Measurement of Yield Distributions:
Time-Varying Mixture Distribution Models

Tsung Yu Yang

North Carolina State University
Department of Economics

Campus Box8110, Raleigh
NC27695-8110
tyang4@ncsu.edu

Selected Paper prepared for presentation at the Agricultural & Applied Economics Association's 2011 AAEA & N AREA Joint Annual Meeting, Pittsburgh, Pennsylvania, July 24-26, 2011

Abstract

Regarding the nature of yield data, there are two basic characteristics that needs to be accommodated while we are about to model a yield distribution. The first one is the nonstationary nature of the yield distribution, which causes the heteroscedasticity related problems. The second one is the left skewness of the yield distribution. A common approach to this problem is based on a two-stage method in which the yields are detrended first and the detrended yields are taken as observed data modeled by various parametric and nonparametric methods. Based on a two-stage estimation structure, a mixed normal distribution seems to better capture the the secondary distribution from catastrophic years than a Beta distribution. The implication to the risk management is the yield risk may be underestimated under a Beta distribution. A mixed normal distribution under a time-varying structure, in which the parameters are allowed to vary over time, tends to collapse to a single normal distribution. The time-varying mixed normal model fit the realized yield data in one step that avoid the possible bias caused by sampling variability. Also, the time-varying implies that the premium rates can be adjusted to represent the most recent information and that lifts the efficiency of the insurance market.

Key Words: Time-Varying Distribution, Mixture Distribution, Crop Insurance

1. Introduction

A precise yield risk assessment depends on the accuracy of modeling the distribution. Regarding the nature of yield data, there are two basic characteristics needed to be accommodated while we are about to model a yield distribution. The first one is the nonstationary nature of the yield distribution. This is primarily due to the technological progress and changing environmental conditions. That is, the estimated residuals derived from a trend model are conditional on that model, and thus the heteroscedasticity would arise and lead to misspecification to the data. The heteroscedasticity is quite a concern to the selection of modeling approach (Goodwin and Ker, 2001).

The second one is the negative skewness of the yield distribution, which might be caused by catastrophic events. Ker and Goodwin (2000) inferred to this issue and suggested that the conditional yield may best be modeled as a mixture of two unknown distributions where the secondary distribution from catastrophic years lives on the lower tail of the primary distribution from non-catastrophic years and has less mass. That is because the realized yields are far less in catastrophic years and catastrophic events are realized with less frequency.

There are lots of methods to modeling the characteristics of a yield distribution. The common catalog to distinguish these methods is parametric, nonparametric. Most of those who adopt the parametric methods encounter the model selection problem. Because the space of parametric models is dense, the probability of assuming the correct parametric model is zero. Sufficiently rich data will be needed to reduce the set of candidates to the point of economic invariance (Goodwin and Ker, 2001). Alternatively, nonparametric methods could mitigate the model selection problem, such as Ker and Goodwin (2000), and Turvey and Zhao (1993) use nonparametric kernel method to avoid the possible statistical inconsistency. However, nonparametric methods tend to be inefficient relative to maximum likelihood methods when the assumed parametric model is correct. Based on the aspect of view, there exist some other methods, which either combine the advantages of parametric and nonparametric estimators or use extraneous yield data to improve the efficiency of the nonparametric estimator.

Among the variety of estimation methods, a very common approach of “two-stage” estimation is widely applied. In the first stage, the technique of eliminating the “trend” effect of yield data is very crucial for establishing the correct distributional properties of the detrended data in the second stage. It has been recognized that the resulting estimated residuals, representing the detrended yields, are subject to the estimation uncertainty associated with sampling variability in the first stage estimates of trend and thus may not necessarily provide an accurate representation of the actual yield distribution (Zhu et al., 2010; Robinson, 1987).

Some conditional parameters methods are proposed to fit the realized data in one step and avoid the possible bias in the two-stage estimation. For example, Zhu et al. (2010) proposed a flexible parametric model, which simultaneously and coherently specifies the first four moments using suitable polynomials of time. Goodwin et al. (2000) used a conditional heteroskedasticity model to characterize the nonconstant variances of crop prices. The model parameters are then estimated simultaneously by maximizing the resulting likelihood function.

The same feature among the estimation models is that an appropriate postulated distribution is needed to capture the characteristics of the data. Beta distribution is the common postulated distribution for yield data. Nelson and Preckel (1989), Hennessy et al. (1997), and Borges and Thurman (1994) found sufficient evidence of skewness and / or kurtosis in their yield data and use the Beta distribution instead of Normal distribution. However, the Beta distribution cannot accommodate one of the main possible distributional structures (Goodwin and Ker, 2001). Therefore, a mixed normal distribution may be considered as another candidate for capturing the skewness and kurtosis of yield data.

In this study, I would first compare the specifications of Beta distribution and a mixed normal distribution model based on the same detrended yields. Additionally, a time-varying model will be adopted to allow the conditional parameters to vary over time. The specifications based on different models and postulated distributions will then be examined. Finally, a simulation practice for insurance premium rates will be used to assess the influences on pricing an insurance contract.

2. A conventional Two-Stage Estimation

The first stage of conventional two-stage estimation is to detrend the time series of yield. Based on various assumptions to a trend model, there are lots of methods for the detrending purpose. A quadratic trend model would be applied as the basic assumption to process the first stage. A comprehensive survey of other possible detrending methods is not in the scope of this study. The main purpose is to process two-stage estimation as the benchmark for the the time-varying model.

Consider the following trend model:

$Y_t = m(X_t) + \varepsilon_t$

where Y_t is the observed crop yield in year t, (t = 1, ..., T), $m(X_t)$ denotes the regression function $E(Y_t \mid X_t = x)$, X_t represents linear or nonlinear time indexes representing trend, and ε_t represents residuals that are assumed to be independently distributed with mean zero. The regression function $m(X_t)$ can be estimated nonparametrically using kernel methods or smoothing spline methods. Alternatively, if we assume a parametric functional form for $m(X_t)$, then the regression coefficients can be obtained by using ordinary least squares (OLS). In either case, the residuals are obtained as $\widehat{\varepsilon_t} = Y_t - \widehat{m}(X_t)$.

I adopt the corn yield data of Kossuth county from year 1926 to 2009. The empirical data are county-level yields applied from National Agricultural Statistics Services (NASS). Comparing the R square in the trend models, the R square, 0.9126, in the cubic trend model does not significantly increase. The significance of the coefficients in the quadratic trend model also suggests the quadratic trend stands for a better specification of the realized data (See Table 1).

$Y_t = a + bt + ct^2 + \varepsilon_t$ .... (1)

$\varepsilon_t \sim N(0, \sigma^2)$

$E(Y_t) = m(X_t = t) = a + bt + ct^2$,

$Y_t \sim N(E(Y_t), \sigma^2)$

Figures 1 to 4 demonstrate the possible temporal heteroscedasticity, and the regression of the square error on time also verifies the significant temporal pattern of residuals (See Table 2). In order to specify the temporal heteroscedasticity effect, a rescaled form of the deviations from a trend-based equation is applied (Miranda and Glauber, 1997; Atwood et al., 2003):

$\widetilde{Y_t} = \widehat{Y_{09}}\,(1 + \frac{\widehat{\varepsilon_t}}{\widehat{Y_t}})$

$\widehat{Y_{09}}$ refers to the predicted yield in the last year. The residual of year t is scaled to $\widehat{Y_{09}}$ by dividing it correspondent predicted yield in year t and hence to obtain the detrended yield, $\widetilde{Y_t}$. We are looking forward to preserving as much information as possible other than the trend effect in this detrending process. So that, in the second stage an appropriate distribution would be applied to the specification for detrended yields.

Figure 5-1 demonstrates that the detrended data are more negatively skewed than the data applied under normal distribution. Q-Q plots (Figure 5-2, 5-3) based on residuals, ε_t, and the detrended data, $\widetilde{Y_t}$, also verify the skewness. These examinations suggest that the distribution selections other than normal distribution, which could well specify the skewness of data set, may be an executable candidate for the detrended data. Those candidate selections may include Beta, Mixture distributions, or Kernel (non-parametric). In this study, I adopt a mixed-normal distribution in order to capture the potential cluster with less frequency.

In the second stage, the parameters of a two-component mixture normal distribution based on the detrended yield would be estimated by maximum likelihood estimation. The log-likelihood function is given by,

LLF (λ, μ_1, μ_2, σ_1, σ_2 | $\widetilde{Y_t}$, t =1, 2, ..., T)

$$= \sum_{t=1}^{T} log \left\{ \lambda \left[\frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(\frac{(\widetilde{Y_t} - \mu1)^2}{2\sigma_1^2}\right) \right] + (1-\lambda) \left[\frac{1}{\sqrt{2\pi\sigma_2^2}} \exp\left(\frac{(\widetilde{Y_t} - \mu2)^2}{2\sigma_2^2}\right) \right] \right\},$$

where, $\widetilde{Y_t} = \widehat{Y_{09}}\ (1 + \frac{\widehat{\varepsilon_t}}{\widehat{Y_t}}) = \widehat{Y_{09}}\ (1 + \frac{Y_t - E(Y_t)}{\widehat{Y_t}}) = \frac{\widehat{Y_{09}}}{\widehat{Y_t}} Y_t$. Let $\frac{\widehat{Y_{09}}}{\widehat{Y_t}} = \ell_t$, and the detrended yield, $\widetilde{Y_t}$, can be represented as the ratio of predicted yield in year T (=2009) to the predicted yield in year t times the realized yield, Y_t . Therefore, the pseudo log-likelihood function of the parameters based on the original data Y_t is LLF (λ, μ_1, μ_2, σ_1, σ_2 | Y_t , t =1, 2, ..., T)

$$= \sum_{t=1}^{T} log \left\{ \lambda \left[\frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(\frac{(\ell_t Yt - \mu1)^2}{2\sigma_1^2}\right) \right] + (1-\lambda) \left[\frac{1}{\sqrt{2\pi\sigma_2^2}} \exp\left(\frac{(\ell_t Yt - \mu2)^2}{2\sigma_2^2}\right) \right] \right\},$$

3. Time-Varying Mixed Normal Distribution Model

The basic assumption of the time varying model is that the parameters of the distribution follow a specific temporal pattern, such that the whole temporal changes of the yield distribution can be captured by the time-varying parameters. The time-varying model accounts for parameter uncertainty by maximizing the time-varying likelihood function, which is to estimate time-trend parameters and the distributional parameters in one step.

Zhu et al. (2010) use an exponential form to estimate the variation of the parameters in terms of a trend model to the Beta distribution. The exponential functional form ensures that the Beta shape, scale, and location parameters are positive at every observation. Thus, the log-likelihood of the time-varying Beta distribution model is identical to that of the constant Beta distribution model. The results suggest that the time-varying Beta model could better specify the yield data than a conventional two-stage Beta distribution model; particularly, in the specification to the trend of yield.

In this study, I accept the idea to allow the means and variances varying over time in a mixed normal distribution model. Because the increasing yield and the heteroscedasicitty of the error terms, we can anticipate the means and variances have significant trend over time and thus we set up the means and variances as :

$\mu_1 = \exp(a+b*t);$

$\mu_2 = \exp(c+d*t);$

$\sigma_1 = \exp(e+f*t);$

$\sigma_2 = \exp(g+h*t);$

$\lambda = 1 + (0-1) / (1+\exp(w))$

The temporal pattern for means and variances of the density functions may be better specified by a quadratic trend model. However, the quadratic trend does not obviously raise the log-likelihood value. Hence, I simply set up a linear trend model to the means and variances. The lambda is bounded between 0 to 1 in order to estimate the possible secondary distribution from catastrophic years on the lower tail of the primary distribution. We should note that the optimality is not global for this non-linear log-likelihood function. It is restricted to a local, but reasonable location, which allows us to specify the realized yield.

The time-varying model not only addresses the dynamic characteristics of yield distributions, but also provides a more flexible specification of the heteroscedasticity and higher order moments (e.g. skewness and kurtosis). Ideally, the time-varying model could well specify the trend and heteroscedasticity due to the avoidance of possible bias from the estimation uncertainty in the two-stage estimation.

4. Specification Tests and Model Performance

4.1 The Vuong Test

Vuong's non-nested specification test (Vuong, 1989) is a likelihood-based test for model selection. The Vuong's test statistic is given by :

$$v = n^{-1/2} LR_n(\hat{\theta}_n, \tilde{\theta}_n) / \hat{w}_n$$

where n is the number of observations in the sample, $LR_n(\hat{\theta}_n, \tilde{\theta}_n)$ = the difference between the maximum log-likelihood values of the competing models. $\hat{w}_n$ is defined as :

$$\hat{w}_n = \text{sqrt}\{1/n \sum_{t=1}^{n}(log\frac{f(Y_t|X_t;\hat{\theta})}{g(Y_t|X_t;\tilde{\theta})})^2 - (1/n\sum_{t=1}^{n} log\frac{f(Y_t|X_t;\hat{\theta})}{g(Y_t|X_t;\tilde{\theta})})^2\}$$

Because of the Vuong statistic's sensitivity to the number of estimated coefficients in each model, the test will be corrected for the model dimensionality (Clarke, 2007). A correction suggested by Vuong (1989) corresponds to Schwarz's (1978) Bayesian information criteria. The adjusted statistic becomes

$$\widetilde{LR}_n(\hat{\theta}_n, \tilde{\theta}_n) = LR_n(\hat{\theta}_n, \tilde{\theta}_n) - [(\frac{p}{2}) \ln(n) - (\frac{q}{2}) \ln(n)],$$

where p and q are the number of estimated coefficients in the competing models. The test statistic is asymptotically normally distributed, and the actual test is therefore

$$H_0: \frac{1}{n}\widetilde{LR}_n(\hat{\theta}_n, \tilde{\theta}_n) \xrightarrow{a.s.} E[\ln\frac{f(Y|\hat{\theta}_n)}{g(Y|\tilde{\theta}_n)}] = 0$$

As specified, if v > c, where c is Normal (0, 1) critical value for some significant level, we reject null that the models are the same in favor of the model ($f(Y \mid \hat{\theta}_n)$). Alternatively, if v ≤ -c, we would reject the null in favor of the alternative model ($g(Y \mid \tilde{\theta}_n)$).

4.2 The Goodness of Fit

The goodness of fit for the competing models is evaluated based on Akaike Information Criterion (AIC), AICC, and the Bayesian Information Criterion of Schwarz (BIC). The idea of AIC (Akaike Information Criterion) is to maximize the "goodness of fit" minus "complexity." The expression of AIC is -2L (θ) +2K; the BIC is -2L (θ) +Kln(n) where L(θ) is the log-likelihood of the model evaluated at the MLE, n is the number of observations in the sample and K is the number of parameters in the model. AICC is AIC with a second order correction for small sample sizes, which is AIC+2K (K+1) / (n-K-1). AICC converges to AIC as n gets large. The smallest AIC, AICC, or BIC gives the best model.

5. Empirical Results

In the first part of this study, I compare the specifications of Beta, mixed normal and normal distributions in the conventional two-stage model. In the second part, I test and compare the performance of a mixed normal in the conventional two-stage model and in the time-varying model. The data applied are the corn yield data of Kossuth County, Iowa from 1926 to 1990. It is widely recognized that yield data have characteristics of the nonstationary nature, so does the data I apply.

5.1 Two- Stage Estimation

The initial analyses of the two-stage model are presented in figure 1 to 5. Figure 1 to 4 demonstrates a significant trend and the heteroscedasticity of the error term, which is estimated from the assuming temporal process (equation 1) . Figure 5-1demonstrates that the detrended yields represent negative skewness. Q-Q plots based on the residuals and predicted yields (from equ. 1) also suggest that both the residuals and predicated yields are more negatively skewed than what would be implied by the normal distribution. This implies that a distribution selection other than normal distribution, such as Beta or mixed normal might be under the consideration.

5.1.1 MLE – Beta vs. Mixture Normal

From previous specification tests, we have strong evidence to prove that the normality of the detrended yield is not supported. The questions we are going to ask next are -- which distribution candidate would better capture the left-tailed events, and what is the implication on the insurance premium rate?

Before examining the performances of more flexible distribution assumptions of Beta and mixed normal, we would like to identify the parameters of individual distribution. Unfortunately, some regularity conditions do not hold with the mixed normal model, i.e., constrained mixing parameter (λ) could cause some parameters cannot be identified when it lie on the boundary (at 0 or 1). That is, a standard nested hypothesis test statistic, such as likelihood ratio, no longer follows its asymptotic null distribution (chi-squared). Therefore, we need a non-nested test for determining the number of components. This is equivalently to test a null hypothesis that mixing parameter is 0 (or 1), or statistical difference between mixed normal and normal distribution.

Vuong test is applied to test the significance of the mixing parameter of the mixed normal. The statistic of 1.454 enables us to distinguish the mixed normal is statistically better (under 10% level of significance) than normal distribution. That implies the alternative component with a smaller mean and variance can be significantly identified. Moreover, we can also identify the less frequent low-mean and low-variance regime from the difference between the log-likelihood values, AIC, AICC, and BIC of the mixed normal, normal, and Beta distribution. Most of those criteria suggest that the mixed normal distribution may have better specification on the detrended data.

As would be expected from Figure 1, the premium rate based on mixed normal distribution is higher than that based on Beta and normal distribution. This reflects the thicker left tail of the mixed normal distribution. Also, this implies that the risk estimation with a Beta distribution and a normal distribution may underestimate the premium rate. The implication of the underestimation might need more consideration since the extreme events occur in catastrophic years tend to cause larger and larger loss.

The simulated premium rate is base on a guarantee of 75 percent of the expected annual yields with a predetermined price of 5 dollars. An insurance premium rate in this simulation practice is given by expected loss over total liability. The expected loss is defined as the fair premium of the insurance contract and takes the form as

$$E(Loss) = E[(\lambda Y^e - Y) \mid (Y \leq \lambda Y^e)]P = E[(\lambda Y^e - Y)^+]P,$$

where the contract offers a guarantee of λ ($\lambda \in (0,1)$), which is a portion of the expected yield (Y^e) . The expected loss states that if $Y \leq \lambda Y^e$, the insurer will pay $(\lambda Y^e - Y)*P$ as an indemnity. P denotes the predetermined price.

5.2 Time-Varying Mixed Normal Distribution Model

The Vuong test statistic is -0.47026 (see, Table 5), which is larger than the critical value of 10 percent level of significance, -1.282, suggests that (1-λ) is not significantly different from zero. With a correction corresponding to Schwarz's Bayesian information criteria, the adjusted Vuong test suggests that a single time-varying normal is in favor (i.e., the mixing parameter λ is significantly equal to 1). Therefore, we can conclude that under the estimation of a time-vary mixture normal, the density collapses into a time-varying single normal distribution. The maximized log-likelihood function is nearly identical to that of the time-varying single normal distribution (the -2 Log Likelihood values are identically equal to 672.9).

As to this time-varying model, we can think of each annual yield is drawn from an independent normal distribution and the parameters (mean and standard deviation) of the normal distributions are conditioned by linear temporal processes. That is to say, the normal distributions would be characterized by different parameters annually. The estimated model demonstrates increasing mean and variance, which reflects the technology improvement and the raising instability of crop yields.

As would be expected, insurance premium rates based on higher expected yields and lower variances are relatively lower. The lower rate in part reflects the higher forecasted yields and lower expected loss. Based our case, the model represents increasing estimates of mean and variance. The premium rate therefore is not necessarily decreasing year by year as expected. Nevertheless, a simulation practice is adapted to confirm the descending premium rate in this time-varying model. This result may illustrate the technology improvement effect is larger than the variability of crop yields.

Based on results of the Vuong test and other information criterions, the time-varying model is significantly better than the two-stage estimation with mixed normals. The simulated premium rate based on the time-varying model in 2009 is 0.0041 and the rate based on the two-stage estimation with mixed normals under the same coverage is 0.012. Thus, the time-varying model may offer another approach that could improve the efficiency of pricing the insurance contracts.

6. Conclusion

Due to the nonstationary characteristics of the yield data, an appropriate postulated distribution is needed to well specify the realized yields. This study compares the performance of Beta distribution and mixed normal distribution model given the same detrended yield, and the time-varying mixed normal distribution model. We actually exam the specification abilities of Beta and mixed normal distribution to the skewness and kurtosis. Also, we exam whether the time-varying model has a superior fit than the conventional method.

Most criteria of goodness of fit suggest that mixed normal distribution provide a superior fit to the detrended yields than Beta distribution in the conventional two-stage structure. However, the result of Vuong's nonnested specification test does not suggest a significantly superior model between them. Overall, under the conventional two-stage structure, the mixed normal distribution slightly outperforms the Beta distribution at specification to the skewness and kurtosis based on the yield data I use.

While the parameters of mixed normal distribution are allowed to vary over time, its goodness of fit and performance of specification are all superior to the conventional two-stage methods. This model simultaneously and coherently specifies nonstationary nature and the parameters of the yield distribution and therefore overcomes possible drawbacks of treating the detrended yields as observed data.

The simulation practice of insurance premium rates based on the two-stage model and the time-varying model may result in several suggestions for pricing crop insurance contracts. First, the premium rate might be underestimated due to the underestimation of the probability of yields in catastrophic years. Second, when the parameters evolve over time, we find this time-varying model can more accurately capture the yield risk, which implies that the premium rates can be adjusted to represent the most recent information.

In conclusion, although the results of statistical tests may be different according to different yield data sets, the suggestions to the candidate selection should be consistent while abound data are applied. The superior specification of the time-varying model implies the potential to improve the accuracy of models used in rating crop insurance contracts and thus may improve risk management mechanisms to protect producers from risk.

Reference

Atwood, J., S. Shaik, and M. Watts. 2003. "Are Crop Yields Normally Distributed? A Re-examination." *American Journal of Agricultural Economics* 85(4):888-901.

Borges, R., and W. Thurman. 1994. "Marketing Quotas and Subsidies in Peanuts." *American Journal of Agricultural Economics* 76: 801-809.

Goodwin, B. K.., M. C. Roberts, and K. H. Coble. 2000. " Measurement of Price Risk in Revenue Insurance: Implications of Distributional Assumptions." *Journal ofAgricultural and Resource Economics* 25(1): 195-214.

Goodwin, B.K., and A.P. Ker. 2001. "Modeling Price and Yield Risk." Chapter in A Comprehensive Assessment of the Role of Risk in U.S. Agriculture, pp. 289-323.

Clarke, K. A. 2007. "A Simple Distribution-Free Test for Nonnested Model Selection." *Political Analysis* 15:347–363.

Feng, Z. D., and C. E. McCulloch. 1994. "On the Likelihood Ratio Test Statistic for the Number of Components in a Normal Mixture with Unequal Variances." *Biometrics* 50:1158-62.

Fujisawa, H., and S. Eguchi. 2006. "Robust estimation in the normal mixture model." *Journal of Statistical Planning and Inference* 136 : 3989 – 4011.

Hennessy, D.A., B.A. Babcock, and D.J. Hayes. 1997. "Budgetary and Producer Welfare Effects of Revenue Insurance." *American Journal of Agricultural Economics* 79: 1024- E034.

Ker, A.P., and B.K. Goodwin. 2000. "Nonparametric Estimation of Crop Insurance Rates Revisited." *American Journal of Agricultural Economics* 83: 463-478.

McLachlan, G. J. 1987. "On Bootstrapping the Likelihood Ratio Test Statistic for the Number of Components in a Normal Mixture." *Journal of the Royal Statistical Society. Series C (Applied Statistics),* 36(3): 318-324.

Miranda, M.J., and J.W. Glauber. 1997. "Systemic Risk, Reinsurance, and the Failure of Crop Insurance Markets." *American Journal of Agricultural Economics* 79:206-215.

Nelson, C.H., and P.M. Preckel. 1989. "The Conditional Beta Distribution As a Stochastic Production Function," *American Journal of Agricultural Economics* 71:370-378.

Robinson, P. 1987. "Asymptotically efficiency estimation in the presence of heteroscedasticity of unknown form." *Econometrica* 55:875-891.

Turvey, C.G., and C. Zhao. 1993. "Parametric and Nonparametric Crop Yield Distributions and their Effects on All-Risk Crop Insurance Premiums." Working Paper, Department of Agricultural Economics and Business, University of Guelph (May).

Vuong, Q. H., 1989. "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses." *Econometrica* 57(2) : 307-333.

Ying Zhu, B.K. Goodwin, and S.K. Ghosh. 2010. "Modeling Yield Risk Under Technological Change: Dynamic Yield Distributions and the U.S. Crop Insurance Program." Dissertation Paper, Department of Agricultural, and Resource Economics, North Carolina State University.

APPENDIX

Table 1

Dependant Variable : Realized Yields				
	Cubic Trend		Quadratic Trend	
Variable	Estimate	Std. Error	Estimate	Std. Error
Intercept	30.42808	6.53611*	25.01456	4.81314*
t	0.32440	0.66200	1.06679	0.26135*
t^2	0.03091	0.01804	0.00920	0.00298*
t^3	-0.00017	0.00013	-	-
R^2	0.9126		0.9110	
N	84		84	

Note : An asterisk(*) denotes statistical significance at the α = 0.05 or small level.

Table 2

Dependant Variable : Square Error				
	Linear Trend		Quadratic Trend	
Variable	Estimate	Std. Error	Estimate	Std. Error
Intercept	59.49537	36.29338	-21.446	54.24905
t	1.76807	0.74593*	7.41598	2.94545*
t^2	-	-	-0.0664	0.03358*
R^2	0.06		0.11	
N	84		84	

Note : An asterisk(*) denotes statistical significance at the α = 0.05 or small level.

Table 3 : Two-Stage Estimation

Parameter	Beta	Mixed Normal	Mixed Normal (at λ =0.1007)	Normal
α	15.3154 (2.3502*)			
β	10.0810 (1.5339*)			
λ		0.1007 (0.2391)	0.1007	
μ_1		126.99 (76.3546)	126.98 (15.5343*)	180.00 (3.1264*)
σ_1		27.4647 (30.8233)	27.4621 (10.9761*)	28.6540 (2.2107*)
μ_2		185.94 (7.5479*)	185.94 (2.7998*)	
σ_2		21.8728 (4.3990*)	21.8730 (2.1133*)	
-2Log Likelihood	799.8	792.5	792.5	802.1
AIC	803.8	802.5	800.5	806.1
AICC	804.0	803.3	801.0	806.2
BIC	808.7	814.7	810.2	810.9

Note: Numbers in parentheses are standard errors. An asterisk(*) denotes statistical significance at the α = 0.05 or small level.

Table 4 : Two-Stage Estimation (Specification Tests)

	Mixed Normal vs. Beta	*Mixed Normal vs. Normal*	Beta vs. Normal
Vuong statistic	1.19	1.45*	2.45*
Adjusted Vuong	-0.97	-0.57	2.45*

Note: An asterisk(*) denotes statistical significance at the α = 0.1 or small level.

Table 5 : Time-Varying Model

Parameter	*Mixed Normal*		*Normal*	
	Estimate	Std. Error	Estimate	Std. Error
a	3.4583	0.04113*	3.4583	0.04109*
b	0.02197	0.000816*	0.02197	0.000816*
c	1.7575	0.1842*	1.7552	0.1840*
d	0.01950	0.003932*	0.01956	0.003931*
e	1.7719	313.96		
f	0.2804	64.3667		
g	3.1307	3.9730		
h	0.2578	0.05985*		
w	8.9636	9.6984		
-2 Log Likelihood	672.9		672.9	
AIC	690.9		680.9	
AICC	693.4		681.4	
BIC	712.8		690.6	
Vuong statistic	-0.47026			
Adjusted Vuong	-491.688*			

Note: An asterisk(*) denotes statistical significance at the α = 0.05 or small level.

Figure 1

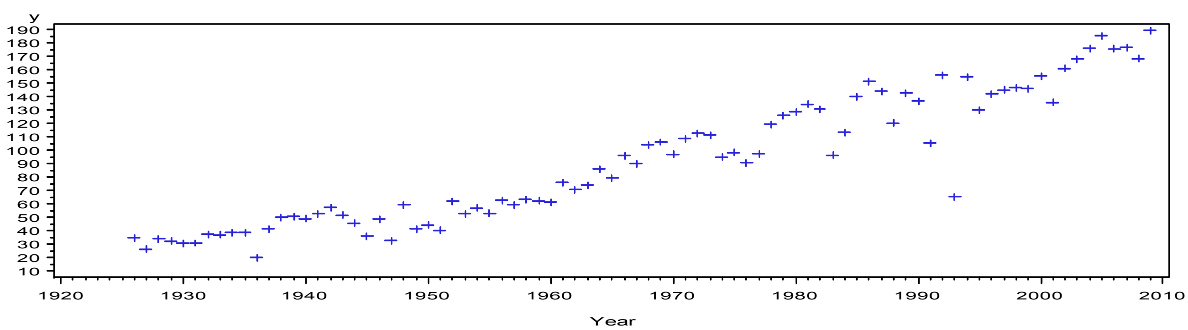

Figure 2

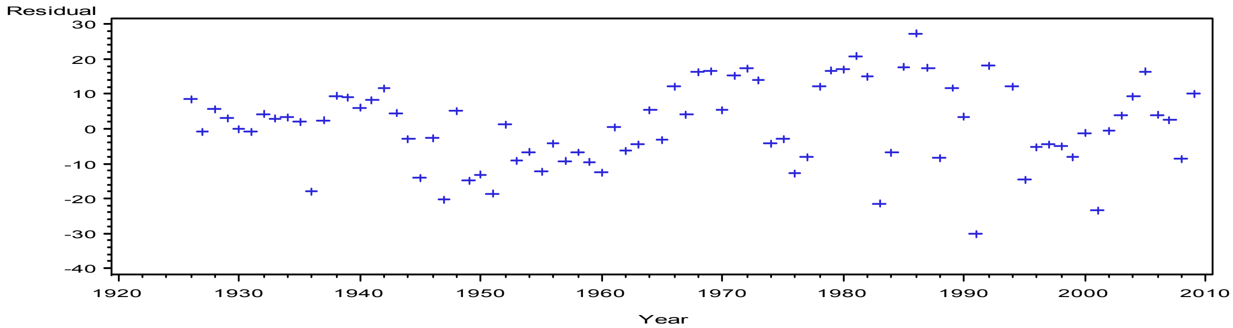

Figure 3

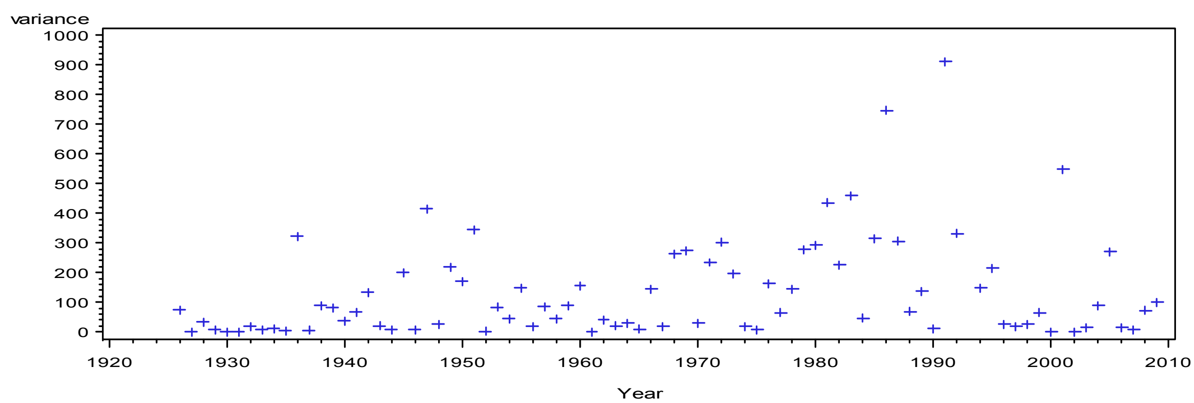

Figure 4

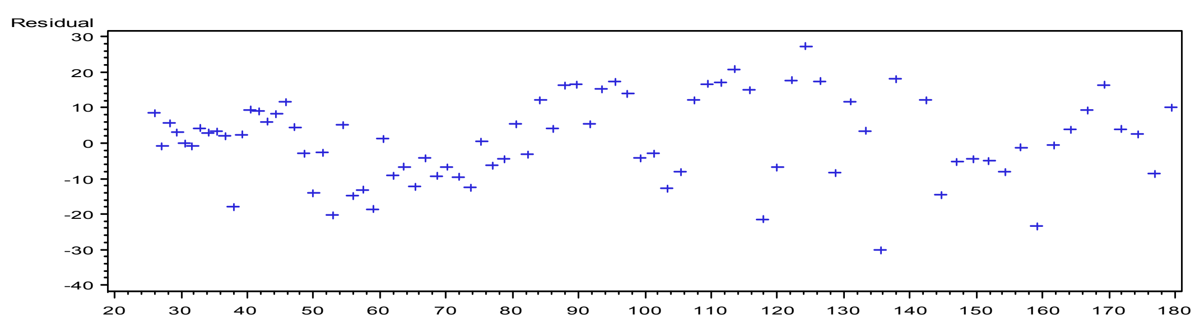

Figure 5-1

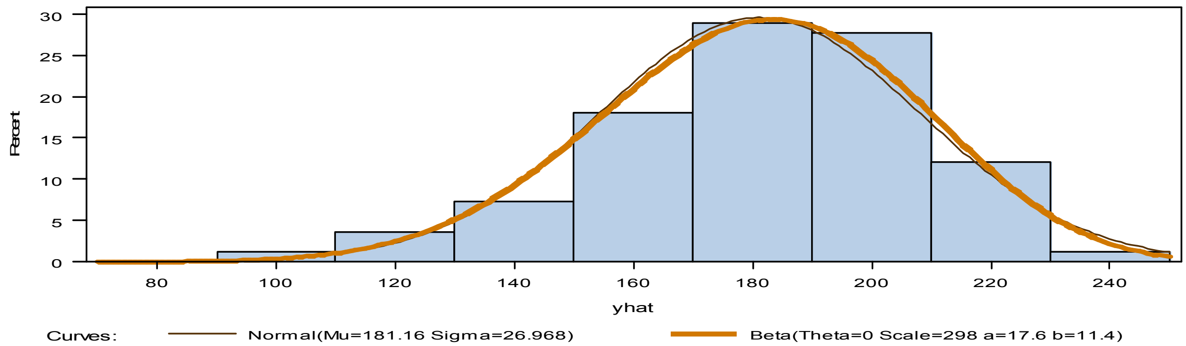

Figure 5-2

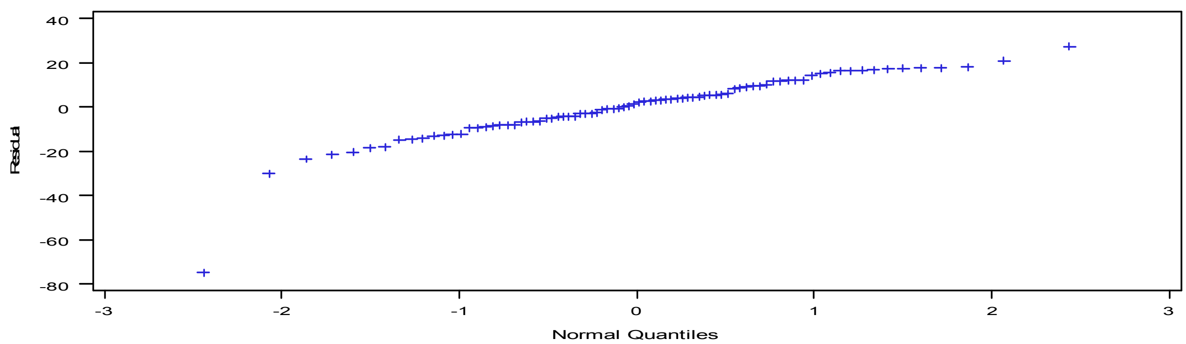

Figure 5-3

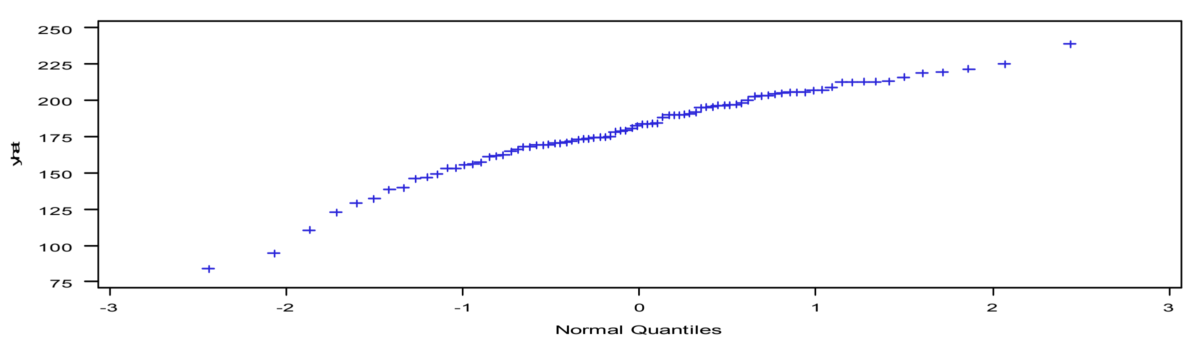

Table 6. Simulation Premium Rates

Two-Stage	
Normal	0.0054545
Beta	0.0059174
Mixed Normal	0.0120156
Time-Varying Normal	
t=85	0.0040772
t=84	0.0041224
t=74	0.0045961
t=64	0.0051088
t=54	0.0056637
t=44	0.0062626
t=34	0.0069079
t=24	0.0076006
t=14	0.0083435
t=4	0.0091383
t=1	0.0093872