# CREDIBILITY FOR REGRESSION MODELS WITH APPLICATION TO TREND (REPRINT)

*Charles A. Hachemeister,*
*with discussion by Al Quirin*

Reprinted from:
CREDIBILITY
Theory and Applications

ACADEMIC PRESS, INC.
New York San Francisco London

# Credibility for Regression Models with Application to Trend

Charles A. Hachemeister
Prudential Reinsurance Company

## Introduction

Inflation has moved from a minor annoyance to a major element in Casualty insurance rate making. Twenty years ago it was sufficient to adjust automobile rate levels without any trend of loss severity or frequency. Presently, this minor annoyance has become a major element in the rate making process. This development has led to the necessity of estimating these trends by state. However, no standards have been specifically developed for evaluating credibility of state trend line versus country wide trend lines.

Standards for developing credibility adjusted state trend lines are developed in this paper. The general approach is a direct extension of the Bühlmann & Straub (1970), "Credibility for Loss Ratios." The results obtained apply to much more general models than simple linear trend. In fact, credibility standards have been developed for arbitrary linear regression models.

## Expected Severity Over Time

To put our thoughts into perspective, let us consider a concrete example of estimating expected severity over time for total private passenger BI total limits severity.[1]

[1] The Automobile Bodily Injury data in this paper has been supplied by the Insurance Services Office.

FIGURE 1

State #1
Private Passenger
Bodily Injury
Total Limits Severities

| Time Period | t | # of Claims $P_{t1}$ | Observed Severity $x_{t1}$ |
|---|---|---|---|
| 7-9/70 | 12 | 7861 | 1738 |
| 10-12/70 | 11 | 9251 | 1642 |
| 1-3/71 | 10 | 8706 | 1794 |
| 4-6/71 | 9 | 8575 | 2051 |
| 7-9/71 | 8 | 7917 | 2079 |
| 10-12/71 | 7 | 8263 | 2234 |
| 1-3/72 | 6 | 9456 | 2032 |
| 4-6/72 | 5 | 8003 | 2035 |
| 7-9/72 | 4 | 7365 | 2115 |
| 10-12/72 | 3 | 7832 | 2262 |
| 1-3/73 | 2 | 7849 | 2267 |
| 4-6/73 | 1 | 9077 | 2517 |

Figure 1 shows Private Passenger Automobile data from a particular state giving a number of claims in each calendar quarter along with the observed severity. Time is denoted by an index, $t$, for which observations are available from time $n$ to time 1. Time runs backwards for reasons of computational ease below. In figure 1, we also introduce notation $P_{ts}$ as the number of claims, and $x_{ts}$ as the observed severity in time period $t$ and state $s$.

It is our objective to estimate the expected value of $x$ over time given $s$:

$$E(x_{ts}) = \mu_{ts}$$

Two competing choices for a model to estimate $\mu_{ts}$ are time series analysis, where the major emphasis lies on the interdependence of the $x_{ij}$ for various $i$ and $j$, and the regression model, where $\mu_{ts}$ is considered a linear combination of other observed variables. These two approaches are not entirely independent since it is possible to create a model which contains both the elements of interdependence of the $x_{ij}$ and also a mean value $\mu_{ts}$ which is dependent upon observed values of other variables. The problem of dealing with such a model is the practical one of producing estimates of the auto-covariance function of the $x_{ij}$ for different $i$ and $j$ at the same time as estimating the regression coefficients. However, the results of the analysis below will follow in large measure for either choice of model.

The Classical Trend and Regression Model

We will make the particular choice to model this expected value as a linear trend:

$$\mu_{st} = a_s + b_s t$$

If we introduce the two column matrices,

$$\beta_s = \begin{pmatrix} a_s \\ b_s \end{pmatrix}; \quad Y_{ts} = \begin{pmatrix} 1 \\ t \end{pmatrix}$$

then we will be able to write the expected value of $x_{ts}$ in matrix form,

$$\mu_{ts} = Y'_{ts}\beta_s$$

Notice that this matrix formulation of $\mu_{ts}$ is not limited to a simple trend, but would apply also for models where

$$\mu_{ts} = \sum_{i=1}^{r} \beta_{si} y_{sti}$$

In this case,

$$\beta_s = \begin{pmatrix} \beta_{s1} \\ \beta_{s2} \\ \vdots \\ \beta_{sr} \end{pmatrix}$$

and the $r$ by $1$ matrix of independent variables is

$$Y_{ts} = \begin{pmatrix} y_{st1} \\ y_{st2} \\ \vdots \\ y_{str} \end{pmatrix} .$$

While we will only discuss the trend model in the numerical example given below, all the theoretical results follow for this more general model.

For development of the classical regression results, it will be necessary to deal with our data in matrix formulation. We will refer to the column matrix of severities for a given state as

$$X_s = \begin{pmatrix} x_{ns} \\ x_{n-1,s} \\ \vdots \\ x_{1s} \end{pmatrix} .$$

For each state we will also refer to the $n$ by $r$ matrix of independent variable observations over time as

$$Y_s = \begin{pmatrix} Y'_{ns} \\ Y'_{n-1,s} \\ \vdots \\ Y'_{1,s} \end{pmatrix} .$$

For our trend model this is a 12 by 2 matrix. The first column of which is all 1's; the second column of which has entries which go from 12 to 1.

With regard to the number of claims, it will be valuable to introduce an $n$ by $n$ square matrix with zeros in the nondiagonal elements and with the number of claims for each time period going down the main diagonal:

$$P_s = \begin{pmatrix} P_{ns} & & & & \bigcirc \\ & P_{n-1,s} & & & \\ & & \cdot & & \\ & & & \cdot & \\ \bigcirc & & & & \cdot \quad P_{1s} \end{pmatrix}$$

We will also find it necessary to refer to the mean value of the process for various time periods for a given state,

$$\mu_s = \begin{pmatrix} \mu_{ns} \\ \mu_{n-1,s} \\ \vdots \\ \mu_{1s} \end{pmatrix}$$

for which

$$\mu_s = Y_s\beta_s$$

now follows.

Time Series Implications

In a time series model one does not usually consider that the mean value $\mu_{ts}$ as dependent upon other variables, $Y_{ts}$. The direction of the investigation in such models is concerned with the $n$ by $n$ autocovariance matrix

$$C_s = E[X_sX_s'] - \mu_s\mu_s'$$

It is not the intention of this paper to pursue the time series direction of analysis. However, the results developed in this paper hold in large measure with an arbitrary autocovariance matrix.

We will follow the Bühlmann, Straub formulation in which the variance of $x_{ts}$ is proportional to the number of claims:

$$E(x_{ts}^2) - \mu_{ts}^2 = \frac{\sigma_s^2}{P_{ts}}$$

and the severity $x_{ts}$ is independent from time period to time period:[2]

$$E(x_{is}x_{js}) - \mu_{is}\mu_{js} = 0 \qquad i \neq j$$

This is, of course, an over simplification of the real world. With these assumptions we find the $n \times n$ auto-covariance matrix in terms of matrix $P_s$, defined above, as

$$C_s = \sigma_s^2 P_s^{-1}$$

<u>Basic Summary Statistics</u>

There will be certain statistics which will arise frequently in our discussion of the trend example. Figure 2 defines the summary statistics that we will need below. Note, of course, that only those statistics which involve $x_{ts}$ are random variables.

FIGURE 2

Basic Summary Statistics

$$P_{.s} = \sum_{t=1}^{n} P_{ts} \qquad P_{..} = \sum_{s=1}^{N} P_{.s}$$

$$\bar{t}_s = \sum_{t=1}^{n} P_{ts}t/P_{.s} \qquad \bar{t} = \sum_{s=1}^{N} P_{.s}\bar{t}_s/P_{..}$$

$$\overline{t^2_s} = \sum_{t=1}^{n} P_{ts}t^2/P_{.s} \qquad \overline{t^2} = \sum_{s=1}^{N} P_{.s}\overline{t^2_s}/P_{..}$$

$$\bar{x}_s = \sum_{t=1}^{n} P_{ts}x_{ts}/P_{.s} \qquad \bar{x} = \sum_{s=1}^{N} P_{.s}\bar{x}_s/P_{..}$$

$$\overline{xt}_s = \sum_{t=1}^{n} P_{ts}tx_{ts}/P_{.s} \qquad \overline{xt} = \sum_{s=1}^{N} P_{.s}\overline{xt}_s/P_{..}$$

---

[2]Note particularly that this last assumption implies that there are no seasonal factors affecting the data.

FIGURE 2 (continued)

$$\sigma^2_{ts} = \overline{t^2_s} - \bar{t}^2_s \qquad \sigma^2_t = \overline{t^2} - \bar{t}^2$$

$$\sigma_{txs} = \overline{xt}_s - \bar{x}_s\bar{t}_s \qquad \sigma_{tx} = \overline{xt} - \bar{x}\,\bar{t}$$

$$\sigma^2_{xs} = \overline{x^2_s} - \bar{x}^2_s \qquad \sigma^2_x = \overline{x^2} - \bar{x}^2$$

State Wide Full Credibility Trend Estimates

Were we to follow the classical generalized least squares estimation procedures for $\beta_s$, we would find in terms of the matrices defined above

$$\hat{\beta}_s = (Y_s' C_s^{-1} Y_s)^{-1} Y_s' C_s^{-1} X_s$$

For our particular trend example these results become:

$$\hat{a}_s = \bar{x}_s - \bar{t}_s \hat{b}_s$$

and

$$\hat{b}_s = \sigma_{txs}/\sigma^2_{ts}$$

Pooled Data

Figure 3 compares the private passenger BI severity experience from state to state. Figure 4 contains the values for the summary statistics needed to calculate the estimates of slopes and intercepts contained on Figure 3. For our purposes we will consider that these five states make up the entire country. However, the analysis can be generalized to any number of states. Accordingly, we will refer below to N states. The right-hand two columns of this figure show the pooled data being the sum of the data elements from the five states for comparable time periods.

## FIGURE 3

**Private Passenger**
**Bodily Injury**
**Total Limits Severities**
**by State**

| | | 1 | | 2 | | 3 | | 4 | | 5 | | "Countrywide" | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Time Period | $t$ | # of Claims $P_{t1}$ | Severity $x_{t1}$ | # of Claims $P_{t2}$ | Severity $x_{t2}$ | # of Claims $P_{t3}$ | Severity $x_{t3}$ | # of Claims $P_{t4}$ | Severity $x_{t4}$ | # of Claims $P_{ts}$ | Severity $x_{ts}$ | # of Claims $P_{t.}$ | Severity $x_{t.}$ |
| 7-9/70 | 12 | 7861 | 1738 | 1622 | 1364 | 1147 | 1759 | 407 | 1223 | 2902 | 1456 | 13939 | 1623 |
| 10-12/70 | 11 | 9251 | 1642 | 1742 | 1408 | 1357 | 1685 | 396 | 1146 | 3172 | 1499 | 15918 | 1579 |
| 1-3/71 | 10 | 8706 | 1794 | 1523 | 1597 | 1329 | 1479 | 348 | 1010 | 3046 | 1609 | 14952 | 1690 |
| 4-6/71 | 9 | 8575 | 2051 | 1515 | 1444 | 1204 | 1763 | 341 | 1257 | 3068 | 1741 | 14703 | 1882 |
| 7-9/71 | 8 | 7917 | 2079 | 1622 | 1342 | 998 | 1674 | 315 | 1426 | 2693 | 1482 | 13545 | 1827 |
| 10-12/71 | 7 | 8263 | 2234 | 1602 | 1675 | 1077 | 2103 | 328 | 1532 | 2910 | 1572 | 14180 | 2009 |
| 1-3/72 | 6 | 9456 | 2032 | 1964 | 1470 | 1277 | 1502 | 352 | 1953 | 3275 | 1606 | 16324 | 1836 |
| 4-6/72 | 5 | 8003 | 2035 | 1515 | 1448 | 1218 | 1622 | 331 | 1123 | 2697 | 1735 | 13764 | 1853 |
| 7-9/72 | 4 | 7365 | 2115 | 1527 | 1464 | 896 | 1828 | 287 | 1343 | 2663 | 1607 | 12738 | 1893 |
| 10-12/72 | 3 | 7832 | 2262 | 1748 | 1831 | 1003 | 2155 | 384 | 1243 | 3017 | 1573 | 13984 | 2024 |
| 1-3/73 | 2 | 7849 | 2267 | 1654 | 1612 | 1108 | 2233 | 321 | 1762 | 3242 | 1613 | 14174 | 2027 |
| 4-6/73 | 1 | 9077 | 2517 | 1861 | 1471 | 1121 | 2059 | 342 | 1306 | 3425 | 1690 | 15826 | 2157 |
| Intercept | $\hat{a}_s$ | | 2470 | | 1621 | | 2096 | | 1538 | | 1676 | | 2148 |
| Slope | $\hat{b}_s$ | | - 62.39 | | - 17.14 | | - 43.32 | | - 27.81 | | - 11.87 | | - 43.35 |

FIGURE 4

Values of Summary
Statistics by State

| State: | 1 | 2 | 3 | 4 | 5 | "Countrywide" |
|---|---|---|---|---|---|---|
| $P_{.s}$ | 100,155 | 19,895 | 13,735 | 4,152 | 36,110 | 174,047 |
| $\bar{t}_s$ | 6.54972 | 6.41171 | 6.69982 | 6.66089 | 6.43725 | 6.52511 |
| $\overline{t^2}_s$ | 54.88889 | 53.22398 | 56.91824 | 56.79143 | 53.75876 | 54.66964 |
| $\bar{x}_s$ | 2,060.92 | 1,511.22 | 1,805.84 | 1,352.98 | 1,599.83 | 1,865.40 |
| $\overline{xt}_s$ | 12,750.36 | 9,481.90 | 11,577.80 | 8,666.54 | 10,152.19 | 11,647.75 |
| $\sigma^2_{ts}$ | 11.99009 | 12.11393 | 12.03068 | 12.42402 | 12.32061 | 12.09264 |
| $\sigma_{txs}$ | -748.09102 | -207.62975 | -521.01641 | -345.04749 | -146.30085 | -524.21257 |
| $\sigma^2_{xs}$ | 55,881. | 18,725. | 60,776. | 68,275. | 7,573. | 99807 |

Just as we have a need to be able to refer to all the data within a state in a concise fashion, we will have a need to refer to all of the data country wide in a concise fashion. To this end for severities we define the $n \times N$ by 1 column of severities as

$$X = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{pmatrix},$$

the $n \times N$ by $r$ matrix of independent variable observations as

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_N \end{pmatrix},$$

and the super matrix of numbers of claims matrices as the $n \times N$ matrix

$$P = \begin{pmatrix} P_1 & & & \bigcirc \\ & P_2 & & \\ & & \ddots & \\ \bigcirc & & & P_N \end{pmatrix}.$$

Also, we will consider the $n \times N$ by 1 column matrix of mean values:

$$E(X) = \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_N \end{pmatrix}$$

It will also be necessary for us to use the autocovariance matrix of all of the severities country wide:

$$E[XX'] - \mu\mu' = C = \begin{pmatrix} C_1 & & & \bigcirc \\ & C_2 & & \\ & & \ddots & \\ \bigcirc & & & C_N \end{pmatrix}$$

It is important to note that since this "super" autocovariance matrix is made up of the state autocovariance matrices down the super diagonal with zero elements elsewhere, this model specifically considers that the observations from one state are independent of those from another state.

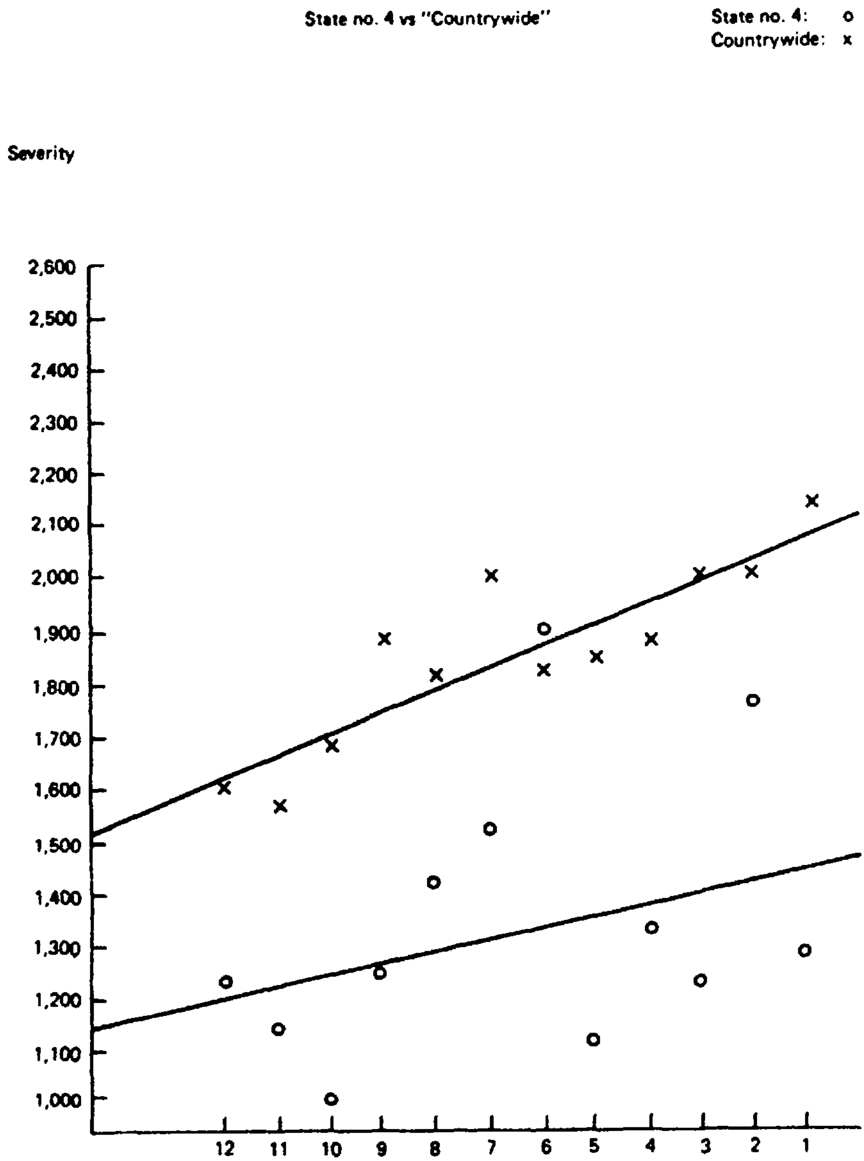
In terms of these super matrices, the pooled "country wide" estimates of $\beta$ become:

$$\hat{\beta} = (Y'C^{-1}Y)^{-1}Y'C^{-1}X$$

State Versus "Countrywide" Trend

The estimates of the intercept and slope of the trend line shown on figure 3 vary substantially from state to state. Without credibility the only two alternatives available to the decision maker is whether to consider the data from the other states to be from the same basic population as the state in question, and therefore use the country wide estimate; or to consider that the state data was sufficiently different, and therefore throw out the data from other states using only the state estimate. Figure 5 compares the country wide severity data with that of state #4. Notice that the country wide data lies more closely about the least squares trend line, although the country wide line lies substantially above the state line. One is not exactly happy with the trend line estimate for the state because of the very wide variation of the data points about that line. In this instance, one might be more ready to accept the country wide versus the state trend.

Figure 5


State no. 4 vs "Countrywide"
State no. 4: o
Countrywide: x
Severity
2,600
2,500
2,400
2,300
2,200
2,100
2,000
1,900
1,800
1,700
1,600
1,500
1,400
1,300
1,200
1,100
1,000
12
11
10
9
8
7
6
5
4
3
2
1

However, state versus country wide are not the only two choices. If one were to believe that the distribution of $x_{ts}$ varied from state to state and had to choose an optimal decision over all of the states, a compound decision problem, then it is not clear whether the choice should be a state wide or a country wide trend. The exact solution of this problem, produces a credibility weighting between the two trends, as will be seen below.

Alternatively, if one is only making a single decision for one state but if it is believed that the distribution of $x$ is a random pick from some set of distributions governed by an index, say $\theta_s$, then the result is the same as the compound decision.

Figure 6 contains the estimated trend lines for each of our five states and the heavier line as that for country wide. It is clear from looking at this figure that the slopes and intercepts vary from state to state. In the compound problem of trying to choose a set of trend lines for all of the states to optimize the total trend choice, one should act as if the slopes and intercepts do have a distribution which is reflected in these differences.

With the introduction of an index $\theta_s$ to describe these distributions, we need to reformulate the state data in terms of this index. First of all, the $\beta_s$ become functions of $\theta_s$

$$\beta_s = B(\theta_s)$$

as does the expected value of $x_{ts}$ given $\theta_s$

$$E[x_{ts}|\theta_s] = \mu_{ts}(\theta_s) = Y'_{ts}B(\theta_s)$$

The autocovariance matrix is in general a matrix function of $\theta_s$

$$C_s = C_s(\theta_s)$$

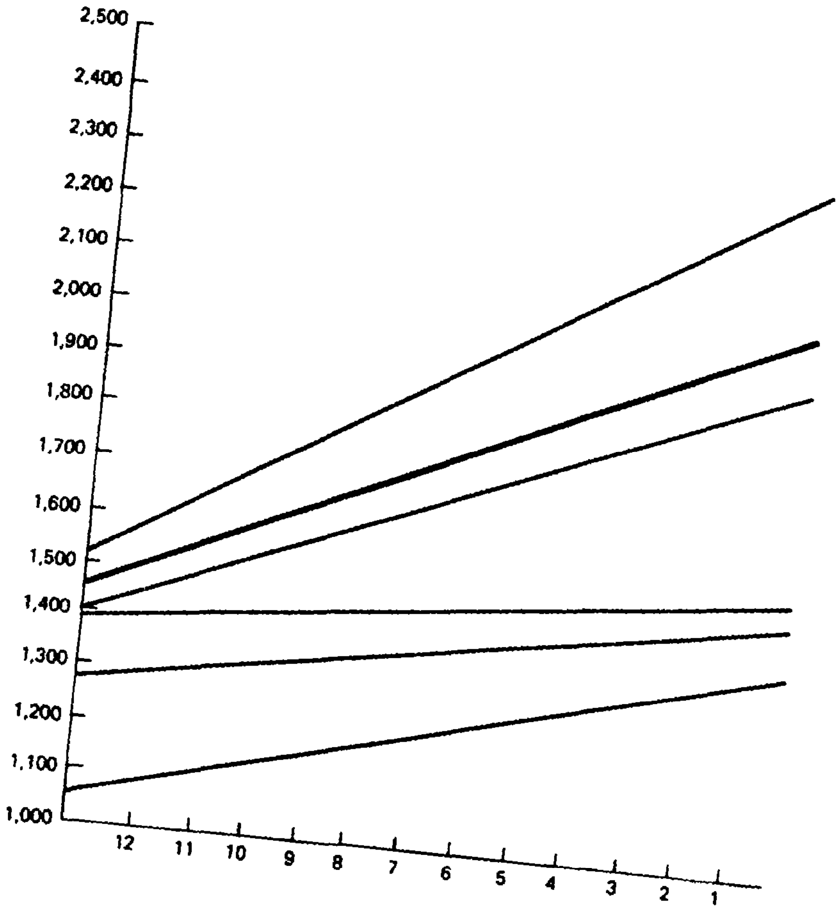



**Figure 6**

Comparison of Observed State Trends

In this paper we will pursue the case of where the autocovariance matrix is known up to a scalar multiplier, the variance of $x_{ts}$ which is a function of $\theta_s$:

$$C_s(\theta_s) = \sigma^2(\theta_s)P_s^{-1}$$

<u>Expected Values Over $\theta$</u>

It will be necessary below to take expected values of various functions of $\theta$.

$B(\theta)$:

The expected value of the column matrix $B$ is equal to a column matrix $\beta$ without subscripts

$$E[B(\theta)] = \beta \text{ .}$$

The covariance matrix of the $B(\theta)$ will be denoted by the $r$ by $r$ matrix:

$$E[B(\theta_s)B'(\theta_s)] - \beta\beta' = \Gamma_{rxr} \text{ .}$$

$\mu$:

The expected value of $\mu_{ts}$ is now:

$$E[\mu_{ts}(\theta_s)] = Y'_{ts}\beta$$

With a natural extension to the column matrix $\mu_s$ within a state and then country wide to $\mu$ as:

$$E[\mu_s(\theta_s)] = Y_s\beta \quad \text{and} \quad E[\mu(\theta_1,\ldots,\theta_N)] = Y\beta$$

We will also find it necessary to refer below to the column matrix of autocovariances between a particular mean value and that of all other mean values:

$$E[\mu\mu_{tk}(\theta_k)] - Y_h\beta\beta' Y_{tk} = \begin{pmatrix} Y_1\Gamma Y_{tk}\delta_{1k} \\ Y_2\Gamma Y_{tk}\delta_{2k} \\ \vdots \\ Y_N\Gamma Y_{tk}\delta_{Nk} \end{pmatrix}$$

where $\delta_{ij}$ is the Kronecker delta:

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

The autocovariance matrix of the mean values is a super matrix of $n \times n$ matrices down the super diagonal with zero elements elsewhere:

$$E[\mu\mu'] - E[\mu]E[\mu'] = \begin{pmatrix} Y_1\Gamma Y_1' & & & \bigcirc \\ & Y_2\Gamma Y_2' & & \\ & & \ddots & \\ \bigcirc & & & Y_N\Gamma Y_N' \end{pmatrix}$$

$\sigma_s^2(\theta_s)$:

The state variance is also a variable now, which depends upon $\theta_s$. The expected value of the autocovariance matrix for a given state is denoted by:

$$E[C_s(\theta_s)] = V_s$$

However, in our case we will take:

$$V_s = \sigma^2 P_s^{-1}$$

The extension of this to the country wide autocovariance matrix is:

$$E[C] = V = \begin{pmatrix} V_1 & & & \bigcirc \\ & V_2 & & \\ & & \ddots & \\ \bigcirc & & & V_N \end{pmatrix} = \sigma^2 P^{-1}$$

<u>Estimation of</u> $\mu_{ij}(\theta_j)$

With this preliminary background, it is now possible to consider estimates of the mean value of the trend line at any point of time. We take the usual conditions of unbiasedness and minimum variance:

$$E\hat{\mu}_{ij} = E\mu_{ij}(\theta_j)$$
$$E\{(\hat{\mu}^*_{ij} - \mu_{ij}(\theta))^2\} \leq E\{\hat{\mu}_{ij} - \mu_{ij}(\theta))^2\} \qquad (1)$$

where we will accept the estimator $\mu^*_{ij}$ as the optimal estimator, if (1) holds for all possible estimators $\hat{\mu}_{ij}$.

Following Bühlmann and Straub, we will consider estimators of the form:

$$\hat{\mu}_{ij} = \alpha_0 + \sum_{s=1}^{N} \sum_{t=1}^{n} \alpha_{ts} x_{ts} = \alpha_0 + X' A$$

Where we introduce the column vector of coefficients for state and country wide as

$$A_s = \begin{pmatrix} \alpha_{1s} \\ \alpha_{2s} \\ \vdots \\ \alpha_{ns} \end{pmatrix} \quad \text{and} \quad A = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_N \end{pmatrix}$$

While we require our estimator to be unbiased, this will happen automatically because of the inclusion of the additive constant of $\alpha_0$ in the estimator. Accordingly, to determine our estimator we will minimize:

$$\Phi_{ij} = E[(\alpha_0 + X' A - \mu_{ij}(\theta_j))^2]$$

To do this, we take the partial derivative of $\Phi_{ij}$ with respect to $\alpha_0$ set to 0

$$\frac{\partial \Phi_{ij}}{\alpha_0} = 2E[\alpha_0 + X' A - \mu_{ij}(\theta_j)] = 0$$

to find:

$$\alpha_0 = E[\mu_{ij}(\theta_j)] - E[\mu' ]A = \beta' [Y_{ij} - Y' A]$$

The column vector of partial derivatives of $\Phi_{ij}$ with respect to $A$ is set equal to 0,

$$\frac{\partial \Phi_{ij}}{\partial A} = 2E[XX' A + X(\alpha_0 - \mu_{ij}(\theta_j))] = 0$$

finding:

$$E[(C + \mu\mu')A + \mu\alpha_0] = E[\mu\mu_{ij}(\theta_j)]$$

after taking conditional expectations holding the $\theta_s$ for $s = 1$ to $N$, constant and rearranging terms. Carrying out the expectation over the $\theta_s$, we find:

$$[V + E(\mu\mu') - E(\mu)E(\mu')]A = E[\mu\mu_{ij}(\theta_j)] - E[\mu]E[\mu_{ij}(\theta_j)]$$

To this point the analysis has been quite general without depending upon the form of $V$ or of the form of the autocovariance matrix of the $\mu$. To proceed it is necessary for us to assume $V$ and the autocovariance matrix of $\mu$ to be comprised of $n$ by $n$ matrices of state data down the super diagonal with zeros elsewhere. If this is the case for each state, we may now write:

$$(V_s + Y_s\Gamma Y_s')A_s = Y_s\Gamma Y_{ij}\delta_{sj} \qquad (2)$$

which immediately indicates that

$$A_s = 0 \quad \text{for} \quad s \neq j$$

If we premultiply (2) for state $j$ by $Y_j'V_j^{-1}$, we find:

$$(I + Y_j'V_j^{-1}Y_j\Gamma)Y_j'A_j = Y_j'V_j^{-1}Y_j\Gamma Y_{ij}$$

Anticipating later results, let us pause for a moment to define:

$$K_j = P_{.j}(Y_j'V_j^{-1}Y_j\Gamma)^{-1}$$

and the credibility matrix:[3]

$$Z_j = P_{.j}(P_{.j}I + K_j)^{-1}$$

[3] The $K_j$ matrix only exists if $\Gamma$ is positive definite. However, the $Z_j$ matrix always exists even when $K_j$ does not; and may be written in the form:

$$Z_j = Y_j'V_j^{-1}Y_j\Gamma(I + Y_j'V_j^{-1}Y_j\Gamma)^{-1}$$

This immediately yields:

$$Y_j' A_j = Z_j Y_{1j}$$

Combining this with (2), we now find:

$$A_j = V_j^{-1} Y_j \Gamma [I - Z_j] Y_{1j}$$

Premultiplying this by $X_j'$ and rearranging terms, since

$$Y_j' V_j^{-1} Y_j \Gamma [I - Z_j] = Z_j$$

we find:

$$X_j' A_j = \hat{\beta}_j' Z_j Y_{1j}$$

for the case where $C_j$ is known up to a scalar multiplier[4] which depends upon $\theta_j$. Recall that in the case of greatest interest to us $C_j = \sigma^2(\theta_j) P_j^{-1}$. Now since

$$\alpha_0 = \beta' [I - Z_j] Y_{1j}$$

we can finally write our estimator as:

$$\hat{\mu}_{1j} = [\hat{\beta}_j' Z_j + \beta' (I - Z_j)] Y_{1j}$$

It is particularly interesting and satisfying to note that this estimator holds for any $Y_{1j}$. In other words, we have credibility adjusted the regression coefficients.

<u>Relation to the Bühlmann, Straub Model</u>

The form of the estimator in the Bühlmann, Straub model was:

$$\hat{\mu}_{1j} = X' A$$

---

[4] If $C_j$ is some more complex function of $\theta_j$, $\hat{\beta}_j$ becomes a function of $\theta_j$ such that in general

$$E\hat{\beta}_j \neq (Y_j' V_j^{-1} Y_j)^{-1} Y_j' V_j^{-1} X_j$$

without an additive constant. If this model were followed through for the regression case, one would find:

$$\hat{\mu}_{ij} = [\beta_j' Z_j + d\beta'(I - Z_j)]Y_{ij}$$

which is the same as the estimator above, except for $d$, which is equal to the expression:

$$d = \sum_{s=1}^{N} \hat{\beta}_s' Z_s \Gamma^{-1}\beta / \sum_{s=1}^{N} \beta' Z_s \Gamma^{-1}\beta \ .$$

In the univariate case of Bühlmann and Straub the parameter equivalent of $\beta$ cancelled entirely from the estimator. However, in the multivariate case, this is not so; so that there is no benefit to using the estimator without the additive constant.

Parameter Estimation

To apply our credibility model to real data, we need to be in a position to estimate the various elements which are not directly observable within it. Up to this point we have been able to be very general in the form of the autocovariance matrix within a given state. At this point, we sacrifice this generality to be able to produce unbiased estimators of the parameters in question. The easiest parameter to deal with is the column matrix $\beta$. The least squares estimate of $\beta$, using pooled data, is unbiased:

$$E(\hat{\beta}) = E[(Y'PY)^{-1}Y'PX] = \beta$$

For an estimator of expected value of the state variance $\sigma^2$, let us consider the mean square error for a given state:

$$\hat{\sigma}_s^2 = \frac{1}{n - r} \sum_{t=1}^{n} P_{ts}(x_{ts} - \hat{\mu}_{ts})^2 \ .$$

In matrix terms this becomes:

$$\hat{\sigma}_s^2 = \frac{1}{n - r} (X_s' P_s X_s - X_s' P_s Y_s (Y_s' P_s Y_s)^{-1} Y_s' P_s X_s)$$

Following the classical evaluation of the expected value of the mean square error as outlined in Goldberger,[5] we note that the above matrix is a 1 by 1 matrix and further that the trace of any two matrices is independent of the order of multiplication:

$$tr(AB) = tr(BA)$$

so that we may evaluate the expected value of $\hat{\sigma}_s^2$ as:

$$(n - r)E(\hat{\sigma}_s^2) = E\ tr[P_s(I - Y_s(Y_s' P_s Y_s)^{-1} Y_s' P_s)X_s X_s']$$

since

$$I - Y_s(Y_s' P_s Y_s)^{-1} Y_s' P_s \quad \text{annihilates} \quad Y_s B(\theta_s) B'(\theta_s) Y_s'$$

this becomes:

$$(n - r)E(\hat{\sigma}_s^2) = tr[P_s(I - Y_s(Y_s' P_s Y_s)^{-1} Y_s' P_s)V_s]$$

or

$$E(\hat{\sigma}_s^2) = \frac{1}{n - r}\{trI_{nxn} - tr\ I_{rxr}\}\sigma^2 ,$$

so that $\hat{\sigma}_s^2$ is an unbiased estimator of $\sigma^2$. We shall take the unweighted average of these state mean square errors as our overall estimator of $\sigma^2$:

$$\hat{\sigma}^2 = \frac{1}{N} \sum_{s=1}^{N} \hat{\sigma}_s^2$$

which is clearly unbiased.

The estimator of the covariance matrix of the $B(\theta)$ is somewhat more difficult to find an estimator for. First of all, consider:

$$G = \sum_{s=1}^{N} (Y' PY)^{-1}(Y_s' P_s Y_s)(\hat{\beta}_s - \hat{\beta})(\hat{\beta}_s - \hat{\beta})' .$$

To evaluate the expected value of $G$, let us first consider expected values of matrices of estimators of the $\hat{\beta}_s$. In

[5]"Econometric Theory"; John Wiley & Sons, Inc. - Page 166

particular, we note:

$$\hat{\beta}_j\hat{\beta}_s' = (Y_jP_jY_j)^{-1}Y_j'P_jX_jX'P_sY_s(Y_s'P_sY_s)^{-1} ,$$

so that:

$$E(\hat{\beta}_j\hat{\beta}_s') = \beta\beta' + [\Gamma + \sigma^2(Y_s'P_sY_s)^{-1}]\delta_{js} .$$

At this point we now wish to consider the expected value of $\hat{\hat{\beta}}\hat{\beta}_s'$. To evaluate this expected value, we will assume:

$$\hat{\hat{\beta}}\hat{\beta}_s' = \sum_{j=1}^{N} (Y'PY)^{-1}(Y_j'P_jY_j)\hat{\beta}_j\hat{\beta}_s'$$

Using this relationship, we find:

$$E(\hat{\hat{\beta}}\hat{\beta}_s') = \beta\beta' + (Y'PY)^{-1}(Y_s'P_sY_s)\Gamma + (Y'PY)^{-1}\sigma^2$$

Using a similar analysis for $\hat{\hat{\beta}}\hat{\hat{\beta}}'$ yields:

$$\hat{\hat{\beta}}\hat{\hat{\beta}}' = \sum_{j=1}^{N} \hat{\hat{\beta}}\hat{\beta}_j'(Y_j'P_jY_j)(Y'PY)^{-1} \quad \text{and}$$

$$E(\hat{\hat{\beta}}\hat{\hat{\beta}}') = \beta\beta' + \sum_{j=1}^{N} (Y'PY)^{-1}(Y_j'P_jY_j)\Gamma(Y_j'P_jY_j)(Y'PY)^{-1} + (Y'PY)^{-1}\sigma^2$$

Combining our results we find:

$$E(G) = \left[I - \sum_{s=1}^{N} (Y'PY)^{-1}(Y_s'P_sY_s)(Y'PY)^{-1}(Y_s'P_sY_s)\right]\Gamma + (N-1)(Y'PY)^{-1}\sigma^2$$

If we introduce the $r$ by $r$ matrix

$$\Pi = I - \sum_{s=1}^{N} (Y'PY)^{-1}(Y_s'P_sY_s)(Y'PY)^{-1}(Y_s'P_sY_s) ,$$

an unbiased estimator for $\Gamma$ is

$$H = \Pi^{-1}(G - (N-1)(Y'PY)^{-1}\hat{\sigma}^2) .$$

However, since $\Gamma$ is symmetric we will take our estimator as

$$\hat{\Gamma} = \tfrac{1}{2}(H + H')$$

Form of the Estimators for the Trend Example

To put the above theoretical results into perspective, let us translate them into the trend example. The 2 by 2 matrix of weighted independent variables becomes:

$$Y_s' P_s Y_s = P_{.s}\begin{pmatrix} 1 & \bar{t}_s \\ \bar{t}_s & \overline{t^2_s} \end{pmatrix}$$

The slope and intercept are:

$$\hat{\beta}_s = \begin{pmatrix} \hat{a}_s \\ \hat{b}_s \end{pmatrix} = \begin{pmatrix} \bar{x}_s - \bar{t}_s\, \sigma_{txs}/\sigma^2_{ts} \\ \sigma_{txs}/\sigma^2_{ts} \end{pmatrix}$$

The estimate of average variance is:

$$\hat{\sigma}^2 = \frac{1}{N(n-2)} \sum_{s=1}^{N} P_{.s}(\sigma^2_{xs} - \sigma^2_{txs}/\sigma^2_{ts})$$

The elements of $\hat{\Gamma}$ are denoted as:

$$\hat{\Gamma} = \begin{pmatrix} \hat{\sigma}^2_a & \hat{\sigma}_{ab} \\ \hat{\sigma}_{ab} & \hat{\sigma}^2_b \end{pmatrix}$$

The K matrix within the credibility form then becomes:

$$\hat{K}_s = \begin{pmatrix} \hat{k}_{s11} & \hat{k}_{s12} \\ \hat{k}_{s21} & \hat{k}_{s22} \end{pmatrix}$$

$$= \frac{\hat{\sigma}^2}{\sigma^2_{ts}(\hat{\sigma}^2_a\hat{\sigma}^2_b - \hat{\sigma}^2_{ab})}\begin{pmatrix} \hat{\sigma}^2_b\overline{t^2_s} + \hat{\sigma}_{ab}\bar{t} & -\hat{\sigma}^2_b\bar{t} - \hat{\sigma}_{ab} \\ \hat{\sigma}_{ab}\overline{t^2} - \hat{\sigma}^2_a\bar{t} & \hat{\sigma}_{ab}\bar{t} + \hat{\sigma}^2_a \end{pmatrix}$$

Thus the credibility formula becomes:

$$Z_s = \frac{P_{.s}}{P^2_{.s} + (\hat{k}_{s11} + \hat{k}_{s22})P_{.s} + \hat{k}_{s11}\hat{k}_{s22} - \hat{k}_{s12}\hat{k}_{s21}} \times$$

$$\times \begin{pmatrix} P_{.s} + \hat{k}_{s22} & -\hat{k}_{s12} \\ -\hat{k}_{s21} & P_{.s} + \hat{k}_{s11} \end{pmatrix}$$

Using the data shown in figure 4 these estimators take on the values as shown in figure 7.

Figure 7

Numerical Value of the estimates

$$\Pi = \begin{pmatrix} .61017 & -.00468 \\ -.00066 & .60537 \end{pmatrix}$$

$$\hat{\Gamma} = \begin{pmatrix} 241,550 & -13,819 \\ -13,819 & 805 \end{pmatrix}$$

$$K_1 = \begin{pmatrix} -49,179 & 9,073 \\ -874,219 & 160,327 \end{pmatrix}$$

$$K_2 = \begin{pmatrix} -48,080 & 9,097 \\ -854,430 & 160,691 \end{pmatrix}$$

$$K_3 = \begin{pmatrix} -49,479 & 8,914 \\ -879,957 & 157,592 \end{pmatrix}$$

$$K_4 = \begin{pmatrix} -47,466 & 8,664 \\ -844,260 & 153,154 \end{pmatrix}$$

$$K_5 = \begin{pmatrix} -47,194 & 8,923 \\ -838,835 & 157,632 \end{pmatrix}$$

$$G - (N-1)(Y'PY)^{-1}\hat{\sigma}^2 =$$

$$= \begin{pmatrix} 147,451 & -8,415.88 \\ -8,544.26 & 496.3438 \end{pmatrix}$$

$$\hat{\sigma}^2 = 44,057,744$$

$$Z_1 = \begin{pmatrix} 1.2489 & -.0435 \\ 4.0219 & .2444 \end{pmatrix}$$

$$Z_2 = \begin{pmatrix} 1.3871 & -.0699 \\ 6.4852 & -.2165 \end{pmatrix}$$

$$Z_3 = \begin{pmatrix} 1.3680 & -.0712 \\ 7.0261 & -.2854 \end{pmatrix}$$

$$Z_4 = \begin{pmatrix} 1.1083 & -.0610 \\ 6.0202 & -.3052 \end{pmatrix}$$

$$Z_5 = \begin{pmatrix} 1.2376 & -.0570 \\ 5.5842 & -.0708 \end{pmatrix}$$

Using these numerical values, we find the credibility adjusted slopes and intercepts. These are compared with the state and country wide slopes and intercepts on figure 8.

FIGURE 8

| State | | | State Data | Credibility Adjusted Data | Countrywide Data |
|---|---|---|---|---|---|
| 1 | Intercept: | a | 2470 | 2473 | 2148 |
| | Slope: | b | -62.39 | -61.98 | -43.35 |
| 2 | Intercept: | a | 1621 | 1587 | 2148 |
| | Slope: | b | -17.14 | -12.19 | -43.35 |
| 3 | Intercept: | a | 2096 | 2077 | 2148 |
| | Slope: | b | -43.31 | -39.64 | -43.35 |
| 4 | Intercept: | a | 1538 | 1566 | 2148 |
| | Slope: | b | -27.81 | -10.85 | -43.35 |
| 5 | Intercept: | a | 1676 | 1740 | 2148 |
| | Slope: | b | -11.87 | -18.68 | -43.35 |

Figure 9 compares the state trend line denoted by S and the country wide trend line denoted by C with the credibility adjusted trend line denoted by A. In all of the states, except state # 4, the credibility adjusted trend line is virtually the same as the state trend line. However, in state #4, with a smaller claim volume, the credibility adjusted trend line is such different from the state trend line. State #4 trend lines clearly point out a distressing aspect of the credibility adjusted trend line. The credibility adjusted trend line has a lower trend than both the country wide and state trend lines. In fact, a closer examination of the other state trend line graphs will show that the credibility adjusted trend for state #2 is also lower than both state and country wide. In state #1 the credibility adjusted slope is less than for the state but the credibility adjusted trend line lies above both the state and country wide lines for the

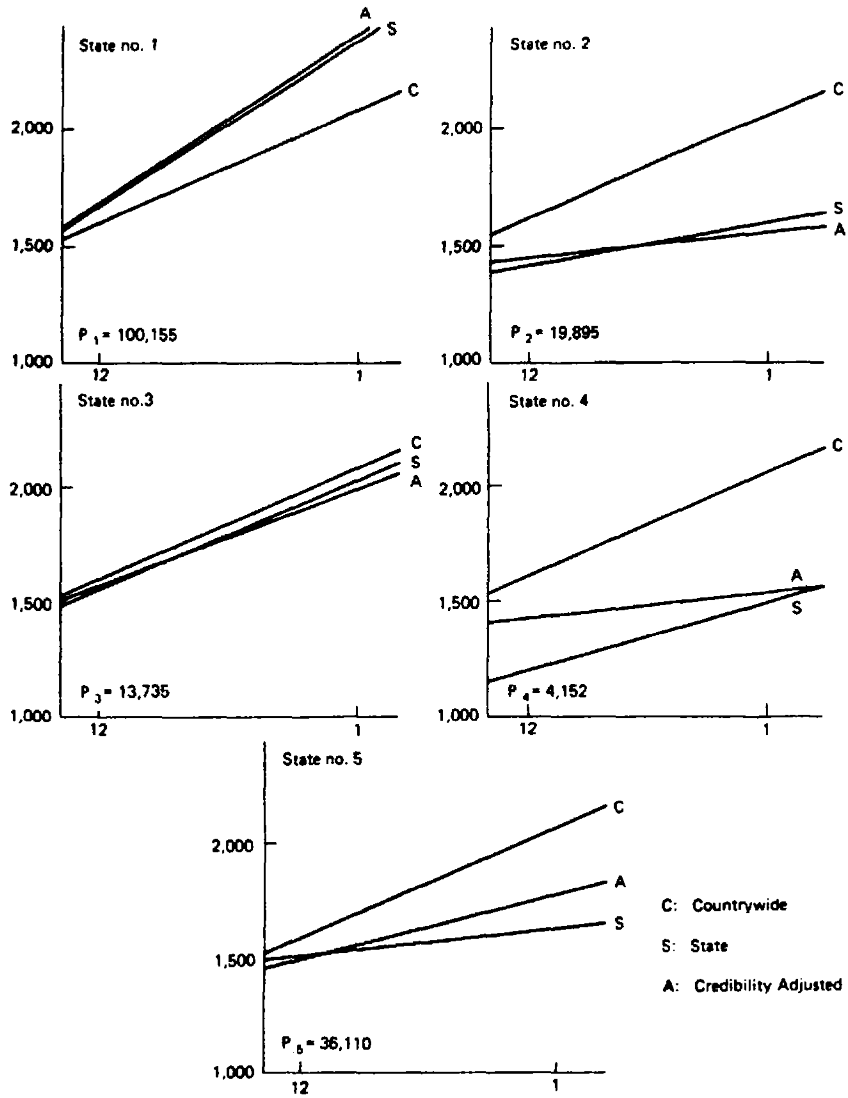



**Figure 9**
Comparison of
Credibility Adjusted Trend Lines
with State and Countrywide Lines

time period from our observed values were taken.

These strange results arise from our choice of model. That is, we have assumed that not only can the trend for a given state be considered as being a pick from a distribution of trends, but also that the level of severity for a random pick over some distribution of average severity levels. However, if we were to reflect upon what a proper model for trend would be, it is fairly easy to conclude that the level of severity as embodied by the intercept, $a_s$ in the trend line, is distinctly different from state to state and should not be credibility adjusted for.

It is possible to alleviate this defect by changing the basic credibility model. In order to more adequately discuss this, it is necessary for us to first discuss the effect of linear transformations of the independent variables on our credibility estimate, $\hat{\mu}_{ij}$.

<u>Invariance of $\hat{\mu}_{ij}$ Under Transformations of the Independent Variables</u>

The column matrix $Y_{ts}$ describes the values of $r$ variables which are observed at time $t$. Such that

$$\mu_{ts} = Y'_{ts}\beta_s$$

This mean value could just as well be described by a linear combination of transformed variables $Y^*_{ts}$

$$\mu_{ts} = Y^{*\prime}_{ts}\beta^*_s$$

The easiest example of this is simple scaling and translation of each of the independent variables. In our case we would define time about an origin and with a scale such that the weighted average of the scaled times was zero and the sample variance of the scaled times was equal to one. This transformation would be accomplished by a matrix:

$$T_s = \begin{pmatrix} 1 & 0 \\ -\bar{t}_s/\sigma_{ts} & 1/\sigma_{ts} \end{pmatrix}$$

This matrix can be considered a mapping of $Y_{ts}$ to $Y^*_{ts}$:

$$Y_{ts} = \begin{pmatrix} 1 \\ t \end{pmatrix} \xrightarrow[T_s]{} Y^*_{ts} = \begin{pmatrix} 1 \\ \dfrac{t - \bar{t}_s}{\sigma_{ts}} \end{pmatrix}$$

However, it is not necessary to merely consider simple locations scaling transformations; but any arbitrary linear transformation on $Y_{ts}$ will not affect the credibility estimate $\hat{\mu}_{ts}$.

An arbitrary transformation $T_s$ will generate:

$$Y^*_{ts} = T_s Y_{ts}$$

from which

$$Y^*_s = Y_s T'_s$$

and

$$Y^{*\prime}_s P_s Y^*_s = T_s Y'_s P_s Y_s T'_s$$

follow immediately.

In order that the mean value estimate still holds, the inverse transformation must be applied to $\beta_s$

$$\mu_{ts} = Y'_{ts}\beta_s = Y^{*\prime}_{ts}\beta^*_s \Rightarrow \beta^*_s = T'^{-1}_s \beta_s$$

Similarly, if the mean value were to hold using the countrywide $\beta$, this same transformation needs to be applied:

$$\beta^*_s = T'^{-1}_s \beta$$

With regard to the transformed estimates of $\beta_s$, it follows from the above that:

$$\hat{\beta}^*_s = T'^{-1}_s \hat{\beta}_s$$

With regard to the countrywide estimates $\hat{\beta}$, a transformed estimate will be denoted as:

$$\hat{\beta}_s^* = T_s'^{-1}\hat{\beta}$$

The transformed $\beta_s$ will now generate a transformed $\Gamma$ matrix which varies by state, denoted by:

$$\Gamma_s^* = T_s'^{-1}\Gamma T_s^{-1}$$

This will lead to a transformed credibility matrix:

$$Z_s^* = T_s Z_s T_s^{-1}$$

combining these elements to find the transformed estimate:

$$\mu_{ts}^* = [\hat{\beta}_s^* Z_s^* + \hat{\beta}_s^*(I - Z_s^*)]Y_{ts}^*$$

It is immediately clear that this estimate is identical with the original untransformed estimate.

Origin and Scale Transformations for the Trend Model

One of the immediate implications of the above results is that the credibility results found above would have been the same if our time data had been transformed to have zero mean and unit variance. Using the result of this transformation

$$Y_{ts}^* = \begin{pmatrix} 1 \\ \dfrac{t - \bar{t}_s}{\sigma_{ts}} \end{pmatrix}$$

simplifies the credibility form since

$$Y_s^{*\prime} P_s Y_s^* = P_{.s} I$$

However, now the $\Gamma$ matrix varies from state to state. Explicitly

$$\Gamma_s^* = \begin{pmatrix} \sigma_{as}^2 & \sigma_{abs} \\ \sigma_{abs} & \sigma_{bs}^2 \end{pmatrix}$$

$$= \begin{pmatrix} \sigma_a^2 + 2\sigma ab\bar{t}_s + \bar{t}_s^2\sigma b^2 & \sigma_{ts}(\sigma_{ab} + \bar{t}_s\sigma b^2 \\ \sigma_{ts}(\sigma_{ab} + \bar{t}_s\sigma b^2) & \sigma_{ts}^2\sigma b^2 \end{pmatrix}$$

The transformed credibility constant $K_s^*$ now takes on the simple form:

$$K_s^* = \sigma^2\Gamma_s^{*-1}$$

The transformed credibility matrix:

$$Z_s^* = P_{.s}(P_{.s}I + K_s^*)^{-1}$$

still has the same general form as in the untransformed case. The $\hat{\beta}_s^*$, $\hat{\hat{\beta}}_s^*$ and estimated values of $\Gamma_s^*$, $K_s^*$ and $Z_s^*$ are shown in figure 10 by state for the scale and location transformation.

Mixed Models

The upsetting results for the credibility adjusted trend line shown above in figure 9 came about because the mean value $\mu_{ts}$ is modeled in the same fashion for each state, specifically assuming that both slopes and intercepts were distributed about some mean value slope b and mean value intercept a. If we were to pause for a moment to think about our personal model of the trend situation; we would be more inclined to believe that while the avarage dollar at any point and time would vary substantially from state to state, the rate of change in the average dollar would tend to be the same from state to state. The modeling implication of this is, first of all, not to use a trend line; but to use an exponential trend. We will not pursue this direction in this paper. However, this analysis will be carried out in further research on this subject.

## FIGURE 10

### Estimates for Scaled t

| State | Transformed State Coefficients | Transformed Countrywide Coefficients | Transformed $\Gamma$ | Transformed $K$ | Transformed Credibility Matrix |
|---|---|---|---|---|---|
| $s$ | $\hat{\beta}^*_s$ | $\hat{\hat{\beta}}^*_s$ | $\Gamma^*_s$ | $K^*_s$ | $Z^*_s$ |
| 1 | $\begin{pmatrix} 2{,}061 \\ -216.04 \end{pmatrix}$ | $\begin{pmatrix} 1{,}864 \\ -150.11 \end{pmatrix}$ | $\begin{pmatrix} 95{,}058 & -29{,}596 \\ -29{,}596 & 9{,}651 \end{pmatrix}$ | $\begin{pmatrix} 10{,}244 & 31{,}415 \\ 31{,}415 & 100{,}904 \end{pmatrix}$ | $\begin{pmatrix} .9494 & -.1483 \\ -.1483 & .5213 \end{pmatrix}$ |
| 2 | $\begin{pmatrix} 1{,}511 \\ -59.66 \end{pmatrix}$ | $\begin{pmatrix} 1{,}870 \\ -150.88 \end{pmatrix}$ | $\begin{pmatrix} 97{,}433 & -30{,}135 \\ -30{,}135 & 9{,}751 \end{pmatrix}$ | $\begin{pmatrix} 10{,}244 & 31{,}661 \\ 31{,}661 & 102{,}367 \end{pmatrix}$ | $\begin{pmatrix} .9068 & -.2348 \\ -.2348 & .2235 \end{pmatrix}$ |
| 3 | $\begin{pmatrix} 1{,}806 \\ -150.21 \end{pmatrix}$ | $\begin{pmatrix} 1{,}858 \\ -150.36 \end{pmatrix}$ | $\begin{pmatrix} 92{,}511 & -29{,}227 \\ -29{,}227 & 9{,}684 \end{pmatrix}$ | $\begin{pmatrix} 10{,}244 & 30{,}919 \\ 30{,}919 & 97{,}868 \end{pmatrix}$ | $\begin{pmatrix} .8911 & -.2469 \\ -.2469 & .1915 \end{pmatrix}$ |
| 4 | $\begin{pmatrix} 1{,}353 \\ -98.01 \end{pmatrix}$ | $\begin{pmatrix} 1{,}860 \\ -152.80 \end{pmatrix}$ | $\begin{pmatrix} 93{,}168 & -29{,}811 \\ -29{,}811 & 10{,}000 \end{pmatrix}$ | $\begin{pmatrix} 10{,}244 & 30{,}539 \\ 30{,}539 & 95{,}443 \end{pmatrix}$ | $\begin{pmatrix} .8251 & -.2530 \\ -.2530 & .1193 \end{pmatrix}$ |
| 5 | $\begin{pmatrix} 1{,}600 \\ -41.68 \end{pmatrix}$ | $\begin{pmatrix} 1{,}869 \\ -152.16 \end{pmatrix}$ | $\begin{pmatrix} 96{,}991 & -30{,}318 \\ -30{,}318 & 9{,}917 \end{pmatrix}$ | $\begin{pmatrix} 10{,}244 & 31{,}320 \\ 31{,}320 & 100{,}194 \end{pmatrix}$ | $\begin{pmatrix} .9222 & -.2119 \\ -.2119 & .3136 \end{pmatrix}$ |

Restricting our thinking to the trend line model, the credibility model which is most meaningful would be one in which only the slope is considered to be a variable from state to state, but where the intercept is a constant:

$$\mu_{ts}(\theta_s) = a_s + b(\theta_s)t$$

This sort of model is directly analogous to the Bühlmann, Straub introduction of treaty conditions in their paper, which allow the severities to be modified by some function before entering the credibility formula.

We have shown above that scale and translation formulation will not affect our final credibility estimate. For ease of exposition in this section, we will assume that the time values in our trend line have been chosen so that the weighted average of observed times is zero and the weighted sample variance is equal to one. The modifications to our basic credibility model, because of the constant values $a_s$ within the mean value $\mu_{ts}$ formula, are fairly simple. For the regular credibility model $\beta_s$ was the same function of $\theta_s$ for all states. In our mixed model this function varies from state to state:

$$B_s(\theta_s) = \begin{pmatrix} a_s \\ b(\theta_s) \end{pmatrix}$$

The expected value of this function varies from state to state:

$$E[B_s(\theta_s)] = {}_s\beta = \begin{pmatrix} a_s \\ b \end{pmatrix}$$

We have chosen to denote this expected value as ${}_s\beta$ to avoid confusion with the function of $\theta$, $\beta_s$. The covariance matrix $\Gamma_s^*$ is:

$$E[B_s(\theta_s)B'(\theta_s)] - {}_s\beta_s\beta' = \Gamma_s^* = \begin{pmatrix} 0 & 0 \\ 0 & \sigma_{bs}^2 \end{pmatrix}$$

with the only non-zero entry being $\sigma_b^2$.

If we introduce for state $j$

$$K_{bj} = \sigma^2/\sigma_{bj}^2$$

to define:

$$Z_{bj} = P_{.j}/(P_{.j} + K_{bj})$$

The credibility matrix for our mixed model becomes:

$$Z_j = \begin{pmatrix} 0 & 0 \\ 0 & Z_{bj} \end{pmatrix}$$

Using the same theoretical development as in the regular credibility model, for the mixed model leads to:

$$\hat{\mu}_{ij} = [\hat{\beta}_j' Z_j + {}_j\beta'(I - Z_j)]Y_{ij}$$

The only difference is this estimate is that ${}_j\beta'$ replaces $\beta'$.[6] This estimate may be written for the trend case without recourse to matrices simply as:

$$\hat{\mu}_{ij} = a_j + [\hat{b}_j Z_{bj} + \hat{b}(1 - Z_{bj})]i$$

Using the formulas for the mixed model, the constant $K$, the credibility and finally the credibility adjusted slopes are shown on figure 11. For this mixed model, our credibility results are much more pleasing since the credibility adjusted

[6] It is important to note that this result holds for any mixed model, not just for out trend case. The most general mixed model, of course, allows arbitrary elements of $\beta_s$ to be considered independent of $\theta_s$.

FIGURE 11.

Credibility Adjusted Slopes
Without Intercept Adjustments

| State | Number of Claims Over 3 Years | | Credibility | Transformed State Slope | Transformed Credibility Adjusted Slope | Transformed Countrywide Slope |
|---|---|---|---|---|---|---|
| $s$ | $P_{.s}$ | $K_{bs}$ | $Z_{bs}$ | $\hat{b}^*$ | $\tilde{b}^*$ | $\hat{\hat{b}}^*$ |
| 1 | 100,155 | 4,565 | .9564 | -216.04 | -213.17 | -150.11 |
| 2 | 19,895 | 4,518 | .8149 | - 59.66 | - 76.54 | -150.88 |
| 3 | 13,735 | 4,550 | .7512 | -150.21 | -150.21 | -150.36 |
| 4 | 4,152 | 4,406 | .4852 | - 98.01 | -126.22 | -152.80 |
| 5 | 36,110 | 4,443 | .8904 | - 41.68 | - 53.79 | -152.16 |

slope must lie between the state and countrywide slopes. Further, some general observations can be made concerning the relative size of credibility to be given to state data. With this five state base as countrywide for most states, the number of claims that are observed show extremely high credibility. Only for the smallest state #4, with 4,152 claims observed over three years is credibility lower than .5. Of course, for practical application, the credibility standard should be developed using all of the states not just five.

Discussion by Al Quirin of <u>Credibility for Regression Models with Application to Trend</u>

This paper considers an arbitrary linear regression model, incorporates the Bühlmann Straub formulation of the model, extends the estimator form considered in the Bühlmann Straub model, exhibits the relationship between the least squares estimators, and finally derives computational results involving simple linear trend.

| | | |
|---|---|---|
| Arbitrary Linear Regression Model Considered | $E(x_{ts}) = u_{ts} = y'_{ts}\beta_s$ | (1) |
| Bühlmann-Straub Formulation Incorporated | $E(x^2_{ts}) - u^2_{ts} = \sigma^2_s/\rho_{ts}$ | (2) |
| | $E(x_{is}x_{js}) - u_{is}u_{js} = 0,\ i \neq j$ | (3) |
| Bühlmann-Straub Estimator Form | $\hat{u}_{ts} = x'A$ | (4a) |
| Extended to | $\hat{u}_{ts} = \alpha_0 + x'A$ | (4b) |

Relationship of Least Squares Estimators

using (4b) $\qquad \hat{u}_{ts} = [\hat{\beta}'_s z_s + \beta'(I - z_2)]y_{ts}$

using (4a) $\qquad \hat{u}_{ts} = [\hat{\beta}'_s z_s + \beta'(I - z_2)]y_{ts}$

$$\text{Where } d = \frac{\sum_s \beta'_s z_s \Gamma^{-1}\beta}{\sum_s \beta' z_s \Gamma^{-1}\beta}$$

Adequate accountability for inflation has become the single most important need in Property and Casualty insurance ratemaking today. In response to this need, Mr. Hachemeister's paper developing credibility standards for arbitrary linear regression models and in particular, developing credibility adjusted state trend lines, should prove to be invaluable.

In his Introduction, the author mentions that "no standards have been specifically developed for evaluating (the) credibility of state trend lines vs. countrywide trend lines." Although not specifically developed for analyzing trend, a credibility procedure has been used for some time by the Insurance Services Office (ISO) in their trend calculations, at least in private passenger automobile insurance. In each state, the determination of the average annual change in paid claim costs and claim frequencies is accomplished by credibility weighing the state and countrywide average annual changes. These average annual changes are taken from linear and exponential least squares trend lines for paid claim costs and claim frequencies, respectively. The credibility weights assigned are based on the latest year ending number of claims. Unfortunately, the theoretical justification for this approach is no deeper than assuming the number of claims has a Poisson distribution, and approximating probabilities by the use of the normal distribution. The standard for full credibility is 10,623 claims and reflects a probability of .99 that the number of claims will be within ±2.5% of the expected number of claims (on the assumption that the mean is equal to the variance). Partial credibilities are obtained using the formula $Z^2 = \frac{P}{10,623}$, where P is the latest year ending number of claims needed for partial credibility Z. The theoretical soundness of this procedure has been proven deficient by several authors, but up until this point in time, the theoretical advantages of alternative procedures do not seem to outweigh the practical advantage of simplicity (both in explanation to state insurance departments and in mathematical computation) present in the current procedure. From my own point of view, even though I feel that simplicity is a much overrated virtue in the very technical business of insurance ratemaking and that theoretical soundness should be of primary importance, I am convinced that <u>any alternative</u> credibility procedure will face the rather strict test of practical expediency before being implemented by those in the business of pricing insurance. With regard to Mr. Hachemeister's paper, it is precisely its simplicity in practical application (as well as its theoretical validity) which leads me to believe that it will someday soon become extensively utilized in calculating trend.

In the first half of the paper, the author states the problem of state vs. countrywide trend, introduces notation, displays data for a computational example, presents basic summary statistics, and reviews the classical and generalized linear regression model. Although the author has made mention to the point, it should be reiterated that even though the form of the estimator

$$\hat{\beta}_s = (y_s' c_s^{-1} y_s)^{-1} y_s' c_s^{-1} x_s$$

follows that obtained in classical <u>generalized</u> least squares estimation and that the theoretical results hold in general for the positive-definite matrix $C_s$, the assumption made regarding autocorrelation in deriving numerical results is not that of generalized least squares. In particular, recall that the classical generalized least squares formulation of the state $s$ trend model is

i) $E(x_{ts}) = u_{ts} = a_s + b_s t \qquad t = 1,\dots,n$

ii) $E(x_{ts}^2) - u_{ts}^2 = c_{ts} = \sigma_s^2/\rho_{ts} \qquad s = 1,\dots,N$

The $n \times n$ positive definite matrix $C_s$ allows for both heteroscedasticity and autocorrelation, i.e., for both

iii) $E(x_{ts}^2) - u_{ts}^2 = \sigma_s^2, \qquad \forall t$

and

iv) $E(x_{is}x_{js}) - u_{is}u_{js} = 0, \qquad i \neq j$

not holding. However, in deriving numerical results, Hachemeister disallows autocorrelation by assuming that iv) holds. In other words, should these problem be found to occur in trend data, further computational refinements will become necessary in practical application.

An approach to the solution of the problem of state vs. countrywide trend, is then formulated as a compound decision problem. In particular, the mean value $\mu_{ts}$ of a "credibility adjusted state $s$ trend line" is modeled as

v) $\mu_{ts}(\theta_s) = a_t(\theta_s) + b_t(\theta_s)t$

where for each state $s$ and each time period $t$, one acts as if the slopes and intercepts were distributed about some mean slope $E[b_t(\theta_s)]$ and some mean intercept $E[a_t(\theta_s)]$. Best linear unbiased estimators (BLUE) are then considered of the form

vi) $\hat{\mu}_{ts} = \alpha_0 + \sum_{s=1}^{N} \sum_{t=1}^{n} \alpha_{ts}x_{ts} = \alpha_0 + x'A$

and are found to be

vii) $\hat{\mu}_{ts} = [\hat{\beta}_s z_s + \beta'(I - z_s)]y_{ts}$

The application of this result to real data requires that estimates be made of various parameters not directly observable within the credibility model (e.g. $z_s$ in vii) is a function of $K_s$ which in turn depends on estimates of $\sigma^2$, $v$, and $\Gamma$). Because of the need for these estimates, assumptions iii) and iv) are made to simplify the derivation of numerical results.

The invariance property of $\hat{\mu}_{ts}$ for <u>any linear transformation</u> of the independent variables follows in a straightforward manner. Using this result, Hachemeister performs a scaling and translation on the linear trend model so that the weighted average (using # claims as weights) of scaled times in zero and the sample variance of scaled times is equal to unity. Finally, a mixed model is employed, to avoid the distressing results obtained when state intercepts are credibility adjusted, so that the final model chosen is

$$\mu_{ts}(\theta_s) = a_s + b(\theta_s)t \ .$$

Note that in this model the intercept varies by state but is assumed constant over all time periods. For each state $s$ and each time period $t$ the slope is still considered to be distributed about some mean slope. The effect of the estimated form in this mixed model is that slopes are credibility adjusted while intercepts are not.

To investigate the credibility standards developed and to evaluate the procedure finally decided upon in credibility adjusting state trend lines, consider the transformed simple linear trend model which credibility adjusts slopes without intercept adjustments.

Here,

$$u_{ts}(\theta_s) = a_s + b(\theta_s)t$$

where

$$\beta_s(\theta_s) = \begin{pmatrix} a_s \\ b(\theta_s) \end{pmatrix}$$

and

$$E[\beta_s(\theta_s)] = {}_s\beta = \begin{pmatrix} a_s \\ b \end{pmatrix} \ .$$

The estimator becomes for state $s$

$$\begin{aligned} \hat{u}_{ts} &= [\hat{\beta}'_s z_s + {}_s\beta'(I - z_s)]y_{ts} \\ &= a_s + [\hat{b}_s z_{bs} + \hat{b}(1 - z_{bs})]t \end{aligned}$$

where

$$z_{bs} = P_{.s}/(P_{.s} + K_{bs})$$

and

$$K_{bs} = \sigma^2/\sigma^2_{bs} \ .$$

Note that the credibility parameter $K_{bs}$ satisfies the general definition demonstrated by Bühlmann that it be equal to

$$\frac{\text{expected value of process variance } (= \sigma^2)}{\text{variance of the hypothetical means } (= \sigma^2_{bs})} \ .$$

The $K_{bs}$'s vary by state but a single constant K value could be adopted should the $K_{bs}$'s developed for all states show the same stability (centered around 4,500) as those developed for the five selected states.