Using the ODP Bootstrap Model: A Practitioner's Guide
Model Instructions

In the "ODP Bootstrap Practitioners Guide.zip" file you will find all of the files described in the Supplementary Materials section of the "Using the ODP Bootstrap Model: A Practitioner's Guide" paper. The primary ODP bootstrap modeling files you can use to review the full details of the calculations of the modeling framework are:

1. Model Instructions.pdf (this PDF document)
2. Industry Data.xlsm (Schedule P data as of 12/31/08)
3. Bootstrap Models.xlsm
4. Best Estimate.xlsm
5. Aggregate Estimate.xlsm
6. Correlation Ranks.xlsm

The last four Excel files contain the bootstrap models described in the paper. Reading the paper will guide you when exploring the educational content of the model files, but not all of the issues and model variations discussed in the paper are reproduced in these Excel files.

All of the sheets in each file are protected (without a password) to maintain the integrity of the calculations, but the unprotected data entry cells are noted in **blue**. You may unprotect the sheets to review the calculations, but you may not use a password when protecting the sheets again as this will cause an error in the routines. You must set the Macro Security level low enough to enable macros before any of these files are opened. The Bootstrap Models file is set up to run a 10 x 10 triangle or less, but NO larger size. All data must be annual and no data cells can be missing. Finally, you must not change any of the range names stored in the files as this may create problems with the routines included with the files.

Bootstrap Models:

The **Bootstrap Models** file is the first file you should open. There are nine visible sheets in the file:

A. *Inputs* – on this sheet you can enter data for a business segment and choose model options.

B. *Bootstrap* – on this sheet you will find the calculations for the paid and incurred bootstrap models as well as the substitute steps for the Bornhuetter-Ferguson, Cape Cod and GLM Bootstrap models. You can also select hetero groups and outliers for each model.

C. *Diagnostics* – on this sheet you will find the diagnostic graphs described in section 5 of the paper. When the "GLM Bootstrap" model is selected on the Input sheet, the parameters for that model are selected on this sheet.

D. *Groups* – on this sheet you will be able to calculate optimal hetero groups which can be entered as a parameter in the *Bootstrap* sheet.

E. *GLM* – on this sheet you will be able to calculate parameters for the GLM Bootstrap model.

F. *Calcs* – on this sheet you will find most of the details of the calculations related to the *Bootstrap* and *Diagnostics* sheets.

G. *Residuals* – on this sheet you will find arrays of the residuals used for sampling.

H. *Hat Matrix* – on this sheet you will find the details of the hat matrix calculations used in the *Bootstrap* sheet.

I. *Results* – on this sheet the results of your simulations will be stored and summarized.

There is also one hidden sheet in the file where the random numbers are stored.

Start with the *Inputs* sheet and enter information about your company name and LOB/segment, then enter paid and incurred loss triangles, earned premium, exposures (if available) and your assumptions for the tail factors, Bornhuetter-Ferguson and Cape Cod models. [**Note:** If you copy and paste we recommend that you use paste special, values so that you do not accidentally erase any range names.] The default model options are set up to match the current modeling framework described in the paper, but they can be changed to see how they impact the final results. [**Note:** To help the interested reader in following the evolution of the ODP Bootstrap model, the model options for earlier papers are summarized Bootstrap Model Options section at the end of these instructions.]

On the *Bootstrap* sheet, you should set all hetero group values to zero (Step 9) and make sure all of the cells in the outliers triangle are set to zero (Step 10) – for both the Paid and Incurred models. [**Note:** Clicking on the Reset Assumptions button will do this for you.]

Once the data has been entered, hitting F9 (recalculation) will set up the graphs on the *Diagnostics* sheet so that you can assess the assumptions of the model. Clicking on the buttons on this sheet will adjust the scales on the graphs to make them easier to read. You can also change the column used in the graphs at the bottom of the sheet, but you will need to hit F9 to update the graphs.

Using the graphs on the *Diagnostics* sheet, you can manually set up other hetero groups and remove outliers (in the *Bootstrap* sheet) to adjust the model fit to the underlying assumptions (See section 5 of the paper for more guidance). For the hetero groups you should start with group zero and number the other groups sequentially using integers for group numbers (e.g., 0, 1, 2, etc.). For the outlier triangle, a one is used to indicate a cell that will be removed from the calculations (i.e., given zero weight), otherwise use a zero.

Instead of just using the graphs on the *Diagnostics* sheet, you can use the results of the algorithm shown in the *Groups* sheet to select the hetero groups. Start by changing the Auto Recalc option to True and hitting F9 to recalculate. You can also change other options on this sheet to adjust the algorithm and hit F9 to update the results. As statistics for the best sets of hetero groups are often quite close, you can review the groups in the Actual Codes columns compared to the Plot of Residuals by Development Column to select the option that best explains the pattern in the observed residuals. Once you select an option, you will need to manually type the codes from the Actual Codes column in the *Groups* sheet to the appropriate row in the *Bootstrap* sheet. After hitting F9, the statistics in the *Diagnostics* sheet should match the statistics from the selected row in the *Groups* sheet. [**Note**: After using the optimal hetero group algorithm, change the Auto Recalc option back to False as leaving it on True will greatly slow the simulation speed.]

After you are satisfied with the model assumptions, you can select the data type you wish to use (e.g., paid or incurred) and the model type (e.g., chain ladder, Bornhuetter-Ferguson or Cape Cod) on the *Inputs* sheet. These six different combinations give you six different ODP Bootstrap models you can run for each LOB/segment. You can also

use the "GLM Bootstrap" model option on the Inputs sheet and then select parameters on the *Diagnostics* sheet. The paid or incurred "GLM Bootstrap" models can also be included when combining multiple models.[1]

Instead of just using the graphs on the *Diagnostics* sheet, you can use the results of the algorithm shown in the *GLM* sheet to select the GLM parameters. Start by changing the Auto Recalc option to True and hitting F9 to recalculate. You should also reset the hetero groups and outliers as these may be different for the GLM Bootstrap compared to the ODP Bootstrap. As statistics for the best sets of GLM parameters are often quite close, you can review the parameters in the Alpha Codes, Beta Codes and Gamma Codes columns. Note that the parameters from the algorithm may not be optimal as the algorithm does not test all parameters being set to zero and Calendar Year trends may not be calculated if there are too many Alpha and Beta parameters. Once you select an option, you will need to manually type the codes from the Alpha Codes, Beta Codes and Gamma Codes columns in the *GLM* sheet to the appropriate row in the *Diagnostics* sheet. After hitting F9, the statistics in the *Diagnostics* sheet should match the statistics from the selected row in the *GLM* sheet. [**Note:** After using the optimal GLM algorithm, change the Auto Recalc option back to False as leaving it on True will greatly slow the simulation speed.]

Set the number of iterations you want to use in the simulation process (from 10 to 10,000) and use the same number of iterations for every model you run for each LOB/segment. [**Note:** While the model will accommodate up to 10,000 iterations it is not designed for speed so you should consider limiting the number to say 100 or less.] You can also input the four percentiles you wish to see in the results summaries (the percentiles can be changed at any time).

To run each model, use the following steps:

1. Click on the Generate Random Values button to generate all the random numbers for the model. You should generate new random numbers for each model to avoid creating correlation between the models.

2. Click on the Run Simulations button to run all of the simulations and save the results to the *Results* sheet.

[1] The total number of models that can be combined is eight. There is no technological reason why more models are not being combined, it is simply a matter of convenience.

3. Review the summaries on the *Results* sheet to make sure you don't need to adjust the model assumptions. Occasionally, adjusting the "Negative Incremental" constraints on the *Inputs* sheet will help the model. [**Note:** You can drill down to any individual iteration in the model by entering the iteration number on the Inputs sheet (in the cell two to the right of the Iterations cell), hitting F9 and then reviewing the calculations on the *Bootstrap* sheet.]

4. Save the **Bootstrap Models** file in a common data directory using a different file name for each model and LOB/segment. [**Note:** To save file size you can click on the Clear Random Values button, which will remove the data generated in Step 1. Similarly, clicking on the Clear Results button will delete the simulation output on the *Results* sheet, but you should not do this for any file you want to combine for the next model.]

Repeat these steps for each model and LOB/segment.

Best Estimate:

After each model has been run and saved for a LOB/Segment, you can open the **Best Estimate** file and combine the results. There are two visible sheets in the file:

A. *Inputs* – on this sheet you can import or enter data for a LOB/Segment and select model weights.

B. *Results* – on this sheet the results of your simulations will be stored and summarized.

There is also one hidden sheet in the file where the random numbers are stored.

Start with the *Inputs* sheet and enter the directory path where you saved the **Bootstrap Models** files for the LOB/Segment. [**Note:** All files for a LOB/Segment must be in the same directory and the path should end with a backslash.] Next, select the number of files you want to combine, enter the file name for each of the **Bootstrap Models** files next to the Model File labels and select the Model Type for each model. [**Note:** the default model types are the six ODP Bootstrap and two GLM Bootstrap variations, but you can use a model type more than once.] After this information has been entered, you can click on the Import Raw Data from Model File 1 button which will import all of the rest of the data on the *Inputs* sheet, except the model weights.

Once the raw data has been imported, you must enter weights in the Model Weights by Accident Year table. Ideally, you should select the weights such that for each year the total of the weights by model sum to 100%, but if the weights do not sum to 100% each model will be selected in proportion to the sum. You can also input the four percentiles you wish to see in the results summaries (the percentiles can be changed at any time).

To combine the model results into your best estimate, use the following steps:

1. Click on the Generate Random Values button to generate all the random numbers for the combination process.

2. Clicking on the Combine Results button will open each model sequentially; combine the results based on the weights you specified by year and save the combined results to the *Results* sheet.

3. Review the summaries on the *Results* sheet to make sure you don't need to adjust the model weights.

4. Save the **Best Estimate** file in a common data directory using a different file name for each LOB/segment. [**Note:** To save file size you can click on the Clear Random Values button, which will remove the data generated in Step 1. Similarly, clicking on the Clear Raw Data and/or Clear Results buttons will delete the raw data on the *Inputs* sheet and simulation output on the *Results* sheet, respectively, but you should not do this for any file you want to combine with a **Best Estimate** file for another LOB/Segment.]

Repeat these steps for each LOB/segment.

Aggregate Estimate:

After each LOB/Segment has been combined and saved, you can open the **Aggregate Estimate** file and aggregate the results. There are two visible sheets in the file:

A. *Inputs* – on this sheet you can import or enter data for three LOB/Segments and select the correlation assumptions.

B. *Results* – on this sheet the results of your simulations will be stored and summarized.

There are also four hidden sheets in the file where the correlation ranks are stored and results for three different LOB/Segments can be found.

Start with the *Inputs* sheet and enter the directory path where you saved the **Best Estimate** files for each of the three LOB/Segments. [**Note:** All LOB/Segment files must be in the same directory and the path should end with a backslash.] Next, enter the file name for each of the **Best Estimate** files next to the appropriate LOB file number. After this information has been entered, you can click on the Import Raw Data from LOB Files button which will import all of the rest of the data on the *Inputs* sheet and calculate the correlations between the residuals in the LOBs.

Once the raw data has been imported, you must select a correlation matrix assumption. You can use the Select Correlation button to select one of the calculated correlation matrices or to enter a constant value, or you can enter a unique value for each LOB pair. You can also input the four percentiles you wish to see in the results summaries (the percentiles can be changed at any time).

Calculating the correlation ranks to match the selected correlation assumption using the Iman-Conover method can be done by clicking on the Generate Rank Values button. If you would like to use a different method, you can generate the ranks (based on your correlation assumption and number of iterations) using a separate process (e.g., using a Copulas), paste them in the "Correlation Ranks.xls" file and then save the file using another name. [**Note:** The **Correlation Ranks** file has some ranks generated for the figures used in the paper but these should be deleted and replaced.]

To combine the best estimate results into your aggregate estimate, use the following steps:

1. Click on the Import Results Data from LOB Files button to import the results from each of the best estimate files into their respective hidden LOB sheet.

2. Click on the Generate Rank Values button to create rank values or the Import Rank Values button to import the rank values you created outside of this file for the aggregation process.

3. Click on the Aggregate Results button to combine the results from each of the LOB files using the correlation ranks and save them to the *Results* sheet.

4. Review the summaries on the *Results* sheet to make sure you don't need to adjust the correlation assumption.

   Save the **Aggregate Estimate** file using a different file name. [**Note:** To save file size you can click on the Minimize Result Data and Clear Rank Values buttons, which will remove the all data from Step 1 not used in the Results sheet and the rank values from Step 2, respectively. Similarly, clicking on the Clear Raw Data, Clear Result Data and/or Clear Results buttons will delete the raw data on the *Inputs* sheet, all data in the hidden LOB sheets and simulation output on the *Results* sheet, respectively, but you should not do this if you want so save everything on the *Results* sheet.]

Bootstrap Model Options:

In the **Bootstrap Models** file, the Model Options in the *Inputs* and *Bootstrap* sheets can be changed to test out different variations on the ODP bootstrap algorithm. Below is a brief description of each option, then a table which summarizes the options to replicate the results in prior papers.

Inputs Sheet Options:

1) *Tail Factors* – This option allows the specification of a *mean* and *standard deviation* for the tail factors, as well as the *distribution* used to simulate the tail factors. The tail factors can also be *extrapolated* and/or *limited*.

2) *GLM Error Distribution* – This option only applies when Model Type is set to "GLM Bootstrap". Changing the error distribution changes the z parameter in the GLM framework. [Default: Poisson or z=1]

3) *Use Exposure Adjustment* – This option is used to adjust the input data for exposure growth. When "Yes" is selected, all triangle data is divided by the exposures by year. [Default: No]

4) *Don't Sample Zero Residuals* – This option will either include or exclude the zero residuals (typically in the corners of the triangle) during the simulation process. [Default: Yes]

5) *Include Zero Residuals in Stats* – This option is used to either include or exclude the zero residuals (typically in the corners of the triangle) in all of the Diagnostics. [Default: Yes]

6) *Include Hat Matrix Factors* – This option is used to determine whether the hat matrix adjustment factors are used to standardize the residuals. [Default: Yes]

7) *Include DoF Adjustment* – This option is used to determine whether the Degrees of Freedom adjustment factor is used to scale the residuals. [Default: No]

   Note: Only one of the two options (5 & 6) above should be set to Yes.

8) *Include Process Variance* – This option is used to either include or exclude process variance for the sample future incremental values during the simulation process. [Default: Yes]

9) *Include Scale Parameter* – This option is used to either include or exclude the Scale Parameter (times the sample future incremental values) when calculating the variance of the Gamma distribution used for the process variance. [Default: Yes]

10) *Calc Hetero-Adj Factors Using* – This option is used to calculate the Hetero Adjustment factors based on either the Standard Deviation (StdDev) or Scale Parameter (ScalParm) methods described in the paper. [Default: ScalParm]

11) *Set Negative Incrementals to Zero* – This option is used by development period and data type to limit any incremental values to a positive number (i.e., >= 0). [Default: No]

12) *Triangle Size* – This value is used to set up the triangle size. It can be any value from 5 (5 x 5 triangle) to 10 (10 x 10 triangle). [Default: 10]

Bootstrap Sheet Options (separate for Paid and Incurred Data):

13) *Heteroscedasticity Group* – In Step 9, the hetero groups can be selected using the same integer for each development period in the same group. The group numbers must be integers starting with zero and increasing by one per group. [Default: 0]

14) *Outliers* – In Step 10, any individual incremental cell can be excluded from the model parameterization (i.e., given zero weight) by changing the outlier triangle cell to one. [Default: 0]

The ODP Bootstrap model was developed and described in several papers, including but not limited to:

[12] England, Peter D. and Richard J. Verrall. 1999. Analytic and Bootstrap Estimates of Prediction Errors in Claims Reserving. Insurance: Mathematics and Economics, 25: 281-293.

[13] England, Peter D. and Richard J. Verrall. 2002. Stochastic Claims Reserving in General Insurance. British Actuarial Journal, 8-3: 443-544.

[31] Pinheiro, Paulo J. R., João Manuel Andrade e Silva and Maria de Lourdes Centeno. 2001. Bootstrap Methodology in Claim Reserving. ASTIN Colloquium: 1-13.

To replicate the models described in these papers, use the following model options. The model options noted above which are not included in the table below are excluded as they were not mentioned in these papers. Thus, the options excluded from the table below should be set to their default values.

	Paper Reference		
Model Option	**[13]**	**[14]**	**[31]**
Don't Sample Zero Residuals	No	No	Yes
Include Hat Matrix Factors	No	No	Yes
Include DoF Adjustment	No	Yes	No
Include Process Variance	No	Yes	Yes
Include Scale Parameter	No	Yes	Yes