REGULATORY SOLVENCY PREDICTION IN PROPERTY-LIABILITY INSURANCE: RISK-BASED CAPITAL, AUDIT RATIOS, AND CASH FLOW SIMULATION

By

J. David Cummins, Martin F. Grace, and Richard D. Phillips

December 16, 1997

J. David Cummins
Wharton School
3641 Locust Walk
Philadelphia, PA 19104
Phone: 215-898-5644
Fax: 215-898-0310
e-mail:
cummins@wharton.upenn.edu

Martin F. Grace
College of Business
Administration
PO Box 4036
Atlanta, GA 30302-4036
Phone: 404-651-2789
Fax: 404-651-4219
email: mgrace@gsu.edu

Richard D. Phillips
College of Business
Administration
PO Box 4036
Atlanta, GA 30302-4036
Phone: 404-651-3397
Fax: 404-651-4219
e-mail: rphillips@gsu.edu

REGULATORY SOLVENCY PREDICTION IN PROPERTY-LIABILITY INSURANCE: RISK-BASED CAPITAL, AUDIT RATIOS, AND CASH FLOW SIMULATION

ABSTRACT

This paper analyzes the accuracy of the principal models used by U.S. insurance regulators to predict insolvencies in the property-liability insurance industry and compares these models with a relatively new solvency testing approach — cash flow simulation. Specifically, we compare the risk-based capital (RBC) system introduced by the National Association of Insurance Commissioners (NAIC) in 1994, the "FAST" audit ratio system used by the NAIC, and a cash flow simulation model developed by the authors. Both the RBC and FAST systems are static, ratio-based approaches to solvency testing, whereas the cash flow simulation model implements dynamic financial analysis. Logistic regression analysis is used to test the models for a large sample of solvent and insolvent property-liability insurers, using data from the years 1990-1992 to predict insolvencies over three-year prediction horizons. We find that the FAST system dominates RBC as a static method for predicting insurer insolvencies. Further, we find the cash flow simulation variables add significant explanatory power to the regressions and lead to more accurate solvency prediction than the ratio-based models taken alone.

1. Introduction*

Increases in the frequency and severity of insurer insolvencies in the mid-1980s led to concern about the adequacy of state insurance regulation and the accuracy of the methods used by regulators to provide early warning of insurer insolvencies.[1] The National Association of Insurance Commissioners (NAIC) responded by adopting a "solvency policing agenda" in 1989. The agenda resulted in a number of changes in state solvency regulation including the adoption of the Financial Analysis and Surveillance Tracking (FAST) solvency monitoring system and risk-based capital (RBC) requirements for both life and property-liability insurers.[2] FAST was implemented in 1993, and the property-liability insurance RBC system went into effect in 1994.

Well-designed solvency monitoring systems should identify a high proportion of troubled companies early enough to permit regulators to take prompt corrective action and should minimize the number of financially sound insurers that are identified as being troubled. Earlier research has called into question the effectiveness of the NAIC's RBC system in accomplishing these objectives. Grace, Harrington, and Klein (1993) (GHK) find that, although the ratio of actual capital to RBC is negatively

---

*The authors would like to thank Richard Derrig and the participants of the 5th International Conference on Insurance Solvency and Finance for their helpful comments. In addition, the authors are grateful to Dr. Charles Metz of the University of Chicago for making his receiver operating characteristic (ROC) software available to us.

[1]State regulators were criticized for insufficient solvency monitoring and exercising regulatory forbearance (see U.S. House of Representatives, 1990, U.S., General Accounting Office, 1991). The criticisms led to proposals for federal insurance solvency regulation. Causes of the worsening insolvency experience include unexpected growth in claim costs and interest rate volatility in the early and mid-1980s, as well as moral hazard induced by risk-insensitive guaranty fund assessments (Cummins, Harrington, and Niehaus, 1993, A.M. Best Company, 1990).

[2]For more information on the FAST system, see Klein (1995). An older system, the Insurance Regulatory Information System (IRIS), which tests insurer solvency based on twelve audit ratios, is also still in use by the NAIC. Although not all of the IRIS ratios appear in precisely the same form in the FAST system, FAST can be considered a super-set that encompasses nearly all of the information conveyed by IRIS. Hence, the analysis in this paper focuses on the FAST ratios.

and significantly related to the probability of subsequent failure, relatively few companies that later failed had ratios of actual capital to RBC within the NAIC's ranges for regulatory action. Cummins, Harrington, and Klein (1995) (CHK) confirm that the predictive accuracy of the RBC ratio is very low, even when the components of the ratio, rather than the overall ratio, are used as predictors.[3]

The only prior tests of the FAST system were performed by GHK (1993, 1995). They test the overall FAST score, a univariate summary statistic compiled by the NAIC based on the approximately thirty-one financial ratios comprising the FAST system. The NAIC assigns scores corresponding to a company's ratios based on a subject evaluation of the importance of the ratios and their relationship to solvency, and the scores are summed to obtain the company's overall FAST score. Financial strength is considered to be inversely related to the overall FAST score. In their tests, the overall FAST score performs considerably better than RBC in predicting insolvencies, and the addition of the RBC ratio to the FAST-ratio prediction models leads to only modest improvements in predictive accuracy.

A limitation of both the RBC and FAST systems is that they are based on a "snapshot" of the firm at a given point in time, i.e., they are static rather than dynamic approaches to solvency testing (Cummins, Harrington, and Niehaus, 1993, 1995). The more modern approach to solvency testing is *dynamic financial analysis* (DFA), usually implemented using cash flow simulation.[4] DFA has the ability to capture information that is not utilized in the static systems and thus to lead to more accurate solvency prediction. For example, a cash flow model can take into account patterns of loss reserve runoffs and asset cash flows and can incorporate external economic information such as yield curves and inflation

[3]CHK also report that predictive power can be significantly improved by adding controls for insurer size and organizational form.

[4]For further discussion of the cash flow simulation approach, see Casualty Actuarial Society (1996) and Hodes, et al. (1996).

rates. Thus, DFA can provide information on a company's ability to withstand potentially adverse economic developments that cannot be captured by a static system.

The present paper extends the existing research on regulatory solvency prediction in two major ways: The first major contribution is to test the accuracy of a DFA model in predicting insurer insolvencies using a cash flow simulation model developed by the authors.[5] Variables based on cash flow simulations are tested by themselves and in combination with the RBC and FAST scores to measure the potential incremental power of DFA.

The second major contribution of the paper is motivated by a limitation of GHK's analysis that is likely to have produced an upward bias in their estimates of the predictive accuracy of the FAST system. The problem arises because the FAST scores used in their tests were optimized by the NAIC to accurately predict the insolvencies that actually occurred in 1993, while also considering information on insolvencies occurring in 1991 and 1992. In addition, the NAIC changed ratios over their sample period based on the characteristics of the insolvencies that occurred in each year, and GHK use the set of ratios chosen by the NAIC in 1993. Thus, the variables used in their tests incorporate information that would not have been known in 1989, 1990, and 1991, the three base years for their solvency tests.[6] Although the degree of bias is difficult to determine without further analysis, it seems clear that GHK did not perform a true ex ante test of the FAST system's predictive accuracy. We perform an ex ante test of FAST in this paper by measuring the predictive accuracy of the set of nineteen FAST ratios that were

---

[5]The modeling project was sponsored by the Alliance of American Insurers. The authors developed the model in collaboration with Douglas Hodes and Shalom Feldblum of Liberty Mutual Insurance Group. Although cash flow simulation has not been used by the NAIC in solvency testing, the cash flow simulation approach has been proposed as a regulatory system by industry groups and implemented for solvency testing by the New York Insurance Department for certain life insurance products.

[6]GHK use data from each base year to forecast insolvencies over a three-year horizon, e.g., the 1989 data were used to predict insolvencies occurring in 1990-1992.

used *consistently* throughout our sample period. We compare the predictive performance of the consistent FAST ratios with the FAST variables used by GHK (1993, 1995), i.e., the overall FAST score and the full set of thirty-one ratios, both optimized by the NAIC for 1993. Thus, our tests avoid any bias caused by the 1993 optimization problem and the NAIC's tendency to change the ratios after the fact to increase predictive accuracy.

An additional contribution of our paper is to introduce to the insurance insolvency literature a new approach to comparing the performance of the alternative solvency monitoring models — receiver operating characteristic (ROC) analysis. ROC analysis has its roots in engineering (Petersen et al., 1954) and is the generally accepted methodology for evaluating diagnostic performance of competing models in such fields as psychometrics, medical imaging, and weather forecasting (Swets, 1996). We use ROC analysis in this paper to determine whether the cash flow variables add significantly to RBC's and FAST's ability to predict the solvency of insurers.

The sample of firms used in our analysis consists of all property-liability insurers that became insolvent during the period 1991 through 1995 that were reported to the NAIC as well as a sample of solvent property-liability insurers.[7] The predictions are conducted using financial statement data from three years — 1990, 1991, and 1992 — with three-year prediction horizons.

By way of preview, our results confirm the GHK (1993, 1995) and CHK (1995) finding that the risk-based capital formula and its components are not very effective in predicting insolvencies. We find that predictive accuracy is significantly improved through the use of variables based on the FAST system. As expected, the GHK (1993, 1995) FAST variables perform better than the set of FAST ratios used

[7] All insurance insolvencies of any meaningful size are reported to the NAIC by the state insurance commissioners.

consistently throughout the sample period. The cash flow simulation results add significant incremental explanatory power to the RBC and FAST variables taken alone or in combination.

The remainder of the paper is organized as follows: In section 2 we describe the NAIC's RBC and FAST systems and our cash flow simulation model. Section 3 presents the methodology and discusses our data base. The results are presented in section 4, and section 5 concludes.

2. The Solvency Prediction Methods

The NAIC Property-Liability Risk-Based Capital System

The NAIC property-liability risk-based capital system consists of a series of ratios that are multiplied by various balance sheet and income statement variables to compute RBC "charges" for the principal risks facing insurers. The sum of the charges, reduced by a covariance adjustment, equals the insurer's risk-based capital. The insurer's actual capital is divided by its risk-based capital to obtain the RBC ratio, and regulatory action is prescribed for insurers whose RBC ratios fall below specified thresholds (see below).

The RBC formula assesses charges for four major types of risks — asset risk, credit risk, underwriting risk, and growth and other forms of off-balance sheet risk (see NAIC, 1993, and Cummins, Harrington, and Niehaus, 1995, for more details). The provision for underwriting risk applies separate risk factors to loss and loss adjustment expense reserves and to net premiums written for each line of business. We refer to the resulting RBC charges as "loss reserve RBC" and "written premium RBC," respectively. The loss reserve charges accounted for about 41 percent of total industry RBC in 1992, while written premium RBC accounted for about 27 percent. The loss reserve and written premium risk factors are based on regulatory judgment and analysis of the industry's worst accident year development and worst accident year loss ratio over the previous ten years.[8] The factors are discounted to present value to reflect the fact that reserves are generally reported on a non-discounted basis. Charges are based on a weighted average of industry and company experience so that an insurer with worse than average loss development and loss ratios will have risk factors higher than the industry factors.

Asset charges ("investment RBC") accounted for about 21 percent of the total property-liability insurer RBC in 1992. The bond and preferred stock (for non-affiliates) factors are based on NAIC valuation categories, which generally parallel Moody's and Standard & Poor's. The bond factors range from 0 for Treasury bonds to 30 percent for bonds in or near default. They are adjusted upward (downward) if the number of issuers reflected in its bond portfolio is less (more) than 1,300 to reflect the diversification of credit risk across issuers (Klein, 1995). There is a 15 percent charge for common stocks of non-affiliated corporations. An asset concentration factor increases the RBC charges for the 10 largest asset exposures grouped by issuer.

The credit component of the formula ("credit RBC") applies a 10 percent charge to reinsurance recoverable from non-affiliates and affiliated alien insurers and smaller charges to various other receivables. Additional RBC charges are given to insurers with three-year average growth in gross premiums written in excess of 10 percent and for off-balance sheet liabilities. We refer to these two components as "growth RBC." In 1992, credit RBC accounted for 10 percent of total industry RBC and

---

[8]Accident year development" refers to the ratio of developed (i.e., estimate of ultimate) incurred losses and allocated loss adjustment expenses evaluated at the current year to the sum of the initial evaluations of these incurred losses and allocated loss adjustment expenses. Positive development indicates that initial estimates of ultimate losses were too low. "Accident year loss ratio" refers to the ratio of developed incurred losses to net premiums earned. Under "accident year" reporting, all losses are assigned to the year in which the event occurred that triggered coverage (e.g., date of an accident).

growth RBC accounted for 1 percent of the total. The sum of the RBC charges is reduced through a covariance adjustment to reflect the effects of diversification across risk factors.[9]

The RBC system requires regulators to take specified actions if an insurer's actual capital falls below certain thresholds. The "authorized control level" (ACL), which is equal to the final RBC formula result, is used as the primary point of reference. Other levels are calculated as percentages of the ACL:

(1) Company Action Level. An insurer with capital below 200 percent of the ACL must file a plan with the insurance commissioner that explains its financial condition and how it proposes to correct its deficiency.

(2) Regulatory Action Level. When an insurer's capital falls below 150 percent of the ACL, the commissioner is required to examine the insurer and institute corrective action, if necessary.

(3) Authorized Control Level. If an insurer's capital falls below 100 percent of its ACL, the commissioner has the legal grounds to rehabilitate or liquidate the company.

(4) Mandatory Control Level. If capital is less than 70 percent of the ACL, the insurance commissioner is required to seize the company.

The accuracy of the RBC system is of great importance not only to avoid costs to the guaranty fund system arising from insolvencies that are not identified in time to avoid large deficits but also to avoid imposing unnecessary regulatory costs on financially sound insurers.

The FAST System

The NAIC's financial analysis and surveillance tracking (FAST) system and the older insurance regulatory information system (IRIS) were designed to prioritize insurers for further regulatory action. The IRIS system consists of twelve audit ratios with published ranges that were deemed acceptable by the regulators. The FAST system consists of approximately thirty ratios and corresponding scores for each ratio (Klein, 1995). The ultimate output from the FAST system is the overall FAST score equal to

---

[9]For 1994, the RBC charge obtained from application of the covariance adjustment formula was multiplied by 0.4 to calculate the benchmark RBC levels (the final formula result) reported in insurer annual statements. This scale factor was increased to 0.45 in 1995 and to 0.5 for years after 1995.

the sum of the individual insurer's audit ratios multiplied by the corresponding scores. Companies performing poorly in terms of the IRIS and FAST test results are given a higher priority by regulators in deciding upon subsequent regulatory attention.

The FAST system was introduced in part as a result of the allegation that insurers were able to "game" the IRIS system because it is based on only a few ratios, for which the regulatory action cutoffs are specified in advance and rarely changed (Klein, 1995). In contrast to the IRIS system, the FAST scores are not revealed by the NAIC, and both the ratios and the scores change over time as new information becomes available. Thus, the FAST system is expected to provide more accurate solvency predictions than the IRIS system. Even though not all of the IRIS ratios appear in the FAST system in precisely the same form, nearly all of the relevant information captured by IRIS is also incorporated in FAST, and FAST captures a significant amount of information not reflected in IRIS (Grace, Harrington, and Klein, 1995). Accordingly, in this paper we focus on the FAST system rather than the IRIS system.

In Grace, Harrington, and Klein (1995), FAST was tested against alternative specifications and with additional scoring methods. After an exhaustive investigation, the authors concluded that changes in the scoring methodology and other alternative specifications did not lead to better predictions than a logistic regression model based solely on the FAST ratios and other firm characteristic variables (such as total assets and a mutual versus stock dummy variable). Thus, the authors concluded that there are diminishing returns to examining additional audit ratios based on financial statement data and that other approaches that add new types of information to solvency analysis, such as cash flow simulation, should be explored. However, as mentioned above, their tests are subject to potential bias because the scores and ratios they used were modified after the fact by the NAIC and thus contain information that would not have been known in a true ex ante test of predictive accuracy. Our methodology corrects this problem, as explained below.

The Cash Flow Simulation Model

In contrast to the NAIC's RBC and FAST systems, cash flow simulation is a dynamic approach to solvency testing. Rather than evaluating the financial health of an insurer using a snapshot of the company's financial condition at a point in time, the cash flow simulation approach projects the company's financial condition over a period of time under alternative economic scenarios. This section provides a brief overview of the cash flow simulation model. For more details, the reader is referred to Cummins and Phillips (1994).

General Principles. Cash flow simulation is based on the fundamental principle of asset pricing that the value of any asset is determined by its cash flows. A pro-forma cash flow statement for the property-liability insurance industry is shown in Table 1. The principal cash inflows are premiums and investment income, and the principal cash outflows are losses, expenses, and Federal tax payments. The net cash flow reflects the industry's net cash position for the year. Our model utilizes the same basic cash flows but is much more complex, containing separate modules for eighteen lines of insurance and several categories of assets.

The starting point for the simulation is the company's financial position at the end of a specified year (1990, 1991, or 1992 in this study), as reflected in its balance sheet, income statement, and other financial accounts. The cash inflows and outflows implied by the company's beginning financial condition are then simulated over a twenty-year time horizon. At the end of the twenty-year projection period, the present values of the company's remaining cash flows are computed and added to its net resources at the end of the period. If the net resources are positive, the company is considered to have survived, but if the net resources are zero or negative, it is considered to have become insolvent. Companies are classified as solvent or insolvent on the basis of this test.

The model utilized in this paper is intended as a practical tool for regulatory solvency testing rather than a managerial decision making model. Because the regulator is primarily concerned with whether the company's current resources are adequate to pay its obligations, the model is designed as a **runoff** model rather than a **going concern** model. The objective of developing a simulation model that could be used in practical regulatory solvency testing also drives other important characteristics of the model: (1) The data used by the model are from the company's regulatory annual statement, which can be read into the program automatically from the computer diskettes filed by the insurers with the NAIC.[10] (2) The model utilizes a non-stochastic, scenario-testing approach rather than stochastically simulating investment, loss, and other important flows. Seven scenarios are programmed into the model, which can be changed by an advanced user. The scenarios include a baseline scenario, where cash flows are based on expected values, and six progressively more adverse scenarios, involving elements such as higher than expected losses, adverse reserve development, and lower investment income. The reason for choosing a scenario testing rather than a stochastic approach is that accurate stochastic modeling requires careful estimation of probability distributions, which are likely to vary by company. Such a detailed analysis would not be feasible in a regulatory context where approximately two thousand companies are to be analyzed within a few months time.[11] The model contains separate modules for each of eighteen lines of insurance as well as various categories of investments and other insurer accounts.[12] The remainder of this section discussion provides brief descriptions of the premium, loss, and investment modules.

---

[10]The model is designed to operate on a personal computer using the Microsoft Excel® spreadsheet program can be run automatically with little or no operator intervention.

[11]The scenario approach is also much easier to explain to non-technical users and thus would be more likely than the stochastic approach to gain widespread acceptance among regulatory personnel.

[12]For more details on the model see Cummins and Phillips (1994).

In order to provide a robust test of the simulation model, we chose not to optimize the model to perform well in predicting the insolvent firms in our sample. The scenarios were designed a priori based on actuarial and financial theory rather than through an analysis of financial statement data. Thus, any incremental predictive power provided by the cash flow model probably could be improved upon by adopting a more aggressive modeling strategy.

The Premium Module. In keeping with the runoff approach, the cash flows are simulated on the assumption that the company does not continue to write new business over most of the simulation period. However, the company is assumed to write new business for eighteen months following the starting date of the simulation, reflecting the amount of business that would be written prior to the next audit following the financial statement used to begin the simulation.[13] New premium writings are projected as the company's current premium writings increased by a growth rate equal to the company's average annual premium growth over the prior five years.[14]

The Loss Modules. The loss modules are used to generate the company's loss cash flows from each of eighteen lines of insurance. The input data for the loss module are taken from Schedule P of the insurer's regulatory annual statement, which provides information on the history of the company's loss reserves. The loss cash flows are estimated by multiplying the company's loss reserves as of the starting date of the model by a sequence of loss cash flow factors projecting the amount of the loss reserve paid out over each of the next twenty years.

[13]That is, the company is assumed to be solvent as of the date of the financial statements used to start the simulation. We assume that the company will be audited again eighteen months following the starting date of the simulation, because NAIC regulatory statements are filed annually (in March) and it takes the NAIC several months to analyze the data.

[14]This discussion oversimplifies the treatment of premium and expense flows in the interests of brevity. For more details, see Cummins and Phillips (1994).

Two methods are used to estimate the loss payout proportions for each line of business — a **premium-based** method and a **loss-based** method. The final payout proportions are based on a weighted average of the factors generated by the two methods.[15] The premium-based method involves computing the proportion of incremental paid losses relative to earned premiums. Define $C_{y,d}$ to be the cumulative paid losses and allocated expenses for accident year y paid after d years of development. Define P_y to be the net earned premium for accident year y. Define $g_{y,d}$ to be the proportion of accident year y premiums paid out as losses during development year d. Then the incremental payout proportions under the premium-based method, $g_{y,d}$, are equal to:

$$g_{y,d} = \frac{C_{y,d}}{P_y} \quad \text{for } d = 1 \text{ , and } g_{y,d} = \frac{C_{y,d+1} - C_{y,d}}{P_y} \text{ , for } d > 1 \text{ .} \tag{1}$$

Once the $g_{y,d}$'s have been estimated, the model computes the average amount of the net earned premium that is paid out as losses in each development year d over all the accident years. The payout proportion for development year d is the average of the $g_{y,d}$ or $g_{\bullet,d}$, where

$$g_{\bullet,d} = \frac{1}{N-d+1} \sum_{y=s-N+1}^{s-d+1} g_{y,d} \quad \text{for } d \in [1,N]. \tag{2}$$

where N = the number of years of available data, and s = the final year for which data are available.

The loss-based method for estimating loss payout proportions computes the proportion of the remaining reserve for accident year y as of development year d that is paid during development year d+1. Define $R_{y,d}$ to be the total reported incurred losses and loss adjustment expenses for accident year y reported at the end of development year d. Define $R_{y,d}$ to be the proportion of the age y+d reserve paid

---

[15] The program gives more weight to the premium-based method in early development years and more weight to the loss-based method in the later development years when the book of business is more mature and therefore loss reserve estimates are more accurate.

during year d+1. Then, using the notation developed for the premium-based method, the incremental loss payout proportions are equal to:

$$\ell_{y,d} = \frac{C_{y,d+1} - C_{y,d}}{R_{y,d} - C_{y,d}}. \tag{3}$$

As with the premium based method, once the $\ell_{y,d}$s have been estimated, the model computes the average incremental loss payout proportion $\ell_{\bullet,d}$ as follows:

$$\ell_{\bullet,d} = \frac{1}{N-d+1} \sum_{y=s-N+1}^{s-d+1} \ell_{y,d} \quad for \quad d \in [1,9]. \tag{4}$$

The final loss-based payout proportions used in the model are equal to a weighted average of the company's and the industry's loss-based payout proportions, and the final premium-based payout proportions are obtained similarly. The weighting factor in the model has been set equal to ½.

The model utilizes the premium-based and loss-based payout proportions to simulate the loss cash flows attributable to each accident year. Define $r_{y,d}$ to be the amount of the reserve remaining to be paid out as losses for accident year y at the beginning of year y+d, i.e., it is the amount of accident year y's reserve left to be paid after d development years and is equal to

$$r_{y,d} = R_{y,d} - C_{y,d}. \tag{5}$$

Define $L_{y,d}$ to be the estimated loss cash flow from accident year y in development year d. Then $L_{y,d}$ is equal to

$$L_{y,d} = w_d \times g_{\bullet,d} \times P_y + (1 - w_d) \times l_{\bullet,d} \times r_{y,d-1}, \tag{6}$$

The weight w_d, $0 \leq w_d \leq 1$, is an industry-wide estimate of the proportion of total losses unpaid by the end of development year d.

Property-liability insurers face the risk that ultimate loss payments will exceed current loss reserves. This type of risk, known as reserving risk, is incorporated into the model through the scenario definitions. The baseline scenario assumes that a given company's reserves are understated by a percentage based on the company's prior loss reserve development and the industry's prior loss reserve development, where loss development is defined as the ratio of the estimated losses incurred for a given accident year as of the most recent reporting date to the estimated losses incurred as of the initial reporting date for that accident year, minus 1. The moderately and severely adverse scenarios assume that the company's loss reserves have been understated by the weighted average adverse development factor plus 1 and 2 loss development standard deviations respectively.[16]

The model distinguishes between lines of business that pay nearly all of their losses in the first three years of development (short-tail lines), the first ten years of development (intermediate-tail lines), and lines of business where significant payments still remain to be made after ten development years (long-tail lines). For short-tail (intermediate-tail) lines all losses not paid out by the end of the third (tenth) development year are assumed to be paid out in the fourth (eleventh) year and no further loss cash flows are generated from these lines after the fourth (eleventh) development year.

Long-tail lines of business continue to generate loss cash flows for the entire twenty-year simulation period. The estimates of the future loss payments for the first 10 development years are generated as described above. However, because Schedule P contains only ten years of loss development history, starting with development year 11 the model estimates loss payments by fitting an inverse power

[16]As an element of conservatism, no credit for over-reserving is given when estimating the average adverse development for the company or the industry. The relevant adverse reserve development percentage(s) is (are) set equal to 0 for the baseline scenario, and underreserving in the moderately and severely adverse scenarios is reflected by increasing the company's stated loss reserves by one and two standard deviations of the reserve development factor.

curve to the first 10 years of loss payments and then extrapolating the curve to determine the loss payments for the next ten years. In development year 21, a lump sum payment is made to pay off any reserves that may remain (for the details, see Cummins and Phillips, 1994). Once the model has estimated the loss cash flows for each line of business, the flows from the individual lines are added to determine the total loss payments made by the company for each of the next twenty years.

The Investment Module. The investment module starts with the company's current asset holdings and estimates the cash flows that will be generated from these investments over the twenty-year projection period. The starting value of the equity portfolio is equal to the market value of equities reported in the company's annual statement. The annual statement gives data on the bonds held by insurers categorized by quality (default risk) and maturity. The bonds are treated by the model as cash flow vectors, with the cash flows generated by any given bond equal to its coupon payments and projected payment of principal, subject to adjustments for default risk and interest rate risk, discussed below. Insurers report bond maturities in the NAIC regulatory annual statement in multi-year bands, and we assume that bonds in a given maturity band mature uniformly over the years covered by the band.[17]

The model simulates stock price risk by accumulating the company's initial stock holdings using capital gains and dividend accumulation factors. Define S_i to be the market value of the equities held by the firm, where i ranges from 0 (the starting year of the simulation) to 20. Stock price risk is simulated by varying the vectors of capital gains and dividends used to accumulate the value of the insurer's stock holdings over the projection period. Define CAP_i to be the amount of the capital gain for simulation year

[17]Insurers report the value of bonds maturing in less than 1 year, greater than or equal to 1 year but less than 5 years, greater than or equal to 5 years but less than 10 years, greater than or equal to 10 years but less than 20 years, and in 20 or more years.

i. Define $CG_{i,x}$ to be the capital gain factor used to generate year i's capital gains. Then, the formula used to determine the amount of capital gains for any year is

$$CAP_i = S_{i-1} \times CG_{i,x} \qquad for\ i \in [1,20]. \tag{7}$$

The baseline scenario assumes that all capital gains factors are equal to the long-term historical average (1926-1990) capital gain factors reported by Ibbotson Associates (1995) for the Standard & Poor's 500 Stock Composite Index. The moderately adverse and severely adverse scenarios shock the equity portfolio of the firm by supposing the company suffers a large capital loss in the first year, followed by average capital gains in the subsequent years. The capital gain factors for the first year for the moderately and severely adverse scenarios are equal to the average capital gain factor minus 1 and 2 standard deviations, respectively. Other capital gains assumptions could be used to generate alternative scenarios.

The design of the dividend factors is very similar to the capital gains factors. However, dividends are assumed to be received currently as investment income cash flows, whereas capital gains are held on the books until the insurer needs to sell assets (see below). The baseline scenario assumes the company's equity portfolio receives dividends equal to the long-term average dividend yield on the Standard & Poor's 500 Stock Composite Index for the period 1926-1990. The moderately and severely adverse scenarios assumes the first year dividend rate is one and two standard deviations below the average long-term dividend yield, respectively.

The model incorporates bond default risk using the bond mortality loss tables presented in Altman (1992). The bond mortality loss rates measure the estimated proportion of total book value that is lost due to default in each year of the simulation period. The scheduled bond cash flows for each payment year and quality class are proportionately reduced by the mortality charges. The mortality charges are higher for bonds in lower NAIC quality classes. The same mortality rates are applied to both corporate and state/municipal bonds. The baseline scenario assumes the bond mortality charges are equal to Altman's average bond mortality rates for each bond quality class. The moderately and severely adverse scenarios assume the bond mortality loss rates are equal to the Altman averages plus 1 and 2 standard deviations, respectively.

In addition to the risk of bond default, insurers are subject to fluctuations in bond market values due to unexpected changes in interest rates. The model's bond cash flows consist of coupon and principal payments. If the insurer does not have to sell bonds, these cash flows will not be affected by fluctuations in interest rates. Accordingly, as a simplifying assumption, the model assumes that stocks are sold to meet loss payments until the stock portfolio has been exhausted. At that point, further uncovered loss cash flows must be met by selling bonds. To determine the proceeds from bond sales, the market values of the bonds must be determined. This is done by maturity and quality class using estimates of the market yield rates for bonds in various quality and maturity classes. The bonds' cash flows are discounted using the estimated yield rates to obtain market values.[18] If the bonds must be sold, the sale proceeds are equal to the estimated market values.

Interest rate risk is incorporated in the simulation scenarios by allowing for yield curve shifts. The default version of the model incorporates a flat yield curve, and adverse yield curve shifts reflect parallel changes in the yield curve to a higher yield.[19] The effect of a yield curve shift is to alter the interest rate at which bond cash flows are discounted to obtain bond market values. Define r_f to be the risk-free interest rate at the beginning of the simulation, i.e., simulation year 0. Define $s_{i,x}$ to be the amount the term-structure shifts in time period i relative to time 0 for interest rate scenario x, x e [1, 7]. Define d_c

[18]The interest rates used to discount bond cash flows to obtain market values are estimated as the risk-free rate of interest (one year Treasury bill rate) plus a default risk premium.

[19]The model is sufficiently general to allow for yield curves that are not flat and non-parallel yield curve shifts.

to be the default risk premium for bonds of quality class c, where $c \in [g, 1, \ldots, 6]$. Finally, define $r_{i,c,x}$ to be the yield to maturity for bonds of class c in simulation year i under scenario x. Then $r_{i,c,x}$ is equal to

$$r_{i,c,x} = r_f + d_c + s_{i,x}. \tag{8}$$

These yield to maturity rates are used to determine the market value of the bond through time.

Other Components of the Model. In addition to the primary modules discussed above, the model also recognizes risks from other insurer cash flows. Like the NAIC's RBC formula, the cash flow model incorporates credit risk — the risk that the insurer will not be able to collect the full amount of receivables owed by agents and reinsurers. The model allows for the possibility that the agents and/or the reinsurers will default on payments due to the insurer, and the default charges are varied by scenario. The default charges are applied against cash flows anticipated from agents and reinsurers in each simulation period. In addition to the usual balances due from agents, the model also incorporates a module for accrued retrospective premiums.[20] Insurers with substantial amounts of retrospective business may face significant risk that the retrospective charges will not be paid. The model allows for this by making charges against retrospective premium flows that are varied by scenario.

Net Cash Flow. The net cash flow for simulation year i, NCF_i, is the sum of the cash inflows and outflows of the company. The principal cash outflows are loss and expense payments. The principal cash inflows consist of premiums and investment cash flows, including bond coupon payments and proceeds from asset sales and maturities, payments received from agents and reinsurers, cash flows from accrued retrospective premiums, and miscellaneous flows. The net cash flow for simulation year i is equal to:

[20] Retrospective premiums are additional premiums owed to the insurer by policyholders under retrospectively rated policies, which adjust premiums after the coverage period has ended to reflect the policyholder's actual losses during the period, usually subject to a maximum and minimum premium.

$$NCF_i = P_i + ACF_i + RRCF_i + DIV_i + CPN_i + MAT_i + SALE_i - L_i - EX_i \pm MS_i \tag{9}$$

where P_i = total premiums collected by the company for simulation year i,

ACF_i = total agent's balances and retrospective premiums collected during simulation year i,

$RRCF_i$ = total reinsurance recoverables collected during simulation year i,

DIV_i = total dividends received by the company on its equity portfolio for simulation year i,

CPN_i = total bond coupon payments received by the company in simulation year i,

MAT_i = total bond principal payments received by the company for simulation year i,

$SALE_i$ = proceeds from asset sales in year i,

L_i = total losses paid by the company for simulation year i

EX_i = total expenses paid by the company for simulation year i, and

MS_i = miscellaneous cash flows for simulation year i.

Because asset sales are used to offset otherwise uncovered cash outflows, a negative net cash flow is almost always equivalent to insolvency.

3. Methodology and Data

The Logistic Regression Methodology

We compare the accuracy of the NAIC's risk-based capital formula, the FAST system, and the cash flow simulation model by estimating logistic regression models and use the results to compare the predictive abilities of the various solvency assessment technologies. The logistic regression model takes the following general form:

$$y_{it}^{*} = \alpha_0 + \alpha_1 \frac{RBC_{1jt}}{S_{jt}} + \ldots + \alpha_k \frac{RBC_{kjt}}{S_{jt}} + \beta_1 F_{1jt} + \ldots + \beta_n F_{njt}$$
$$+ \gamma_1 C_{1jt} + \ldots + \gamma_m C_{mjt} + \delta_1 X_{1jt} + \ldots + \delta_r X_{rjt} + \epsilon_{jt} \qquad \textbf{(10)}$$

where RBC_{ijt} = risk-based capital variable i for insurer j in year t,

S_{jt} = the actual surplus of insurer j in year t,[21]

F_{ijt} = FAST variable i for insurer j in year t,

C_{ijt} = cash flow variable i for insurer j in year t,

X_{ijt} = control variable i for insurer j in year t, and

e_{jt} = a random error term, assumed to follow a logistic distribution.

The dependent variable y_{jt}^{*} is the propensity for the insurer to fail subsequent to year t. We do not observe y_{jt}^{*} but instead observe $y_{jt} = 1$ if the insurer fails and $y_{jt} = 0$ if it remains solvent. The model is estimated using the method of maximum likelihood. Because the number of insolvent firms was not large enough to use a hold-out sample, we estimated the model using an approximate jackknife procedure as a control for within-sample prediction bias.[22]

Specification of Independent Variables

With respect to the risk-based capital variables (the RBC_{ijt} in equation (10)), we follow CHK (1995) in testing both total risk-based capital and its major components as possible predictors of insurer insolvency. The decomposition of the overall ratio leads to five variables — the ratio to actual surplus

[21]Consistent with the NAIC formula, the surplus variable we use here is the insurer’s "adjusted surplus," i.e., its reported surplus reduced by any deductions from reported reserves due to reserve discounting.

[22]The procedure uses a one-step approximation to the coefficient vector that would be obtained if the observation were excluded from the sample (see Pregibon, 1981).

of the risk based capital charges for asset risk, credit risk, loss reserve risk, written premium risk, and growth and other off-balance sheet risk. Using the ratios of the RBC components to surplus facilitates the decomposition.

The specification of FAST variables is somewhat more problematical than the other variables tested in our research. We identify two problems with the use of the FAST results. The first problem causes a bias and the second causes multicollinearity.

Recall that the FAST system is a series of financial ratios and accompanying scores for ratios falling in various ranges. Each company is given a score for each of its ratios depending upon the ratio's value. The scores are then summed to give the overall FAST score for the company. Companies with *high* overall FAST scores are given the highest priority for regulatory scrutiny.

The first problem we encountered in testing the FAST system can be called "look-ahead bias" (see GHK, 1993). The look-ahead bias problem arises because the NAIC chooses ratios to include in the system (the ratios vary somewhat by year) and assigns scores to the ratios so that the system performs well in "predicting" observed insolvencies after the fact. Specifically, the NAIC scores and ratios available for our study (and also for GHK, 1993, 1995) were chosen to give the highest scores to the companies that actually failed during 1993, with consideration also given to the insolvencies occurring in 1991 and 1992, i.e., the FAST scores and ratios were specified *after* the regulators observed the outcome of the 1991-1993 experience. As a result, the FAST variables available for this study are biased towards accuracy in tests of insolvency prediction for years prior to 1994 because they incorporate information that was not available ex ante in these years. Using this ex post information would not provide an accurate comparison between the alternative solvency prediction models since both the risk-based capital and the cash flow simulation results are based solely on ex ante information.[23]

To correct for the bias in the reported FAST ratios and scores, we use as predictors the nineteen FAST ratios (not scores) common to the FAST system during the entire period under investigation. (This is a subset of the thirty-one ratios included in the 1993 FAST system.) We refer to these ratios as the *consistent FAST ratios.* Because the ratios in this set were not modified after the fact in light of observed insolvency experience, this set of ratios does not suffer from the look-ahead bias problem.[24] For purposes of comparison, we also predict insolvencies using the FAST ratios and the overall FAST score based the NAIC's 1993 optimization of the system.[25] We refer to the 1993 ratios as the *ex post FAST ratios.* The FAST score and ex post FAST ratios are the same variables used by GHK (1993, 1995).

The second methodological complication presented by the FAST ratios is that the number of ratios is quite large. Including all of the variables in our logistic regressions would result in an unmanageable degree of multicollinearity, and including arbitrary combinations of the variables is likely to understate the accuracy of the FAST ratios if the omitted ratios convey information about insurer financial condition. To solve this problem, we conduct a factor analysis on the FAST ratios and use as

[23]Note that it would be appropriate to use the scores and ratios optimized to a given year's insolvencies in predicting insolvencies for *subsequent* years. We could not conduct this type of prediction exercise because the scores and ratios provided by the NAIC for use in this study were based on 1993. GHK (1993, 1995) faced the same limitation.

[24]We are comfortable using the ratios alone as Grace, Harrington, and Klein (1995) conclude that adding the scores to the analysis does not significantly improve the logistic regression model based solely on the FAST ratios and other firm characteristic variables.

[25]Neither the individual ex post FAST scores nor the overall FAST score optimized for other years were available in the data base provided to us by the NAIC.

variables those factors with eigenvalues greater than one.[26] We employ a varimax rotation to orthogonalize the factors, eliminating the problem of multicollinearity among the factors.

We test several variables based on the cash flow simulation model. The most straightforward variable is the ratio of predicted surplus at the end of the simulation period relative to the level of surplus at the beginning of the simulation. We expect this variable to be inversely related to insolvency.

The second cash flow variable is the logarithm of the estimated time to failure. This is defined as the actual time to failure for insurers that are predicted to fail by the model under the baseline scenario, meaning that their simulated resources become negative with claims still outstanding sometime prior to or including simulation year twenty. For those companies having positive surplus after 20 years, we use a hazard rate model to estimate the expected time to failure, as follows:

$$E(t|t>20) = \int_{20}^{\infty} t \, \frac{h(t)}{1 - H(20)} \, dt \qquad \textbf{(11)}$$

where $E(t|t > 20)$ = the expected time to failure conditional on the insurer having survived to time 20,

h(t) = the density function of time to failure, and

$H(t)$ = the distribution function of the time to failure.

The hazard rate model is estimated using a lognormal time-to-failure distribution with two exogenous covariates: a size variable equal to the natural logarithm of the insurer's assets and a mutual/stock organizational form dummy variable. We expect the time to failure variable to be inversely related to the probability of insolvency.

---

[26]The set of factors included in the final version of the logistic models is further reduced by eliminating factors that were not statistically significant in earlier runs. Factors were eliminated one at a time with the least significant eliminated first.

Both the simulated ending-to-beginning surplus ratio and the time to failure variable incorporate information only from the baseline scenario. To capture information conveyed by the other six more adverse scenarios, we conduct a factor analysis on the ratios of predicted ending surplus to initial surplus and on the logarithm of the time to failure estimates under all seven scenarios and use the significant factors as alternative independent variables to capture the outcome of the cash flow simulations.[27] The fourteen cash flow variables are highly collinear, but the use of factor analysis enables us to reduce the set of information to that contained in only one or two factors, depending upon the analysis year.

Our principal control variables are those used by CHK (1995) — a size dummy variable set equal to one for small companies and zero otherwise and an organizational form dummy variable set equal to one for mutual insurers and to zero otherwise.[28] The motivation for the use of these variables is the CHK finding that they significantly improve solvency prediction as well as earlier univariate analyses that provides evidence of higher insolvency rates for small firms and lower failure rates for mutual firms than for non-mutual firms (e.g., A.M. Best Company, 1990).[29]

Model Evaluation

To compare the predictive accuracy of logistic models containing alternative sets of variables, we employ three approaches. Since logistic regression maximum likelihood techniques do not generate

[27]In this factor analysis we follow the same procedure outlined above for the FAST scores. We employ those factors with eigenvalues greater than or equal to one after a varimax rotation, and insignificant factors have been eliminated from the models shown in the tables.

[28]Our definition of a small company is the same as that used by CHK (1995). Specifically, the size variable is set equal to 1 if a firm's assets are less than $100 million. This criterion was chosen judgmentally, but our tests revealed that it performs better than a continuous size variable equal to assets or the logarithm of assets.

[29]Other things being equal, claim costs tend to be more volatile for small firms. Small firms also might have relatively lower franchise values and thus less incentive to reduce insolvency risk than large firms. Mutual insurers may be less prone to moral hazard in the presence of guaranty funds. They also tend to specialize in the sale of less risky coverages than non-mutuals (see Lamm-Tennant and Starks, 1993).

standard measures of goodness of fit, we examine the pseudo R^2 (likelihood ratio index), the Type I/Type II error trade off, and receiver operating characteristic (ROC) curves. The pseudo R^2 or likelihood ratio index is equal to one minus the ratio of the estimated log likelihood function value relative to the value of the likelihood function when the coefficients of the model are constrained to be zero (see Greene, 1990, p. 682). The Type I error rate is defined as the probability that a firm which subsequently fails is predicted to remain solvent, and the Type II error rate is the probability that a firm which remains solvent is predicted to fail. To evaluate the Type I/Type II error trade-off, we calculate the Type I error rates for various levels of the Type II error rate.[30] Models with relatively low Type I error rates conditional on the Type II error rate are deemed to be superior.

Although ROC analysis is widely used in other disciplines, it does not appear to have a strong presence in economics. Accordingly, we provide a brief discussion of ROC analysis here and more details in the Appendix.

The goal of ROC analysis is to provide a statistical test of whether a given model outperforms an alternative model in a binary prediction exercise (i.e., in categorizing observations into two mutually exclusive groups) for various Type II error rates. ROC analysis is usually summarized graphically by plotting a ROC curve in a two dimensional plane where the Type II error rate is plotted along the X-axis and the complement of the Type I error rate (1 minus the Type I error rate) is plotted along the Y-axis. In our analysis, we assume that two alternative models, X and Y, will yield unique ROC curves, and the parameters of the curves are estimated using the maximum likelihood technique developed by Metz, Wang, and Kronman (1984).

---

[30]The model produces a fitted value of y_{jt}^*, y_{jt}^f for each firm. Assume that we are interested in a Type II error rate of z percent, i.e., an error rate such that z percent of the solvent firms are classified as insolvent. We find the cutoff value of y_{jt}^*, y_{jt}^c such that $y_{jt}^f > y_{jt}^c$ for z percent of the solvent firms in the sample. The proportion of insolvent firms with $y_{jt}^f < y_{jt}^c$ then equals the Type I error rate.

A useful statistic that summarizes the accuracy of a particular model is the area below the ROC curve. This statistic, known as the area index, is denoted A_z where the subscript represents the particular model (i.e., X or Y) being summarized. A model that perfectly discriminates between the insolvent and solvent companies will have an area index equal to 1.0 and a model with no discriminatory power will result in an area index of 0.50. Using the results of the maximum likelihood estimation, we test the null hypothesis of equal areas under the two estimated ROC curves by calculating

$$z = \frac{A_x - A_y}{\sqrt{\sigma_x^2 + \sigma_x^2 - 2\rho\sigma_x\sigma_y}} \tag{12}$$

where A_i = the area under ROC curve i, σ_i^2 is the standard error of A_i, i = x, y, and ρ is the correlation coefficient between A_x and A_y. The test statistic z is distributed as a standard normal variate, and the null hypothesis is rejected for large values of z.

Study Design and Data

Our study design is very similar to that used by GHK (1993, 1995) and CHK (1995). Specifically, we predict failure rates over three-year prediction horizons using data from three base years — 1990, 1991, and 1992. Thus, the 1990 data are used to predict insolvencies over the period 1991-1993, the 1991 data are used to predict insolvencies for the period 1992-1994, and the 1992 data are used in insolvency prediction for the period 1993-1995. Our sample of insolvent companies consists of all property-liability insurers that were reported to the NAIC by state insurance regulators as becoming insolvent during the period 1991-1995.[31] Our insolvent firm sample consists of 44 companies that

[31]The insolvency data come from the NAIC *Contact Person Reports*. The reports contain information regarding single and multistate insolvencies brought to the attention of the NAIC by state regulators for 1990-

became insolvent during the 1991-1993 period, 48 during the 1992-1994 period, and 27 during the 1993-1995 period; and these insolvent companies were used in the analyses from the 1990, 1991, and 1992 base years, respectively.[32] In addition to the insolvent companies, our sample also included 244 solvent insurers for the 1990 analysis, 215 for the 1991 analysis, and 226 for the 1992 analysis.[33] Comparison of our sample of solvent firm with industry-wide data reveals that our sample is representative of the industry.

4. Results

Summary Statistics and Univariate Results

Summary statistics for the solvent and insolvent insurers in the 1990 sample are presented in Table 2. The results for the other two years are similar and hence are not shown. Tests of differences between means reveal that the solvent and insolvent insurers differ significantly across several dimensions.

1995. The *Reports* contain information regarding the first public regulatory order involving a company. We use this year of the first public order as the year of "failure." Any formal state regulatory order including restrictions on management, conservation, rehabilitation, or liquidation was treated for the purposes of this paper as a "failure." Almost all companies having a formal regulatory order entered against them are eventually liquidated. Two companies that were classified as "under rehabilitation" and were subsequently successfully rehabilitated were removed from the sample. Several insolvent single-state companies were eliminated from the sample because they did not file financial reports with the NAIC and thus do not appear on the NAIC data tapes. These companies were likely excused by a state from filing an annual statement because they had few assets and/or liabilities, were undergoing liquidation, or were in the process of being sold. A final constraint on the sample of failed companies is due to the way the NAIC determined the sample of companies for which it would calculate risk-based capital. Of the approximately 1,800 companies that filed statements with the NAIC, the NAIC calculated RBC for approximately 1,200. The excluded companies are typically small single-state companies or small companies with exotic organizational forms such as Texas Lloyds or reciprocals.

[32]There is overlap between the three sample of insolvent insurers. For example, companies becoming insolvent during 1993 were included in all three base year analyses.

[33]The solvent firm sample size differs by year due to our sampling approach. For each of the years 1990-1992, we chose a random sample of 300 solvent insurers that had meaningful data in the NAIC data tapes, e.g., companies that did not have negative or zero values for surplus, assets, or premiums. Some of the 300 firms were dropped from the solvent firm sample in each year because they were missing data needed as inputs for the cash flow simulation model and/or were missing RBC scores.

Insolvent insurers are significantly smaller than solvent insurers, are significantly less likely to be mutuals, and have significantly higher ratios of risk-based capital to actual surplus. The ratio of predicted ending surplus from the cash flow model to initial surplus is significantly lower for insolvent insurers than for solvent insurers, and the predicted failure year is significantly lower for insolvent insurers than for solvent insurers. The overall FAST score, most of the FAST and consistent FAST factors, and both cash flow factors differ significantly between the two sets of firms.

The univariate prediction results based on the risk-based capital formula and the cash flow simulation model are presented in Table 3. For the RBC analysis, the Type I error rates represent the percentages of insolvent companies with RBC ratios (ratios of actual surplus to RBC) less than the RBC ratio for the solvent companies that produces the specified Type II error rates. For the cash flow analysis, the results are reported by scenario, with the scenario results grouped by the stringency of the loss reserving scenario definition. Each scenario produces only one Type I-Type II error combination.

The RBC univariate results are consistent with those of CHK for 1990 and 1991, the two years in which the present study and the CHK study overlap. The Type I error rates for these years are quite high for the lower Type II error rate levels. For example, at a 10 percent Type II error rate, the Type I error rates in 1990 are 70.5 percent and 54.76 percent, respectively. Thus, at a Type II error rate that might be considered reasonable by solvent insurers, the RBC formula fails to detect the majority of insolvencies. The RBC results are significantly better for the 1992 tests. In this case the Type I error rate for the 10 percent Type II level drops to around 30 percent.

The cash flow simulation model does better than the RBC formula with the Type II error rate roughly in the 10 percent range for 1990 and 1991 but has a higher Type I error rate in 1992 than the RBC formula (this comparison is based on scenarios 1 and 6). With Type II error rates in the 20 percent range, the cash flow model again performs better than RBC in 1990 and 1991 but performs about the same in 1992. Thus, the cash flow model is more accurate than RBC in 1990 and 1991, but the RBC formula performs better in the 1992 tests.

Logistic Regression Results

The estimation strategy in the logistic regression analysis is to start with separate models that include risk-based capital variables, the overall FAST score, FAST factors estimated from the 1993 FAST ratios, FAST factors based on the ratios that were used consistently throughout the sample period, and cash flow simulation variables. We then analyze models that include combinations of variables from the five sources. All models include the small company and mutual firm dummy variables. Since the results for many of the models are similar across all three years, we present the results based on the 1990 sample in the tables and note any significant differences between the 1990 results and those for 1991 and 1992. The criteria for gauging the accuracy of models at this stage are the Type I - Type II error rate tradeoffs and pseudo R^2 values, which indicate explanatory power. We report the results of the ROC analysis in the next section.

Logistic models based on the solvency prediction methods considered separately are presented in Table 4, which shows the results for the 1990 sample. The best-performing of the cash flow models is shown in the table, i.e., the model based on factor analysis of the results of all seven scenarios.[34] Not surprisingly, the models with look-ahead bias, i.e., the models containing the overall FAST score and FAST factors based on the 1993 FAST ratios, perform the best both in terms of the pseudo-R^2 and Type I - Type II error tradeoffs. Among the models that do not suffer from look-ahead bias, the model with the consistent FAST factors has the highest pseudo-R^2, but the model with the cash flow factors generally

[34]The model with the log of the predicted failure year as the cash flow variable performed almost as well as the model using the cash flow factors. Although the predicted ending/current surplus variable was statistically significant, the model containing this cash flow variable did not perform quite as well as the other two cash flow-based models.

has the best Type I - Type II error performance. The RBC models are the least accurate, confirming earlier findings that RBC does not have much discriminatory power in predicting insolvencies.

The results presented in Tables 5 and 6 provide evidence on the predictive power of the RBC system, the FAST system, and cash flow variables considered in combination with one another. The FAST variables considered in Table 5 are the overall FAST score and the factors based on the 1993 FAST ratios. Because the FAST variables in Table 5 all suffer from look-ahead bias, this provides a strong test of the robustness of the RBC and cash flow results.

Nearly all of the RBC variables in Table 5 are statistically insignificant. Thus, RBC adds little or no explanatory power to FAST. In contrast, all of the cash flow variables in the Table 5 models are statistically significant, and the models containing the cash flow variables generally have better Type I - Type II error performance than the models containing FAST and RBC variables but no cash flow variables. Thus, cash flow analysis adds information to the NAIC's static solvency prediction systems, even when these systems have been optimized based on ex post information.[35] Table 5 also reveals that the FAST factors have better explanatory power than the overall FAST score, suggesting that predictive power can be lost by averaging or summing variables that are not perfectly correlated.

Table 6 is designed to provide evidence on the predictive power of RBC, FAST, and cash flow variables when look-ahead bias is not present. The NAIC variables used in this table consist of RBC and factors based on the consistent FAST ratios. None of the RBC variables in Table 6 is statistically significant, again confirming that RBC adds no predictive power to FAST. All of the cash flow variables are statistically significant, and the models that contain cash flow variables generally have superior Type

[35]The 1992 results are similar. In 1991, the log of the predicted year to failure is statistically significant and improves the Type I - Type II error tradeoffs. However, the cash flow factors for 1991 are often insignificant and produce only modest improvements in the Type I - Type II error tradeoffs.

I -Type II error tradeoffs than models containing only the NAIC variables. This provides further evidence that cash flow analysis has the potential to add power to the NAIC's solvency prediction models.[36]

ROC Results

In this section, we investigate the benefits of adding variables based on the cash flow simulation model to models based on the NAIC's solvency prediction systems by using the receiver operating characteristics (ROC) methodology discussed earlier. To do so, we jointly estimate the parameters of ROC curves that pair each of several solvency prediction models based only on NAIC variables with corresponding models containing the NAIC variables plus variables based on factor analysis of the cash flow simulation results. The resulting area indexes are then calculated and tested to determine whether adding the cash flow variables significantly improves the prediction results.

Table 7 shows the ROC results for the 1990-1992 samples. The first column defines the NAIC solvency models that are employed. The next column shows the estimated area index for the regression model specified in the first column with the standard error of the area index given in parentheses below. Column three displays the area indices and standard errors for the model defined in column one with the addition of the cash flow factors. The column labeled "Correlation A1,A2" shows the estimated correlation between the two area indices. The last two columns show the results of the test of the null hypothesis of equality between the two area indices and the one sided p-value, respectively.

Nearly all of the tests shown in Table 7 suggest that adding the cash flow factors leads to a statistically significant increase in the explanatory power of the models defined in column 1. Of the 24 comparisons shown in the table, 14 are significant at the 5 percent level or better, and 21 are significant at the 10 percent level or better. For 1990 and 1992, the largest and most significant improvements in

[36]The conclusions based on 1992 data are similar. However, the cash flow variables do not perform quite as well for 1991.

explanatory power are obtained by adding the cash flow variables to models containing only the RBC variable(s). For 1991, the improvements in explanatory power obtained by adding cash flow variables to models containing the FAST score, the consistent FAST factors, and RBC plus the consistent FAST factors are largest and most significant. The overall conclusion is that cash flow analysis tends to add significant explanatory power even to models characterized by look-ahead bias.

Another way to view the results of the ROC analysis is to plot the ROC curves for competing models. Recall, the ROC curve is a plot of the probability of correct predictions of insolvent companies, i.e., the complement of the Type I error rate (on the Y-axis), against the probability of incorrectly classifying solvent companies, i.e., the Type II Error Rate (along the X-axis). Figure 1 displays the estimated 1990 ROC curves for the models using the individual RBC components only, the model containing both the individual RBC components plus the consistent FAST factors, and the model containing the RBC components, the consistent FAST factors, and the cash flow factors. The plot shows that the model which performs with the least accuracy is the model containing only the individual RBC components. When the consistent FAST factors are added, the estimated ROC curve moves further into the northeast corner of the graph indicating that, for any given Type II error rate, this model discriminates better than the model based only on the individual RBC components. The ROC curve furthest into the northeast corner adds the cash flow factors to the individual RBC components and the consistent FAST factors. The ROC curves for the other years lead to similar conclusions and thus are not shown.

5. Summary and Conclusions

This paper analyzes three methods of predicting insolvencies in the U.S. property-liability insurance industry – the NAIC's risk-based capital (RBC) system, the NAIC's financial analysis and surveillance tracking (FAST) system, and a cash flow simulation model developed by the authors. The systems are tested individually and jointly to determine predictive performance.

The prediction methodology is logistic regression analysis. The dependent variable is equal to 1 if a company becomes insolvent and equal to zero otherwise. The independent variables are drawn from the three solvency prediction methodologies. To test the RBC system, the ratio of the company's overall risk-based capital to its surplus and the ratios of the five principal risk-based capital components to surplus are used alternatively as independent variables. To test the FAST system, we consider the NAIC's overall FAST score and the FAST ratios, where both the scores and the ratios are optimized to accurately predict the insolvencies that occurred in 1993. To avoid the look-ahead bias inherent in the 1993 scores and ratios and hence to provide a true ex ante test of the FAST system, we also test the subset of FAST ratios that were used consistently throughout our sample period. For each variant of the FAST ratios, we use factor analysis to compute orthogonal factors based on the ratios and use the resulting factors as regressors. To test the cash flow model, we use the ratio of predicted surplus at the end of a twenty-year simulation period to actual surplus at the start of the simulation period and the predicted time to failure, both from the baseline (expected value) simulation, as well as factors estimated from the same two variables generated under all seven scenarios incorporated in the simulation model.

The tests are conducted using data from 1990, 1991, and 1992 to predict insolvencies over three year prediction horizons, 1991-1993, 1992-1994, and 1993-1995, respectively. In conducting the tests, we use all insolvent firms reported by the NAIC for which annual statement data are available as well as a sample of more than 200 solvent firms for each base year.

The results support three principal conclusions: First, we confirm the findings of GHK (1993, 1995) and CHK (1995) that the risk-based capital ratio and its components provide very low explanatory power in predicting insurer insolvencies. Second, we find that the FAST factors tend to dominate the risk-based capital variables. The explanatory power of the FAST models is considerably higher than that of the RBC models, and the RBC variables tend to be insignificant when included in models jointly with the FAST factors. Thus, RBC appears to add no information to the FAST system.

Third, the cash flow simulation variables add significant discriminatory power to the solvency prediction models based on the RBC and FAST systems, even in the presence of look-ahead bias in the FAST variables. Thus, dynamic financial analysis appears to hold significant promise for providing regulators with better predictions of insurer solvency. This conclusion is particularly strong in view of the fact that additional sophistication could easily be incorporated into the cash flow model, such as upward sloping yield curves and non-parallel yield curve shifts. Cash flow simulation also has the advantage of providing other information likely to be useful to regulators, such as the predicted time to failure, that is not provided by any of the existing regulatory information systems.

References

Altman, Edward I., 1992, "Revisiting The High Yield Bond Market," *Financial Management* 21: 78-92.

A.M. Best Company, 1996, *Best's Aggregates and Averages: 1996 Edition* (Oldwick, NJ).

A.M. Best Company, 1990, *Best's Insolvency Study: Property/Casualty Insurers 1969-1990* (Oldwick, NJ).

Casualty Actuarial Society, 1996, *Proceedings of the Casualty Actuarial Society Forum* (Landover, MD, Spring 1996).

Cummins, J.D., S. Harrington, and R. Klein, 1995, "Insolvency Experience, Risk-Based Capital, and Prompt Corrective Action in Property-Liability Insurance," *Journal of Banking and Finance* 19: 511-527.

Cummins, J.D., S. Harrington, and G. Niehaus, 1993, " An Economic Overview of Risk-based Capital Requirements in the Property-liability Insurance Industry, *Journal of Insurance Regulation* 11: 427-447.

Cummins, J.D., S. Harrington, and G. Niehaus, 1995, "Risk-based Capital Requirements for Property-liability Insurers: A Financial Analysis," in Edward Altman and Irwin Vanderhoof, eds., *The Financial Dynamics of the Insurance Industry* (Homewood, IL: Irwin Professional Publishers).

Cummins, J.D. and R.D. Phillips, 1994, "Cash Flow Simulation Model User's Guide," working paper, Wharton School, University of Pennsylvania.

Grace, M., S. Harrington, and R. Klein, 1993, "Risk-based Capital Standards and Insurer Insolvency Risk: an Empirical Analysis," paper presented at the 1993 Annual Meeting of the American Risk and Insurance Association. Center for Risk Management and Insurance Working Paper, Georgia State University.

Grace, M., S. Harrington, and R. Klein, 1995, *An Analysis of the FAST Solvency Monitoring System*, presented to the NAIC's Financial Analysis Working Group. (Kansas City: NAIC).

Greene, William H., 1990, *Econometric Analysis* (New York, Macmillan).

Hodes, D.M., T. Neghaiwi, J. D. Cummins, R.D. Phillips, and S. Feldblum, 1996, "The Financial Modeling of Property/Casualty Insurance Companies," in *Proceedings of the Casualty Actuarial Society Forum* (Ladover, MD: Casualty Actuarial Society, 1996).

Ibbotson Associates, 1995, *Stocks, Bonds, Bills and Inflation: 1995 Yearbook* (Chicago).

Klein, R.W., 1995, "Insurance Regulation in Transition,' *Journal of Risk and Insurance* 62: 363-404.

Lamm-Tennant, J., and Starks, L., 1993, "Stock Versus Mutual Ownership Structures: The Risk Implications," *Journal of Business* 66: 29-46.

Metz C.E., Wang P-L., Kronman H.B., 1984, " A New Approach for Testing the Significance of Difference Between ROC Curves Measured from Correlated Data," in F. Deconinck, ed.., *Information Processing in Medical Imaging: Proceedings of the 8th Conference* (Boston, MA: Martinus Nijhoff), pp. 432-445.

National Association of Insurance Commissioners, 1993, " NAIC Property/Casualty Risk-based Capital Formula Exposure Draft" (Kansas City, MO).

Petersen, W. W., T. G. Birdsall, and W. C. Fox, 1954, IRE Transactions of the Professional Group on Information Theory, PGIT-4, 171.

Pregibon, D., 1981, "Logistic Regression Diagnostics," *Annals of Statistics* 9: 705-724.

Swets, J. A., 1996, Signal Detection Theory in Psychology and Diagnostics: Collected Papers (Lawrence Erlbaum Associates: Mahwah, NJ).

U.S., General Accounting Office, 1991, "Insurance Regulation: State Handling of Troubled Property/casualty Insurers (Washington, D.C.).

U.S., House of Representatives, Subcommittee on Oversight and Investigations of the Committee on Energy and Commerce, 1990, *Failed promises - Insurance Company Insolvencies* (Washington, D.C.: U.S. Government Printing Office).

Appendix

Receiver Operating Characteristic (ROC) Analysis

The goal of receiver operating characteristic (ROC) analysis is to provide a statistical test of whether a given model outperforms an alternative model in a binary prediction exercise (i.e., in categorizing observations into two mutually exclusive groups) for various Type II error rates. ROC analysis is usually summarized graphically by plotting a ROC curve in a two dimensional plane where the Type II error rate is plotted along the X-axis and the complement of the Type I error rate (1 minus the Type I error rate) is plotted along the Y-axis. In the standard analysis, the curve is assumed to have the same functional form as that implied by two normal distributions. This assumption has the convenient property that ROC curves plotted on normal deviate axes are transformed into straight lines. Thus, determining the slope, b, and y-intercept, a, of a transformed ROC curve will completely describe the relative difference between the parameters of the two underlying normal distributions.

In our analysis, we assume that two alternative models, X and Y, will yield unique ROC curves. The goal is to estimate the parameters a and b for each model. To accomplish this we adopt the maximum likelihood technique developed by Metz, Wang, and Kronman (1984).[1] Formally, let there be two decision variables, X and Y, that arise from two bivariate normal joint probability distributions.[2] For our purposes we want to compare the predictive ability of two different logistic regression models, X and Y. For a given base-year, each model produces a probability of insolvency for each insurer in our sample. Define f(x,y|insolvent) to be the joint probability density function of insolvency for a firm that eventually becomes insolvent over the three year prediction horizon, and f(x,y|solvent) to be the joint probability density function for a firm that remains solvent. Next, define a set of cutoffs t_i and u_j such that

(i) if $x < t_1$, then the response is i=1; $y < u_1$, then the response is j=1
(ii) if $x > t_n$ the response is i=n+1; $y > u_n$ the response is j=n+1; and
(iii) if $t_i < x < t_{i+1}$ the response is i for all $i < n$; $u_i < y < u_{i+1}$ the response is j for all $j < n$;[3]

where x and y represent the fitted values x^f and y^f, respectively, from two estimated logistic models following the general specification of equation (10).

---

[1]We used a program called CORROC2 written by Dr. Charles Metz at the University of Chicago School of Medicine to conduct the maximum likelihood estimation. The software can be downloaded via anonymous FTP by accessing ftp://random.bsd.uchicago.edu.

[2]That is, the variable X, for example, arises from one of two normal probability densities, f(x*n) or f(x*s), where n stands for "noise only" and s for "signal present." In terms of our analysis, "noise only" can be taken to refer to solvent firms and "signal present" to refer to insolvent firms.

[3]In our analysis we chose to employ 11 response categories based on cutoffs at 0.025, 0.05, 0.10, 0.20, 0.30, 0.40., 0.50, 0.60, 0.70, 0.80. Metz, Wang, and Kronman (1984) suggest using as many cutoffs as possible to provide the most information about the distribution of predictions. Eleven was the maximum number of different response categories allowable under the CORROC2 program. In addition, they suggest using more finely aggregated cutoffs where large numbers of observations are found. Given the large number of solvent companies in our sample relative to the number of insolvents, we choose to create a larger number of response categories at the lower probabilities of predicted insolvency than at the higher probabilities.

Thus, for each insolvent observation in the sample, the probability of a pair of ratings, i and j, from each competing model will equal

$$p_{ij} = F(t_i, u_j, r_{ins}) + F(t_{i-1}, u_{j-1}, r_{ins}) - F(t_{i-1}, u_j, r_{ins}) - F(t_i, u_{j-1}, r_{ins}) \tag{A.1}$$

where

$$F(x, y, r_{ins}) = \int_x^{\infty} dv \int_y^{\infty} f(v, w, r_{ins})\, dw \tag{A.2}$$

and

$$f(x, y, r_{ins}) = \frac{1}{2\pi\sqrt{1 - r_{ins}^2}}\, e^{-\frac{x^2 - 2 r_{ins} x y + y^2}{2(1 - r_{ins}^2)}}. \tag{A.3}$$

where $F(x,y,r)$ = a bivariate standard normal probability distribution function and $f(x,y,r)$= a bivariate standard normal density function, both evaluated at x and y with correlation coefficient r; and r_{ins} is the correlation coefficient of the probability of insolvency between two models for the sample of insolvent companies.

Similarly, for each solvent firm in the sample, the probability of a pair of ratings i and j from the two models will equal

$$\begin{aligned} p_{ij} = {} & F(b_x t_i - a_x, b_y u_j - a_y, r_{sol}) + F(b_x t_{i-1} - a_x, b_y u_{j-1} - a_y, r_{sol}) \\ & - F(b_x t_{i-1} - a_x, b_y u_j - a_y, r_{sol}) - F(b_x t_i - a_x, b_y u_{j-1} - a_y, r_{sol}). \end{aligned} \tag{A.4}$$

where r_{sol} is the correlation coefficient of the probability of insolvency between two models for the sample of solvent companies.[4] Given these definitions, the likelihood function to be maximized with respect to a_x, a_y, b_x, b_y, r_{ins}, and r_{sol}, is

[4]Without loss of generality, the mean of the joint "signal present" distribution of X and Y is set to (0,0) and the marginal standard deviations of this distribution are set equal to 1. The transformations of t_i and u_i in equation (A.4) then convert the "noise only" variates to standard normal variates. Before the transformations, the marginal distributions of these variates had means (a_x/b_x) and (a_y/b_y), respectively, and and standard deviations $(1/b_x)$ and $(1/b_y)$. Thus, as explained above, the a parameters represent the differences between the means of the marginal "signal present" and "noise only" distributions.

$$Log\,L = \sum_{i=1}^{n+1}\sum_{j=1}^{n+1} n_{ij}^{ins}\log p_{ij} + \sum_{i=1}^{n+1}\sum_{j=1}^{n+1} n_{ij}^{sol}\log \mathrm{p}_{ij} \qquad (A.5)$$

where n_{ij}^{ins} (n_{ij}^{sol})is the number of insolvent (solvent) observations to have the pair of ratings from the two models equal to i and j.

Once the parameters of the ROC curves have been estimated, a useful statistic that summarizes the accuracy of a particular model involves calculating the area below the ROC curve. This statistic, known as the area index, is denoted A_z where the subscript represents the particular model (i.e., X or Y) being summarized. The area index is related to the estimated parameters of the ROC curve for a particular model Z by the expression

$$A_z = \Phi\left(\frac{a_z}{\sqrt{1 + b_z^2}}\right) \qquad (A.6)$$

where $\Phi(\bullet)$ is the commutative normal distribution function. A model that perfectly discriminates between the insolvent and solvent companies will have an area index equal to 1.0 and a model with no discriminatory power will result in an area index of 0.50. Using the results of the maximum likelihood estimation, we test the null hypothesis of equal areas under the two estimated ROC by calculating where SE_i^2 is the standard error for A_i and *r* is the correlation coefficient between A_x and A_y, and the

$$z = \frac{A_x - A_y}{\sqrt{SE_x^2 + SE_x^2 - 2rSE_xSE_y}} \qquad (A.7)$$

statistic z is distributed as a standard normal variate. The null hypothesis is rejected for large values of the test statistic z.

Table 1
Pro-forma Cash Flow Statement
For A Property-Liability Insurer

Cash Flow	**Industry-Wide Total (1995)**
Cash Inflows:	
Premiums	255,655
Investment income	37,561
Other inflows	1,045
Cash Outflows:	
Loss payments	189,496
Underwriting expense payments	66,718
Policyholder dividend payments	3,512
Federal tax payments	4,202
Net Cash Flow:	30,333

Source: *Best's Aggregates and Averages: 1996 Edition.*

Table 2
1990 Summary Statistics: Solvent vs. Insolvent Insurers

Variable	Solvent Mean m_{sol}	Solvent Standard Deviation	Insolvent Mean m_{ins}	Insolvent Standard Deviation	Test Statistic: $m_{sol}=m_{ins}$
Log(Assets)	17.799	1.771	16.844	1.104	4.745
Small Company Dummy Variable	0.648	0.479	0.909	0.291	4.891
Year of Actual Insolvency	0.000	0.000	92.182	0.582	-
Investment RBC Charge over PH Surplus	0.094	0.129	0.149	0.232	1.548
Credit RBC Charge over PH Surplus	0.069	0.141	0.189	0.221	3.475
Loss RBC Charge over PH Surplus	0.326	0.672	0.384	0.439	0.735
Written Premium RBC Charge over PH Surplus	0.334	0.488	0.594	0.423	3.659
Growth RBC Charge over PH Surplus	0.049	0.203	0.120	0.216	2.026
Mutual Organizational Form Dummy	0.242	0.429	0.091	0.291	2.917
Total RBC over PH Surplus	0.229	0.377	0.352	0.266	2.641
Percent of Companies with RBC Ratio < 2	3.69%	18.89%	15.91%	37.00%	2.141
Overall Fast Score	386.066	236.401	758.182	399.878	5.987
Fast Factor 1	-0.148	0.542	0.823	2.049	3.124
Fast Factor 2	-0.055	0.999	0.303	0.961	2.262
Fast Factor 3	-0.047	0.915	0.262	1.363	1.449
Fast Factor 4	0.046	1.038	-0.252	0.717	2.346
Fast Factor 5	-0.033	1.047	0.184	0.665	1.799
Fast Factor 6	-0.046	0.907	0.252	1.394	1.366
Consistent Fast Factor 1	-0.119	0.640	0.662	1.958	2.623
Consistent Fast Factor 2	-0.100	0.826	0.553	1.566	2.698
Consistent Fast Factor 3	-0.036	0.993	0.197	1.028	1.387
Consistent Fast Factor 4	-0.022	0.999	0.123	1.011	0.878
Consistent Fast Factor 5	-0.059	0.987	0.327	1.018	2.329
Predicted Ending Surplus over Current PH Surplus	7.730	7.498	2.312	8.176	4.096
Predicted Year of Failure	4,116.62	6,307.61	1,006.98	1,841.89	6.345
Log(Predicted Year of Failure)	7.327	1.972	4.237	2.935	6.716
Cash Flow Factor 1	0.143	0.953	-0.793	0.887	6.372
Cash Flow Factor 2	0.126	0.893	-0.699	1.256	4.168

Note; 244 Solvent Companies, 44 Insolvent Companies

Table 3
Univariate Prediction Results:
Risk-Based Capital versus Cash Flow Simulation

Risk Based Capital

Type II	1990 Type I Error	1991 Type I Error	1992 Type I Error
5%	84.09%	66.67%	48.15%
10%	70.45%	54.76%	29.63%
15%	61.36%	35.71%	25.93%
20%	52.27%	33.33%	18.52%
25%	43.18%	28.57%	14.81%
30%	34.09%	26.19%	11.11%

Cash Flow Simulation Model

Baseline Reserving Risk Scenarios

	1990 Error Rates		1991 Error Rates		1992 Error Rates	
	Type II	Type I	Type II	Type I	Type II	Type I
Scenario 1	8.61%	45.45%	7.18%	45.24%	7.52%	51.85%
Scenario 6	11.07%	47.73%	9.57%	45.24%	7.08%	44.44%

Moderately Adverse Reserving Risk Scenarios

	1990 Error Rates		1991 Error Rates		1992 Error Rates	
	Type II	Type I	Type II	Type I	Type II	Type I
Scenario 2	22.54%	27.27%	20.10%	28.57%	17.70%	25.93%
Scenario 4	22.54%	27.27%	20.57%	30.95%	17.26%	25.93%

Severely Adverse Reserving Risk Scenarios

	1990 Error Rates		1991 Error Rates		1992 Error Rates	
	Type II	Type I	Type II	Type I	Type II	Type I
Scenario 3	41.80%	18.18%	37.80%	16.67%	39.82%	18.52%
Scenario 5	41.80%	18.18%	36.36%	16.67%	38.94%	18.52%
Scenario 7	40.98%	18.18%	38.28%	16.67%	38.50%	18.52%

Table Notes

1990 error rates are for the prediction horizon 1991-1993; 1991 error rates are for the prediction horizon 1992-1994; and 1992 error rates are for the prediction horizon 1993-1995.

Insolvent sample size: 44 in 1990; 48 in 1991; 27 in 1992.

Solvent sample size: 244 in 1990; 215 in 1991; 226 in 1992.

Cash Flow Type I Error Rate - percentage of insolvent companies incorrectly predicted to remain solvent.

Cash Flow Type II Error Rate - percentage of solvent companies incorrectly predicted to become insolvent.

Risk Based Capital Error Rates - Type I error rate is percentage of insolvent firms with RBC ratios less than the RBC ratio for solvent companies that produces the specified Type II error rate.

Table 4
Logistic Regression Results: 1990 Sample
RBC vs. FAST vs. Consistent FAST vs. Cash Flow Variables Tested Separately

Variable	Total RBC	Individual RBC	Fast Score	Fast Factors	Consistent Fast Factors	Cash Flow Factors
Intercept	-3.395	-4.419	-5.101	-4.176	-3.628	-2.878
	-(5.85)	-(5.71)	-(7.04)	-(5.81)	-(5.95)	-(5.25)
Small Company Dummy Variable	2.006	2.512	1.814	2.840	2.244	0.945
	(3.46)	(3.52)	(3.01)	(3.88)	(3.53)	(1.61)
Mutual Organizational Form Dummy	-1.536	-1.568	-1.206	-1.535	-1.287	-0.327
	-(2.55)	-(2.55)	-(2.00)	-(2.49)	-(2.21)	-(0.54)
Total RBC over PH Surplus	1.088					
	(2.94)					
Investment RBC Charge over PH Surplus		3.070				
		(1.84)				
Credit RBC Charge over PH Surplus		3.524				
		(3.38)				
Loss RBC Charge over PH Surplus		-0.111				
		-(0.27)				
Written Premium RBC Charge over PH Surplus		0.656				
		(2.03)				
Growth RBC Charge over PH Surplus		-0.406				
		-(0.35)				
Fast Score			0.0039			
			(5.92)			
Fast Factor 1				1.290		
				(4.42)		
Fast Factor 2				0.515		
				(2.43)		
Fast Factor 3				0.259		
				(1.81)		
Fast Factor 4				-0.689		
				-(2.28)		
Fast Factor 5				0.796		
				(1.99)		
Fast Factor 6				0.296		
				(2.11)		
Consistent Fast Factor 1					0.713	
					(3.56)	
Consistent Fast Factor 2					0.538	
					(3.02)	
Consistent Fast Factor 3					0.447	
					(1.78)	
Consistent Fast Factor 4					0.387	
					(1.84)	
Consistent Fast Factor 5					0.721	
					(3.21)	
Cash Flow Factor 1						-0.937
						-(4.28)
Cash Flow Factor 2						-0.591
						-(3.10)
Log Likelihood Function Value	-108.78	-98.88	-89.35	-83.18	-90.269	-94.675
Psuedo R2	11.6%	19.7%	27.4%	32.4%	26.7%	23.1%
Type I Error Rates						
5 Percent Type II Error Rate	89%	68%	52%	64%	66%	55%
10 Percent Type II Error Rate	66%	52%	45%	36%	57%	48%
15 Percent Type II Error Rate	50%	48%	39%	32%	41%	41%
20 Percent Type II Error Rate	39%	39%	34%	27%	36%	36%
25 Percent Type II Error Rate	27%	30%	25%	25%	30%	27%
30 Percent Type II Error Rate	25%	25%	23%	14%	30%	27%

Table 5
1990 Logistic Regression Results: Combining RBC, Ex Post FAST Variables, and Cash Flow Variables

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12
Intercept	-5.153	-2.494	-4.632	-5.372	-2.566	-4.879	-4.255	-1.168	-3.733	-3.612	-0.853	-3.357
	-(6.98)	-(2.63)	-(6.26)	-(6.48)	-(2.47)	-(5.77)	-(5.54)	-(1.15)	-(5.00)	-(4.01)	-(0.80)	-(3.91)
Small Company Dummy Variable	1.850	1.083	1.111	1.974	1.207	1.359	2.849	1.804	1.834	2.425	1.497	1.595
	(3.03)	(1.72)	(1.77)	(2.79)	(1.67)	(1.87)	(3.89)	(2.54)	(2.53)	(3.13)	(1.97)	(2.06)
Mutual Organizational Form Dummy	-1.241	-0.489	-0.446	-1.346	-0.664	-0.588	-1.527	-0.847	-0.744	-1.520	-0.961	-0.882
	-(2.03)	-(0.72)	-(0.66)	-(2.09)	-(0.94)	-(0.83)	-(2.48)	-(1.23)	-(1.06)	-(2.35)	-(1.35)	-(1.22)
Total RBC over PH Surplus	0.286	-0.155	0.227				0.291	-0.974	-0.137			
	(0.46)	-(0.20)	(0.32)				(0.32)	-(0.75)	-(0.11)			
Investment RBC Charge over PH Surplus				-0.368	0.750	1.167				-4.710	-3.543	-2.974
				-(0.20)	(0.36)	(0.54)				-(1.93)	-(1.44)	-(1.20)
Credit RBC Charge over PH Surplus				0.325	0.565	0.804				-0.827	-0.957	-0.551
				(0.24)	(0.39)	(0.54)				-(0.33)	-(0.38)	-(0.22)
Loss RBC Charge over PH Surplus				0.251	-0.179	0.124				-0.042	-0.539	-0.291
				(0.55)	-(0.36)	(0.23)				-(0.07)	-(0.77)	-(0.39)
Written Premium RBC Charge over PH Surplus				0.418	0.354	0.395				0.479	0.370	0.449
				(1.30)	(1.04)	(1.17)				(1.31)	(0.92)	(1.17)
Growth RBC Charge over PH Surplus				-0.855	-1.161	-1.093				0.456	-0.237	-0.020
				-(0.72)	-(0.92)	-(0.87)				(0.31)	-(0.16)	-(0.01)
Fast Score	0.004	0.003	0.003	0.004	0.003	0.003						
	(5.54)	(4.58)	(4.19)	(4.20)	(3.20)	(2.82)						
Fast Factor 1							1.263	1.333	1.300	1.813	1.782	1.666
							(4.16)	(3.98)	(3.65)	(3.00)	(2.90)	(2.75)
Fast Factor 2							0.491	0.504	0.473	0.573	0.585	0.549
							(2.21)	(2.03)	(1.97)	(2.39)	(2.25)	(2.14)
Fast Factor 3							0.258	0.254	0.194	0.282	0.279	0.222
							(1.81)	(1.61)	(1.26)	(1.94)	(1.74)	(1.42)
Fast Factor 4							-0.686	-0.566	-0.529	-0.921	-0.679	-0.629
							-(2.27)	-(1.79)	-(1.71)	-(2.58)	-(1.90)	-(1.79)
Fast Factor 5							0.828	0.692	0.798	1.170	0.942	1.030
							(2.00)	(1.52)	(1.75)	(2.43)	(1.89)	(2.09)
Fast Factor 6							0.289	0.338	0.345	0.308	0.368	0.372
							(2.04)	(2.23)	(2.29)	(2.09)	(2.38)	(2.39)
Log(Predicted Year of Failure)		-0.304			-0.313			-0.334			-0.316	
		-(3.81)			-(3.76)			-(3.89)			-(3.62)	
Cash Flow Factor 1			-0.532			-0.535			-0.514			-0.491
			-(2.24)			-(2.22)			-(1.99)			-(1.89)
Cash Flow Factor 2			-0.560			-0.593			-0.683			-0.636
			-(2.82)			-(2.76)			-(2.94)			-(2.65)
Log Likelihood Function Value	-89.262	-82.117	-82.413	-88.365	-81.264	-81.451	-83.136	-75.514	-75.032	-80.206	-73.705	-73.557
Psuedo R2	27.5%	33.3%	33.1%	28.2%	34.0%	33.8%	32.5%	38.7%	39.1%	34.9%	40.1%	40.3%
				Type I Error Rate								
5 Percent Type II Error Rate	52%	55%	59%	55%	55%	61%	61%	52%	57%	73%	66%	70%
10 Percent Type II Error Rate	45%	41%	41%	48%	45%	48%	36%	32%	27%	50%	36%	39%
15 Percent Type II Error Rate	41%	36%	34%	39%	36%	30%	32%	25%	23%	34%	23%	25%
20 Percent Type II Error Rate	34%	23%	23%	36%	27%	25%	27%	20%	20%	23%	20%	20%
25 Percent Type II Error Rate	25%	18%	20%	30%	23%	23%	25%	18%	20%	16%	18%	20%
30 Percent Type II Error Rate	25%	18%	18%	25%	20%	20%	14%	18%	18%	14%	16%	18%

Table 6
1990 Logistic Regression Results: Combining RBC and Consistent FAST with Cash Flow Variables

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Intercept	-3.634	-2.791	-0.874	-3.187	-3.975	-3.045	-1.340	-3.905
	-(5.61)	-(3.74)	-(0.94)	-(5.02)	-(4.46)	-(3.40)	-(1.28)	-(4.28)
Small Company Dummy Variable	2.245	1.964	1.340	1.332	2.367	1.954	1.576	1.656
	(3.52)	(3.01)	(2.09)	(2.08)	(3.29)	(2.67)	(2.19)	(2.27)
Mutual Organizational Form Dummy	-1.288	-0.893	-0.638	-0.516	-1.374	-0.936	-0.750	-0.630
	-(2.20)	-(1.49)	-(1.00)	-(0.80)	-(2.27)	-(1.46)	-(1.13)	-(0.93)
Total RBC over PH Surplus	0.023	-1.136	-0.686	-0.218				
	(0.03)	-(0.96)	-(0.60)	-(0.20)				
Investment RBC Charge over PH Surplus					0.723	3.188	1.717	2.093
					(0.38)	(1.55)	(0.87)	(1.02)
Credit RBC Charge over PH Surplus					1.849	3.289	2.547	3.236
					(0.69)	(1.15)	(0.94)	(1.17)
Loss RBC Charge over PH Surplus					-0.391	-1.658	-0.738	-0.548
					-(0.46)	-(2.22)	-(0.91)	-(0.55)
Written Premium RBC Charge over PH Surplus					0.434	0.306	0.268	0.363
					(1.15)	(0.74)	(0.60)	(0.85)
Growth RBC Charge over PH Surplus					-0.722	-0.422	-1.237	-1.308
					-(0.50)	-(0.27)	-(0.81)	-(0.87)
Consistent Fast Factor 1	0.711	0.900	0.729	0.670	0.567	0.605	0.520	0.434
	(3.27)	(3.33)	(2.82)	(2.50)	(1.93)	(1.79)	(1.57)	(1.27)
Consistent Fast Factor 2	0.537	0.621	0.441	0.429	0.313	0.166	0.099	0.010
	(2.97)	(3.07)	(2.21)	(2.16)	(0.85)	(0.42)	(0.26)	(0.03)
Consistent Fast Factor 3	0.447	0.342	0.235	0.281	0.506	0.331	0.303	0.350
	(1.78)	(1.27)	(0.81)	(0.97)	(1.94)	(1.17)	(1.02)	(1.19)
Consistent Fast Factor 4	0.386	0.479	0.286	0.326	0.520	0.751	0.431	0.468
	(1.74)	(1.85)	(1.14)	(1.36)	(1.77)	(2.42)	(1.39)	(1.42)
Consistent Fast Factor 5	0.721	0.700	0.646	0.599	0.653	0.537	0.524	0.467
	(3.21)	(3.05)	(2.72)	(2.50)	(2.78)	(2.27)	(2.10)	(1.86)
Predicted Ending Surplus over Current PH Surplus		-0.079				-0.120		
		-(2.84)				-(3.49)		
Log(Predicted Year of Failure)			-0.313				-0.326	
			-(3.78)				-(3.87)	
Cash Flow Factor 1				-0.556				-0.566
				-(2.25)				-(2.23)
Cash Flow Factor 2				-0.558				-0.631
				-(2.58)				-(2.70)
Log Likelihood Function Value	-90.269	-86.113	-83.121	-82.944	-89.401	-82.995	-81.894	-81.430
Psuedo R2	26.7%	30.1%	32.5%	32.6%	27.4%	32.6%	33.5%	33.9%
				Type I Error Rate				
5 Percent Type II Error Rate	66%	59%	50%	57%	73%	66%	57%	57%
10 Percent Type II Error Rate	59%	43%	41%	41%	64%	41%	41%	45%
15 Percent Type II Error Rate	41%	41%	36%	30%	45%	36%	39%	30%
20 Percent Type II Error Rate	36%	32%	34%	27%	30%	30%	34%	30%
25 Percent Type II Error Rate	30%	32%	30%	27%	30%	30%	30%	30%
30 Percent Type II Error Rate	30%	30%	27%	27%	27%	30%	30%	30%

Table 7
ROC Results for Models with and without Cash Flow Factors

	Area Index - A_1	Area Index after Adding Cash Flow Factors - A_2	Correlation A_1, A_2	Z test $H_0: A_1 = A_2$	One sided p-value
1990 Results					
RBC	0.7502	0.8284	0.4859	2.0637	1.95%
	(0.041)	(0.033)			
Individual RBC	0.7980	0.8537	0.7785	2.4717	0.67%
	(0.036)	(0.030)			
Fast Score	0.8252	0.8693	0.8507	2.2281	1.29%
	(0.037)	(0.029)			
Ex Post Fast Factors	0.8482	0.8741	0.8761	1.5843	5.66%
	(0.034)	(0.030)			
Consistent Fast Factors	0.8399	0.8622	0.8607	1.3025	9.64%
	(0.033)	(0.031)			
RBC+Fast Score	0.8268	0.8656	0.8645	2.0954	1.81%
	(0.037)	(0.030)			
Individual RBC+Fast Score	0.8303	0.8592	0.8827	1.6434	5.02%
	(0.037)	(0.032)			
RBC+Ex Post Fast Factors	0.8503	0.8757	0.8828	1.5905	5.59%
	(0.034)	(0.030)			
Individual RBC+Ex Post Fast Factors	0.8630	0.8843	0.8769	1.3765	8.43%
	(0.032)	(0.029)			
RBC+Consistent Fast Factors	0.8391	0.8635	0.8666	1.4464	7.40%
	(0.034)	(0.031)			
Individual RBC+Consistent Fast Factors	0.8380	0.8687	0.8721	1.8729	3.05%
	(0.034)	(0.029)			
1991 Results					
RBC	0.7959	0.8140	0.9429	1.3104	9.50%
	(0.041)	(0.036)			
Individual RBC	0.8141	0.8336	0.9467	1.4702	7.08%
	(0.041)	(0.037)			
Fast Score	0.8377	0.8650	0.9628	2.394	0.83%
	(0.041)	(0.036)			
Ex Post Fast Factors	0.8386	0.8655	0.9536	2.2498	1.22%
	(0.039)	(0.034)			
Consistent Fast Factors	0.7673	0.8016	0.9488	2.5009	0.62%
	(0.043)	(0.039)			
RBC+Fast Score	0.8489	0.8623	0.9749	1.5188	6.44%
	(0.039)	(0.037)			
Individual RBC+Fast Score	0.8473	0.8695	0.9660	2.1326	1.65%
	(0.039)	(0.035)			
RBC+Ex Post Fast Factors	0.8729	0.8758	0.9707	0.3417	36.63%
	(0.035)	(0.035)			
Individual RBC+Ex Post Fast Factors	0.8683	0.8651	0.9782	0.4116	34.03%
	(0.036)	(0.037)			
RBC+Consistent Fast Factors	0.7770	0.8050	0.9683	2.4599	0.69%
	(0.044)	(0.039)			
Individual RBC+Consistent Fast Factors	0.7820	0.7906	0.9676	0.7892	21.50%
	(0.043)	(0.041)			
1992 Results					
RBC	0.7467	0.8616	0.2597	2.1139	1.73%
	(0.053)	(0.032)			
Individual RBC	0.8156	0.8810	0.6826	1.7353	4.13%
	(0.052)	(0.037)			
Fast Score	0.9215	0.9366	0.9203	1.5885	5.61%
	(0.024)	(0.022)			
Ex Post Fast Factors	0.8385	0.8902	0.8029	1.799	3.60%
	(0.048)	(0.039)			
Consistent Fast Factors	0.8304	0.8738	0.7835	1.3855	8.29%
	(0.049)	(0.046)			
RBC+Fast Score	0.9240	0.9433	0.8700	1.5354	6.23%
	(0.025)	(0.021)			
Individual RBC+Fast Score	0.9328	0.9502	0.8361	1.3494	8.86%
	(0.024)	(0.020)			
RBC+Ex Post Fast Factors	0.8390	0.8907	0.7992	1.799	3.60%
	(0.048)	(0.039)			
Individual RBC+Ex Post Fast Factors	0.9070	0.9383	0.8294	1.6266	5.19%
	(0.033)	(0.023)			
RBC+Consistent Fast Factors	0.8314	0.8789	0.7763	1.5245	6.37%
	(0.048)	(0.045)			
Individual RBC+Consistent Fast Factors	0.8662	0.9007	0.8548	1.4242	7.72%
	(0.046)	(0.044)			

Note: Results generated using CORROC2 written by Dr. Charles Metz, Univerisity of Chicago
Standard errors reported in parentheses

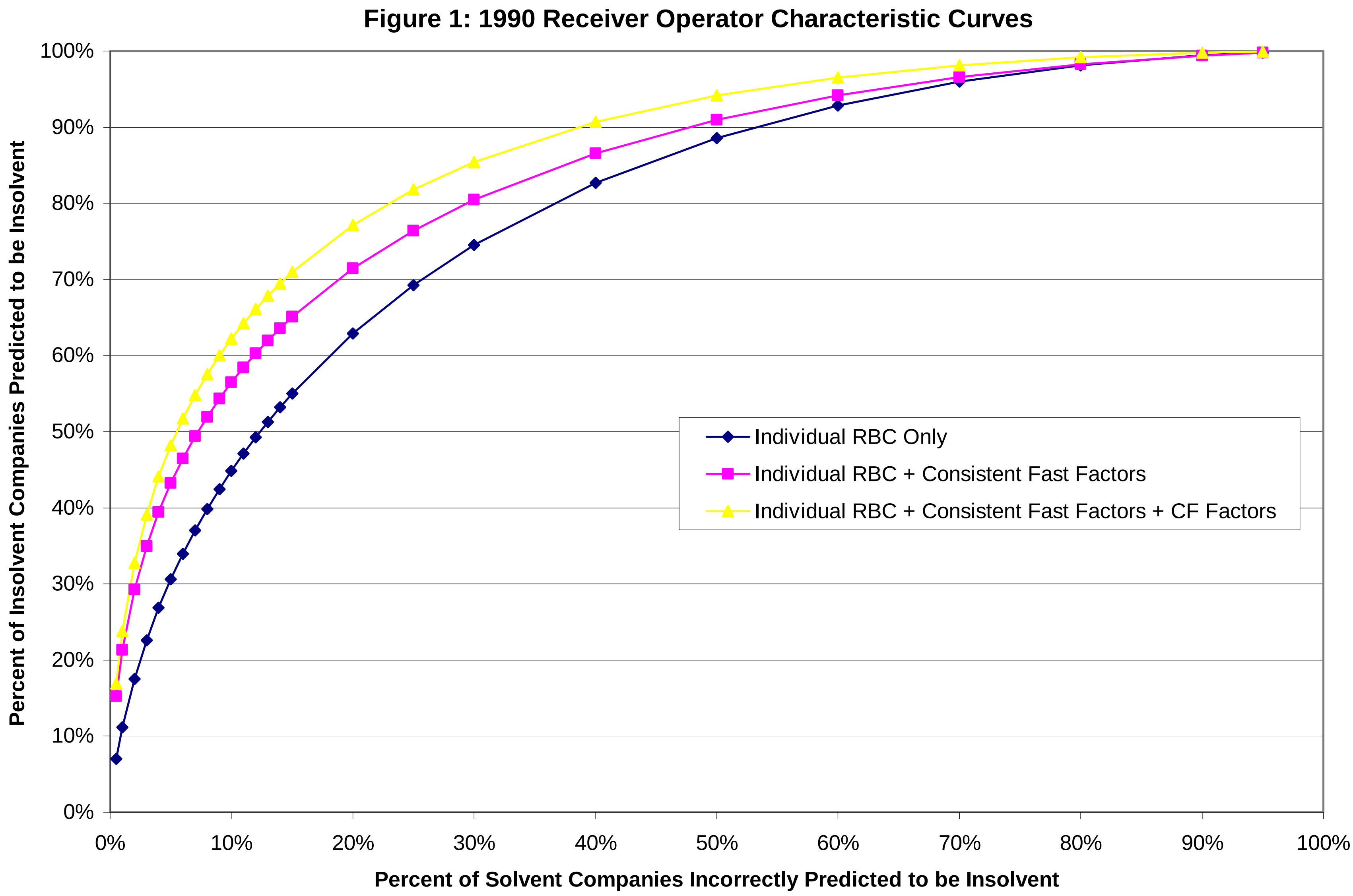
Figure 1: 1990 Receiver Operator Characteristic Curves
Percent of Insolvent Companies Predicted to be Insolvent
Percent of Solvent Companies Incorrectly Predicted to be Insolvent
0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%
Individual RBC Only
Individual RBC + Consistent Fast Factors
Individual RBC + Consistent Fast Factors + CF Factors