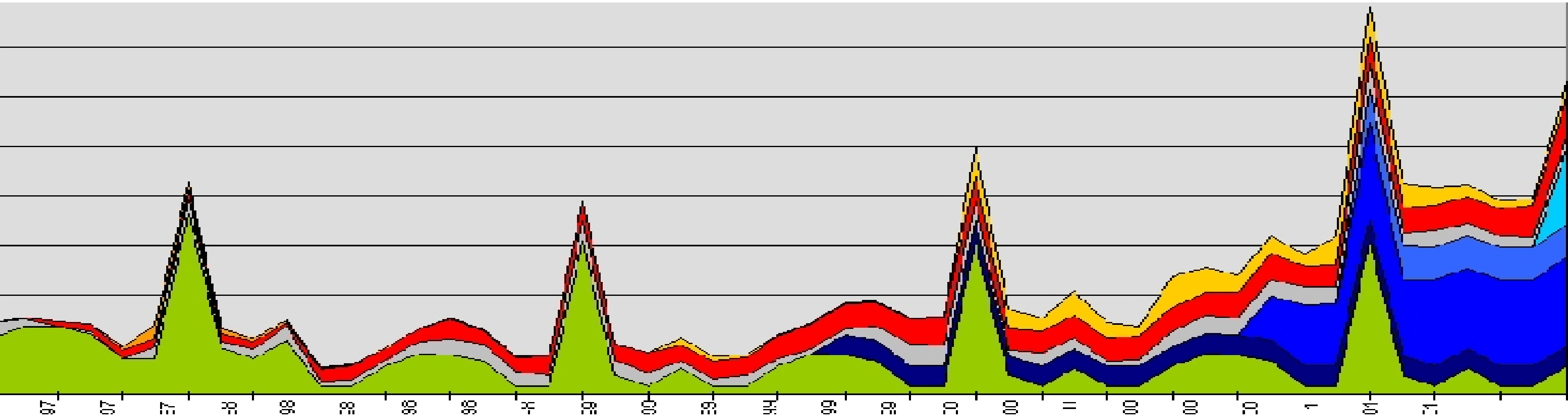
An Extension to the Cape Cod Method with Credibility Weighted Smoothing

Uri Korn, FCAS, MAAA



Because Things Change

Legal Disclaimer

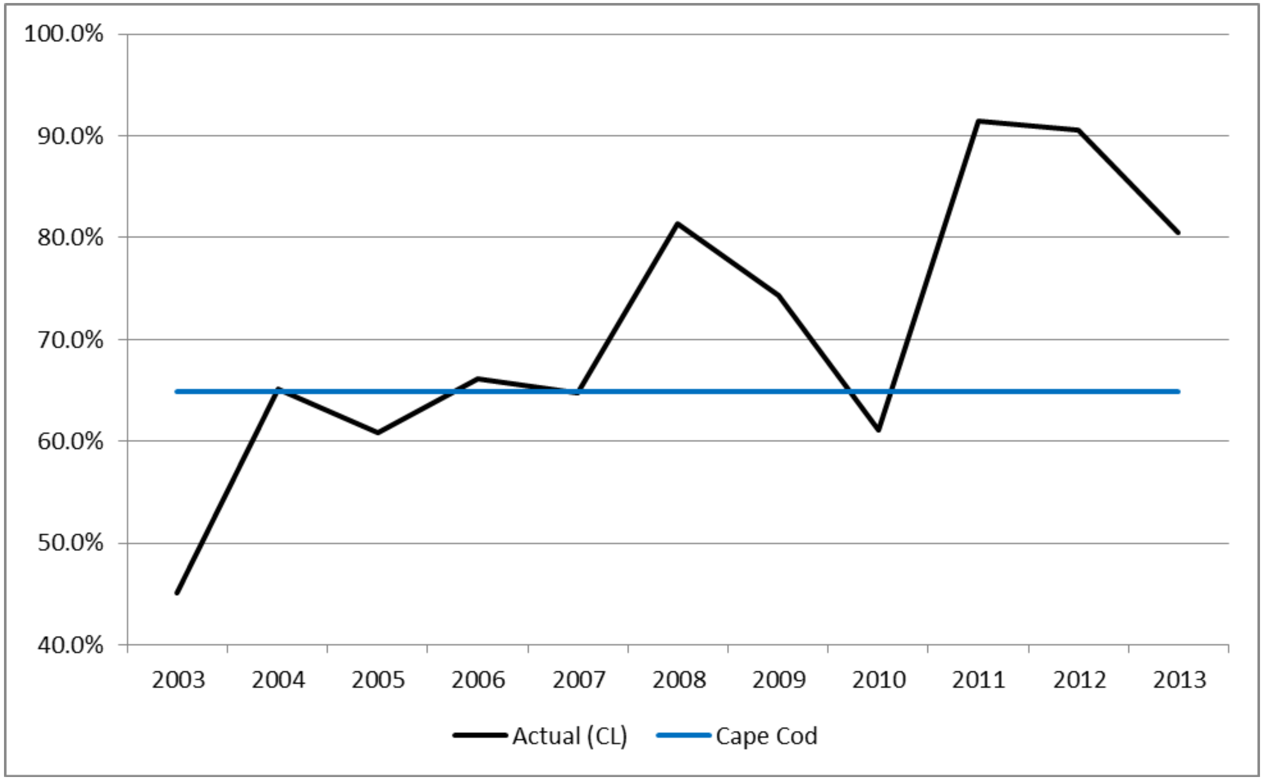
The information in this publication was compiled from sources believed to be reliable for informational purposes only. All sample policies and procedures herein should serve as a guideline, which you can use to create your own policies and procedures. We trust that you will customize these samples to reflect your own operations and believe that these samples may serve as a helpful platform for this endeavor. Any and all information contained herein is not intended to constitute advice (particularly not legal advice). Accordingly, persons requiring advice should consult independent advisors when developing programs and policies. We do not guarantee the accuracy of this information or any results and further assume no liability in connection with this publication and sample policies and procedures, including any information, methods or safety suggestions contained herein. We undertake no obligation to publicly update or revise any of this information, whether to reflect new information, future developments, events or circumstances or otherwise. Moreover, Zurich reminds you that this cannot be assumed to contain every acceptable safety and compliance procedure or that additional procedures might not be appropriate under the circumstances. The subject matter of this publication is not tied to any specific insurance product nor will adopting these policies and procedures ensure coverage under any insurance policy.

Overview

- Difficult to determine how much credibility should be given to Loss Ratio shifts by year
- Extend the Cape Cod method for automatic smoothing by year
  - Relatively simple method
  - Calculates the credibility to give to each year
  - Tricks to make robust in order to be practical
  - Can also be thought of as a credibility weighting between the chain ladder and BF methods
- Adding predictive variables
- Multiple segmentations with credibility

Reserving Challenges - Look Familiar?

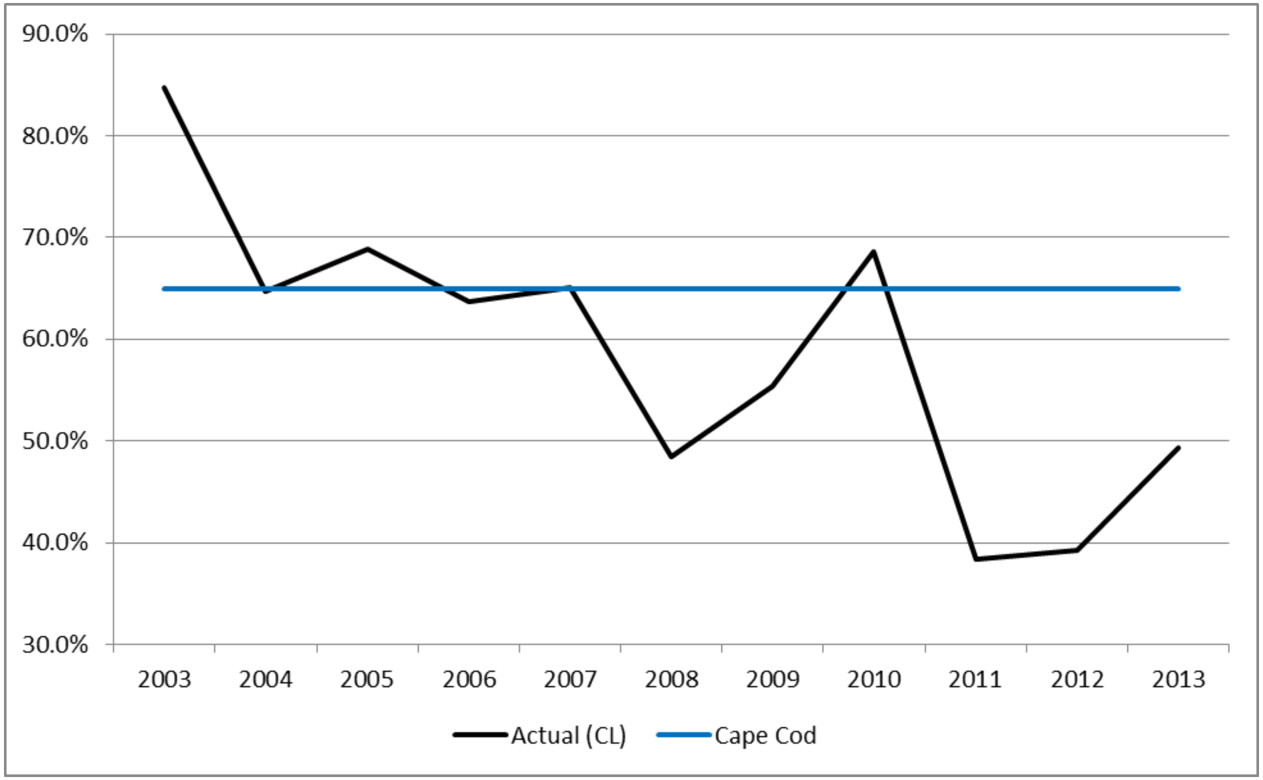
Should you react to this increase?



Simulated Data

Reserving Challenges - Look Familiar?

How about this decrease?



Simulated Data

Cape Cod Method (Review)

- Used Premium =
  Premium x Percent Losses Reported
  or: Premium / LDF

- A priori LR for BF Method =
  Total Reported Losses / Total Used Premium
  or: Weighted average of years by used premium

- (The same a priori loss ratio is selected for all years)

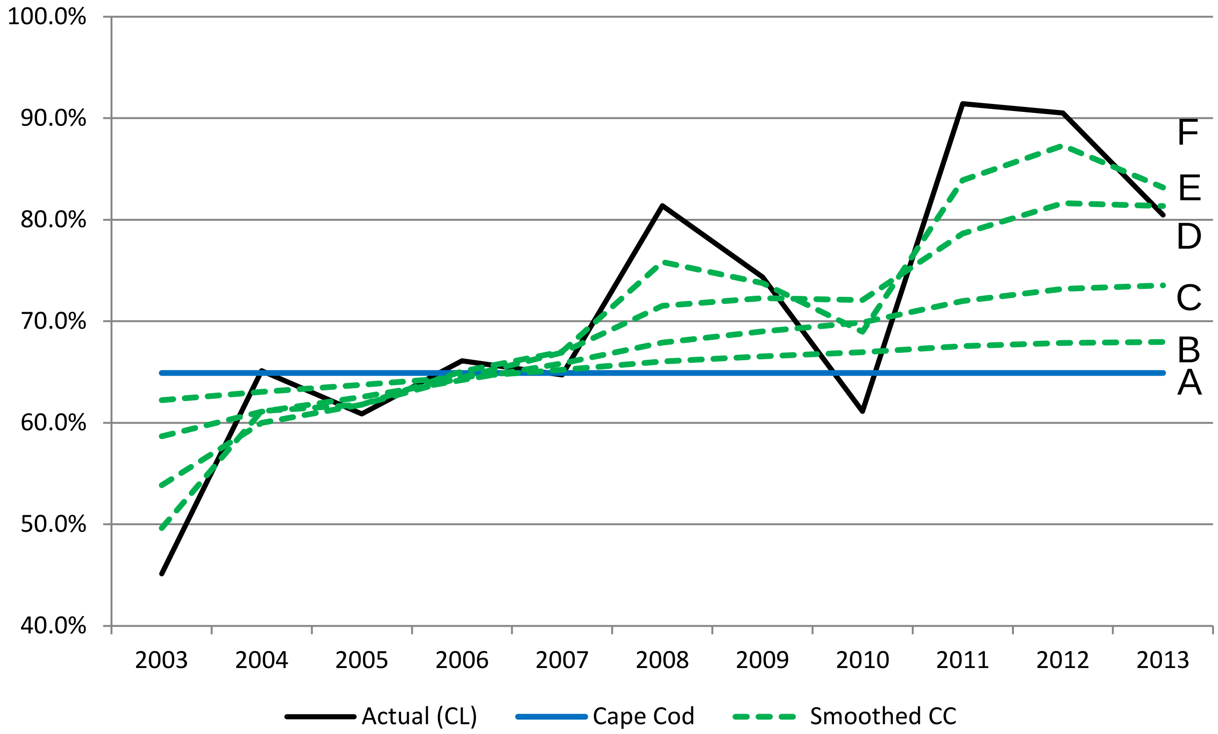
Improvement to Cape Cod Method

- Add a decay factor so that years farther away receive less weight in the a priori loss ratio selection (Gluck 1997)

- This effectively smooths the data

- But little guidance is given as to the amount of credibility/smoothness to use

Which Is Correct?



A) S = 1 (Cape Cod)

B) S = [0.9, 1)

C) S = [0.75, 0.9)

D) S = [0.5, 0.75)

E) S = [0.25, 0.5)

F) S = [0, 0.25) (CL)

Simulated Data

What is the Credibility Method for Yearly Loss Ratios?

- Buhlmann-Straub and related methods are the standard for determining credibility for a segment vs the overall average
- But, here there is an order to the experience (time series data), which is not the case with data segmentations
- What is the standard credibility method for this type of problem?

Time Series Credibility Methods

- Bayesian Approach
  - Pros: (Non-comprehensive)
    - Flexible
  - Cons:
    - Complicated
    - Requires specialized expertise and software
    - Very difficult to implement within a spreadsheet environment

- Kalman Filter
  - Pros:
    - Simpler, can be implemented in spreadsheets
  - Cons:
    - Does not handle varying volume (premium) by year
    - Does not handle non-normal errors or multiplicative changes
    - Formulas can seem complicated

Overview of the Kalman Filter

- Standard econometric method for solving these types of time series problems (in addition to Bayesian methods)
- Is analogous to Buhlmann credibility method for time series
- Developed by Rudolph Kalman in 1960 for use in signal processing
- Also used in radar systems, NASA space shuttles (such as the Apollo program), cruise missiles, and GPS (Wikipedia)

Intuition of the Kalman Filter

Predict the Expected Loss Ratio For Year 2 to be Used in a BF Method
(assuming no rate changes or trend, and that all years have the same loss volatility)

- Loss Ratio Year 1 = 70% (For certain)
- Projected Loss Ratio Year 2 = 80% (Chain Ladder)

- If very low loss volatility
  - Prediction for Year 2: 80%

- If very high loss volatility
  - Prediction for Year 2: 70%

- If loss volatility = volatility of year-to-year changes
  - Prediction for Year 2: 75%

Intuition of the Kalman Filter

Predict the Expected Loss Ratio For Year 2 to be Used in a BF Method (Cont.)

- More generally:
  - Prediction for Year 2:

    $Z = Q / (Q + R)$

    Where
    R = Variance of Experience
    Q = Variance of Year-to-Year Changes

  - Variance of Estimate: P(2)
    $1 / P(2) = 1 / Q + 1 / R$
    (Exact for Gaussian, approximation otherwise)

- Assume Q = 0.5, R = 1.5
  - $Z = 0.5 / (0.5 + 1.5) = 0.25$
  - LR(2) = 70% x 0.75 + 80% x 0.25 = 72.5%
  - Var[Year 2 Estimate] = P(2) = 1 / (1 / 0.5 + 1 / 1.5) = 0.375

Intuition of the Kalman Filter

Predict the Expected Loss Ratio For Year 3

- Projected Loss Ratio Year 3 = 90% (Chain Ladder)

- Variance of using the Year 2 Estimate for Year 3 = Variance of the Year 2 Prediction (Calculated Above) + Volatility of Year-to-Year Changes.
  (P(2) + Q = 0.375 + 0.5 = 0.875)

- Compare this to the volatility of the experience
  - Z = [P(2) + Q] / [P(2) + Q + R] = 0.875 / (0.875 + 1.5) = 0.37
  - LR(3) = 72.5% x 0.37 + 80% x 0.63 = 77%

- Variance:

  1 / P(3) = 1 / [P(2) + Q] + 1 / R

  = 1 / (1 / 0.875 + 1 / 1.5) = 0.553

- And so on…

Intuition of the Kalman Filter

Back-Smoothing

- Now, we can use this year 3 estimate to improve the final year 2 estimate

- Similar formulas are used. Credibility given to Year 3 prediction for Year 2:

  $$Z = P(3) / [\ P(3) + Q\ ] = 0.553 / (\ 0.553 + 0.5\ ) = 0.525$$

  Final LR(2) = 77% x 0.525 + 72.5% x 0.475 = 75%

- And so on…

Solving for the Parameters

- The unknown parameters, Q, R, and the starting LR are all solved via Maximum Likelihood
- When solving, use the estimates for each year before considering the year's experience (which is basically the estimate of the previous year)
  - Otherwise, the method will smooth exactly to the experience
- Calculate the error between this and the actual observed loss ratio for each year
- Variance of Observation = Parameter Variance + Process Variance = P(Previous Year) + Q + R
- Likelihood = Normal(Error, Variance)
  - (Note we will change this part)

Changes to the Formula for Loss Ratios

- Varying Volume (Premium) per Year:
  - Instead of using a parameter for the yearly variance, use a variance factor parameter
  - Variance per Year = R / Volume

- Incomplete Years:
  - Define: Volume = Used Premium = Premium / LDF
  - Also, observed LR = Reported LR x LDF
    - (Note: this is just the input to the method)
  - More incomplete years will have greater variance
  - Consistent with the Cape Cod approach

Changes to the Formula for Loss Ratios

- Non-normal errors and multiplicative changes per year: Do NOT use a log-transformation
  - Difficult to determine the weights per year
  - Would not be consistent with the Cape Cod method
  - Requires a messy bias correction

Changes to the Formula for Loss Ratios

- Non-Normal Errors:
  - Instead of calculating the likelihood using a normal distribution, calculate the likelihood using a Gamma distribution as well. Then take a weighted average of the log-likelihoods.
    - For each year, convert the variance to a coefficient of variation by multiplying by a new parameter
    - Using the expected mean and coefficient of variation for each year, back into the Gamma parameters and calculate the Gamma likelihood
    - final log-likelihood = (p / 2) Gamma log-likelihood + (1 - p / 2) Normal log-likelihood (where p = Tweedie power, often 1.67)
  - This comes close to approximating a Tweedie and is much simpler to implement

Changes to the Formula for Loss Ratios

- Multiplicative Changes per Year:
  - If a normal distribution is used, the year-to-year changes are assumed to be additive
  - If a Gamma distribution is used, the year-to-year changes are assumed to be multiplicative (since the standard deviation, which is related to the magnitude of the error, is proportional to the mean)
  - If a "Tweedie" is used, the year-to-year changes are assumed to be in between additive and multiplicative
  - To make them multiplicative, for the variance of the year-to-year changes, use this instead:

    Q x LR ^ (2 – p)

    Where LR is the loss ratio estimate for each year before considering the year's experience and p is the Tweedie power

Final Formulas

- Refer to the paper
- (Note that K = Z, and is called the Kalman Gain)

- Or just use the spreadsheet on the CAS website

Application of Method

- Apply this method to the Chain Ladder ultimate loss ratios per year
- Volume for each year = Used Premium
- The result is the true expectation for each year (which may be different from what actually occurred, due to random fluctuation)
- Use this result in a BF method, since the expectation is that the remainder of the year will develop according to this expectation

Credibility Weighting Between BF and CL

- Since credibility is given to the year itself as well, this method can also be thought of as a credibility weighting between the BF and Chain Ladder methods

  – When full credibility/smoothness is indicated (Z = 100% for all years)
    – The results of the method will match the Chain Ladder indications
    – Performing a BF using this is equivalent to the Chain Ladder

  – Is no credibility/smoothness is indicated (Z = 0% for all years)
    – The result will be the weighted average of the Chain Ladder loss ratios with weights equal to the used premium
    – This is equivalent to the BF method

This is all great, but it won’t work

- Actual experience is often too volatile to accurately estimate the required parameters (depending on the data…)

- This procedure requires more data points than typically used in the reserving context

Robustifying the Method

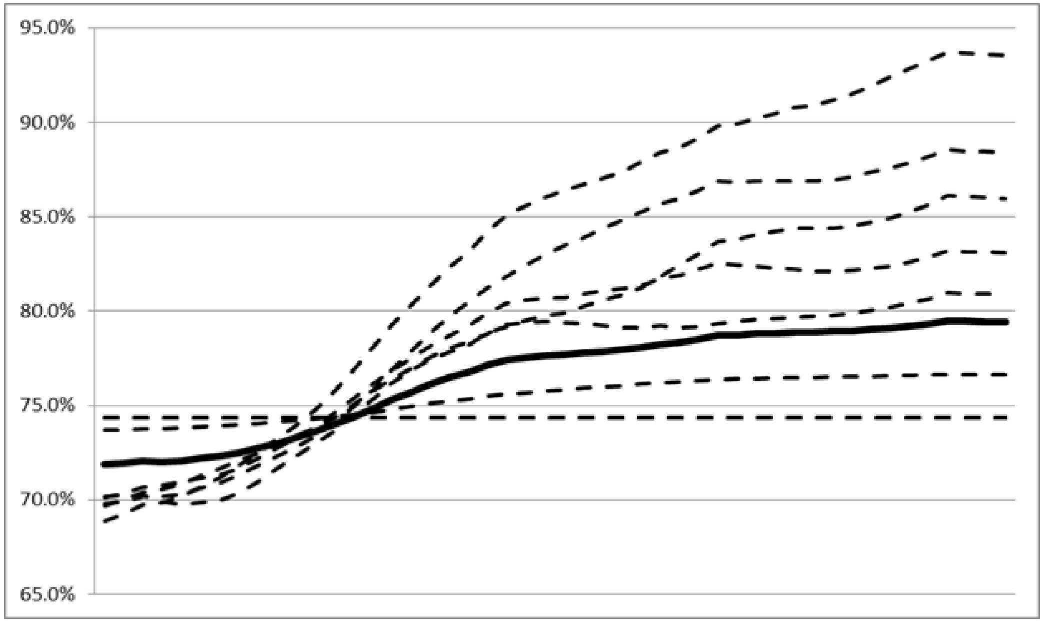
- Simple solution: use quarterly data, especially if < 20 years (?) of data are being used
  - Increases the number of data points four-fold
  - If seasonality by quarter may be a factor, this can be incorporated (and even credibility weighted) similar to a predictive variable (discussed later)

Robustifying the Method

- Use “Bagging” (Bootstrapped Aggregation)
  - Borrowed from machine learning
  - Perform multiple iterations of the method, each time including only a fraction of the data
  - (To skip a point, just set the credibility to 0, but still include in the likelihood)
  - The final indicated a priori loss ratios are then calculated as the average loss ratios across all of the iterations
  - Each iteration will receive a varying amount of smoothness, and averaging across all of these produces a much more stable and reliable result
  - Using 2/3 of the data each iteration and 50 iterations seems to perform well

Bagging Example

- The first 10 runs (from another example) as well as the run that resulted in the most amount of smoothness (highest Z) are shown as dotted lines
- Many (17/50) runs resulted in no smoothness (Cape Cod, Z = 0%)
- The solid black line is the average across all runs
- The key to bagging is that it takes advantage of an unstable method and uses it to produce an overall better result

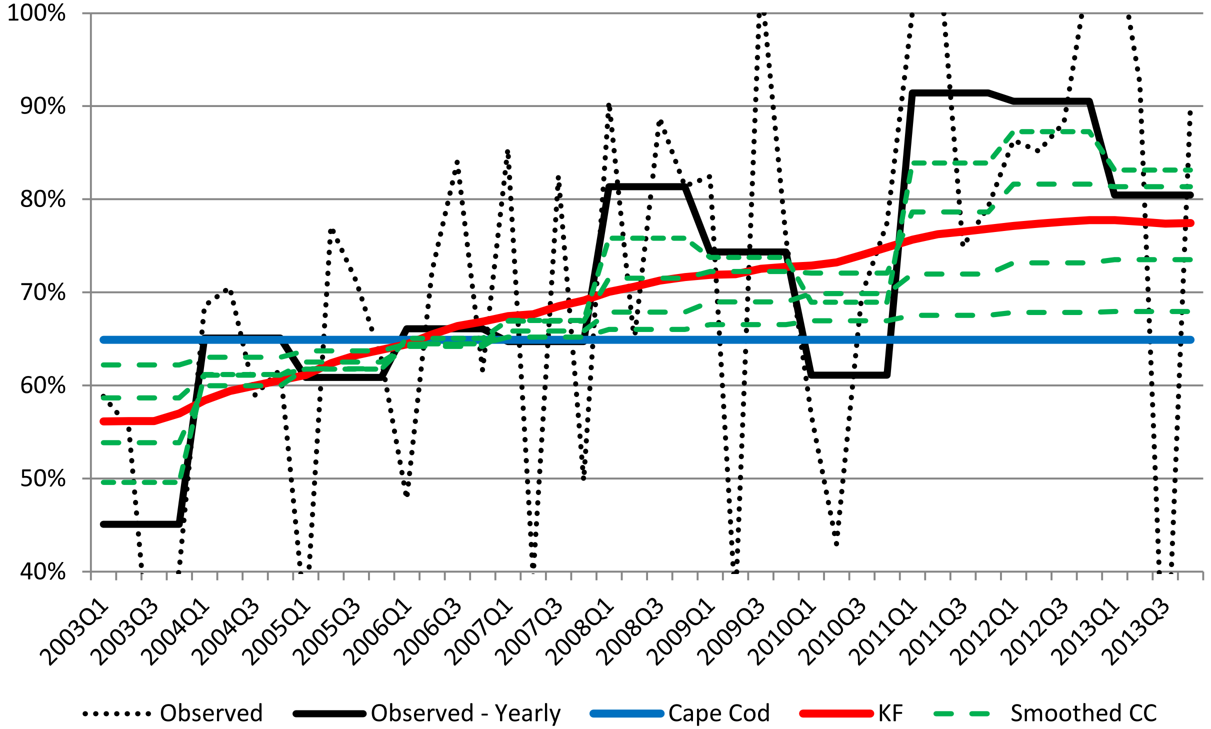


Accident Quarter

Simulated Data

Now We're Ready

Running the Example

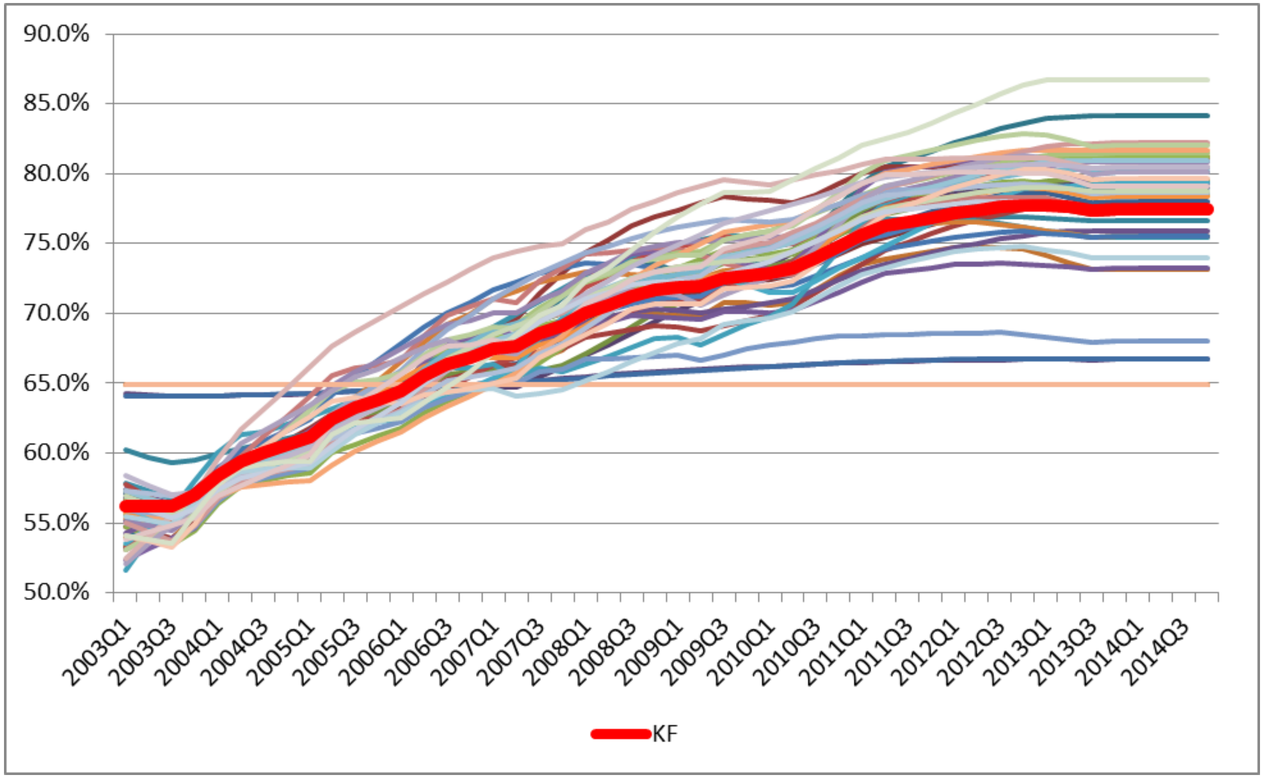


Answer: Cannot Be Determined From the Yearly Data!

But with quarterly data, we decide on about 0.6 (D) for this example (although this method is not really the same as exponential smoothing)

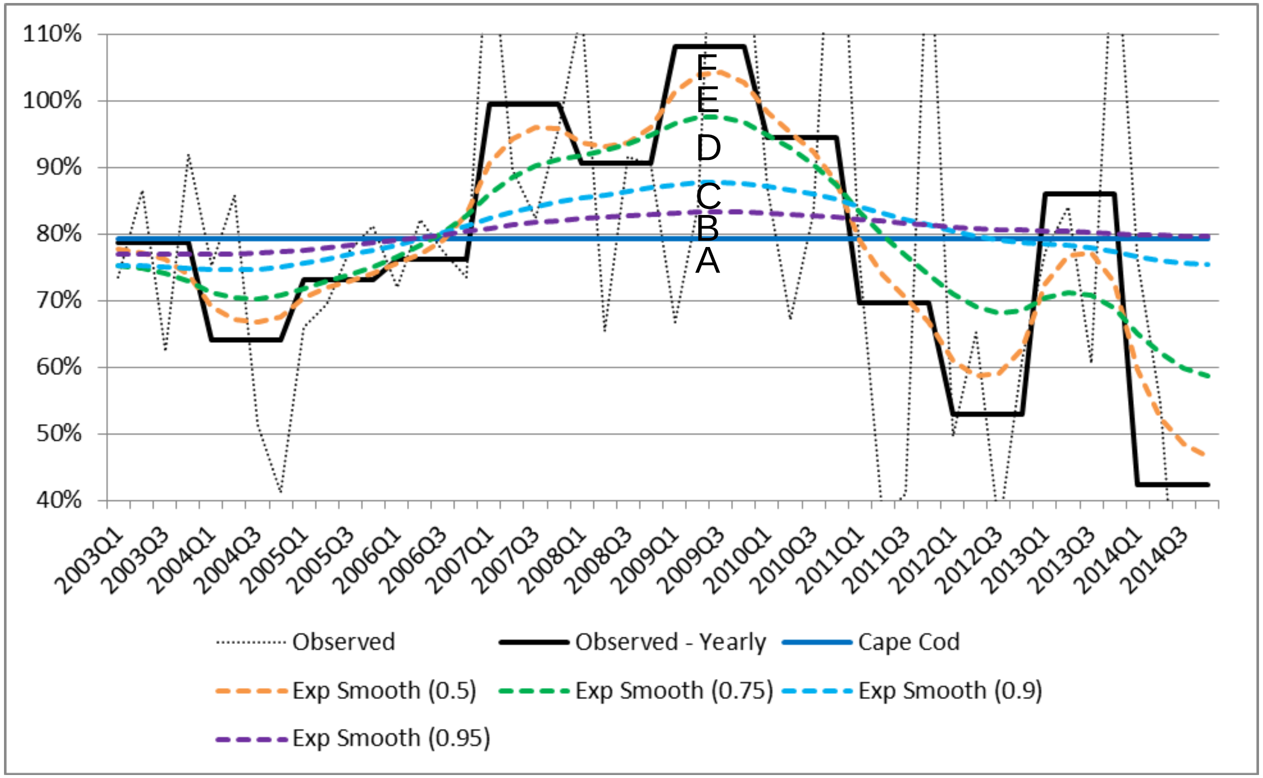
Simulated Data

Bag Results



90.0%
85.0%
80.0%
75.0%
70.0%
65.0%
60.0%
55.0%
50.0%
2003Q1
2003Q3
2004Q1
2004Q3
2005Q1
2005Q3
2006Q1
2006Q3
2007Q1
2007Q3
2008Q1
2008Q3
2009Q1
2009Q3
2010Q1
2010Q3
2011Q1
2011Q3
2012Q1
2012Q3
2013Q1
2013Q3
2014Q1
2014Q3
KF

Example 2



A) S = 1 (CC)

B) S = [0.95, 1)

C) S = [0.9, 0.95)

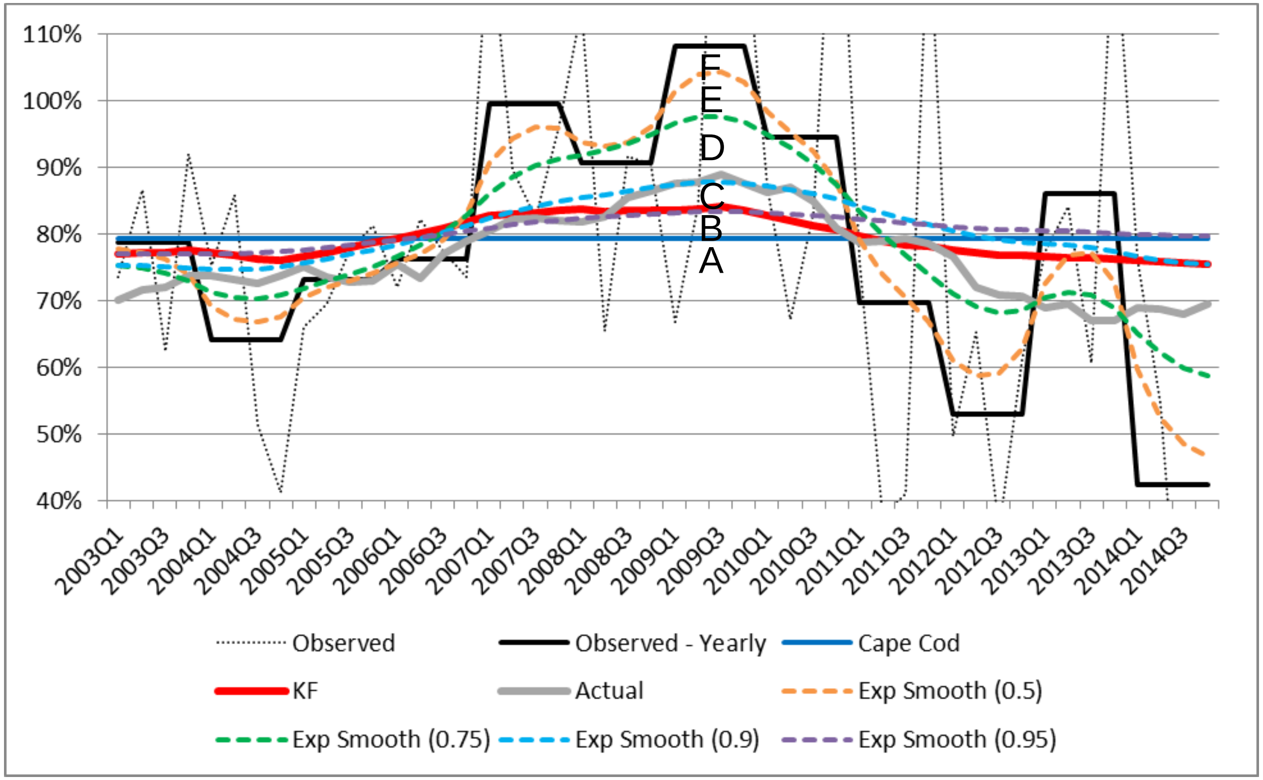
D) S = [0.75, 0.9)

E) S = [0.5, 0.75)

F) S = [0, 0.5) (CL)

Simulated Data

Answer 2

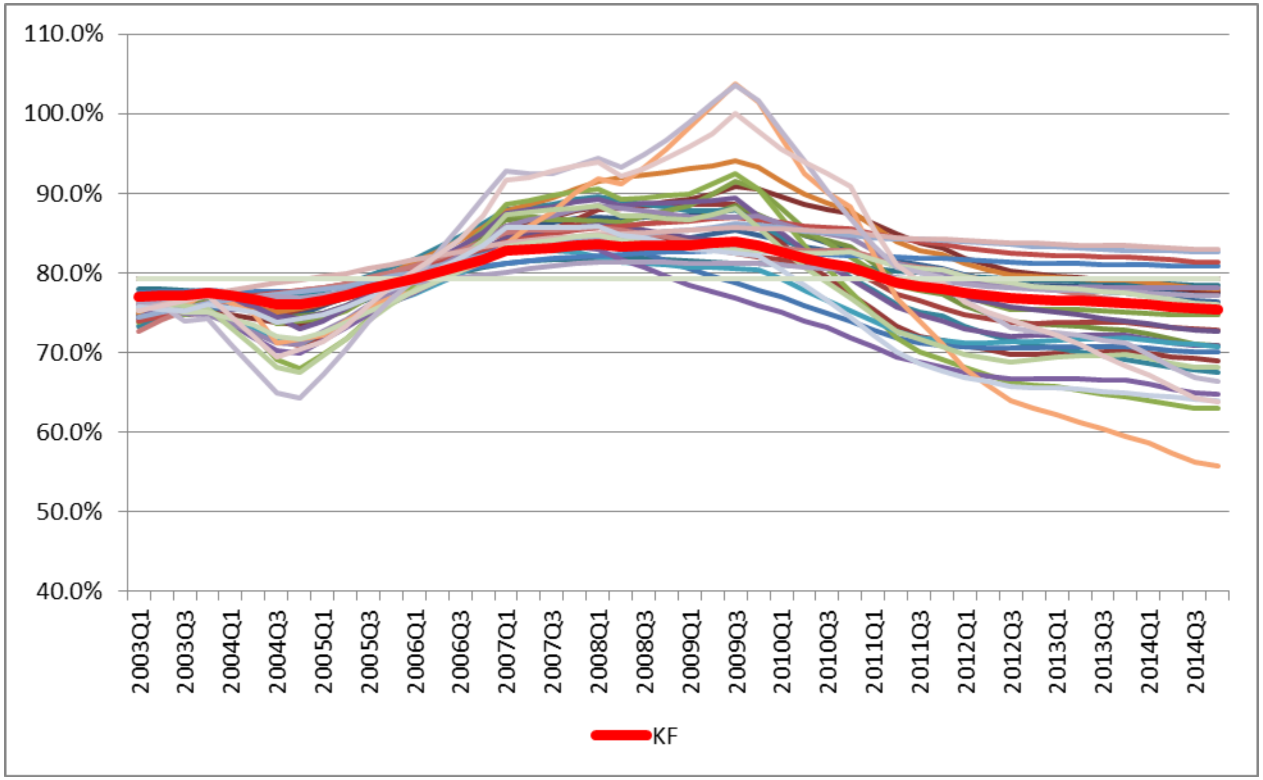


Answer: C or D (C is probably closer)

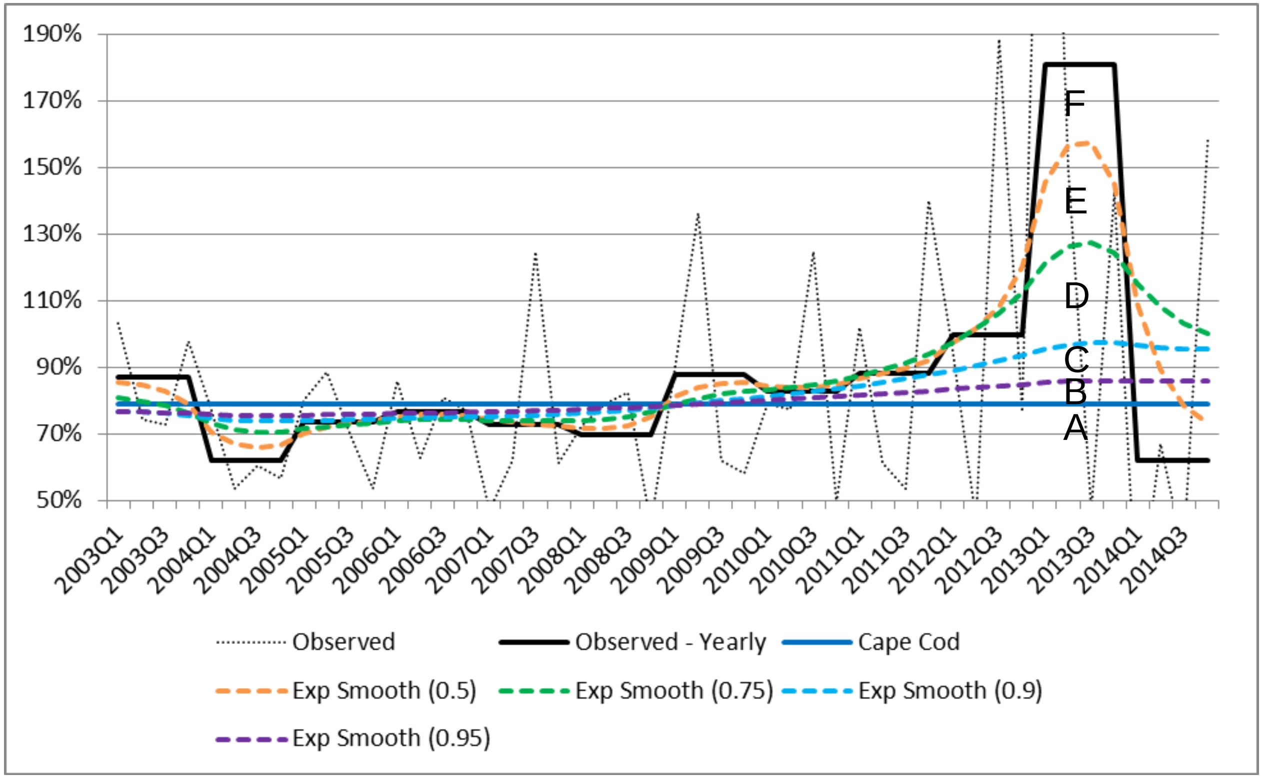
Slightly less than 0.95 until 2010, and then ~ 0.88

Simulated Data

Bag Results 2



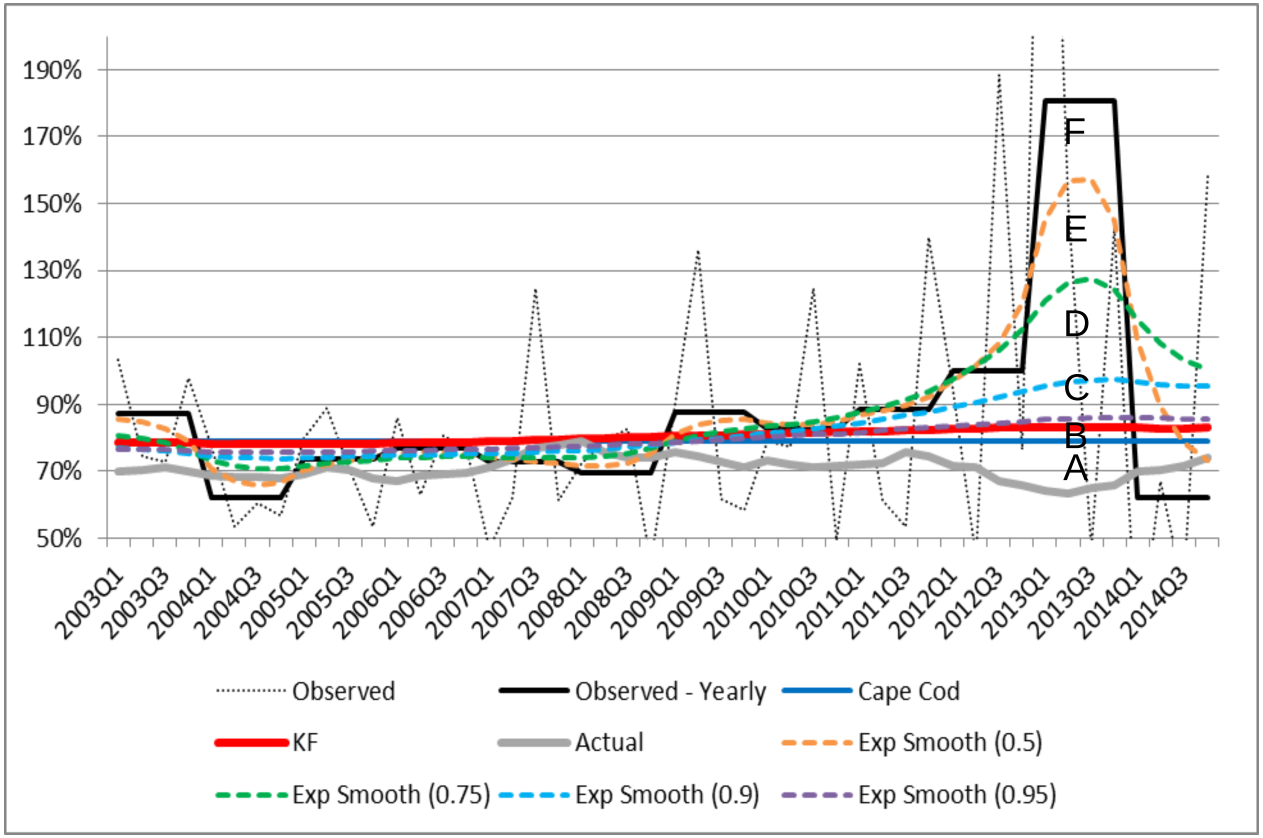
Example 3



A) S = 1 (CC)
B) S = [0.95, 1)
C) S = [0.9, 0.95)
D) S = [0.75, 0.9)
E) S = [0.5, 0.75)
F) S = [0, 0.5) (CL)

Simulated Data

Answer 3



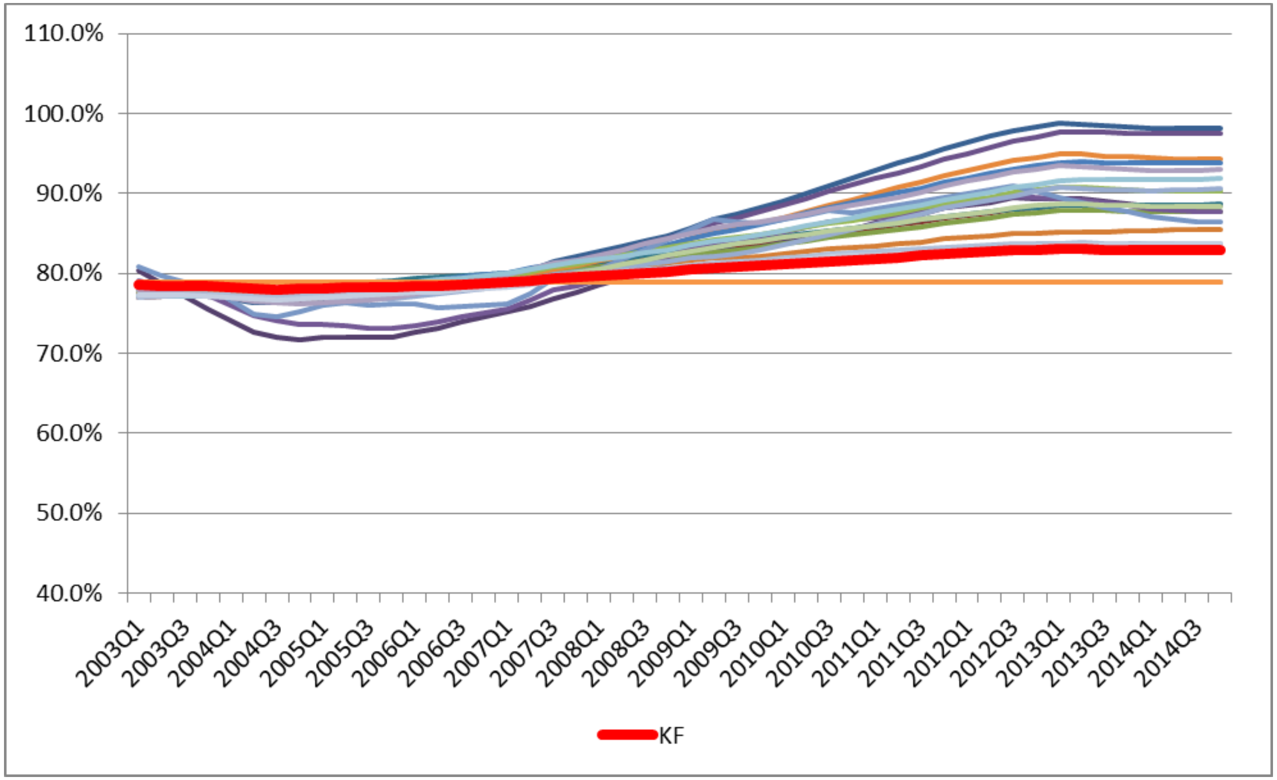
Answer: ~0.97 (B)

(We will consider A correct as well)

Simulated Data

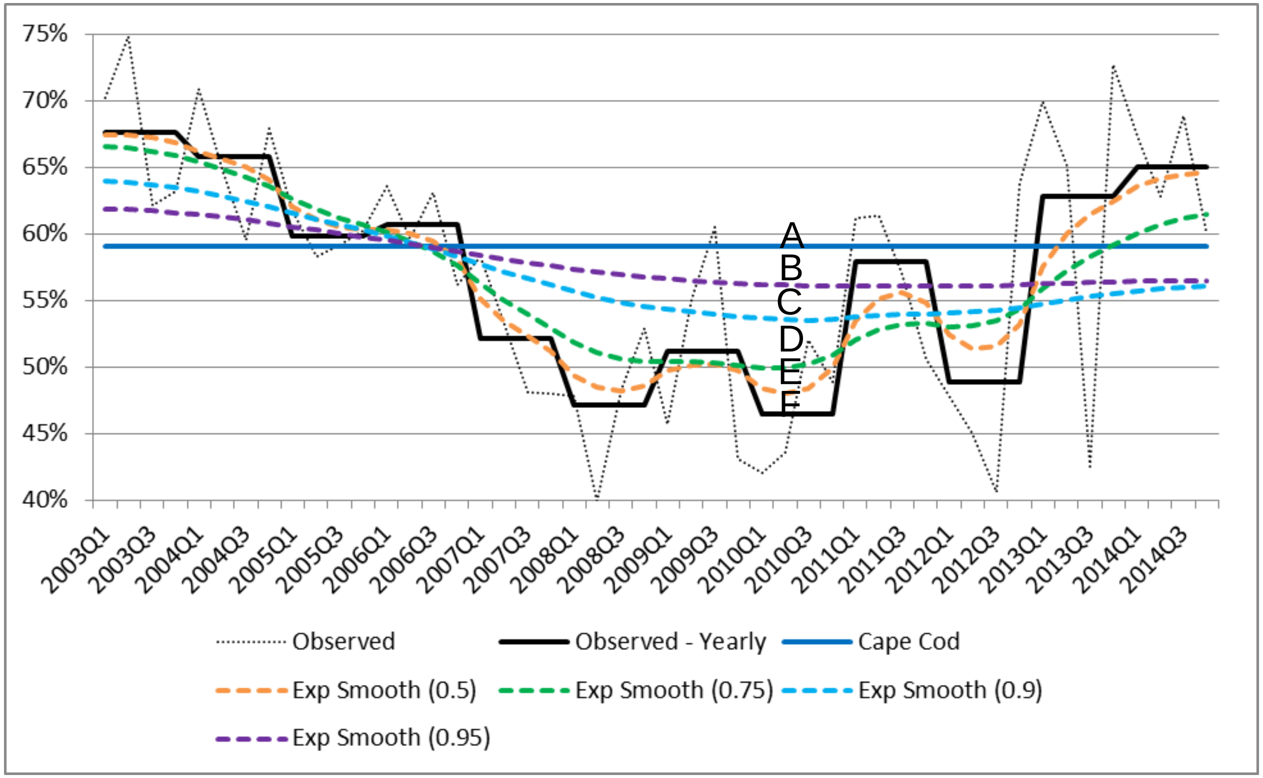
(Note that the gray line should be higher as the simulation used a lognormal distribution and did not correct for the bias)

Bag Results 3



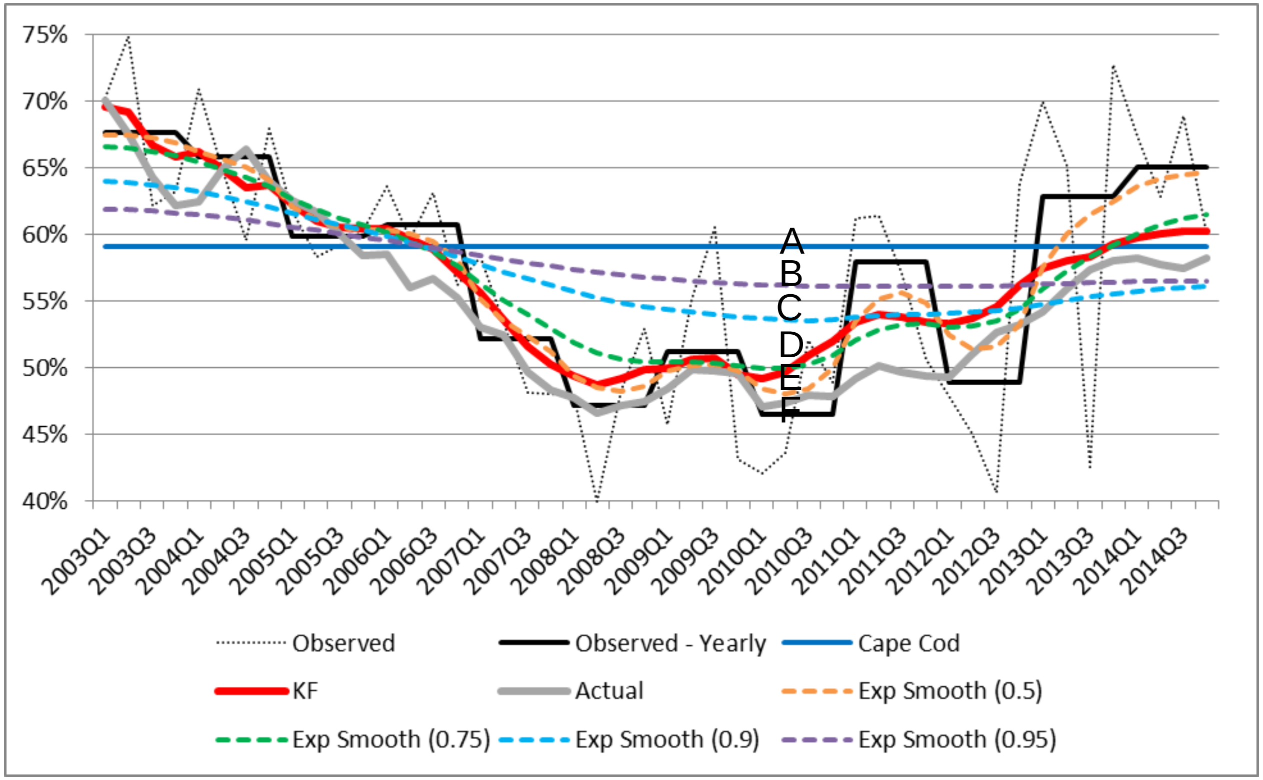
- Cannot see from the picture, but most of the bag iterations produce the Cape Cod (or S = 1)

One Last Example – Lower Volatility



A) S = 1 (CC)
B) S = [0.95, 1)
C) S = [0.9, 0.95)
D) S = [0.75, 0.9)
E) S = [0.5, 0.75)
F) S = [0, 0.5) (CL)

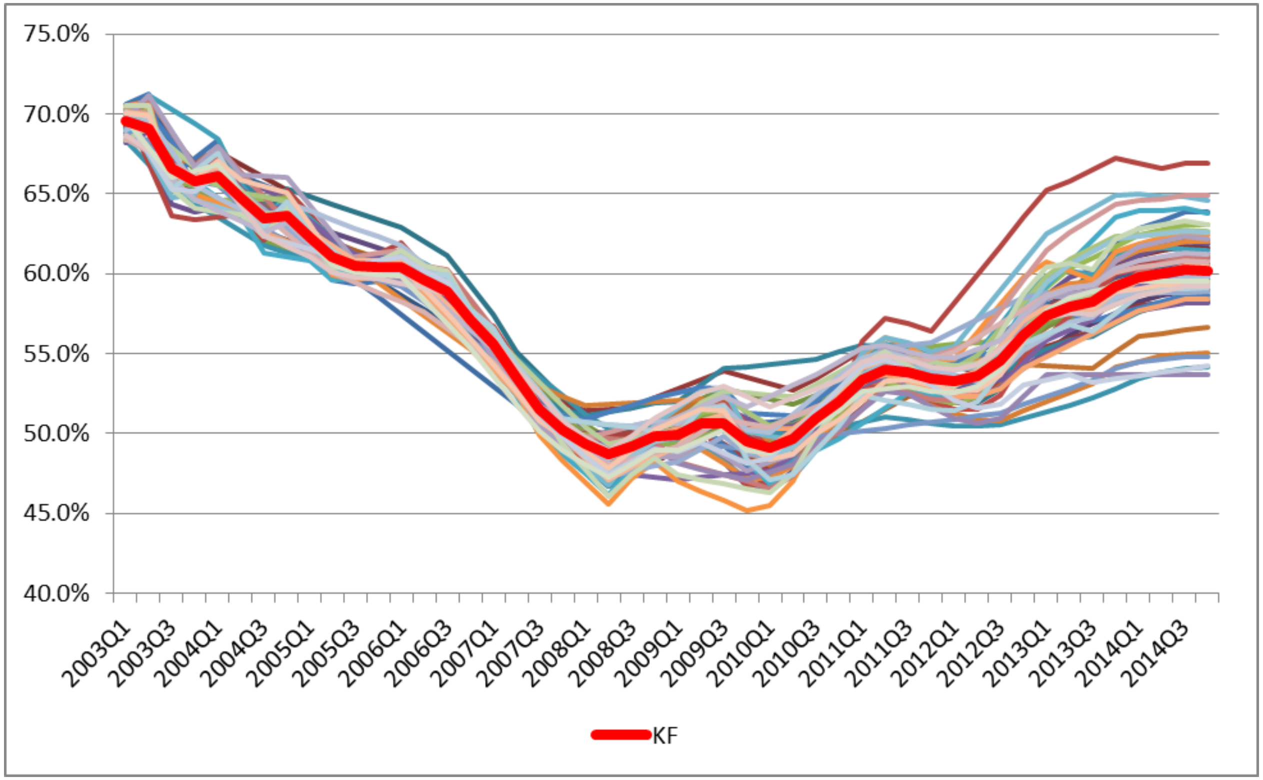
One Last Example – Lower Volatility



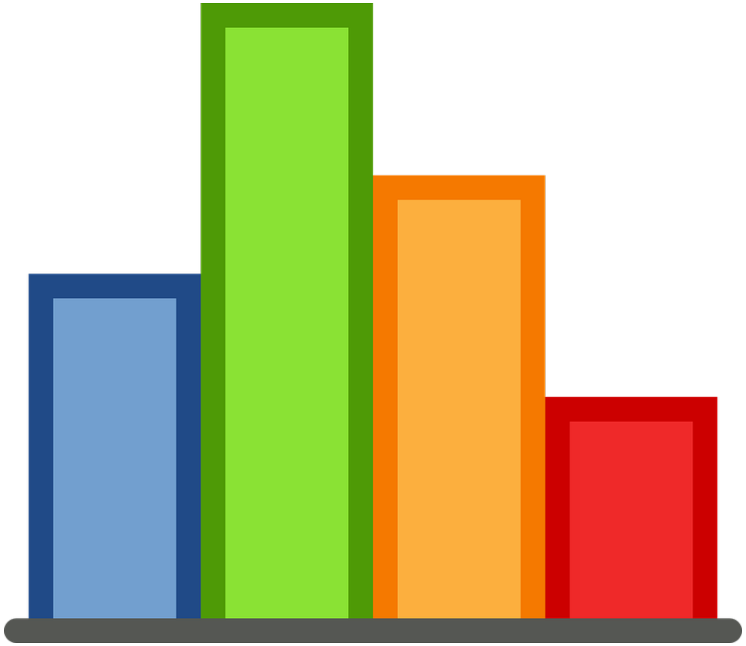
- Answer: E - Smoothing is between 0.5 and 0.75 (orange and light blue)

(Note that the gray line should be higher as the simulation used a lognormal distribution and did not correct for the bias)

Bag Iterations



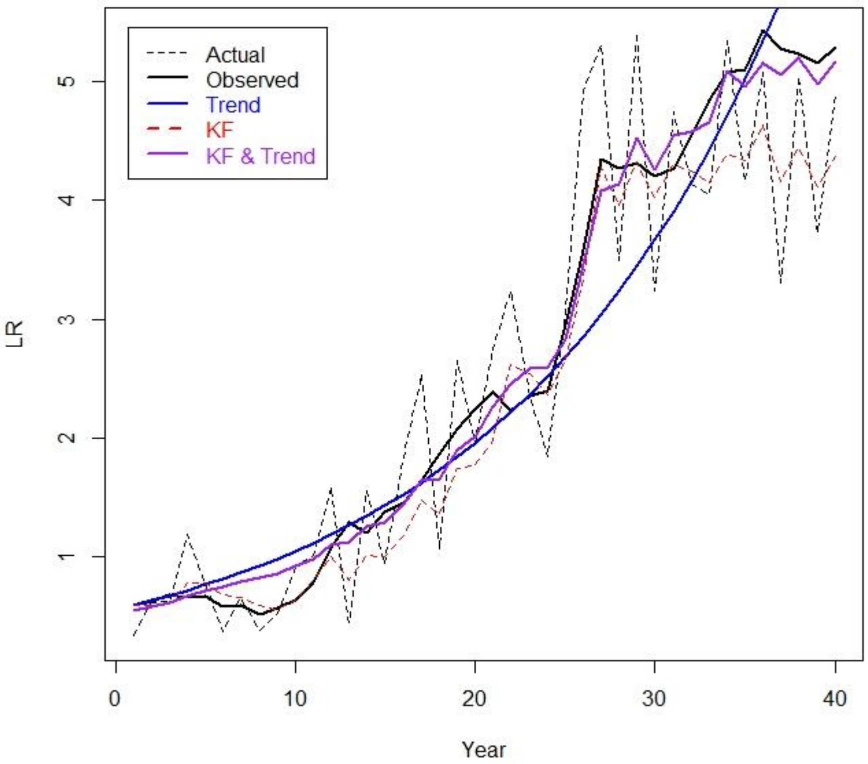
Incorporating Predictive Variables



- The Kalman Filter also allows for predictive variables, such as the state of the economy or the market cycle, to be incorporated in the process
  - Modified formulas are shown in the paper to account for the changes in the expected values and variances
  - Can also be thought of as a regression equation with a varying intercept

Incorporating Predictive Variables

- As a simple example, use the year as a predictive variable to estimate trend, but also allow for additional changes by year



Simulated Data

Credibility/Stability For Predictive Variables

- Add credibility weighting for categorical variables and stability to numerical variables by applying a penalty for larger coefficients
- Use a Ridge Regression type methodology
- Add this to the total log-likelihood:
  - ∑ log(Norm(Coef(i), 0, Penalty Variance))
- Values further away from 0 will lower the likelihood, thus the maximization routine will cause the coefficients to be “pushed” towards 0 based on their amount of credibility
- (With this method, all values of a variable should be assigned coefficients)

Credibility/Stability

- The same Penalty Variance is usually used for all variables
- For this to make sense, all (non-dummy) variables should be standardized first to set them to the same scale
  - (X - Mean) / Standard Deviation (if no dummy variables)
  - Or, if dummy variables are being used as well: X / (2 x Standard Deviation)   (Gelman 2008)
- Using this method improves the performance of predictive variables and also lessens the effect of non-predictive variables that were accidentally included

Credibility/Stability

Solving for the Penalty Variance

- Use cross validation to solve for the Penalty Variance:
  - Test various candidate variance values
  - For each, fit the model on a fraction of data
  - Use the remaining data to calculate the error: (sum of squared error divided by the mean to the Tweedie power)
  - Repeat several times to add stability
  - Use the same fit and test fractions - this greatly decreases the number of iterations needed
  - Hopefully, a nice curve should result - otherwise, more iterations may be needed
  - Select the variance with the lowest test error

Multiple Lines of Business

- Multiple lines of business can be ran together leveraging the same variance parameters, but allowing the starting loss ratios to differ

- Going one step further, credibility weighting between the starting loss ratios can be performed as well

- Similar to credibility for predictive variables, add a penalty term to the total log-likelihood for each line:
  - ∑ log(Norm(LR(i), Complement LR, LR Between Variance))
  - (The Complement LR is added as an additional parameter)

- The “LR Between Variance” can be solved for using Buhlmann-Straub, or more ideally via cross validation (as shown previously)

But, this also doesn't work…

- The problem with this credibility weighting method is that "pushing" the initial loss ratio estimates towards the mean, but then letting them vary freely afterwards often generates results that deviate outwards away from the mean with time, even if this is not the case

Solution

- Instead of applying the credibility penalty to the starting loss ratios, apply it to the ending loss ratios
- These ending loss ratios can be considered the midpoint, since they are the starting point of the back-smoothing iterations
- This is statistically justified by pretending that we are inverting the Bayesian credibility equation…- see the paper
- This method produces better results that do not artificially deviate either towards or away from the mean

Additional Misc. Notes

- Applying this method to frequency and severity separately can often do a better job of capturing the true signal in the data

- The Kalman Filter is able to solve for a vast array of flexible time series models. This method only scratches the surface.

Conclusion

- Estimating expected loss ratios per year with volatile data can often be a confusing and difficult task, subject to a large degree of judgment
- The goal of this paper is to hopefully improve this process by lending support from modern statistical techniques without losing the simple and intuitive nature of the Cape Cod method

Questions?

List of Sources

- [1] Bolstad, W. 2007. Introduction to Bayesian Statistics (Second Edition). Wiley. 2007, p.106-108, 207-209
- [2] De Jong, P. and Zehnwirth, B. 1983. Claims reserving, state-space models and the Kalman filter. Journal of the Institute of Actuaries, 110, pp 157-181.
- [3] Evans, Jonathan P., and Frank Schmid. 2007. Forecasting Workers Compensation Severities and Frequency Using the Kalman Filter, Casualty Actuarial Society Forum, 2007: Winter, pp 43–66
- [4] Gelman, A. 2008. Scaling regression inputs by dividing by two standard deviations. Stat. Med. 27: 2865–2873.
- [5] Gluck, S. M. 1997. Balancing Development and Trend in Loss Reserve Analysis. PCAS 84, pp 482–532.
- [6] Herzog, T. 1989. Credibility: The Bayesian Model Versus Buhlmann Model. Transactions of the Society of Actuaries, 1989, Volume 41.
- [7] Kalman RE. A New Approach to Linear Filtering and Prediction Problems. ASME. J. Basic Eng. 1960;82 pp. 35-45.
- [8] Kim C. and Nelson C. 1999. State-Space Models With Regime Switching. MIT Press, 1999. pp. 29-30, 36-37
- [9] Korn, U. 2015. Credibility for Pricing Loss Ratios and Loss Costs, Casualty Actuarial Society E-Forum, 2015: Fall, pp 327–347
- [10] Schmid, F., Laws, C., & Montero, M. 2013. Bayesian Trend Selection. In Casualty Actuarial Society E-Forum, Spring 2013.
- [11] Stanard, James N. 1985 A Simulation Test of Prediction Errors of Loss Reserve Estimation Techniques. PCAS LXXII, 1985, pp.124–148.
- [12] Taylor, G., & McGuire, G. 2007. Adaptive Reserving using Bayesian revision for the Exponential Dispersion Family. Centre for Actuarial Studies, Department of Economics, University of Melbourne.
- [13] Venter, Gary G. 2003. Credibility Theory for Dummies. Casualty Actuarial Society Forum, 2003: Winter, pp 221-267
- [14] Wüthrich, M.V. and Merz, M. 2008. Stochastic Claims Reserving Methods in Insurance. Wiley.
- [15] Zehnwirth B. 1996. Kalman filters with applications to loss reserving. Research Paper Number 35, Centre for Actuarial Studies, The University of Melbourne.