A quantum leap in benchmarking P&C unpaid claim distributions

Mark Shapland, FCAS, FSA, FIAI, MAAA

If no single estimate is quite as important to an insurance entity as a best estimate of unpaid claims, then the uncertainty around a central point estimate runs a close second. It potentially affects every pivotal decision a carrier makes, from reinsurance needs to risk margins to risk-based capital requirements. For example, if the uncertainty is underestimated, the company could continue to underwrite policies where the expected profit is lower than the risks it continues to assume. If the uncertainty is overestimated, the company may be holding more capital than it really needs to match its reserve risk appetite.

But much like the decisions about a central estimate, quantifying the uncertainty (i.e., determining a loss distribution) is prone to many of the same vulnerabilities of subjectivity and method and/or model error. Recently a set of guidelines was developed to provide a standard for property and casualty (P&C) insurers not only for setting a central estimate of a range, as was discussed in an earlier article in this series, "A Quantum Leap in Benchmarking P&C Reserve Ranges," but also for establishing an objective process for estimating loss distributions.

The introduction of the claims variability guidelines (CVG) is part of an evolutionary process that began with deterministic and stochastic models aimed at understanding an insurance entity's risk. The advent of substantial computing power allowed actuaries to move closer to a reasonable depiction of an entity's risk with the development of sophisticated models that simulate millions of possible outcomes. From there, distributions of the possible outcomes can be used to identify a central estimate and to quantify worst-case scenarios.

Throughout the evolution of these approaches, the objective has always been to develop a reliable methodology for expressing the uncertainty around the mean as a way of understanding the breadth of an entity's reserve risk.

Though an advancement from more rudimentary approaches has occurred, simulation models still fall victim to gaps and aberrations in entity-specific data that can create considerable noise in the loss development patterns used to estimate distributions. Moreover, actuaries have come to realize that simply because a model is theoretically sound it does not necessarily ensure that the model estimates will be reasonable for a given data set. Even robust models can overestimate or underestimate the width of a distribution.

With the introduction of next-generation dynamic benchmarking tools, we are now beginning to see guidelines that provide an objective context for arriving at a more reliable distribution of loss that can add confidence to management's decisions about reserve risk.

Derived from extensive testing, using more than 30,000 data triangle sets involving all long-tail Schedule P lines of business, the guidelines employ models that have been rigorously back-tested and adjusted to compensate for underestimations many commonly used models are prone to make.

In practice, each data set is unique (e.g., entities have different target markets and geographic diversity) so the dynamic benchmarks give actuaries a way of comparing the reasonableness of their model results against a benchmark that can be customized, among other factors, for:

- The size of the company and its exposure base
- The speed at which losses develop

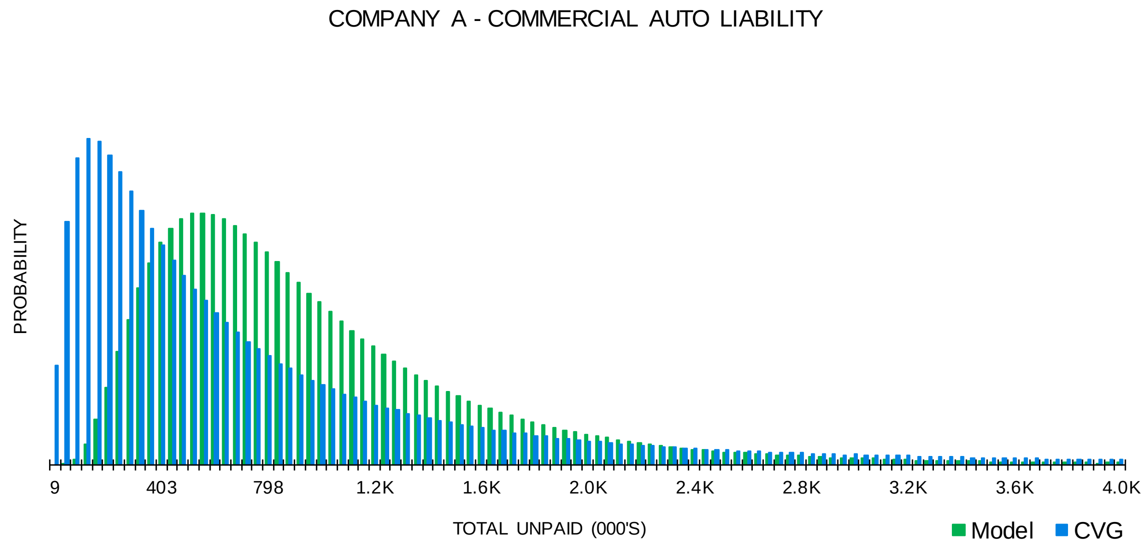
By including options that can approximate the traits of an entity's data set, the guidelines now give actuaries an objective approach for adjusting for anomalies in their carriers' data sets, reinvestigating model assumptions, and revising parameters with a clear guidepost in mind.

Four real-life scenarios

We can see the impact of using such benchmarks by continuing the examples that were introduced in the reserve ranges article for four representative data sets, which represent randomly selected companies of four different sizes: A) small, B) regional, C) small national, and D) large national. Minor changes were made to the data in order to protect the identities of each company. Commercial auto continues to be used in this article as a way of comparing the benchmark distributions for each company, which have been customized using loss development patterns from the reserve ranges article.[1]

[1] While the commercial auto data is the focus of the tables and graphs in that article, the results for all lines are available in the Appendix for the interested reader.

FIGURE 1: COMMERCIAL AUTO DISTRIBUTIONS FOR COMPANY A



	Mean	Std Dev	CoV	75.0%	90.0%	95.0%	99.5%
ODP Pd CL Results	997	665	66.7%	1,249	1,805	2,249	3,956
CVG Benchmark	997	1,555	156.0%	1,138	2,234	3,344	9,403

A common method used to estimate an unpaid claim distribution is to assume that the central estimate from the deterministic process is the mean and then use either the Mack or Over-Dispersed Poisson (ODP) Bootstrap model to estimate the associated uncertainty. To approximate this common method here, the OPD Bootstrap paid chain ladder (ODP Pd CL) model has been used.[2]

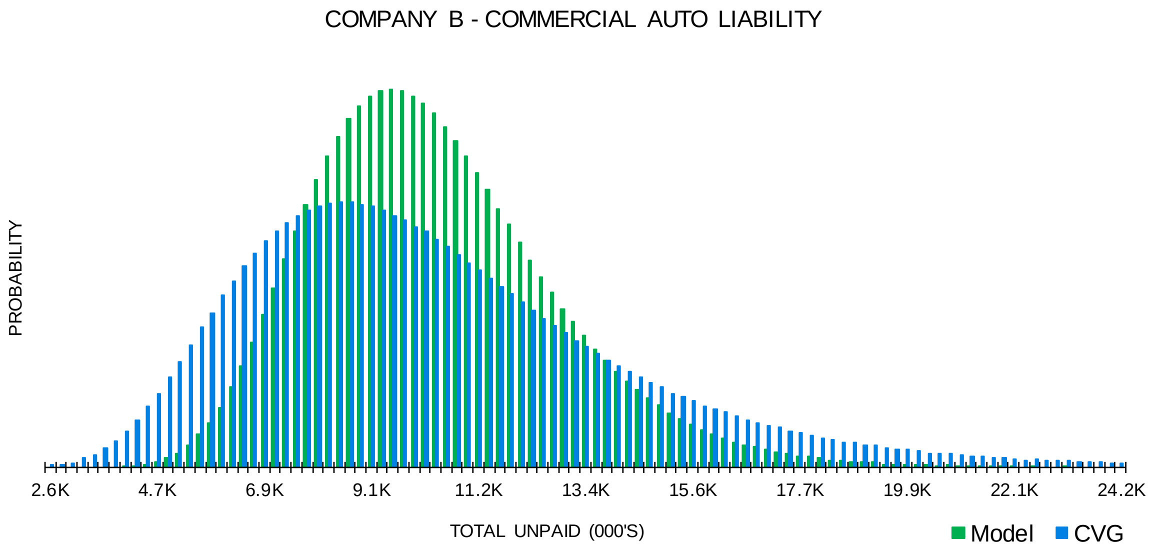
Figure 1, which shows the mean, standard deviation, coefficient of variation (CoV), and results for key percentiles, provides actuaries a way of comparing the results they obtain from their own models, in this case the ODP Pd CL, with those from the customized benchmark. The means for the ODP Pd CL and customized benchmark are the same, but the standard deviation and CoV for the ODP Pd CL model show much less uncertainty than those from the benchmark.

For Company A, the difference is barely noticeable at the 75th percentile, but because of the skewness of the loss distribution, the degree to which the ODP Pd CL appears to underestimate losses, based on the illustrative user's data, becomes apparent at the 90th percentile, where the ODP Pd CL losses are 20% less than those for the benchmark. By the 99.5th percentile, the ODP Pd CL losses are only 42% of the amount estimated by the benchmark. The actuary can still assess whether this difference is due to some characteristic of the portfolio, but without a benchmark a reasonability check is not possible.

Underestimating is a common shortcoming of many single-model approaches that are used today and a central focus of the development of the dynamic benchmark that accommodates for the size of the exposures. Under the guidelines, the exposure base is figured into the benchmark whose assumptions reflect the changes in volatility around the mean based on the many real-world data sets used to calculate the loss distribution. In this sense, the benchmark has been "fitted" to the exposure size of the company.

[2] For all of the models used in these illustrations, only standard model assumptions were used in order to replicate how an actuary might approach these estimates in practice. No attempt was made to calibrate the model assumptions to the benchmarks.

FIGURE 2: COMMERCIAL AUTO DISTRIBUTIONS FOR COMPANY B



	Mean	Std Dev	CoV	75.0%	90.0%	95.0%	99.5%
ODP Pd CL Results	10,428	2,473	23.7%	11,881	13,694	14,909	18,536
CVG Benchmark	10,428	3,878	37.2%	12,460	15,502	17,668	24,700

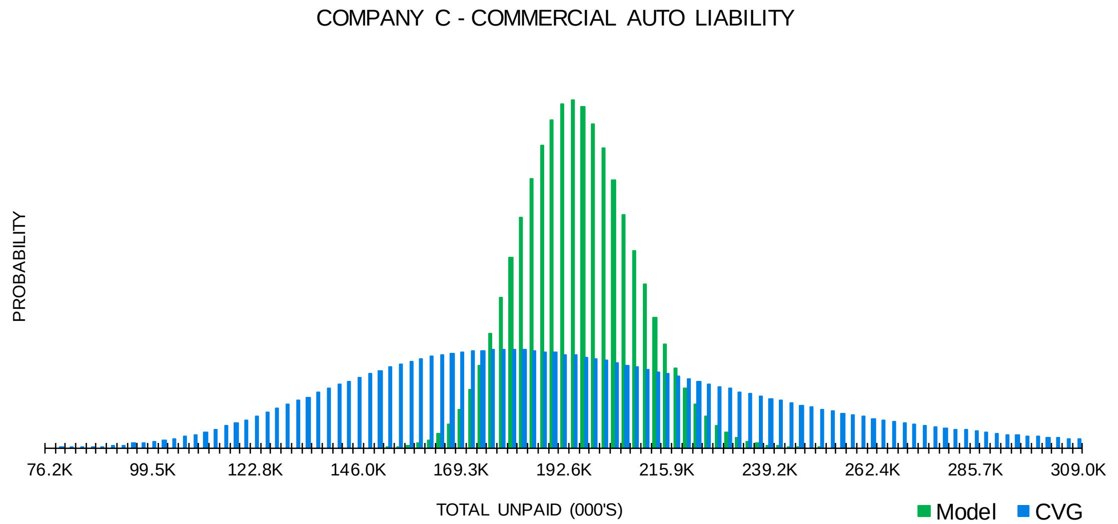
In Figure 2, which provides the same indices for commercial auto but in this case for Company B, the regional carrier, the standard deviation and CoV for both the ODP Pd CL and benchmark narrow when compared with the mean. ODP Pd CL estimates are more closely aligned with the benchmark, but still largely underestimate it, most notably at the 99.5th percentile where the estimate for ODP Pd CL is approximately 25% less than those for the benchmark.

Similar results can be seen in Figures 3 and 4, which provide commercial auto results for Company C, the small national company, and Company D, the large national company, respectively. The standard deviation and CoV for both approaches continue to narrow when compared with the mean, but estimates for the ODP Pd CL for both companies continue to underestimate losses, compared with those for the benchmarks, especially at the 99.5th percentile, where they are short by approximately one-third.

In each case, the single-model approach undershoots benchmark loss expectations at higher percentiles by a significant margin that might not otherwise be evident to actuaries, or if it is, often leaves them debating the merits of different models.

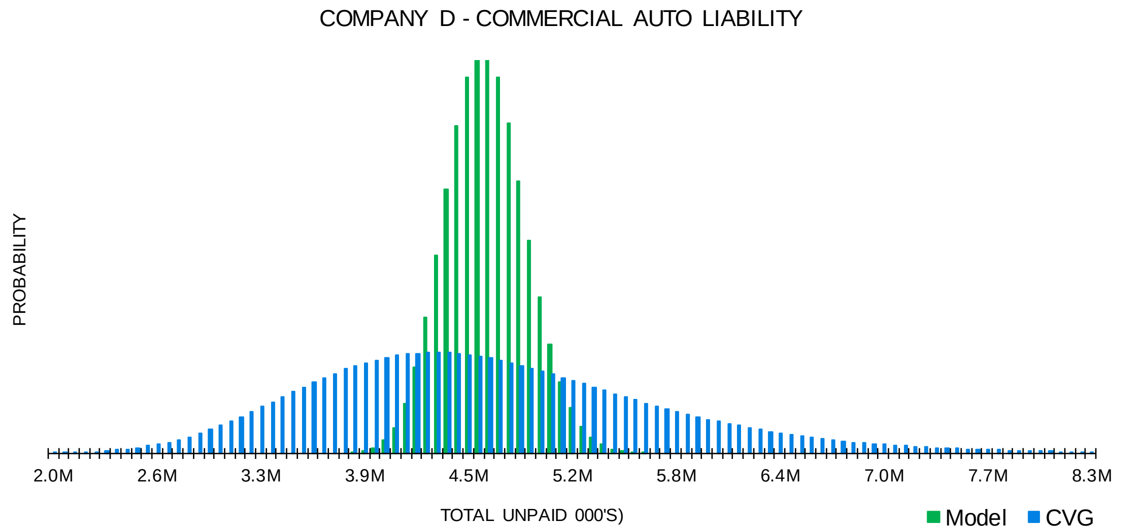
And while actuarial judgment is still a key factor in applying the guidelines, it can now be enhanced by a benchmark that lends objectivity and direction to the process of finding a suitable loss distribution that better characterizes the underlying data and a company's risks, especially for worst-case scenarios. Actuaries can systematically rethink the assumptions they made and the sufficiency or insufficiency of their data sets and explore different options to better gauge whether their model results are reasonable.

FIGURE 3: COMMERCIAL AUTO DISTRIBUTIONS FOR COMPANY C



	Mean	Std Dev	CoV	75.0%	90.0%	95.0%	99.5%
ODP Pd CL Results	197,105	12,992	6.6%	205,609	213,994	219,175	233,030
CVG Benchmark	197,105	49,490	25.1%	225,866	262,444	287,110	361,425

FIGURE 4: COMMERCIAL AUTO DISTRIBUTIONS FOR COMPANY D



	Mean	Std Dev	CoV	75.0%	90.0%	95.0%	99.5%
ODP Pd CL Results	4,674,180	242,389	5.2%	4,833,953	4,988,442	5,083,250	5,334,504
CVG Benchmark	4,674,180	1,039,762	22.2%	5,291,680	6,046,915	6,549,510	8,036,457

A higher order of sophistication

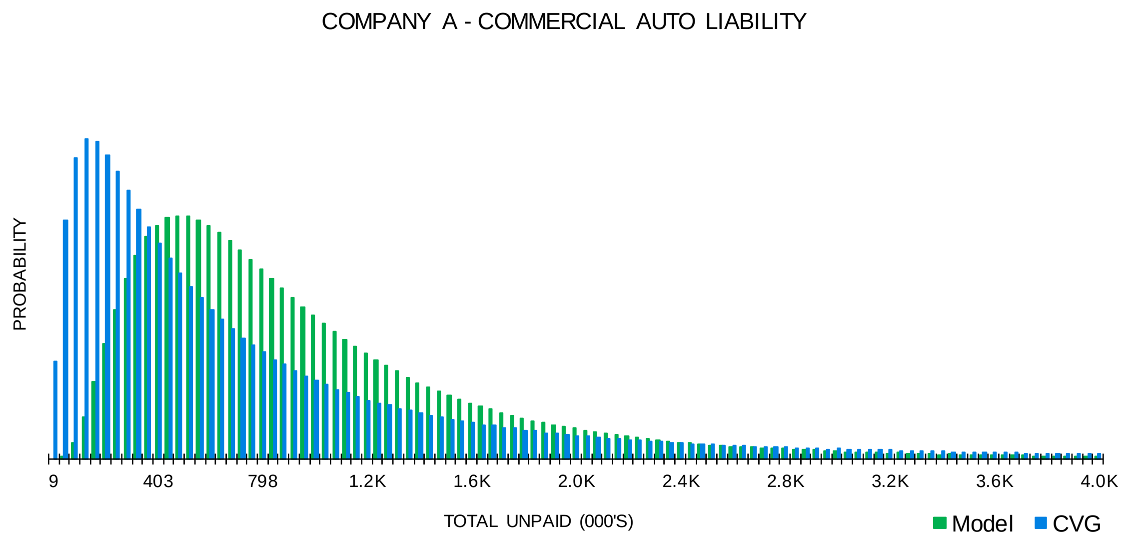
As the above examples illustrate, a single-model approach is often prone to model error, some of which can be overcome with the use of multiple models. For example, Figures 5 through 8 show that the combined use of four models yields results that more closely approximate those for the benchmark.[3] For Companies A and C, losses from the four-model approach are still well short of the benchmark but are definitely closer to the benchmark estimates than the single model. For Companies B and D, they are within 3% of benchmark losses. As noted above, by using standard assumptions for the model results the figures illustrate the value in having the guidelines as a basis for calibrating the model assumptions.

For both the single-model and multiple-model approaches, it is also possible for the model results to indicate a wider distribution than the benchmarks.[4] However, the goal is to gain confidence in the most realistic estimate of the width as both underestimation and overestimation have risk management consequences.

For the actuary whose multiple-model approach still differs from the benchmark, the guidelines offer a pathway for investigating how each model influences results and can inform actuarial decisions on the weights that should be given to each model. For the actuary whose multiple-model approach tracks the benchmark, the guidelines can help to validate the approach and add a level of confidence to the decision making.

But whether a single-model or multiple-model approach is used, actuaries now have a process that helps to ensure trust in the loss distribution and clears the way for a better understanding of the real uncertainties around risk. This understanding can better inform management's thinking about risk transfer and provide a more solid footing for decisions around enterprise risk management, topics to be discussed in subsequent articles.

FIGURE 5: COMMERCIAL AUTO DISTRIBUTIONS FOR COMPANY A



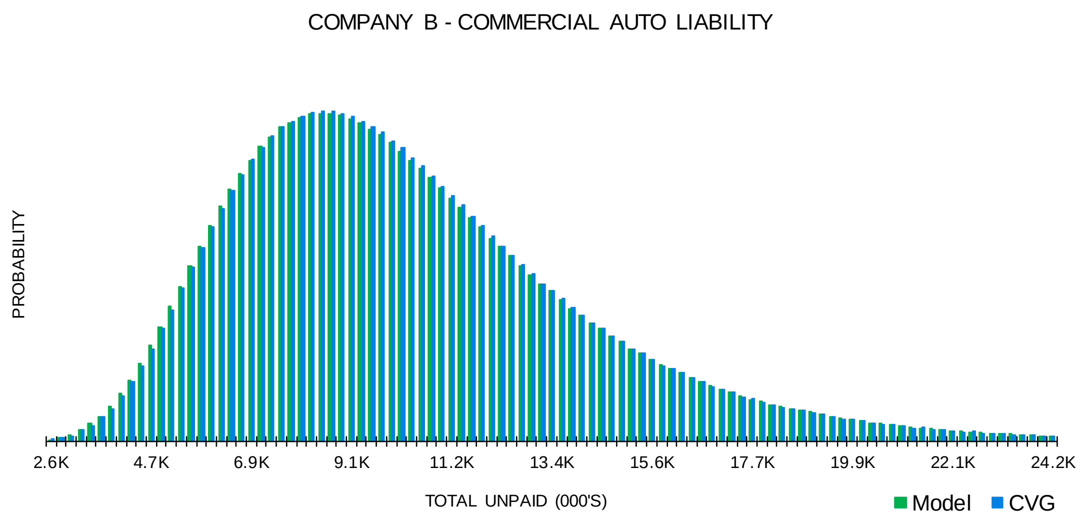
	Mean	Std Dev	CoV	75.0%	90.0%	95.0%	99.5%
Model Results*	997	746	74.8%	1,252	1,876	2,391	4,448
CVG Benchmark	997	1,555	156.0%	1,138	2,234	3,344	9,403

** Model Results based on weighting of 4 different models.*

[3] The four models used are all based on the ODP Bootstrap model framework described in the Shapland monograph for paid and incurred data using the chain ladder and Bornhuetter-Ferguson algorithms. For simplicity all four models were given equal weight for each accident year.

[4] Without calibrating the model assumptions, the data characteristics can also lead to estimated distributions that are wider than the benchmarks. Examples of wider distributions are shown in the Appendix.

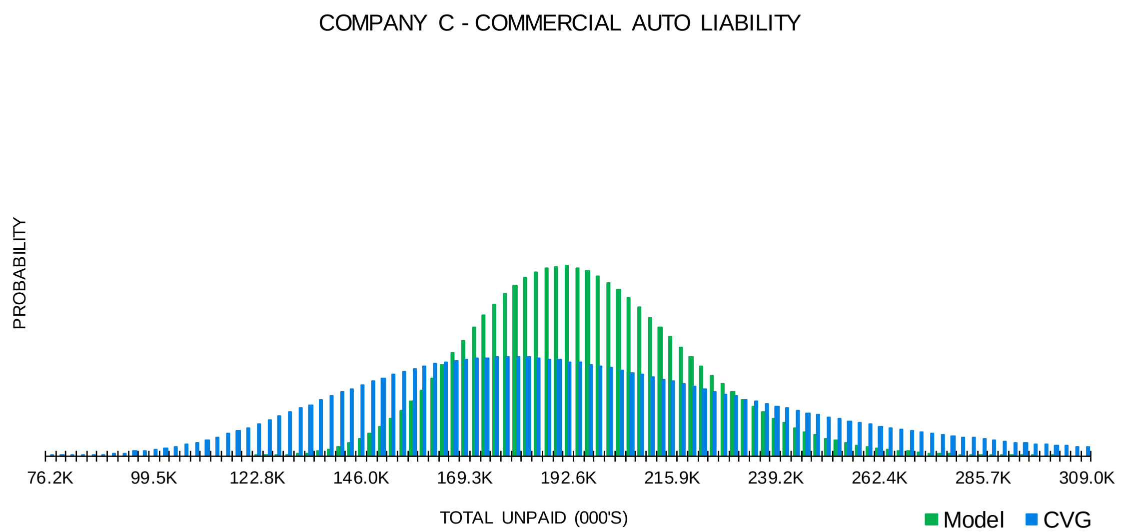
FIGURE 6: COMMERCIAL AUTO DISTRIBUTIONS FOR COMPANY B



	Mean	Std Dev	CoV	75.0%	90.0%	95.0%	99.5%
Model Results*	10,428	3,911	37.5%	12,471	15,543	17,733	24,859
CVG Benchmark	10,428	3,878	37.2%	12,460	15,502	17,668	24,700

** Model Results based on weighting of 4 different models.*

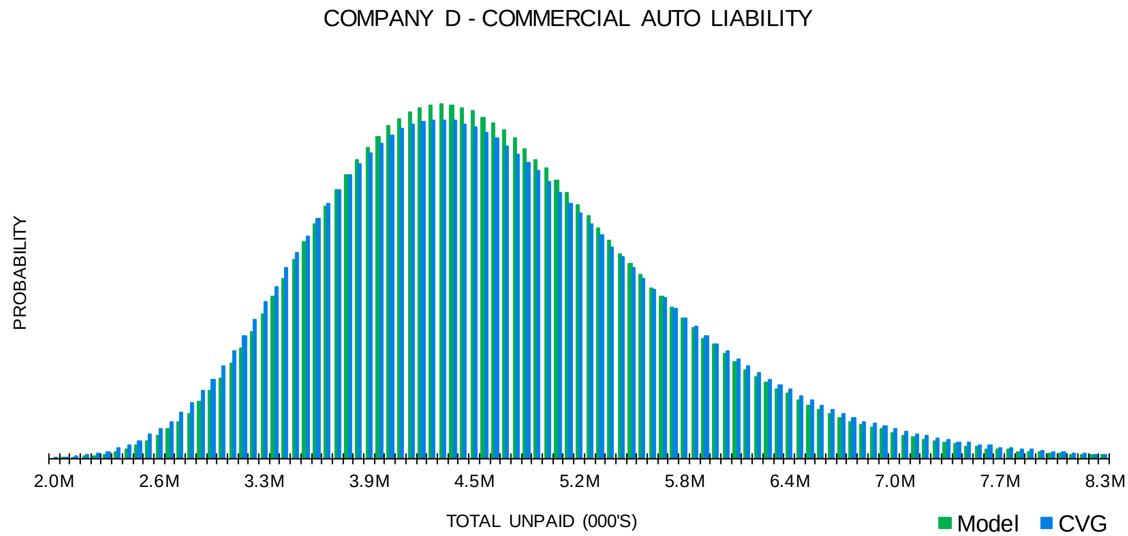
FIGURE 7: COMMERCIAL AUTO DISTRIBUTIONS FOR COMPANY C



	Mean	Std Dev	CoV	75.0%	90.0%	95.0%	99.5%
Model Results*	197,105	24,317	12.3%	212,530	228,994	239,450	268,478
CVG Benchmark	197,105	49,490	25.1%	225,866	262,444	287,110	361,425

** Model Results based on weighting of 4 different models.*

FIGURE 8: COMMERCIAL AUTO DISTRIBUTIONS FOR COMPANY D



	Mean	Std Dev	CoV	75.0%	90.0%	95.0%	99.5%
Model Results*	4,674,180	989,866	21.2%	5,266,643	5,980,746	6,453,616	7,843,130
CVG Benchmark	4,674,180	1,039,762	22.2%	5,291,680	6,046,915	6,549,510	8,036,457

** Model Results based on weighting of 4 different models.*

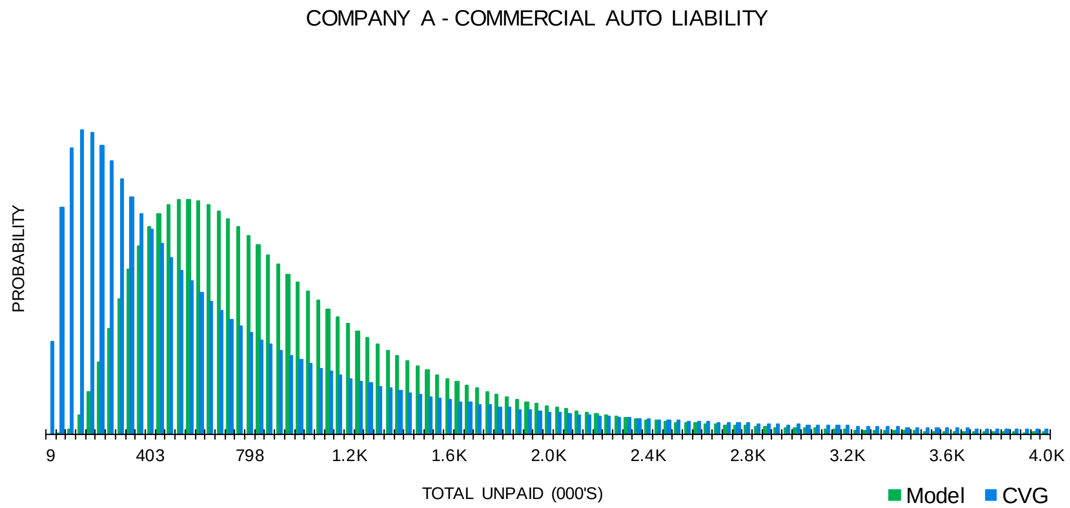
Milliman

Milliman is among the world's largest providers of actuarial and related products and services. The firm has consulting practices in life insurance and financial services, property & casualty insurance, healthcare, and employee benefits. Founded in 1947, Milliman is an independent firm with offices in major cities around the globe.

milliman.com

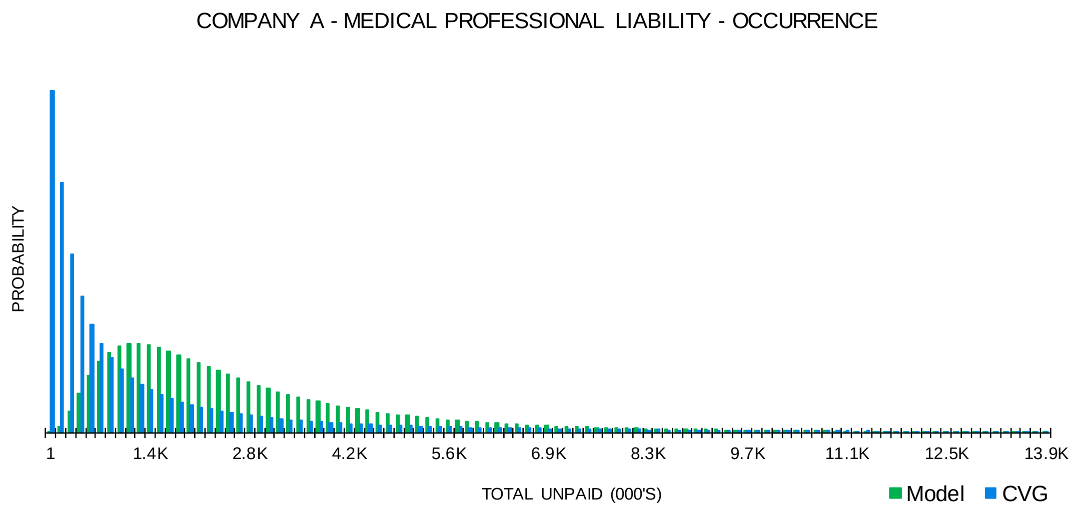
CONTACT
Mark Shapland
mark.shapland@milliman.com

FIGURE 9: COMMERCIAL AUTO



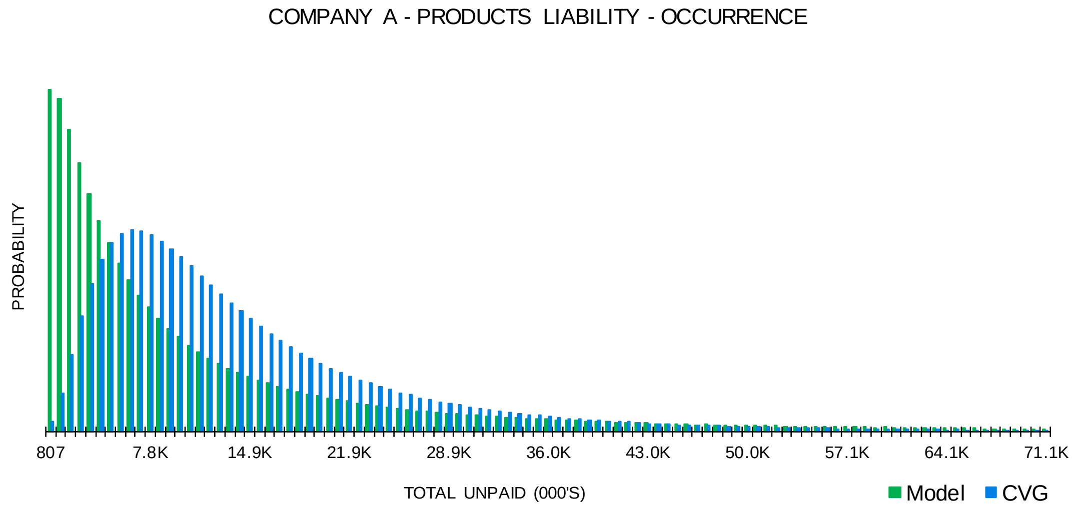
	Mean	Std Dev	CoV	75.0%	90.0%	95.0%	99.5%
ODP Pd CL Results	997	665	66.7%	1,249	1,805	2,249	3,956
CVG Benchmark	997	1,555	156.0%	1,138	2,234	3,344	9,403

FIGURE 10: MEDICAL PROFESSIONAL LIABILITY: OCCURRENCE



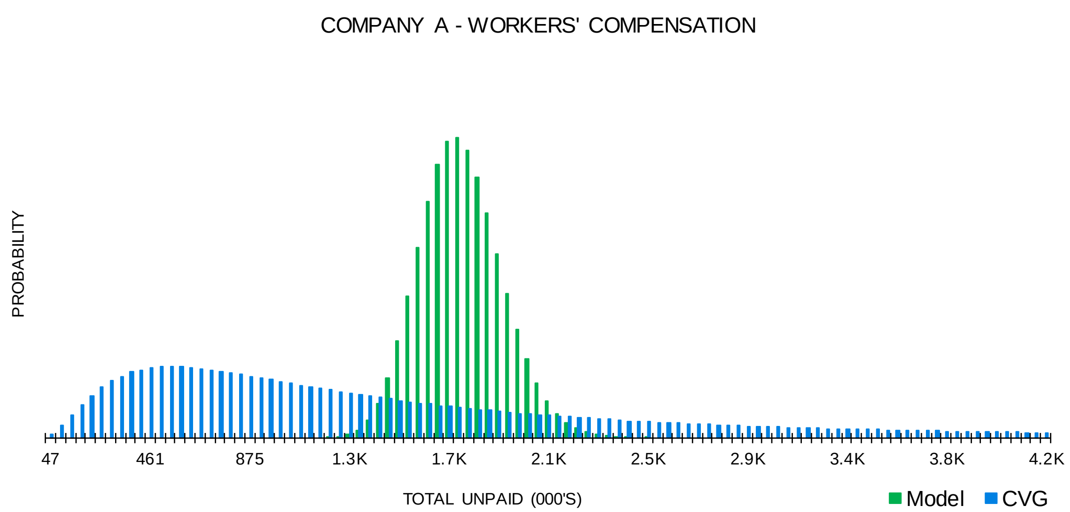
	Mean	Std Dev	CoV	75.0%	90.0%	95.0%	99.5%
ODP Pd CL Results	3,001	2,640	88.0%	3,755	5,945	7,827	15,836
CVG Benchmark	3,001	12,122	403.9%	2,253	6,279	11,597	55,861

FIGURE 11: PRODUCTS LIABILITY: OCCURRENCE



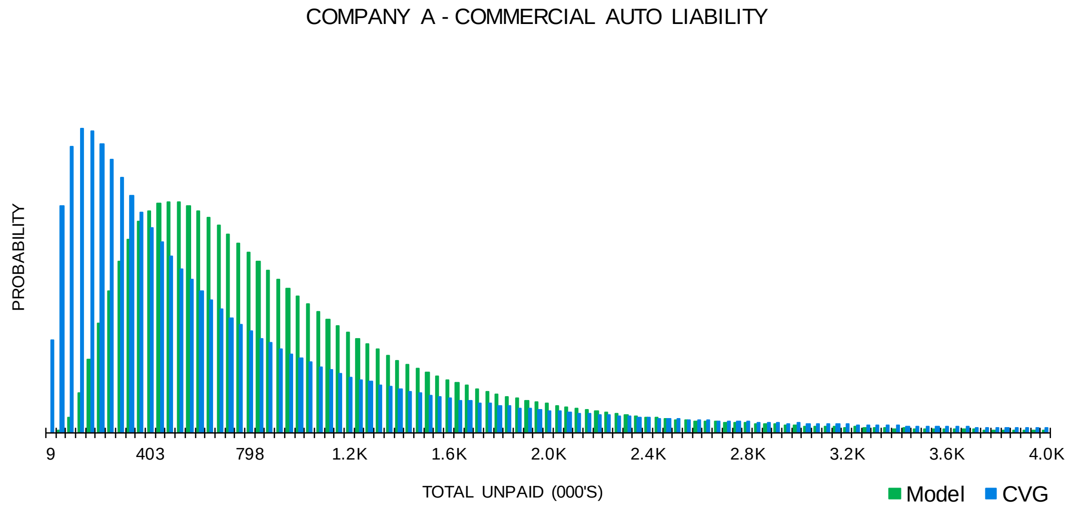
	Mean	Std Dev	CoV	75.0%	90.0%	95.0%	99.5%
ODP Pd CL Results	15,176	30,881	203.5%	15,866	34,498	54,913	180,720
CVG Benchmark	15,176	12,493	82.3%	19,033	29,454	38,251	74,725

FIGURE 12: WORKERS' COMPENSATION



	Mean	Std Dev	CoV	75.0%	90.0%	95.0%	99.5%
ODP Pd CL Results	1,779	172	9.6%	1,890	2,003	2,074	2,269
CVG Benchmark	1,779	1,913	107.5%	2,188	3,726	5,123	11,586

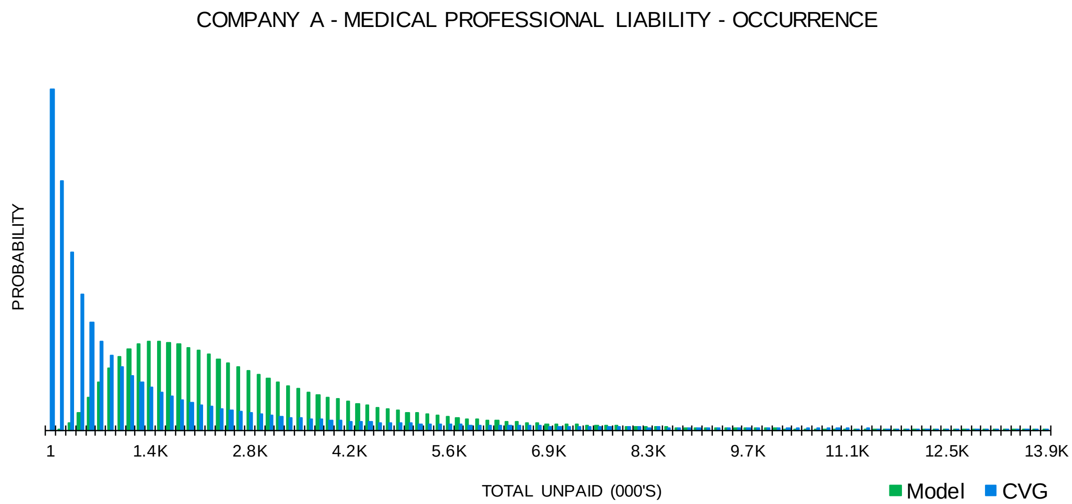
FIGURE 13: COMMERCIAL AUTO



	Mean	Std Dev	CoV	75.0%	90.0%	95.0%	99.5%
Model Results*	997	746	74.8%	1,252	1,876	2,391	4,448
CVG Benchmark	997	1,555	156.0%	1,138	2,234	3,344	9,403

** Model Results based on weighting of 4 different models.*

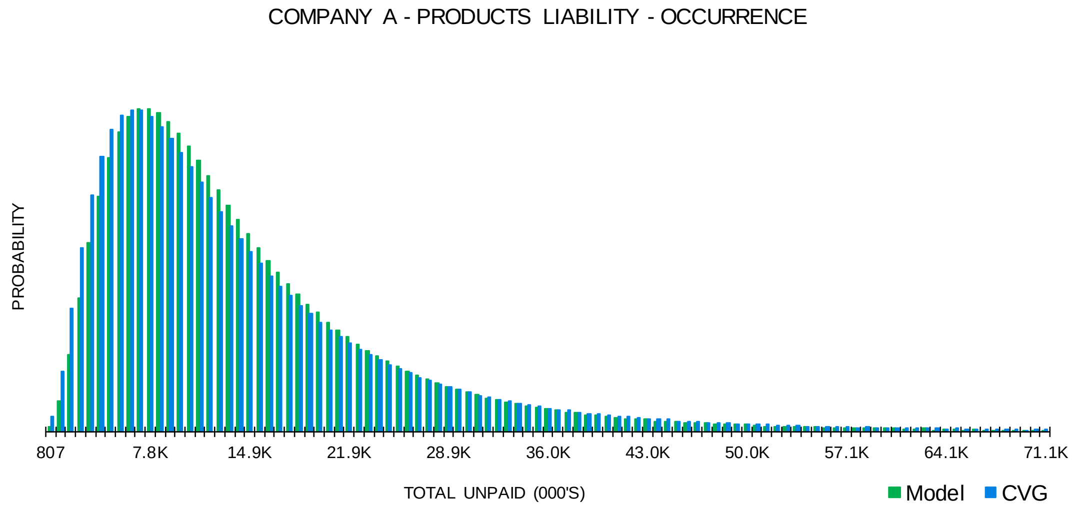
FIGURE 14: MEDICAL PROFESSIONAL LIABILITY: OCCURRENCE



	Mean	Std Dev	CoV	75.0%	90.0%	95.0%	99.5%
Model Results*	3,001	2,201	73.4%	3,767	5,610	7,120	13,116
CVG Benchmark	3,001	12,122	403.9%	2,253	6,279	11,597	55,861

** Model Results based on weighting of 4 different models.*

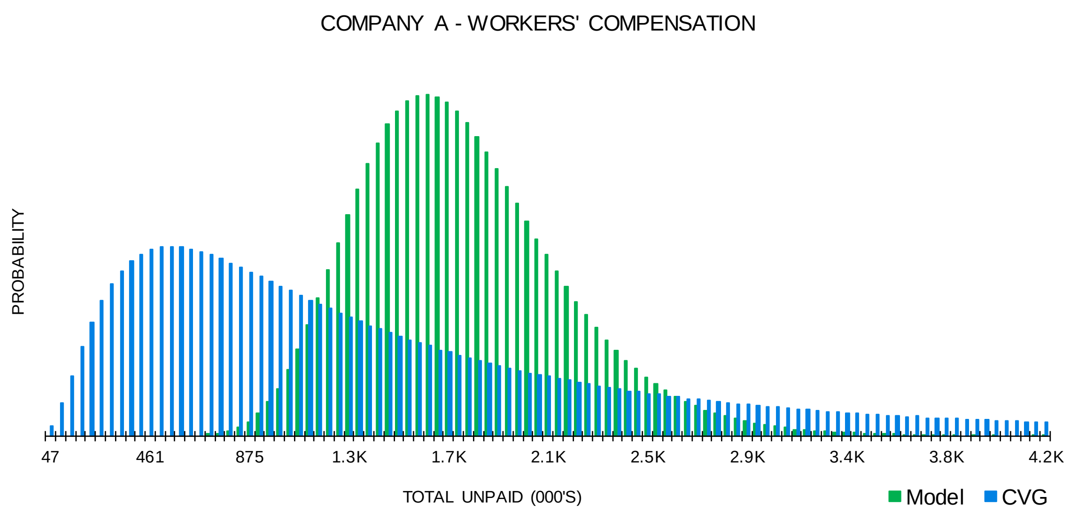
FIGURE 15: PRODUCTS LIABILITY: OCCURRENCE



	Mean	Std Dev	CoV	75.0%	90.0%	95.0%	99.5%
Model Results*	15,176	11,220	73.9%	19,050	28,444	36,156	66,859
CVG Benchmark	15,176	12,493	82.3%	19,033	29,454	38,251	74,725

** Model Results based on weighting of 4 different models.*

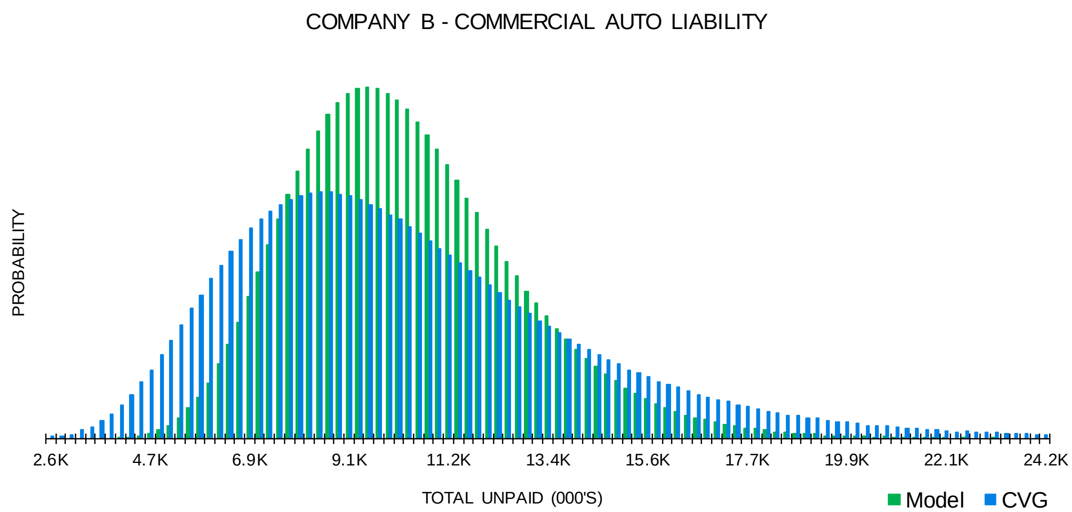
FIGURE 16: WORKERS' COMPENSATION



	Mean	Std Dev	CoV	75.0%	90.0%	95.0%	99.5%
Model Results*	1,779	429	24.1%	2,030	2,346	2,557	3,191
CVG Benchmark	1,779	1,913	107.5%	2,188	3,726	5,123	11,586

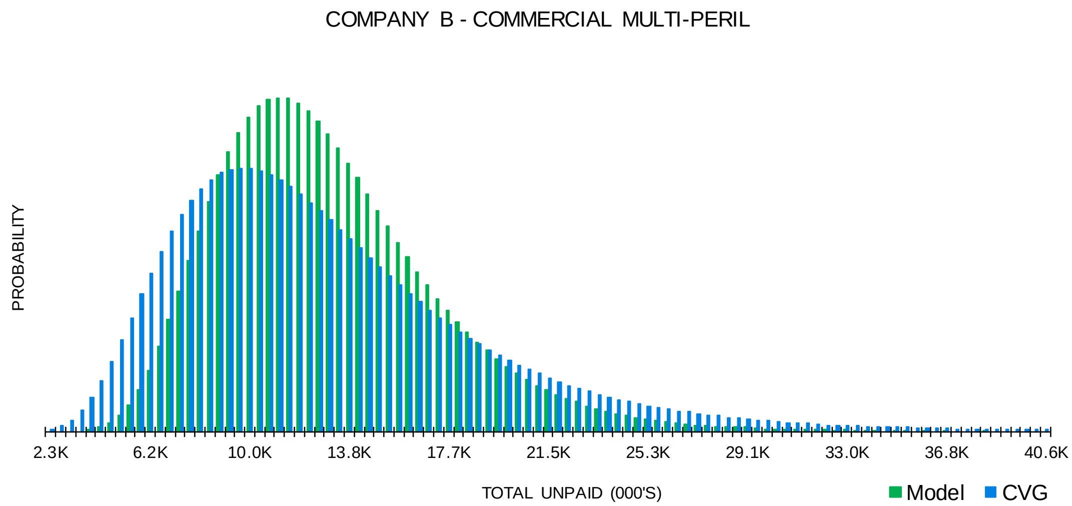
** Model Results based on weighting of 4 different models.*

FIGURE 17: COMMERCIAL AUTO



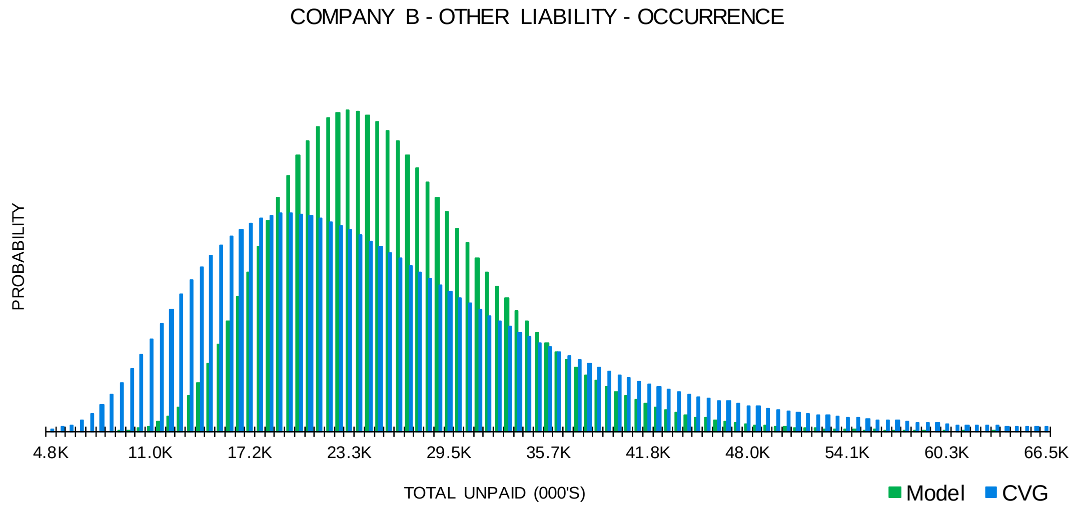
	Mean	Std Dev	CoV	75.0%	90.0%	95.0%	99.5%
ODP Pd CL Results	10,428	2,473	23.7%	11,881	13,694	14,909	18,536
CVG Benchmark	10,428	3,878	37.2%	12,460	15,502	17,668	24,700

FIGURE 18: COMMERCIAL MULTI-PERIL



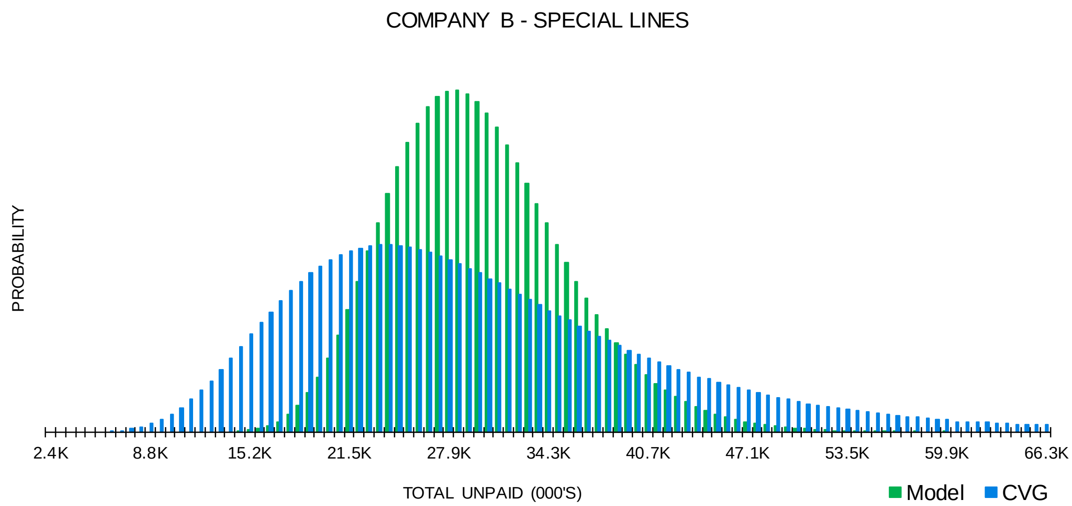
	Mean	Std Dev	CoV	75.0%	90.0%	95.0%	99.5%
ODP Pd CL Results	13,209	4,272	32.3%	15,547	18,828	21,114	28,320
CVG Benchmark	13,209	6,091	46.1%	16,130	21,056	24,698	37,169

FIGURE 19: OTHER LIABILITY: OCCURRENCE



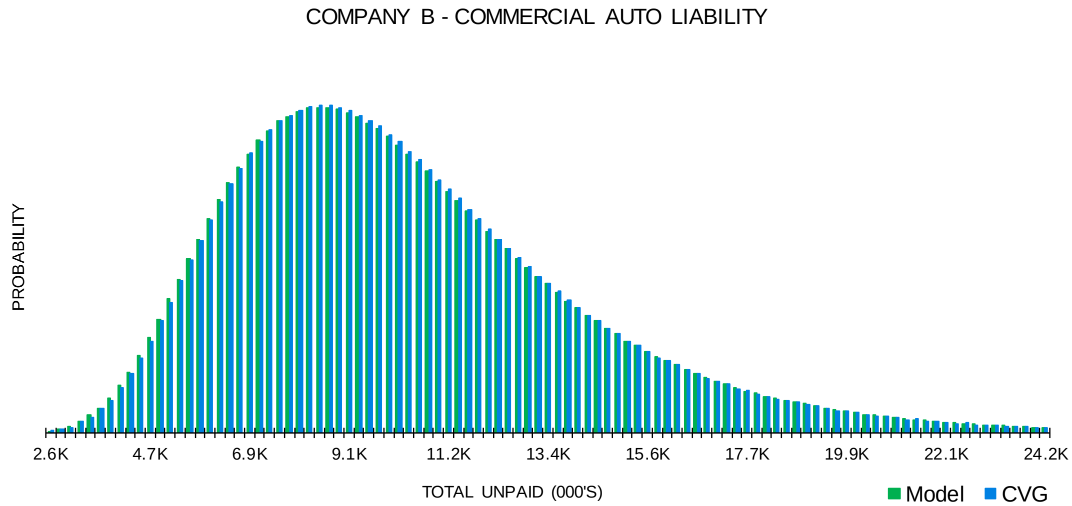
	Mean	Std Dev	CoV	75.0%	90.0%	95.0%	99.5%
ODP Pd CL Results	26,244	6,910	26.3%	30,221	35,365	38,852	49,442
CVG Benchmark	26,244	11,830	45.1%	31,978	41,519	48,541	72,444

FIGURE 20: SPECIAL LINES



	Mean	Std Dev	CoV	75.0%	90.0%	95.0%	99.5%
ODP Pd CL Results	30,278	5,784	19.1%	33,791	37,906	40,605	48,431
CVG Benchmark	30,278	12,246	40.4%	36,496	46,224	53,245	76,499

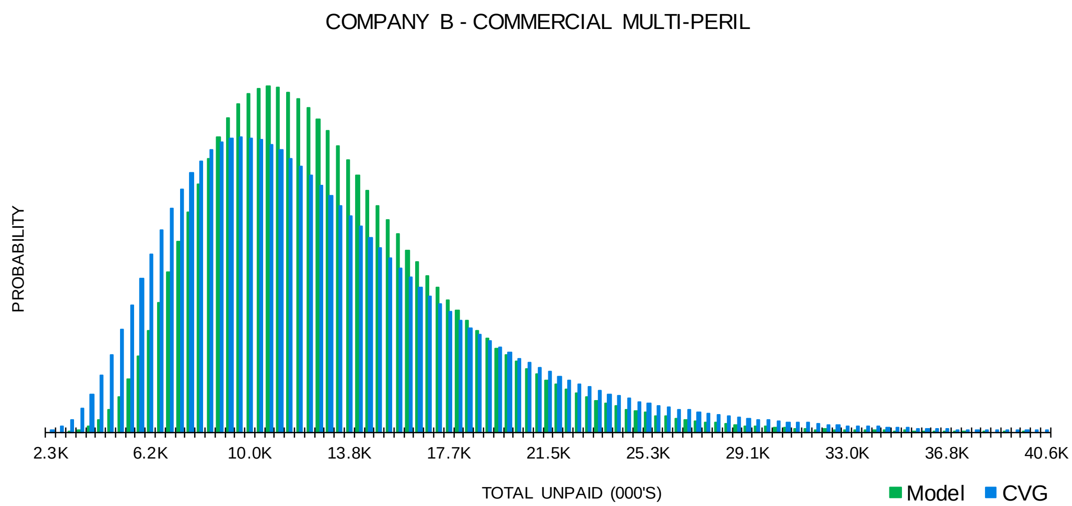
FIGURE 21: COMMERCIAL AUTO



	Mean	Std Dev	CoV	75.0%	90.0%	95.0%	99.5%
Model Results*	10,428	3,911	37.5%	12,471	15,543	17,733	24,859
CVG Benchmark	10,428	3,878	37.2%	12,460	15,502	17,668	24,700

** Model Results based on weighting of 4 different models.*

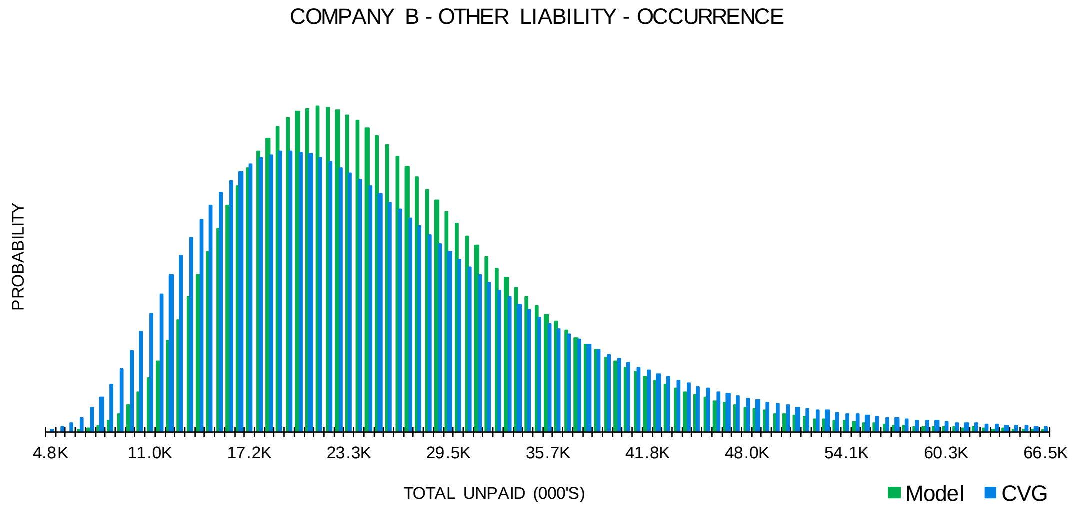
FIGURE 22: COMMERCIAL MULTI-PERIL



	Mean	Std Dev	CoV	75.0%	90.0%	95.0%	99.5%
Model Results*	13,209	4,758	36.0%	15,729	19,443	22,073	30,555
CVG Benchmark	13,209	6,091	46.1%	16,130	21,056	24,698	37,169

** Model Results based on weighting of 4 different models.*

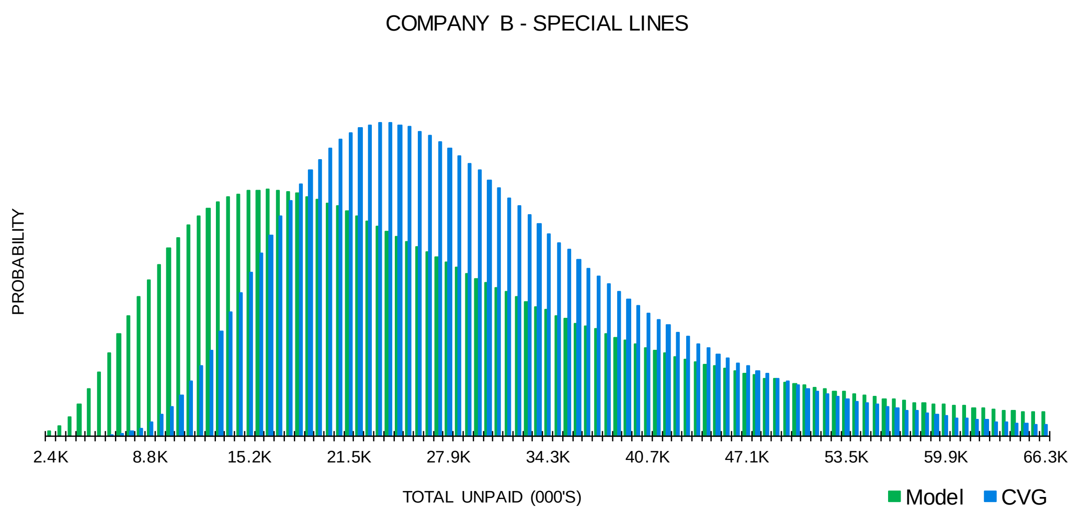
FIGURE 23: OTHER LIABILITY: OCCURRENCE



	Mean	Std Dev	CoV	75.0%	90.0%	95.0%	99.5%
Model Results*	26,244	9,418	35.9%	31,238	38,587	43,789	60,546
CVG Benchmark	26,244	11,830	45.1%	31,978	41,519	48,541	72,444

** Model Results based on weighting of 4 different models.*

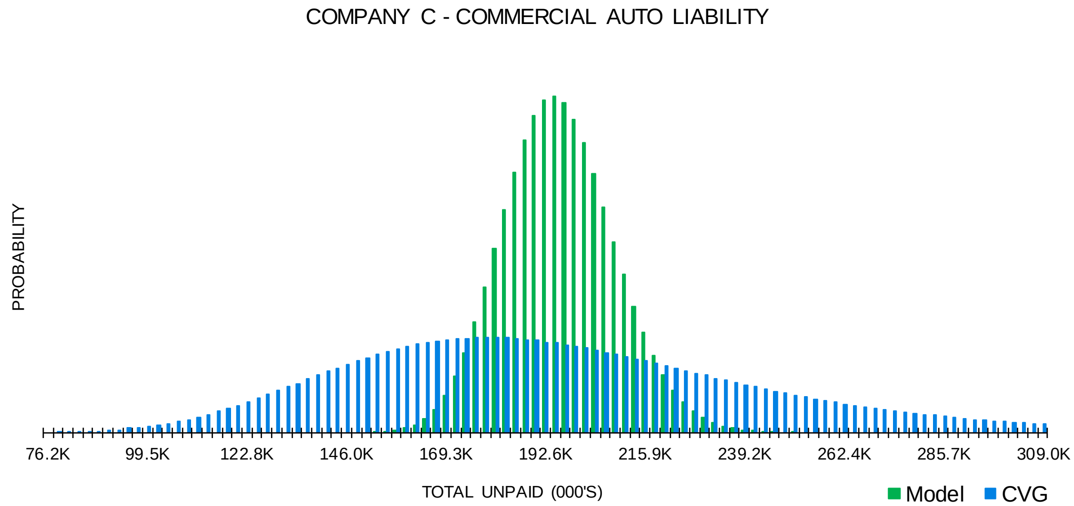
FIGURE 24: SPECIAL LINES



	Mean	Std Dev	CoV	75.0%	90.0%	95.0%	99.5%
Model Results*	30,278	21,237	70.1%	37,978	55,757	70,161	126,428
CVG Benchmark	30,278	12,246	40.4%	36,496	46,224	53,245	76,499

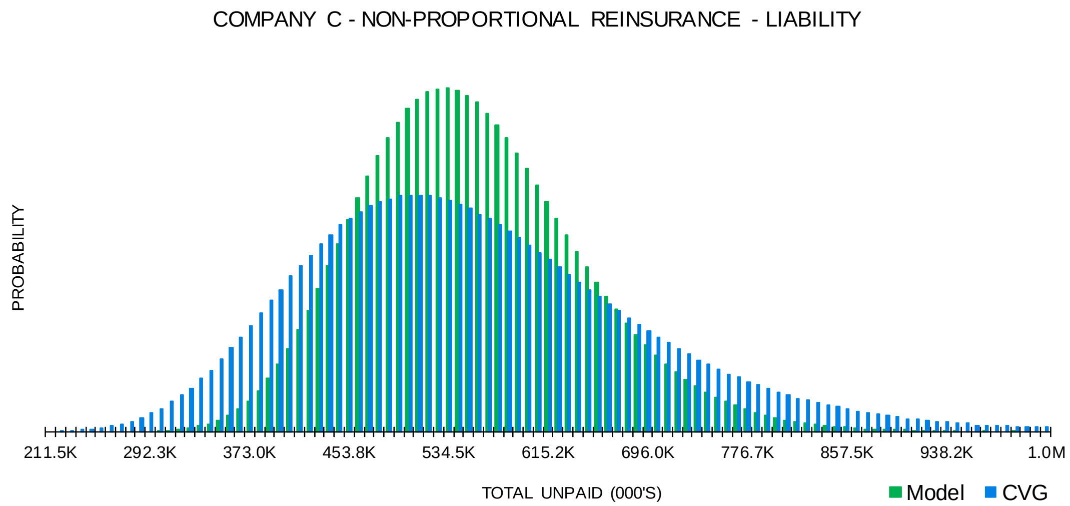
** Model Results based on weighting of 4 different models.*

FIGURE 25: COMMERCIAL AUTO



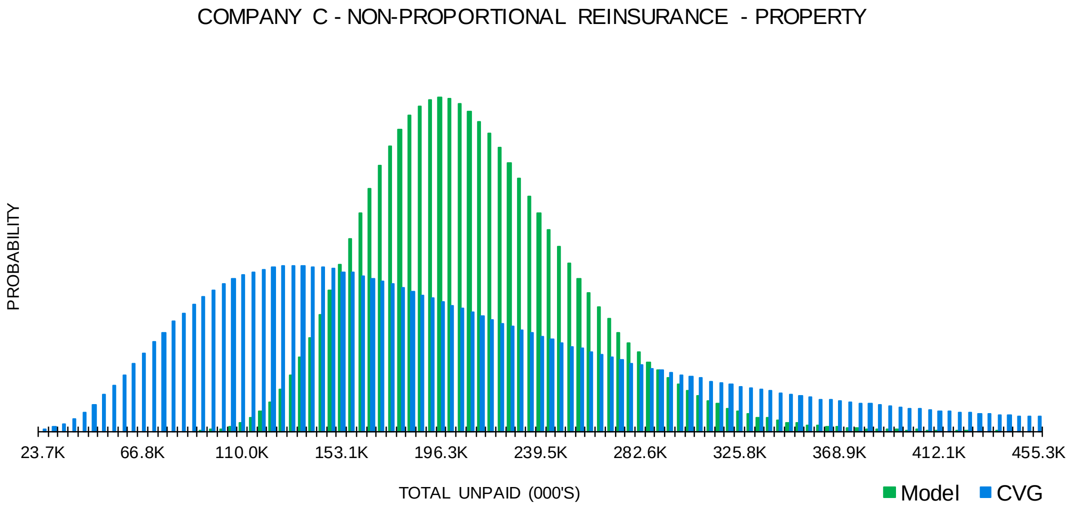
	Mean	Std Dev	CoV	75.0%	90.0%	95.0%	99.5%
ODP Pd CL Results	197,105	12,992	6.6%	205,609	213,994	219,175	233,030
CVG Benchmark	197,105	49,490	25.1%	225,866	262,444	287,110	361,425

FIGURE 26: REINSURANCE: NON-PROPORTIONAL ASSUMED LIABILITY



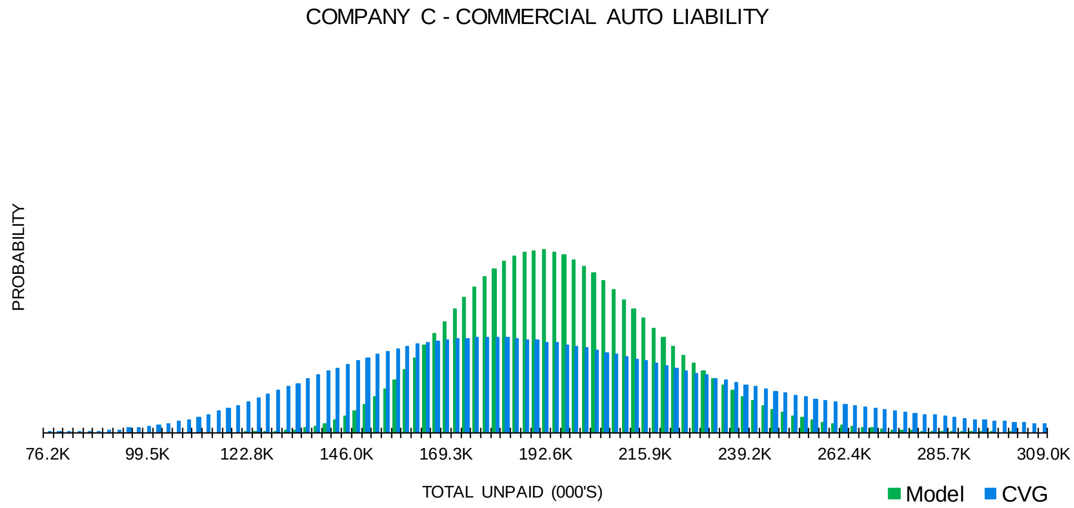
	Mean	Std Dev	CoV	75.0%	90.0%	95.0%	99.5%
ODP Pd CL Results	557,316	89,938	16.1%	613,037	675,709	716,239	831,547
CVG Benchmark	557,316	136,040	24.4%	636,802	736,937	804,245	1,006,137

FIGURE 27: REINSURANCE: NON-PROPORTIONAL ASSUMED PROPERTY



	Mean	Std Dev	CoV	75.0%	90.0%	95.0%	99.5%
ODP Pd CL Results	212,701	45,292	21.3%	239,787	272,485	294,150	357,857
CVG Benchmark	212,701	126,372	59.4%	264,962	369,950	451,748	753,709

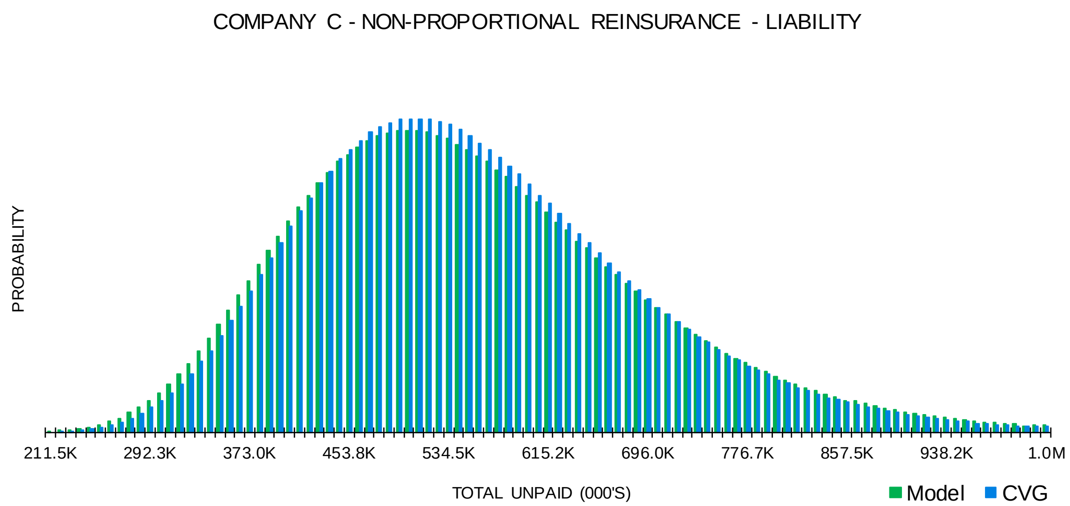
FIGURE 28: COMMERCIAL AUTO



	Mean	Std Dev	CoV	75.0%	90.0%	95.0%	99.5%
Model Results*	197,105	24,317	12.3%	212,530	228,994	239,450	268,478
CVG Benchmark	197,105	49,490	25.1%	225,866	262,444	287,110	361,425

** Model Results based on weighting of 4 different models.*

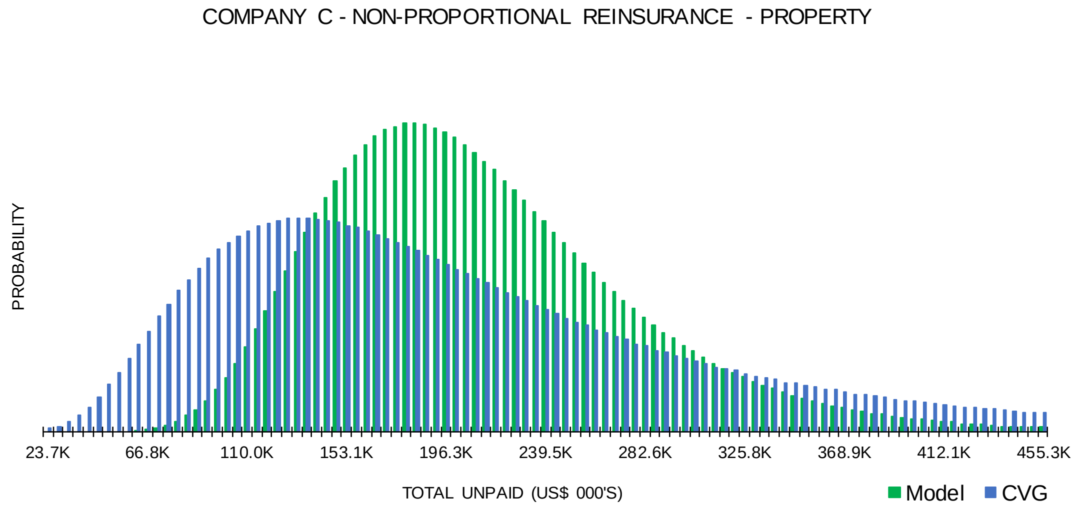
FIGURE 29: REINSURANCE: NON-PROPORTIONAL ASSUMED LIABILITY



	Mean	Std Dev	CoV	75.0%	90.0%	95.0%	99.5%
Model Results*	557,316	142,678	25.6%	639,914	745,673	817,151	1,033,172
CVG Benchmark	557,316	136,040	24.4%	636,802	736,937	804,245	1,006,137

** Model Results based on weighting of 4 different models.*

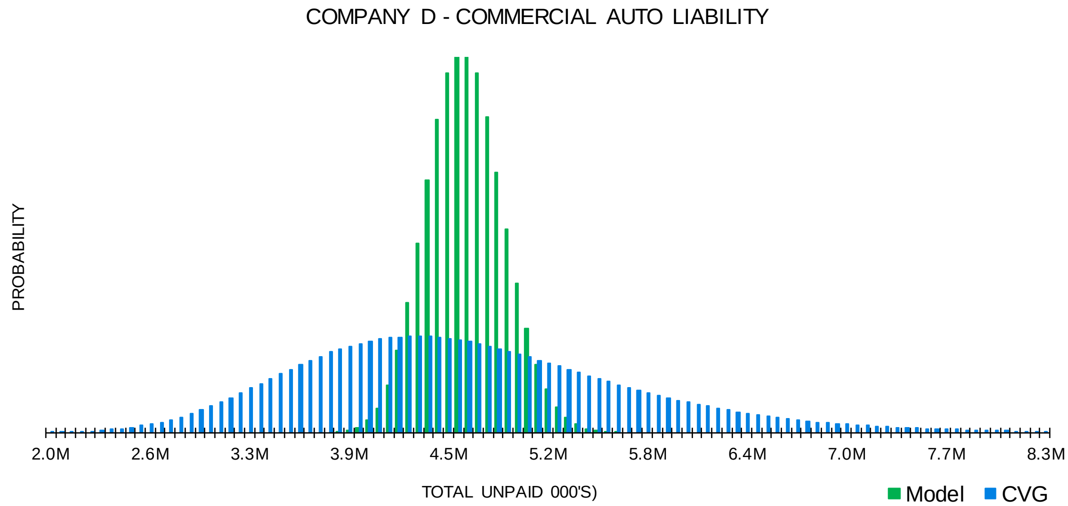
FIGURE 30: REINSURANCE: NON-PROPORTIONAL ASSUMED PROPERTY



	Mean	Std Dev	CoV	75.0%	90.0%	95.0%	99.5%
Model Results*	212,701	67,339	31.7%	249,781	301,330	337,136	449,535
CVG Benchmark	212,701	126,372	59.4%	264,962	369,950	451,748	753,709

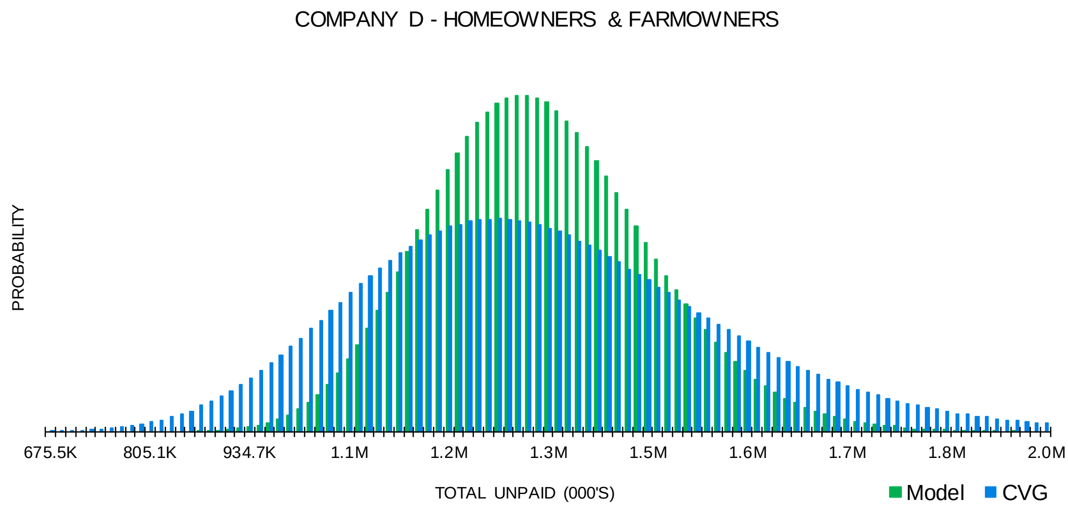
** Model Results based on weighting of 4 different models.*

FIGURE 31: COMMERCIAL AUTO



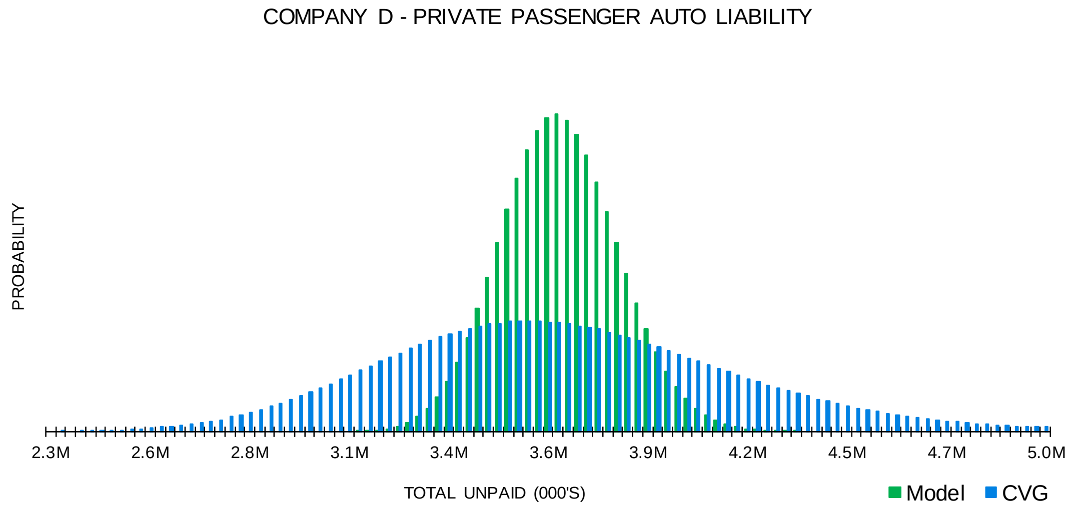
	Mean	Std Dev	CoV	75.0%	90.0%	95.0%	99.5%
ODP Pd CL Results	4,674,180	242,389	5.2%	4,833,953	4,988,442	5,083,250	5,334,504
CVG Benchmark	4,674,180	1,039,762	22.2%	5,291,680	6,046,915	6,549,510	8,036,457

FIGURE 32: HOMEOWNERS AND FARMOWNERS



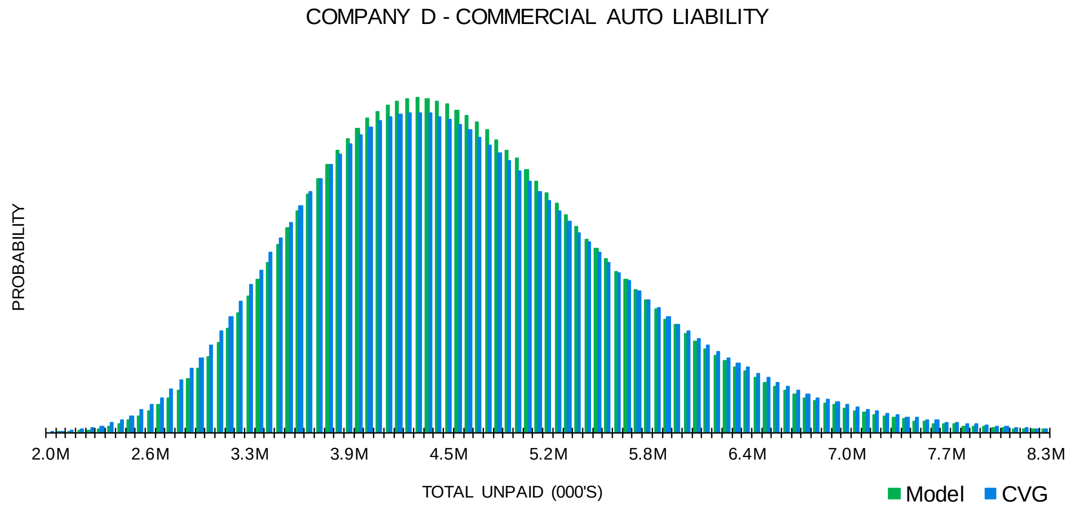
	Mean	Std Dev	CoV	75.0%	90.0%	95.0%	99.5%
ODP Pd CL Results	1,321,026	144,734	11.0%	1,413,573	1,510,489	1,571,639	1,739,877
CVG Benchmark	1,321,026	234,524	17.8%	1,464,786	1,630,109	1,737,843	2,047,547

FIGURE 33: PRIVATE PASSENGER AUTO LIABILITY



	Mean	Std Dev	CoV	75.0%	90.0%	95.0%	99.5%
ODP Pd CL Results	3,692,302	158,689	4.3%	3,797,346	3,897,678	3,958,986	4,120,528
CVG Benchmark	3,692,302	464,946	12.6%	3,986,781	4,302,201	4,502,780	5,060,502

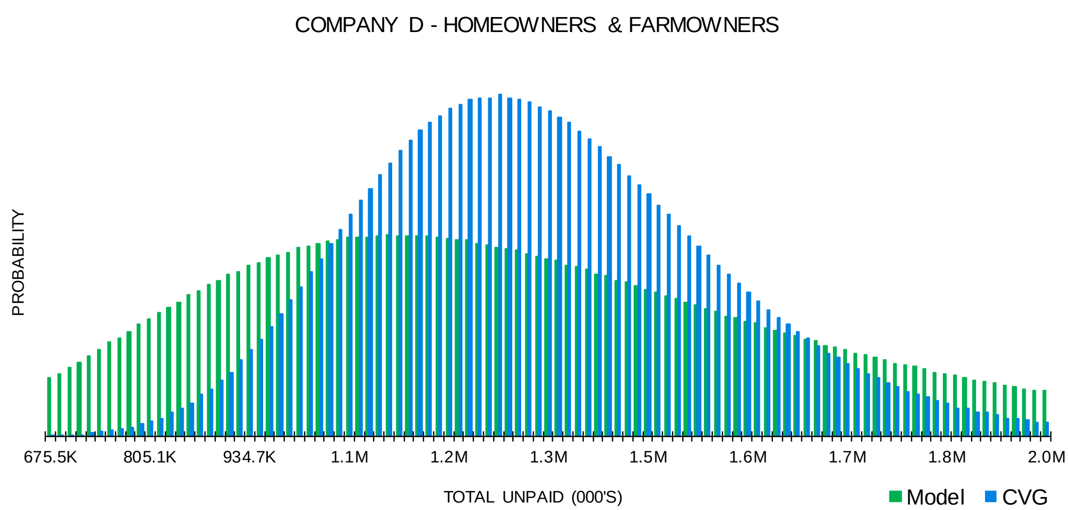
FIGURE 34: COMMERCIAL AUTO



	Mean	Std Dev	CoV	75.0%	90.0%	95.0%	99.5%
Model Results*	4,674,180	989,866	21.2%	5,266,643	5,980,746	6,453,616	7,843,130
CVG Benchmark	4,674,180	1,039,762	22.2%	5,291,680	6,046,915	6,549,510	8,036,457

** Model Results based on weighting of 4 different models.*

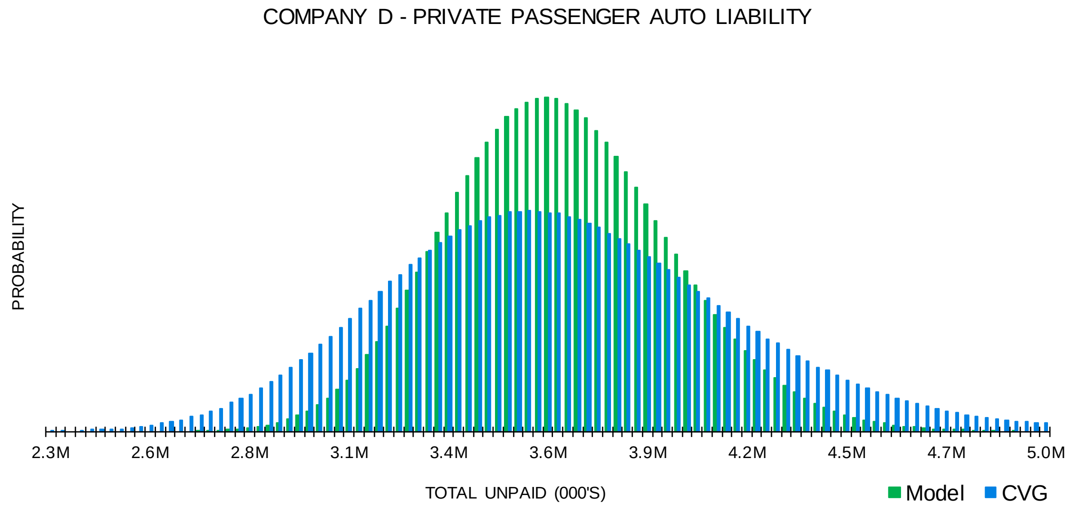
FIGURE 35: HOMEOWNERS AND FARMOWNERS



	Mean	Std Dev	CoV	75.0%	90.0%	95.0%	99.5%
Model Results*	1,321,026	434,146	32.9%	1,557,583	1,891,838	2,125,268	2,863,504
CVG Benchmark	1,321,026	234,524	17.8%	1,464,786	1,630,109	1,737,843	2,047,547

** Model Results based on weighting of 4 different models.*

FIGURE 36: PRIVATE PASSENGER AUTO LIABILITY



	Mean	Std Dev	CoV	75.0%	90.0%	95.0%	99.5%
Model Results*	3,692,302	303,392	8.2%	3,889,242	4,087,821	4,211,479	4,545,703
CVG Benchmark	3,692,302	464,946	12.6%	3,986,781	4,302,201	4,502,780	5,060,502

** Model Results based on weighting of 4 different models.*