Are Property-Casualty Insurance Reserves Biased?

A Non-Standard Random Effects Panel Data Analysis [1]

Herman J. Bierens

Department of Economics, Pennsylvania State University

David F. Bradford [2]

Woodrow Wilson School, Princeton University, and New York University School of Law

(April 13, 2005)

Abstract: In this paper we investigate whether insurance companies systematically over- or underestimate their reserves, in the sense that the reported reserves deviate from the expected future losses (called "bias"), and if so by how much, by focusing on reported reserves for covering losses from previous accident years, in five policy lines. An important part of our analysis consists of formulating conditions on the data generating process, and conditions on the forming of (rational) conditional expectations, such that this test can be conducted on the basis of linear regression analysis. Our conclusion is that in the period 1983-1993 the insurance companies in the US were systematically overstating their reserves.

JEL classification: C11, C12, C23, C51, C52, D84

Key words: Insurance; Reserves; Rational expectations; Bias; Random effects; Panel data model.

---

[1] This joint project was started, on the basis of an earlier paper by David Bradford, while both authors enjoyed the hospitality of the Center for Economic Studies at the University of Munich during August 1998. This paper is a part of the project on the Economics of the Property and Casualty Insurance Industry within the National Bureau of Economic Research, Cambridge, MA, with the support of the Insurance Information Institute. The helpful comments of Christopher A. Sims are gratefully acknowledged. A previous version of this paper was presented by Herman Bierens at Rice University, Texas A&M University, Southern Methodist University, Universite des Sciences Sociales de Toulouse, and the Econometric Society European Meeting 2001, Lausanne. During the preparation of this paper, David Bradford enjoyed the hospitality of the Economic Policy Research Unit at the Copenhagen Business School.

[2] David Bradford passed away on February 22, 2005.

1. Introduction

"Unpaid losses" (known more commonly as loss reserve) is the name for an entry on the balance sheet of a property-casualty insurance company that records the amount the company expects to pay out in the future as a consequence of insured events that have already occurred. Inevitably, because there is no scientific way of knowing future outlays exactly, the loss reserves reported by a company are subject to some element of discretion. Since these balance sheet entries represent important elements of the financial accounts, and translate directly into impacts on income for financial accounting and income tax purposes, companies may have an interest in either overstating or understating them. Overstating a loss reserve can be regarded as conservative behavior, since it increases the chances that the future will bring positive, rather than negative, surprises.

Bradford and Logue (1998) describe the way changes in the U.S. income tax rules over time created incentives for property-casualty insurance companies to exercise varying degrees of conservatism in their reporting of unpaid loss reserves. That paper includes an inspection of data on the U.S. property-casualty insurance industry in the aggregate that the authors found suggestive of behavior consistent with the influence of tax incentives. The authors emphasize, however, the difficulty of drawing firm conclusions about bias in reserve reporting. In this paper we attempt a more formal method to extract information about bias from aggregated accounting data.

We will test the no-bias hypothesis on U.S. industry aggregate data for the reporting years, 1985-94 on loss payments, $l(t,s,i)$, including expenses, during year t with respect to accident year s on policies in line i, and the reported (annual statement) loss reserve, $S(t,s,i)$, carried at the end of year t with respect to unpaid losses attributable to accident year s on policies in line i. The data are aggregated into five lines of insurance, covering virtually all property-casualty policies. See Bradford and Logue (1998, Table 7.1, p. 282). For each reporting year six accident years are included in the sample. (The reports cover identified accident years up to ten years in the past at any time.) All of the variables are expressed as ratios to premiums for the given accident year and line. Thus, the data consist of n (=270) observations on the random variables $l(t,t-m,i)$, $S(t,t-m,i)$ and

$S(t-1,t-m,i)$ for i = 1,..K (=5), m = 1,..,M (=6), and t = 1,..,T (=9).[3]

It is unrealistic to assume that the observations are independent across different lines and accident years. Therefore, effectively our data set consists of nine observations on a 30-variate vector time series process. Actually, we are dealing with a kind of panel data set, but without independence over the cross-sections.

Denoting the expected (by the company) amount remaining to be paid as of time t on policies covering accident year s on policies in line i by $M(t,s,i) = E_t[\sum_{j=1}^{\infty} l(t+j,s,i)]$, called the mean reserve, the problem is to determine the extent of the difference between the reported loss reserve $S(t,s,i)$ and the unobserved mean reserve $M(t,s,i)$. To solve this problem, we will formulate a random effect panel data model for the data generating process of $l(t,s,i)$, together with a behavioral model for the forming of the conditional expectations $E_t[l(t+j,s,i)]$, $j \geq 1$.

2. The problem of inferring bias

2.1 Some terminology

Companies report their balance sheet and operating results to state regulators on standardized "annual statements." The data are broken down by "reporting year" and by individual "lines" of insurance (for example, medical malpractice), and include information on policies classified by "accident year."

For example, suppose a company writes a contract to cover events occurring between September 1, 1987 and August 31, 1988. A premium is fixed and paid on September 1, 1987. This policy applies to parts of two accident years for annual statement purposes, namely, 1987 and 1988. Data relating to this policy would be found in annual statements for the reporting years, 1987, 1988, and subsequently, until all the claims under the contract have been paid. One third of the paid-in premium would be shown as earned in 1987, and a reserve would be shown for losses incurred during the period September 1 through December 31, 1987 and still unpaid as of December 31, 1987. The remainder of the premium would appear on the annual statement for 1988, shown as earned in

---

[3] The data file is available at http://econ.la.psu.edu/~hbierens/BRADBIERDATA.TXT

1988. At that point there would also be a reserve shown for losses incurred during 1988. As of the end of 1988, this policy would be contributing to two loss reserves: one for losses during accident year 1987 that had not been paid as of December 31, 1988; and one for losses during accident year 1988 that had not been paid as of December 31, 1988. Both reserves would be updated annually until there were no further payment liabilities anticipated as a consequence of accidents during 1987 and 1988, respectively.

The total "incurred loss" for an accident year in a given line is, as the term suggests, the sum of claims that have been or will be paid as a result of events during the period covered by the policies in question. This total is unknown until the last payment has been made. Before that time, the incurred loss amounts to a forecast, the sum of "paid losses," which are known, and "unpaid losses," which have to be surmised. The latter may be distinguished from the expected value sum of future loss payments in the minds of the actuaries and other officials who make the forecasts. We refer to this idealized amount as the "expected reserve" or "mean reserve," which we treat as the actuarial best point forecast. This paper concerns possible systematic differences between mean and reported reserves.

2.2 The problem of observing bias

We cannot observe the mean reserves. We do, however, observe the losses paid during each year and the loss reserves reported by the company at the end of each year. To look for biases, the approach taken to date in the literature has been to compare the reported reserves on a portfolio of policies at the end of an accident year with the actual loss payments realized over the lifetime of that portfolio or over the first *x* years in the lifetime of that portfolio.[4]

There are at least three problems with this approach. First is the paucity of long time series. The industry aggregate data used in the present paper, assembled by Bradford and Logue, cover reporting years 1981 through 1994. (Bradford and Logue also collected company-level data that cover reporting years 1984 through 1993.) On the "long-tail" lines (the length of the tail refers to

---

[4] For example, see Weiss (1985), Harrington (1988), and Petroni (1992).

temporal dimension of payouts, rather than to the probability distribution), more than half of the total payment of losses and loss expenses for an accident year typically occurs more than three years after that year.[5] Loss reserves are carried on the balance sheet for these lines even after ten years. For them, then, we do not actually have data on the complete life cycle of any accident year. Even if we cut off the portion of the policy life examined at five years (so we considered, for example, the change in estimated losses for a line with respect to accident year 1985 between the initial report for 1985 through the report for 1990), the aggregate data would permit us to study only nine instances of potential reserve bias (accident years 1981-1989, using data through 1986-1994). The company-level data would include only five instances.

A second problem with using the difference between the initially claimed incurred loss and its value at a subsequent date is sorting out what portion of reserve strengthening or weakening results from the "unwinding" of initial reserve overstatement or understatement and what portion from genuine uncertainty, which the data suggest is considerable, about loss development. Although having separate data as we do on five different lines of insurance increases the amount of useful information (for example, observation of significant reserve strengthening for each of five accident years for all five lines would support the case for downwardly biased initial reserves or upwardly biased current reserves), the true uncertainty in the portfolios renders statistical inference difficult.

A third problem with using initial versus subsequent values of losses incurred is that the rationale for firms' exercising discretion in the reporting of reserves implies that they will want to bias the estimates by different amounts in each reporting year. Suppose, for example, a company wants to "smooth" its underwriting earnings as the loss experience unfolds. Then in an unusually bad year, in terms of current loss payout, the company will want, within the limits of its discretion, to weaken reserves (thereby lowering a balance sheet liability and enhancing reported earnings) on all lines and for all accident years. In an unusually good year, the company will want to strengthen reserves on all lines and for all accident years. This simple theory suggests, therefore, not simply a consistent bias in the reported reserves (relative to an actuarial expectation), but rather a bias that

[5] For descriptive statistics, see Bradford and Logue (1998).

varies over time. The reserve strengthening over the first *n* years of the life of policy portfolios results from the passing of those portfolios through different combinations of reporting years (for example, the 1984 accident year portfolio passes through reporting years 1984 through 1989 in the first five years; the 1985 portfolio through years 1985 through 1990, etc.). It is difficult, therefore, to tell what combination of bias in the initial and subsequent reports of incurred losses results in the difference between the two and thereby to analyze year-to-year variation in the degree of reporting bias.

3. Using annual revisions to infer bias

As an alternative we look for evidence of reporting bias in the annual revisions of reserves. If the company reports, for each insurance line *i* and accident year *s*, its mean reserve as its loss reserve, the ratio of the actual expectation of remaining losses to be paid to reported reserves, $M(t,s,i)/S(t,s,i)$, will be 1. We refer to this ratio as the "adjustment factor" which we call *beta*:

$$\beta(t,s,i) \equiv M(t,s,i)/S(t,s,i). \tag{1}$$

Thus, *beta* expresses the degree of bias in reported reserves.

For the reason given section 2.2, it will be assumed that for given *t* the same *beta* applies to all accident years *s* and lines *i*:

$$\beta(t) = M(t,s,i)/S(t,s,i). \tag{2}$$

The empirical work in this paper is concerned with testing whether the actual reports are consistent with a value of *beta* equal to 1 for all accident years and lines, as would apply in the absence of reporting bias, given a model for the data generating process of $l(t,s,i)$, and a behavioral model for the forming of the conditional expectations $M(t,s,i) = E_t[\sum_{j=1}^{\infty} l(t+j,s,i)]$.

Note that

$$\begin{aligned} M(t-1,s,i) &= E_{t-1}[\textstyle\sum_{j=1}^{\infty} l(t-1+j,s,i)] = E_{t-1}[l(t,s,i)] + E_{t-1}[E_t\{\textstyle\sum_{j=1}^{\infty} l(t+j,s,i)\}] \\ &= E_{t-1}\big[l(t,s,i) + M(t,s,i)\big] = l(t,s,i) + M(t,s,i) - \varepsilon(t,s,i), \end{aligned} \tag{3}$$

say, where

$$\varepsilon(t,s,i) = l(t,s,i) + M(t,s,i) - E_{t-1}\big[l(t,s,i) + M(t,s,i)\big]. \tag{4}$$

For given s and i, $\varepsilon(t,s,i)$ is a martingale difference process:

$$E_{t-1}[\varepsilon(t,s,i)] = 0 \tag{5}$$

with probability 1, so that, in particular, $\varepsilon(t,s,i)$ is uncorrelated with any function of the past values $l(t-j,s,i) + M(t-j,s,i)$, $j = 1,2,...,t-s$.

Under the alternative (2) of reporting bias, equation (3) reads:

$$l(t,s,i) = \beta(t-1)S(t-1,s,i) - \beta(t)S(t,s,i) + \varepsilon(t,s,i). \tag{6}$$

Given that we observe $l(t,t\text{-}m,i)$, $S(t,t\text{-}m,i)$ and $S(t\text{-}1,t\text{-}m,i)$ for $i = 1,..K$, $m = 1,..,M$, and $t = 1,..,T$, and denoting

$$\begin{aligned}
&\beta = (\beta(0),......,\beta(T))' \\
&X(1,m,i) = (S(0,1-m,i),-S(1,1-m,i),0,0,....,0)' \\
&X(2,m,i) = (0,S(1,2-m,i),-S(2,2-m,i),0,....,0)' \\
&.................................... \\
&X(T,m,i) = (0,.........,0,S(T-1,T-m,i),-S(T,T-m,i))', \\
&Y_j = l(t,t-m,i),\quad X_j = X(t,m,i),\quad \varepsilon_j = \varepsilon(t,t-m,i),\ j = 1,...,n = T\times M\times K,.
\end{aligned} \tag{7}$$

where each $X(t,m,i)$ is a $(T+1)\times 1$ vector, we can write (6) more compactly as

$$Y_j = \beta' X_j + \varepsilon_j. \tag{8}$$

At first sight model (8) looks like an ordinary linear regression model. The problem, however, is that it unclear whether for fixed t the errors $\varepsilon(t,t-m,i)$ are uncorrelated, because (5) only guarantees that the $\varepsilon(t,t-m,i)$'s are uncorrelated across the t's for fixed $s = t\text{-}m$ and i. Moreover, (5) does not automatically guarantee that $E[\varepsilon(t,s,i)\,|\,S(t-1,s,i),S(t,s,i)] = 0$, which is a minimum requirement for the validity of a regression model. In particular, (6) reads as

$$E_{t-1}[l(t,s,i) + \beta(t)S(t,s,i)] = \beta(t-1)S(t-1,s,i),$$

so that the actual dependent variable of model (6) is $l(t,s,i) + \beta(t)S(t,s,i)$ rather than $l(t,s,i)$. Only if $E_{t-1}[S(t,s,i)] = S(t,s,i)$ we have $E[\varepsilon(t,s,i)\,|\,S(t-1,s,i),S(t,s,i)] = 0$, hence

$$E_{t-1}[l(t,s,i)] = \beta(t-1)S(t-1,s,i) - \beta(t)S(t,s,i).$$

Therefore, we need to specify first the dependence structure of the variables $l(t,t-m,i)$, $S(t,t-m,i)$, and $S(t-1,t-m,i)$ for different t, m, and i, and the meaning of the conditional expectation operator

E_t, in order to set forth conditions under which model (6), and consequently model (8), can be interpreted as a valid linear regression model.

4. Econometric issues

4.1 A multiplicative random effects panel data model for the data generating process

In order to be able to test the null hypothesis of no bias with this data set, we need to specify the stochastic properties of the variables $M(t,s,i)$ and $l(t,s,i)$. Since $M(t,s,i)$ is the conditional expectation of the sum of future values of $l(t,s,i)$, it suffices to specify the stochastic properties of $l(t,s,i)$ only, together with a specification of the information set of conditional expectation operator $E_t(.)$. The specification of the stochastic properties of the $l(t,s,i)$'s should take into account that this variable is possibly dependent across the t, s and i's, and the conditional expectation operator should be specified such that (3) and (5) hold. Moreover, we have to keep the specification as simple as possible, because our data set is rather small.

In view of these considerations we propose the following (non-standard) random effect panel data model for the data-generating process (DGP) of $l(t,s,i)$:

Assumption 1: $l(t,s,i) = q_{t-s,i}v_s w_i u_{t,s,i}$, *where for* $m = 0,1,....$ *the* $q_{m,i}$*'s are non-negative random variables satisfying* $P(\sum_{m=0}^{\infty}q_{m,i} = 1) = 1$, *and the* v_s*'s and* w_i*'s are positive-valued random variables. The* $u_{t,s,i}$*'s are random variables with conditional expectation* $E_{t-1}[u_{t,s,i}] = 1$ *and finite conditional variance* $E_{t-1}[(u_{t,s,i} - 1)^2] = \sigma^2_{t-1,s,i}$.

Note that Assumption 1 is akin (but not exactly equal) to the assumptions of the chain ladder loss reserving technique. See Taylor (2000).

The random effect w_i represents the risk factor of insurance line i. The random effect v_s represents the total amount of the claims resulting from accidents in year s. For given accident year s and line i the random variables $q_{t-s,i}$ represent the fraction of the total sum of claims paid out in year t, hence $\sum_{t=s}^{\infty}q_{t-s,i} = 1$. The $u_{t,s,i}$'s are the (multiplicative) error terms of the model.

Although the $l(t,s,i)$'s in our data set are all positive valued, it is conceivable that $l(t,s,i)$

takes occasionally a negative value, namely if a company wins a reversal of a lawsuit that has resulted in a payment in the past. Therefore, we will not restrict $u_{t,s,i}$ to be positive with probability 1. There will be no need to specify the distribution of the $q_{s,i}$'s, v_s's, and w_i's further, but we will need to specify the conditional variance $\sigma^2_{t-1,s,i}$ of $u_{t,s,i}$ further later on.

4.2 Conditioning

Next, we need to be more precise about the conditional expectation operator E_t. Since the risk factors for each insurance line will be known to the insurance companies, all the random effects w_i are included in the information set. Moreover, since the random effect v_s represents the total amount of the claims resulting from accidents in year *s*, this variable will be known by the end of year *s*, hence the v_s's for $t \geq s$ will be included in the information set as well. The random variables $q_{m,i}$, $m = 0,1,\ldots\ldots\ldots,$ represent the prior knowledge of the insurance companies as to how the total amount of the claims will be paid out over the years out for each line. These random variables are considered private information of the insurance companies based on past experience, hence they will also be included in the information set. Furthermore, since by the end of year *t* the actual values of $l(t-j,s,i)$ have been observed for $j \geq 0$ and $s \leq t\text{-}j$, the $u_{t\text{-}j,s,i}$'s have then been observed too. Therefore we assume that:

Assumption 2: $E_t[\cdot] = E[\cdot|\mathscr{F}_t]$, *where* $\mathscr{F}_t$ *is the* $\sigma-$*algebra generated by the random variables* $q_{m,i}$, v_s, w_i, $u_{j,s,i}$ *for* $m = 0,1,2,\ldots\ldots$, *all lines i,* $s = 0,\ldots,t$, *and* $j = s,\ldots,t$.

Thus, $\mathscr{F}_t$ is now the information set corresponding to the conditional expectation operator E_t.

4.3 The conditional variance of $u_{t,s,i}$

Since in reality,

$$P[\lim_{t-s\to\infty} l(t,s,i) = 0] = 1, \tag{9}$$

we need to impose further conditions such that (9) is true. Given Assumption 1, condition (9) is true if $P[\lim_{t-s\to\infty} l(t,s,i) = 0 \,|\, \{q_{m,i}\}_{m=0}^{\infty}, v_s, w_i] = 1$, which by the Borel-Cantelli lemma[6] is true if and only if for arbitrary $\delta > 0$, $\sum_{t=s}^{\infty} P[|l(t,s,i)| > \delta \,|\, \{q_{m,i}\}_{m=0}^{\infty}, v_s, w_i] < \infty$ with probability 1. The latter is equivalent to

$$\sum_{t=s}^{\infty} P\left[u_{t,s,i}^2 > \frac{\delta^2}{q_{t-s,i}^2 v_s^2 w_i^2} \middle| \{q_{m,i}\}_{m=0}^{\infty}, v_s, w_i\right] < \infty,$$

which by Chebishev's inequality and the law of iterated expectations holds if

$$\begin{aligned} \sum_{t=s}^{\infty} q_{t-s,i}^2 v_s^2 w_i^2 \, E[u_{t,s,i}^2 \,|\, \{q_{m,i}\}_{m=0}^{\infty}, v_s, w_i] &= \sum_{t=s}^{\infty} q_{t-s,i}^2 v_s^2 w_i^2 \, E[E(u_{t,s,i}^2 | \mathscr{F}_{t-1}) | \{q_{m,i}\}_{m=0}^{\infty}, v_s, w_i] \\ &= \sum_{t=s}^{\infty} q_{t-s,i}^2 v_s^2 w_i^2 \, \sigma_{t-1,s,i}^2 + \sum_{t=s}^{\infty} q_{t-s,i}^2 v_s^2 w_i^2 < \infty \end{aligned} \tag{10}$$

Since by Assumption 1, $P(\sum_{m=0}^{\infty} q_{m,i} = 1) = 1$, we have that $P(0 \le q_{m,i} \le 1) = 1$, hence $P(q_{m,i}^2 \le q_{m,i}) = 1$. Thus, the second term on the second line in (10) is bounded by $v_s^2 w_i^2$. Therefore, a sufficient condition for (9) is that:

Assumption 3: $\sum_{t=s}^{\infty} [q_{t-s,i} v_s w_i]^2 \sigma_{t-1,s,i}^2 < \infty$.

4.5 The regression model errors

The model assumptions 1 and 2 imply that for $j > 0$,

$$E_t\big[l(t+j, s, i)\big] = q_{t+j-s,i} v_s w_i \tag{11}$$

so that

$$M(t,s,i) = \sum_{j=1}^{\infty} E_t\big[l(t+j, s, i)\big] = v_s w_i \sum_{j=1}^{\infty} q_{t+j-s,i}. \tag{12}$$

It follows from Assumption 2 that for $s \le t$-1,[7] (12) is measurable $\mathscr{F}_{t-1}$, hence

$$E_{t-1}[M(t,s,i)] = M(t,s,i) \tag{13}$$

Thus it follows from (12) and (13) that

$$E_{t-1}\big[l(t,s,i)\big] = q_{t-1-s,i} v_s w_i = M(t-1,s,i) - M(t,s,i). \tag{14}$$

---

[6] See for example Bierens (1994, Theorem 2.1.2).

[7] Which is the case in our data set.

Combining (3), Assumption 1, (13) and (14) now yields

$$l(t,s,i) = [M(t-1,s,i) - M(t,s,i)]u_{t,s,i}. \tag{15}$$

Comparing this result with (3), we see that the error term (4) is of the form

$$\varepsilon(t,s,i) = \left(u_{t,s,i} - 1\right)q_{t-s,i}v_s w_i = \left(u_{t,s,i} - 1\right)[M(t-1,s,i) - M(t,s,i)]. \tag{16}$$

which by Assumptions 1-2 has zero conditional expectation:

$$E_{t-1}[\varepsilon(t,s,i)] = \left(E[u_{t,s,i}|\mathscr{F}_{t-1}] - 1\right)[M(t-1,s,i) - M(t,s,i)] = 0 \tag{17}$$

and conditional variance

$$E_{t-1}[\varepsilon(t,s,i)^2] = \sigma^2_{t-1,s,i}[M(t-1,s,i) - M(t,s,i)]^2. \tag{18}$$

So far we have shown that Assumptions 1 and 2 are in accordance with the previous observations in Section 3. However, we also have (implicitly) established that

$$E[\varepsilon(t,s,i)|S(t_1,s_1,i_1),\ \forall\ s_1 \le t-1,\ t_1 \ge s_1+1,\ i_1] = 0. \tag{19}$$

To see this, observe from (12) and Assumptions1- 2 that

$$S(t_1,s_1,i_1) = M(t_1,s_1,i_1)/\beta(t_1) = \textstyle\sum_{j=1}^{\infty}E_t\big[l(t+j,s,i)\big] = v_{s_1}w_{i_1}\sum_{j=1}^{\infty}q_{t_1+j-s_1,i_1}/\beta(t_1)$$

is measurable $\mathscr{F}_{t-1}$ for all $s_1 \le t-1,\ t_1 \ge s_1,\ i_1$, hence the $\sigma-$algebra $\mathfrak{C}_{t-1}$, say, generated by the random variables $\{S(t_1,s_1,i_1),\ \forall\ s_1 \le t-1,\ t_1 \ge s_1,\ i_1\}$ is contained in $\mathscr{F}_{t-1}$. Consequently, it follows from the law of iterated expectations that

$$E[\varepsilon(t,s,i)|\mathfrak{C}_{t-1}] = E[E(\varepsilon(t,s,i)|\mathscr{F}_{t-1})|\mathfrak{C}_{t-1}] = 0.$$

Note that the condition $t_1 \ge s_1+1$ in (19) is automatically fulfilled for our data set, so that the condition $s_1 \le t-1$ in (19) is the actual constraint. Moreover, note that the information set $\mathfrak{C}_{t-1}$ also includes future reserves, as long as the accident years involved are less or equal to $t-1$. At first sight this may look odd, because in time series analysis forecasts are usually functions of past innovations and therefore correlated with these innovations. The reason that we get the result (19) is that our DGP model in Assumption 1 has no feed-back: If we would have specified an AR, MA or ARMA model for $l(t,s,i)$ (or its log), then indeed present and future reserves would have been correlated with $\varepsilon(t,s,i)$.

We will now set forth further conditions such that

$$E[\varepsilon(t,s,i)|S(t_1,s_1,i_1),\ \forall\ t_1,s_1,i_1] = 0, \tag{20}$$

because if so then we may treat the reserves $S(t,s,i)$ as fixed regressors so that the OLS estimators of the betas will be unbiased. Since our data set is small, unbiasedness is preferable if possible.

We cannot assume that the errors $\varepsilon(t,s,i)$ themselves are i.i.d. and independent of all the $q_{m,i}$, v_s, and w_i, because it follows from (14), (18), and Assumption 3 that $\sum_{t=s}^{\infty} E_{t-1}[\varepsilon(t,s,i)^2] < \infty$, which is impossible if the conditional variances involved are constant. At most we may assume that $\varepsilon(t,s,i) = \lambda(t,s,i)\xi(t,s,i)$, say, where the $\xi(t,s,i)$'s are i.i.d. (0,1) and independent of all the $q_{m,i}$, v_s, and w_i, with $\lambda(t,s,i)$ a function of t, s, the $q_{m,i}$'s, v_s, and w_i such that $\sum_{t=s}^{\infty}\lambda(t,s,i)^2 < \infty$. The latter condition is satisfied if we specify

$$\sigma^2_{t-1,s,i} = \left(\frac{\Phi(M(t-1,s,i) - M(t,s,i))}{M(t-1,s,i) - M(t,s,i)}\right)^2, \tag{21}$$

where Φ is a function on $[0,\infty)$ such that

$$\sum_{t=s}^{\infty}\lambda(t,s,i)^2 = \sum_{t=s}^{\infty}\Phi(M(t-1,s,i) - M(t,s,i))^2 = \sum_{t=s}^{\infty}\Phi(q_{t-s,i}v_s w_i)^2 < \infty. \tag{22}$$

This condition is satisfied if for some $\delta > 0$, $|\Phi(x)| \le x^{\delta}$ for $x \downarrow 0$. Thus, if we specify:

Assumption 4: $u_{t,s,i} = 1 + \xi(t,s,i)\Phi(M(t-1,s,i) - M(t,s,i))/(M(t-1,s,i) - M(t,s,i))$, *where the* $\xi(t,s,i)$ *'s are i.i.d.* (0,1) *and independent of all the* $q_{m,i}$, v_s, *and* w_i, *and* Φ *is a non-negative function on* $[0,\infty)$ *such that for some* $\delta > 0$, $\Phi(x) \le x^{\delta}$ *for* $x \downarrow 0$,

then

$$\begin{aligned} &l(t,s,i) = \beta(t-1)S(t-1,s,i) - \beta(t)S(t,s,i) + \varepsilon(t,s,i), \\ &\textit{with}: \\ &E[\varepsilon(t,s,i)|\{S(t_1,s_1,i_1),\ \forall\ t_1,s_1,i_1\}] = 0, \\ &E[\varepsilon(t,s,i)^2|\{S(t_1,s_1,i_1),\ \forall\ t_1,s_1,i_1\}] = \Phi\big(\beta(t-1)S(t-1,s,i) - \beta(t)S(t,s,i)\big)^2. \end{aligned} \tag{23}$$

Consequently, under Assumptions 1-4 model (6) is a valid linear regression model, but with heteroskedastic errors, and the same applies to model (8).

Recall that due to (20) the OLS estimators of the betas are unbiased. If we would assume that

Assumption 5: *The* $\xi(t,s,i)$ *'s in Assumption 4 are i.i.d.* $N(0,1)$, then conditionally on the $S(t,s,i)$ ‘s the OLS estimators are normally distributed. Asymptotic theory then only plays a role in estimating the variance matrix involved.

Without Assumption 5, asymptotic normality has to come from application of the central limit theorem. However, since the number of parameters is T +1 = 10, in assuming that $n = T \times M \times K \to \infty$ we have to assume that T is kept fixed. Since it is unrealistic to assume that the number K = 5 of insurance lines i can be increased, asymptotic results have to come from the dimension $t-s = 1,.....,M$. But M = 6 in our sample, so that the asymptotic normal approximation may be more fiction than fact if Assumption 5 is not true. On the other hand, the degree of freedom in our sample is n-(T+1) = 260, which is not too bad.

Admittedly, the assumption that T is fixed is not attractive, but we will in first instance maintain it. An alternative approach would be to make $\beta(t)$ a deterministic function of t and a fixed number of parameters, which we will consider as well.

As motivated before, it is not impossible that occasionally the loss $l(t,s,i)$ becomes negative. This event is equivalent to $u_{t,s,i} < 0$. Under Assumptions 4-5, the probability that $u_{t,s,i}$ is negative is

$$P(u_{t,s,i} < 0) = P\left[\xi(t,s,i) < -(M(t-1,s,i)-M(t,s,i))/\Phi(M(t-1,s,i)-M(t,s,i))\right]$$
$$\leq \int_{-\infty}^{-\inf_{x \geq 0} x/\Phi(x)} \frac{\exp[-\xi^2/2]}{\sqrt{2\pi}} d\xi. \qquad (24)$$

However, in our data set all the losses $l(t,s,i)$ are positive, hence the probability (24) and its upper bound will be small. Thus $\inf_{x \geq 0} x/\Phi(x)$ must be large, and equivalently, $\sup_{x \geq 0} \Phi(x)/x$ must be small.

5. Empirical results

5.1 Estimates and tests of the betas: Preliminary results

The OLS estimates and test results of the null hypothesis of no bias are presented in Table 1 below. The tests are conducted under the hypothesis of heteroskedasticity: the test statistics are based on White’s (1980) heteroskedasticity consistent asymptotic variance matrix. The tests of a single beta equal to 1 are asymptotic standard normal tests, with two-sided asymptotic p-values, and the joint test is the Wald test of the hypothesis that all the betas are 1, with an asymptotic $\chi^2(10)$ distribution. The Jarque-Bera (1980) normality test has a $\chi^2(2)$ null distribution, and the Breusch-Pagan (1979) homoskedasticity test has a $\chi^2(10)$ null distribution.

Table 1: OLS estimation and test results

t	β	H_0: β = 1 test	p-value
1985	1.079994	4.712	0.00000
1986	1.003663	0.199	0.84247
1987	0.966012	-1.834	0.06669
1988	0.933022	-3.077	0.00209
1989	0.878509	-4.669	0.00000
1990	0.892571	-3.914	0.00009
1991	0.913700	-2.869	0.00412
1992	0.918271	-2.211	0.02702
1993	0.943643	-1.160	0.24617
1994	0.959709	-0.588	0.55628
Joint test:		203.96	0.00000

Standard error of the residuals:	0.02434
R-square:	0.86154
Adjusted R-square:	0.85674

Normality and homoskedasticity tests:	test	p-value
Jarque-Bera normality test:	225.02	0.00000
Breusch-Pagan homoskedasticity test:	114.09	0.00000

Clearly, the no bias hypothesis is strongly rejected by the joint Wald test. For year 1985 the $\beta(t)$ is significantly larger than 1, and for years 1988-1992 significantly less than 1. Recall that a $\beta(t)$ less than 1 constitutes an overstatement of the reserves.

As expected, both normality and homoskedasticity are strongly rejected. However, these results may be related: It is possible that the $\xi(t,s,i)$'s in Assumption 4 are normally distributed (Assumption 5) but that due to the heteroskedasticity of $\varepsilon(t,s,i)$ the Jarque-Bera (1980) normality test rejects.

5.2 Test for model misspecification

The Assumptions 1-4 lead to the key result (20), which allows us (or gives us an alibi) to estimate the model by OLS, but is (20) true? We cannot test (20) entirely, but we can test the weaker hypothesis that

$$E[\varepsilon(t,s,i)|S(t-1,s,i),\ S(t,s,i)] = 0, \tag{25}$$

using the Integrated Conditional Moment (ICM) test of Bierens and Ploberger (1996).

The version of the ICM test we have used is:

$$\hat{F}(c) = \frac{\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\hat{e}_i\hat{e}_j\prod_{k=0}^{1}\frac{\sin\left(c[\arctan(\tilde{x}_{n,i}(k))-\arctan(\tilde{x}_{n,j}(k))]\right)}{c[\arctan(\tilde{x}_{n,i}(k))-\arctan(\tilde{x}_{n,j}(k))]}}{\frac{1}{n}\sum_{j=1}^{n}\hat{e}_j^2}, \tag{26}$$

where the $\hat{e}_j$'s are the OLS residuals, and $\tilde{x}_{n,j}(k)$ is $S(t-k,s,i)$ in deviation of the sample mean and divided by the sample standard errors. As has been shown by Bierens and Ploberger (1997), under the null hypothesis (25), $\hat{F}(c) \to \bar{F}$ in distribution, where

$$P(\bar{F} > 3.23) \leq 0.10, \quad P(\bar{F} > 4.26) \leq 0.05, \tag{27}$$

whereas under the alternative $P(E[\varepsilon(t,s,i)|S(t-1,s,i),\ S(t,s,i)] = 0) < 1$, $\text{plim}_{n\to\infty}\hat{F}(c) = \infty$ for fixed $c \in (0,\infty)$.

The ICM test has been conducted in three-fold, for c = 1,5,10, because the small sample power of the test depends on the choice of c. In particular, $\lim_{c\to\infty}\hat{F}(c) = 0$, so that c should not be chosen too large, whereas

$$\hat{F}(0) = \lim_{c\downarrow 0}\hat{F}(c) = \frac{\left((1/\sqrt{n})\sum_{j=1}^{n}\hat{e}_j\right)^2}{(1/n)\sum_{j=1}^{n}\hat{e}_j^2}, \tag{28}$$

which is only a consistent test against the alternative that the model needs an intercept, but the critical values (27) are then no longer valid.

The ICM test results involved are $\hat{F}(1) = 17.37$, $\hat{F}(5) = 4.88$, and $\hat{F}(10) = 2.77$. Comparing these test results with (27), we see that for c = 1 and 5 the null hypothesis (25) is rejected at the 5% significance level.

The large value of $\hat{F}(1)$ suggests that the misspecification is, at least partly, due to a missing intercept. In order to check this conjecture, we have conducted the same ICM test on the basic of model (23) **with an intercept**. The results are $\hat{F}(1) = 2.53$, $\hat{F}(5) = 2.95$, and $\hat{F}(10) = 2.31$. Indeed, the null hypothesis (25) is now accepted for all three values of c, so we have to conclude that an intercept is required.

5.3 Why is an intercept needed?

Since Assumptions 1-4 do not account for an intercept, the question now arises why an intercept is needed. The reason may be the following:

As is well-known, conditional expectations are the best one-step-ahead forecasts under quadratic loss. But suppose that underestimating the necessary reserves is more costly than overestimating the reserves, or vice versa . Then the conditional expectation is no longer the best one-step ahead forecast.

Combining Assumptions 1 and 4 it follows that

$$\begin{aligned} R(t,s,i) &= \textstyle\sum_{j=1}^{\infty} l(t+j,s,i) = M(t,s,i) + \sum_{j=1}^{\infty} \xi(t+j,s,i)\Phi(M(t+j-1,s,i)-M(t+j,s,i)) \\ &= M(t,s,i) + U(t,s,i), \end{aligned} \tag{29}$$

say, where $R(t,s,i)$ is the total actual loss as of time t+1, given accident year s and line i. Next , suppose that the loss function is $L(u) = u^2 + cI(u \leq 0)$, where $I(.)$ is the indicator function and $c > 0$ is a penalty. Let $M^*(t,s,i)$ be the best one-step ahead forecast of $R(t,s,i)$, given the loss function $L(.)$, i.e., $M^*(t,s,i)$ is chosen such that

$$\begin{aligned} E_t[L(R(t,s,i) - M^*(t,s,i))] &= E_t\Big[\big(U(t,s,i) + M(t,s,i) - M^*(t,s,i)\big)^2\Big] \\ &\quad + cE_t\Big[I(Ut,s,i) \leq M^*(t,s,i) - M(t,s,i))\Big] \\ &= E_t[U(t,s,i)^2] + \big(M^*(t,s,i) - M(t,s,i)\big)^2 + cF_{t,s,i}\big(M^*(t,s,i) - M(t,s,i)\big), \end{aligned} \tag{30}$$

is minimal, where $F_{t,s,i}$ is the conditional distribution function of $U(t,s,i)$. If the insurance companies do not know $F_{t,s,i}$ exactly, and therefore use in (30) a non-random distribution function $\bar{F}_{t,s,i}$ instead of $F_{t,s,i}$ as a proxy, then the solution for $M^*(t,s,i)$ is of the form $M^*(t,s,i) = M(t,s,i) + \gamma_{t,s,i}$, where

$$\gamma_{t,s,i} = \underset{\gamma_*}{\operatorname{argmin}}\Big(\gamma_*^2 + c\bar{F}_{t,s,i}(\gamma_*)\Big) < 0 \tag{31}$$

is non-random. Finally, suppose that the reported reserve $S(t,s,i)$ is based on $M^*(t,s,i)$ rather than on $M(t,s,i)$: $\beta(t) = M^*(t,s,i)/S(t,s,i)$. Then model (6) becomes

$$l(t,s,i) = \beta(t-1)S(t-1,s,i) - \beta(t)S(t,s,i) + \beta(t)\gamma_{t,s,i} - \beta(t-1)\gamma_{t-1,s,i} + \varepsilon(t,s,i). \tag{32}$$

In general the term $\beta(t)\gamma_{t,s,i} - \beta(t-1)\gamma_{t-1,s,i}$ will not be constant, but in view of the ICM test results it is so close to be constant that the variation in $\beta(t)\gamma_{t,s,i} - \beta(t-1)\gamma_{t-1,s,i}$, if any, is not detectable.

5.4 The results for the model with intercept

Including an intercept in the model do not change the previous results in Table 1 too much. See Table 2 below. The main difference concern the betas for 1985 and 1986: The beta for 1985 is no longer significantly different from 1 and the beta for 1986 is now significantly less than 1. As expected from the result of the ICM test, the intercept is significant.

Table 2: OLS estimation and test results for the model with intercept

		$H_0: \beta = 1$	
t	β	test	p-value
1985	1.019714	0.862	0.38856
1986	0.946551	-2.241	0.02501
1987	0.910525	-3.639	0.00027
1988	0.879582	-4.379	0.00001
1989	0.830109	-5.082	0.00000
1990	0.846898	-4.496	0.00001
1991	0.872058	-3.633	0.00028
1992	0.882051	-2.899	0.00374
1993	0.913800	-1.701	0.08894
1994	0.938318	-0.904	0.36612
Joint test:		202.01	0.00000

intercept (=α):		$H_0: \alpha = 0$	
	α	test	p-value
	0.007985	4.356	0.00001

Standard error of the residuals:	0.02395
R-square:	0.86642

Adjusted R-square:		**0.86127**
Normality and homoskedasticity tests:	**test**	**p-value**
Jarque-Bera normality test:	**194.31**	**0.00000**
Breusch-Pagan homoskedasticity test:	**114.09**	**0.00000**

The OLS estimates of the betas, with 95% confidence interval bands, are plotted in Figure 1. As we see from Figure 1, the pattern of $\beta(t)$ is approximately V-shaped, with the kink of the "V" at t =1989. This pattern suggests a reparametrization of the model with $\beta(t)$ replaced by a piecewise linear function of t. We will consider such a reparametrization in the next subsection.

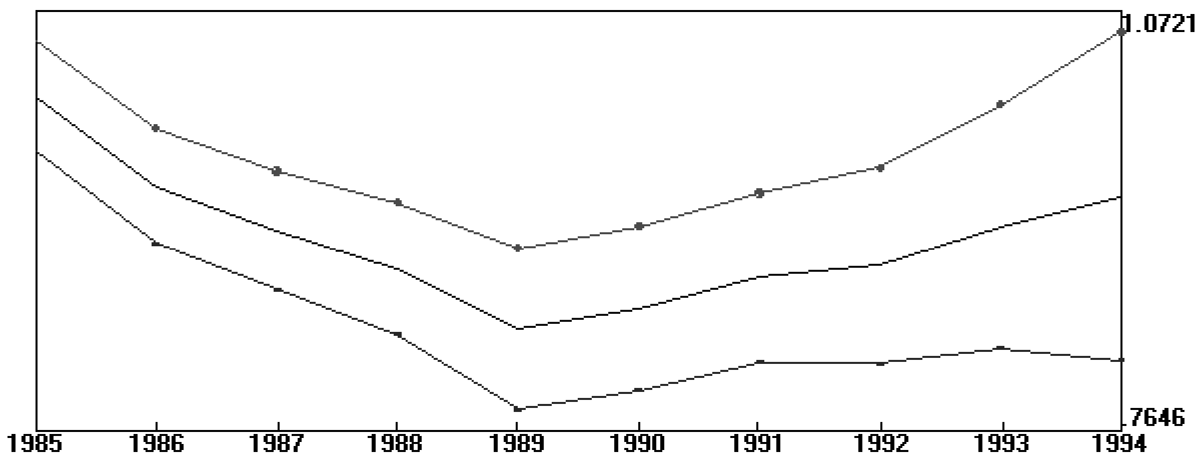

Figure 1: OLS estimates of $\beta(t)$ with 95% confidence interval bands

In order to see how the conditional variance $\Phi\big(M(t-1,s,i) - M(t,s,i)\big)^2$ of the errors look like, we have regressed the squared OLS residuals nonparametrically on the estimated conditional expectation $\hat{\alpha} + \hat{\beta}(t-1)S(t-1,s,i) - \hat{\beta}(t)S(t,s,i)$, where $\hat{\alpha}$ is the OLS estimator of the intercept, using the kernel regression approach in Bierens (1985) and Bierens and Pott-Buter (1990), with standard normal kernel, bias correction, and window width $1.5n^{-1/5}$. The factor 1.5 has been determined by (in-sample) cross-validation over the range [0.5,5], with 10 grid points. The result, i.e., the kernel estimate of $\Phi(x)^2$, is displayed in Figure 2.

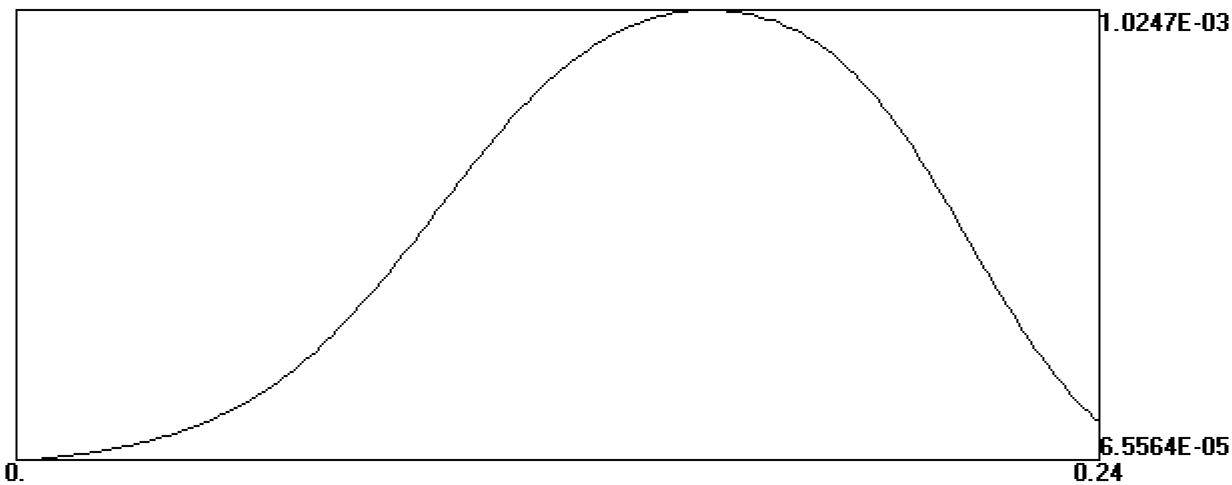

Figure 2: Kernel estimate of $\Phi(.)^2$

Although our sample is actually too small for reliable kernel regression, Figure 2 suggests that $\Phi(x)^2$ is bell-shaped, with zero derivative at zero. Therefore, a possibly parametrization of Φ which resembles the shape in Figure 2 is:

$$\Phi(x) = \gamma x^{1+\delta}\exp[-\theta x], \text{ where } \gamma > 0, \ \delta > 0, \ \theta > 0, \ x \geq 0. \tag{33}$$

As argued before, $\sup_{x\geq 0}\Phi(x)/x$ must be small. Thus in the case (33) $\sup_{x\geq 0}\Phi(x)/x = \gamma(\delta/\theta)^{1+\delta}e^{-\delta}$ must be small.

Given a correct specification of the functional form of Φ, we can improve the asymptotic efficiency of the parameter estimates by conducting (feasible) generalized least squares (GLS) or maximum likelihood (ML). However, the improved asymptotic efficiency comes at price: the GLS and ML estimators are no longer unbiased, although asymptotically the bias will vanish. Since our data set is small, we favor unbiasedness of parameter estimators over asymptotic efficiency. Moreover, as far as the test of no bias of reserves is concerned, we do not need more efficiency: more efficient estimates of the betas would only reject this hypothesis more strongly.

5.5 Reparametrization

As indicated in the previous section, the V-shaped pattern of the betas in Figure 1 suggests that it is possible to reparametrize the model by replacing $\beta(t)$ with a piecewise linear function of $t = 0,...,T = 9$, as follows:

$$\beta(t) = 1 - \frac{1}{4}(1 - \theta_1)t \quad for\ t = 0,1,...,4,$$
$$\beta(t) = \theta_1 + \frac{1}{5}(\theta_2 - \theta_1)(t - 4) \quad for\ t = 5,...,9. \tag{34}$$

Thus, $\beta(t) = 1$ for $t = 0$, $\beta(t) = \theta_1$ for $t = 4$, $\beta(t) = \theta_2$ for $t = 9$, and linear in between for $0 < t \leq 4$ and $4 < t \leq 9$.

Previously, we have assumed that the data dimensions $T = 9$ and $K = 5$ are fixed, so that the asymptotic results in Table 3 are due to letting $n = T{\times}M{\times}K \to \infty$ with $M \to \infty$. Now we can relax this assumption by allowing $T \to \infty$ as well, provided that we make an suitable assumption about how $\beta(t)$ evolves if $T \to \infty$. One option would be to assume that the kink in $\beta(t)$ occurs at time $t_1 = [(4/9)T]$, where $[x]$ is the largest integer $\leq x$, and that for all T, $\beta(T) = \theta_2$. Although such an assumption is not uncommon in econometrics, it implies that the V pattern of $\beta(t)$ is going to be stretched out as $T \to \infty$, which seems pretty implausible. An alternative assumption (which we will adopt) is that the V pattern of $\beta(t)$ will repeat itself as soon as $\beta(t) = 1$ again, at time $t = 5(1-\theta_1)/(\theta_2-\theta_1)$. Also this assumption may be implausible, of course, but we have to assume some regularity in the pattern of the $\beta(t)$'s in order to identify them. For example, if we would assume that $\beta(t)$ stays equal to 1 after $t = 5(1-\theta_1)/(\theta_2-\theta_1)$, or take a different pattern than before, then the parameters θ_1 and θ_2 are only incidental to the fixed period $0 \leq t \leq 5(1-\theta_1)/(\theta_2-\theta_1)$, and therefore asymptotically no longer identified if $T \to \infty$.

It follows from (34) that β can be written as $\beta = \zeta_0 + \zeta_1\theta_1 + \zeta_2\theta_2$, say, where

$$\zeta_0 = (1,0.75,0.5,0.25,0,0,0,0,0,0)',$$
$$\zeta_1 = (0,0.25,0.5,0.75,1,0.8,0.6,0.4,0.2,0)',$$
$$\zeta_2 = (0,0,0,0,0,0.2,0.4,0.6,0.8,1)'.$$

Then model (8), with intercept, becomes:

$$Y_j - \zeta_0' X_j = \theta_0 + \zeta_1' X_j \theta_1 + \zeta_2' X_j \theta_2 + \varepsilon_j . \tag{35}$$

The estimation and test results for this model are presented in Table 3.

Table 3: OLS estimation and test results for the reparametrized model

i	θ	H_0: θ = 1 test	p-value
1	0.807449	-8.958	0.00000
2	0.892529	-2.293	0.02184
joint test:		82.24	0.00000

intercept ($=\theta_0$):	H_0: $\theta_0 = 0$ test	p-value
θ_0	test	p-value
0.009111	5.307	0.00000

Standard error of the residuals:	0.02430
R-square:	0.91583
Adjusted R-square:	0.91520

Normality and homoskedasticity tests:	test	p-value
Jarque-Bera normality test:	173.54	0.00000
Breusch-Pagan homoskedasticity test:	22.23	0.00000

Both θ_1 and θ_2 are significantly different from 1, and the joint test that $\theta_1 = \theta_2 = 1$ strongly rejects this hypothesis. Moreover, normality and homoskedasticity of the errors are still strongly rejected.

Note that with $\theta_1 = 0.807449$ and $\theta_2 = 0.892529$, (34) predicts that $\beta(t) \approx 1$ for $t = 10$. (Actually, $\beta(10.1257) = 1$.)

Substituting the estimated θ_1 and θ_2 in $\beta = \zeta_0 + \zeta_1\theta_1 + \zeta_2\theta_2$ yields the results for the $\beta(t)$'s as functions of θ_1 and θ_2, which are presented in Table 4.

Table 4: Reparametrized betas

t	β
1985	1.00000
1986	0.95186
1987	0.90372
1988	0.85559
1989	0.80745
1990	0.82447
1991	0.84148
1992	0.85850
1993	0.87551
1994	0.89253

In order to test whether the reparametrization (34) is too restrictive or not we have conducted the ICM test in the same way as before on the basis of model (35), with results: $\hat{F}(1) = 1.14$, $\hat{F}(5) = 2.66$, and $\hat{F}(10) = 2.32$. Comparing these test results with (27), we accept, at the 10% significance level, the null hypothesis that model (35) is correctly specified.

6. Conclusions

The conclusion in Bradford and Logue (1998) was that tax laws created a relatively strong incentive to overstate reserves in 1985-1987. The tax incentive is, in general, to overstate reserves (and hence defer income), although, as emphasized by Bradford and Logue, the incentive is theoretically fairly strongly influenced by any anticipated change in the rate of tax.

The empirical results in this paper indicate that during the period 1986-1993, and possibly in 1994 as well, the insurance companies were overstating their reserves, with turning point in 1989. This turning point corresponds to the first year of the Bush (Sr.) administration. It is tempting but speculative to identify this turning point with President Bush' broken campaign pledge "Read my lips! No new taxes".

Summarizing, our empirical finding of systematic overstatement of reserves is (a) generally consistent with tax-motivated influence on reserve accounting and (b) an exciting finding anyway, inasmuch as people need to be able to interpret reserves for all kinds of purposes.

References:

Bierens, H. J. (1987), "Kernel Estimators of Regression Functions", in: Truman F.Bewley (ed.), *Advances in Econometrics: Fifth World Congress, Vol.I*, New York: Cambridge University Press, 99-144.

Bierens, H. J. (1994), *Topics in Advanced Econometrics: Estimation, Testing, and Specification of Cross-Section and Time Series Models*, Cambridge, UK: Cambridge University Press.

Bierens, H. J., and H. Pott-Buter (1990), "Specification of Engel Curves by Nonparametric Regression" (with discussion), *Econometric Reviews*, 9, 123-184.

Bierens, H. J., and W. Ploberger (1997), "Asymptotic Theory of Integrated Conditional Moment Tests", *Econometrica*, 65, 1129-1151.

Bradford, D. F., and K. D. Logue (1998), "The Influence of Income Tax Rules on Insurance Reserves", in K. Froot (ed.), *The Limited Financing of Catastrophe Risk*. Chicago: University of Chicago Press.

Breusch, T. S., and A.R. Pagan (1979), "A Simple Test for Heteroskedasticity and Random Coefficient Variation", *Econometrica*, 47, 1287-1294.

Harrington, S. E. (1988), "Price And Profits In The Liability Insurance Market", in R. E. Litan and C. Winston (ed.), *Liability: Perspectives and Policy*. Washington, D.C.: Brookings Institute.

Jarque, C.M., and A. K. Bera (1980), "Efficient Tests for Normality, Heteroskedasticity, and Serial Independence of Regression Residuals", *Economic Letters*, 6, 255-259.

Petroni, K. R. (1992), "Optimistic Reporting In The Property-Casualty Insurance Industry", *Journal of Accounting and Economics*, 15, 485-508.

Taylor, G. C. (2000), *Loss Reserving*: *An Actuarial Perspective*, Dordrecht, the Netherlands: Kluwer Academic Publishers.

Weiss, M. (1985), "A Multivariate Analysis of Loss Reserving Estimates in Property-Liability Insurance Companies", *Journal of Risk and Insurance*, 52, 199-221.

White, H. (1980), "A Heteroskedasticity-Consistent Covariance Matrix Estimator, and a

Direct Test for Heteroskedasticity", *Econometrica*, 48, 817-838.