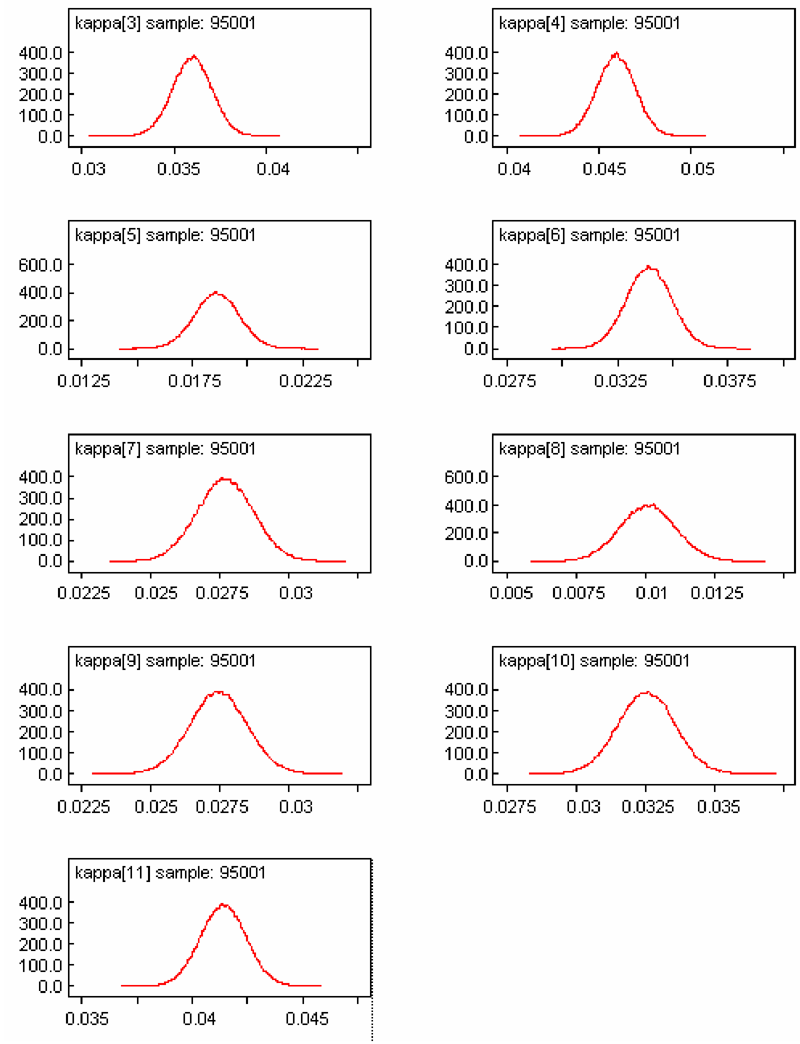
Exposure Year Effect Bayesian Posterior Samples



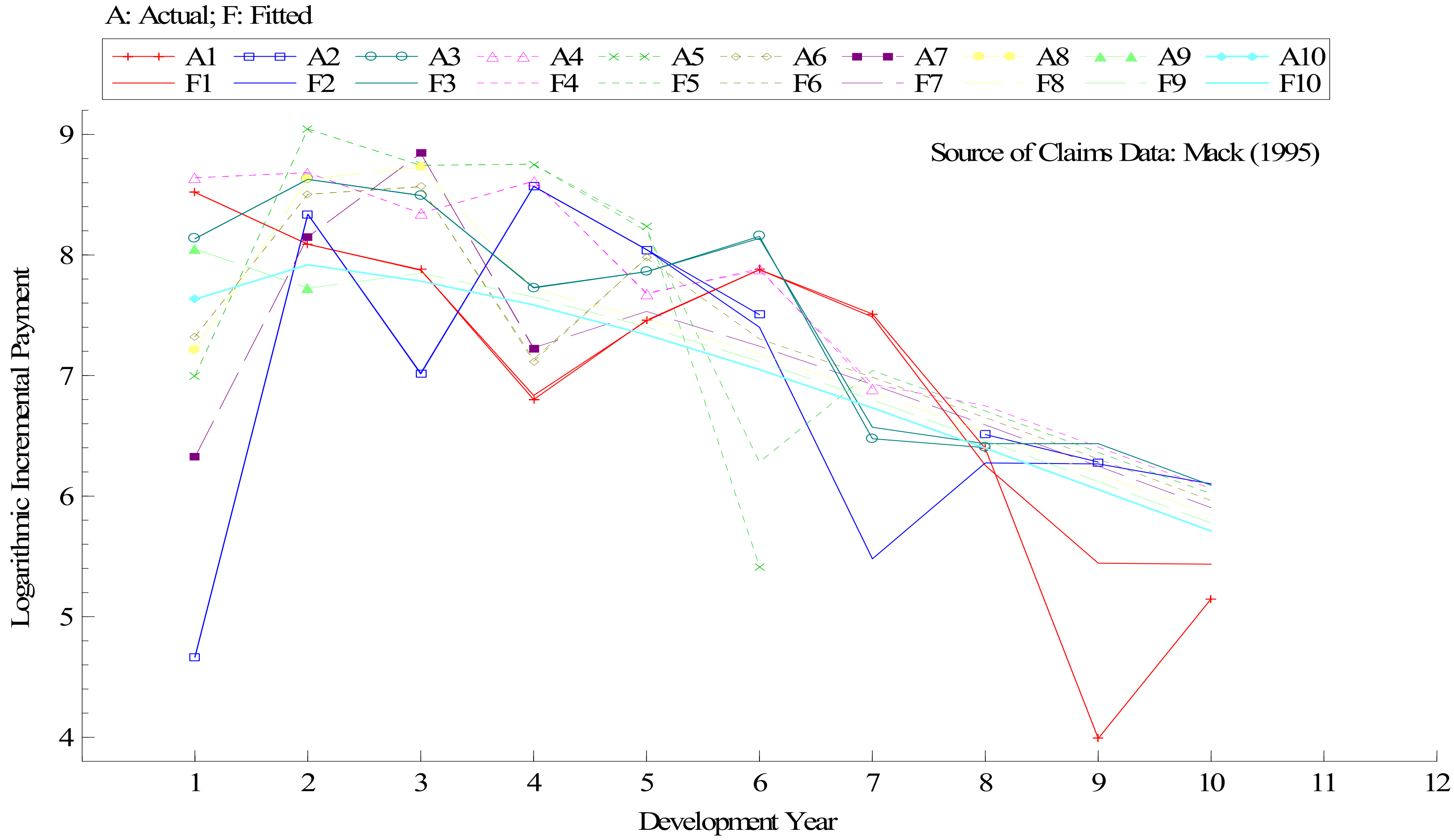
Development Maturity Effect Bayesian Posterior Samples



Calendar Year Effect Bayesian Posterior Samples



Actual Versus Fitted Incremental Payments By Development Maturity



Live WinBUGS Demo on AFG Data Set

(time permitting)