MILLIMAN RESEARCH REPORT

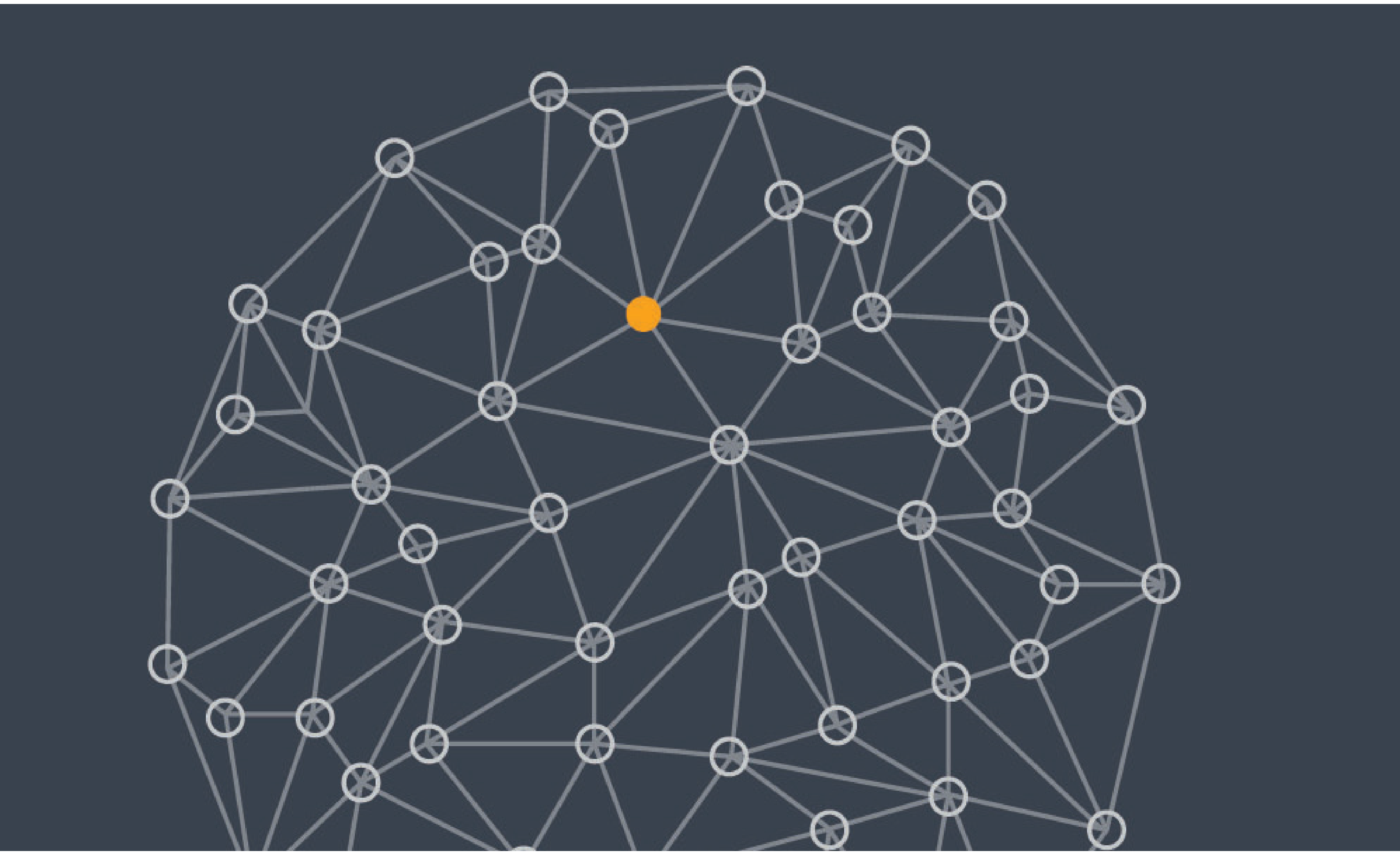
A Comparative Study of the 1-Factor Hull White and the $G2++$ Interest Rate Model

Prepared by:

Marcus Scheffer, Ph.D.
Mario Zacharias, Ph.D.

Peer Reviewed by:

Tigran Kalberer, MA, DAV, SAV



Milliman

Table of Contents

1 Executive Summary

Economic scenarios play a key role in the context of quantitative risk-management and insurance regulatory frameworks such as Solvency II (SII) or the Swiss Solvency Test (SST). The characteristics of these scenarios are heavily affected by the choice of the underlying simulation model in particular with regards to interest rates. Following the criticism of the theoretical possibility of negative rates, risk managers have traditionally used a log-normal convention for calibrating market-consistent scenarios. However, due to the current market environment (high implied volatilities, undulating surface and extremely low (and even negative) nominal interest rates), calibrating the interest rate model has become increasingly difficult. Due to this fact and the growing empirical evidence in favor of normally distributed interest rate changes, among market practitioners recently a shift in convention towards normal models has been observed.

The SII regulations as well as the German supervisor BaFin do not recommend any specific interest rate model. In the German insurance market, the 1-Factor Hull-White model is widely used as it can be considered the simplest of the usual risk-neutral short rate models for arbitrary starting yield curves. In particular, the Association of German Insurers (GDV) provides its members (market share in terms of number of insurance companies 50%-75%) with an Economic Scenario Generator (ESG) based on the 1-Factor Hull-White model and the German Actuarial Association (DAV) publishes periodically a standard calibration of this model. In contrast, the Swiss supervisory authority FINMA has recently included detailed specifications of the mandatory ESG methodology in its 2018 update of the SST standard model. The stipulated methodology is based on a $G2++$ interest rate model, which is equivalent to a 2-Factor Hull-White model.

That is precisely the reason the proposed research focuses on a comparison of these two calibration approaches and the respective underlying short rate models: the 1-Factor Hull White model and the $G2++$ model. Despite the good tractability of the 1-Factor Hull-White model, it has two major drawbacks. Firstly, the model is not capable of reproducing a large volatility surface satisfactorily, due to the lack of free calibration parameters. The $G2++$ model, on the other hand, is much more flexible in terms of fitting different volatility surfaces due to the introduction of a second factor and the mutual correlation parameters. Secondly, the 1-Factor Hull-White model has often been criticized for the number and intensity of the generated highly negative interest rates which, due to the characteristics of the German life insurance products, have a major impact on the valuation of the liabilities and the time value of options and guarantees (TVOGs). Here, the general behavior of the $G2++$ model is much less clear.

We thus investigate the model behavior of the 1-Factor Hull-White model and the $G2++$ model, respectively. We implement both model approaches, calibrate the models to current data and analyze the goodness of fit. Moreover, we provide an extended simulation study, which focusses on the number and intensity of the generated highly negative interest rates and the increased market consistency with respect to the volatility surface.

The remainder of this paper is structured as follows. The calibration methodology of the 1-Factor Hull-White model and the $G2++$ model, respectively, are presented and discussed in Section 2. Section 3 contains the calibration of the models to empirical data and a comprehensive simulation study, which is completed by an extensive discussion of the model characteristics with respect to negative interest rates and the fit to a given volatility surface. Section 4 concludes the report.

2 Calibration of 1-Factor Hull-White and G2++ Model

In the context of using ESGs for risk-management purposes in regulatory frameworks like SII and SST, it is worth mentioning that the characteristics of the produced scenarios crucially depend on the choice of the underlying simulation model, in particular with regards to interest rates. One of the most popular implementation is to take the short-rate as the basis for modeling the term-structure of interest rates.

In this chapter the calibration methodology of two short rate models, which are proposed by the DAV and FINMA, respectively, is presented. At first, the so-called DAV approach, i.e. the calibration methodology of the 1-Factor Hull White model is presented in section 2.1. The prescribed methodology of the FINMA is based on the $G2++$ interest rate model, which is equivalent to a 2-Factor Hull-White model and outlined in section 2.2. Both short rate models are calibrated to observed implied market volatilities of at-the-money swaptions. Swaptions are considered to be the most liquid contracts available, and thereby most representative for the market.

2.1 CALIBRATION OF THE 1-FACTOR-HULL-WHITE MODEL

The 1-Factor Hull-White model assumes that the dynamics of the short rate r_t is given by

$$dr_t = (\theta_t - \beta\, r_t\,)\, dt + \sigma\, dW_t, \qquad r_0 = 0,$$

where β is the mean reversion constant and σ is the volatility parameter. The function θ_t is chosen so that the model fits the current term structure of interest rates exactly and is defined as follows:

$$\theta_t = \frac{\partial F(0,t)}{\partial T} + \beta\, F(0,t) + \frac{\sigma^2}{2\beta}\,(1 - \exp(-2\beta t)).$$

Here $F(0,t)$ denotes the market instantaneous forward rate at time 0 for maturity t, which can be expressed by

$$F(0,t) \;= -\frac{\partial \ln P\,(0,T)}{\partial T},$$

with the market zero-coupon price $P(0,T)$ for maturity T.

For the calibration of the model, apart from θ_t, which can be completely determined by the risk-free yield curve, the model parameters β and σ have to be specified. According to the DAV, the parameter β is assigned by a fixed value (1% and 10% respectively) as this is „consensus among practitioners“.[1] Therefore, the calibration of the 1-Factor Hull-White is reduced to a one-dimensional optimization problem. More precisely the determination of the volatility parameter σ is handled by replicating the price of a (10,10)-swaption.

To put it briefly, the calibration of the 1-Factor Hull-White model according to the DAV approach is implemented by the following two steps:[2]

i. Calculate the corresponding swaption price U for the target at-the-money volatility $\sigma_{10,10}$. The strike equals the forward swap-rate at time 0.

ii. To determine the volatility parameter σ the following steps are conducted iteratively:

   a. For given β and fixed σ

      - determine the critical short-rate of the Jamshidian decomposition via Newton-Iteration and
      - calculate the corresponding swaption price $V(\beta,\sigma)$.

   b. The objective function of the minimization problem is then given by

$$(V(\beta,\sigma) - U)^2.$$

[1] See (DAV - Ausschuss Invest, 2015), p. 9 and (DAV - Ausschuss Invest, 2018), p. 6, respectively.
[2] See (DAV - Ausschuss Invest, 2015), p. 8 ff.

2.2 CALIBRATION OF THE G2++ MODEL

The $G2++$ model specifies the instantaneous spot rate by the sum of two correlated Gaussian processes and a deterministic function. The model is equivalent to the well-known 2-Factor Hull-White model[3], but the formulation with two additive factors leads to less complicated formulas and is easier to implement in practice. This is a particularly crucial point given the fact that the FINMA[4] approves the use of the $G2++$ model for the SST standard model in its version from January 2018.

In the $G2++$ model the dynamic of the short rate r_t is determined by two stochastic factors x_t and y_t

$$r_t = x_t + y_t + \psi_t, \qquad r_0 \in \mathbb{R}$$

and a deterministic time-dependent function ψ_t. The stochastic processes $\{x_t: t \geq 0\}$ und $\{y_t: t \geq 0\}$ are simple 1-factor Hull White processes with zero long term mean, defined as

$$dx_t = -a\, x_t\, dt + \sigma\, dW_{1,t}, \qquad x_0 = 0,$$

$$dy_t = -b\, y_t\, dt + \eta\, dW_{2,t}, \qquad y_0 = 0,$$

where $a, b, \sigma, \eta > 0$ are calibration parameters. The tuple (W_1, W_2) is a two-dimensional Wiener process, where the covariance satisfies

$$Cov\left(dW_{1,t}, dW_{2,t}\right) = \rho dt$$

with $-1 \leq \rho \leq 1$.

The calibration of this interest rate model is performed on the basis of an initial yield curve and a given ATM swaption volatility surface and implemented by the following steps:

i. Use the prescribed swaption parameter (target volatility, volatility type, and strike) and the initial yield curve, the target swaption price for each combination of option term (T_0) and swap tenor (T) is determined. For a normal implied volatility σ^{IV} of an ATM swaption the price is given as

$$P^{target}\,(\sigma^{IV}, T) = \sigma^{IV} \cdot T \cdot \varphi(0) \sum_{i=1}^{n} ZCB_{T_i}(0)$$

Here, φ denotes the normal density function and $ZCB_{T_i}(0)$ is the initial price of a zero coupon bond with maturity T_i. The summation extends over all payment times where $T_n = T_0 + T$ is the final payment.

ii. Leveraging the explicit formula for swaption valuation in the $G2++$ models the parameters a, b, σ, η, ρ are determined via an optimization algorithm.[5] This algorithm minimizes the sum of relative amount differences between the target prices P_j^{target} and the explicit formula prices P_j, i.e. the optimization function is given by

$$f = \sum_{j=1}^{N} \left| \frac{P_j^{target}\left(\sigma_j^{IV}, T_j\right) - P_j\left(a, b, \sigma, \eta, \rho, T_j\right)}{P_j^{target}\left(\sigma_j^{IV}, T_j\right)} \right|,$$

where the sum runs over the N given target swaptions. The minimization is based on the Nelder Mead method.

Note, that besides the process parameters a, b, σ and η also the correlation coefficient ρ of the two Wiener processes is determined during the calibration algorithm.

[3] See (Brigo & Mercurio, 2006), p.159 ff., for a prove of the analogy between these two approaches.
[4] See (FINMA, SST-Standardmodell für Lebensversicherungen, 2018).
[5] For more details see (Brigo & Mercurio, 2006), p.158 f.

3 Calibration Results and Simulation Study

In this section we explore the model behavior of the 1-Factor Hull-White model and the $G2++$ model, respectively. We compare the performance of three different modeling approaches, the 1-Factor Hull-White model with the GDV-benchmark mean reversion speed parametrization of $\beta = 0.1$ and the alternative parametrization of $\beta = 0.01$, respectively, and the $G2++$ model. After calibrating the models in accordance with the methodology described in chapter 2 to current data we analyze the goodness of fit of the models with respect to the volatility surface. Moreover, as the 1-Factor Hull-White model has often been criticized for generating unbounded negative interest rates we provide an extended simulation study, which focusses on the number and intensity of the produced negative interest rates in a well-balanced reflection including the increased market consistency. This is of particular interest for insurance industries, where negative interest rates have a major impact on the valuation of the liabilities and the TVOGs, e.g. Germany[6] or Switzerland.

3.1 CALIBRATION RESULTS

For our simulation study we use the specifications from the current Swiss Solvency Test[7] with respect to the euro area. According to this the implied normal swaption volatility surface at the reference date 31 December 2017 is shown in Table 1.

INITIAL DATA

MATURITY / TENOR	SWAPTION VOLATILITY SURFACE 5	10	15	20	25
5	0.6225	0.6316	0.6008	0.5899	0.5751
10	0.6671	0.6554	0.6071	0.5816	0.5596
15	0.6398	0.6214	0.5701	0.5355	0.5118
20	0.5984	0.5742	0.5244	0.4852	0.4603
25	0.5594	0.5279	0.4804	0.4413	0.4185
30	0.5264	0.4854	0.4414	0.4050	0.3846

Table 1: EUR swaption ATM volatility surface. The table reports the implied normal swaption volatility surface for maturity tenor combinations (5,5), (5,10), …, (30, 25) at 31/12/2017.

The implied volatilities presented in Table 1 are then used to fit the 1-Factor Hull-White model in accordance to the methodology as described in chapter 2.1 for both, the benchmark mean reversion parametrization of $\beta = 0.1$ and the alternative parametrization of $\beta = 0.01$. The corresponding results are presented in Table 2 and Table 3, respectively. After that Table 4 reports the estimates for the $G2++$ model, obtained by using the calibration methodology presented in chapter 2.2.

1-FACTOR HULL-WHITE
$\boldsymbol{\beta = 0.1, \sigma = 0.015}$

MATURITY / TENOR	CALIBRATED SWAPTION VOLATILITY SURFACE 5	10	15	20	25
5	0.9674 (55.4%)	0.7807 (23.6%)	0.6492 (8.1%)	0.5548 (-5.9%)	0.4826 (-16.1%)
10	0.8068 (20.9%)	0.6553 (0.0%)	0.5471 (-9.9%)	0.4654 (-20.0%)	0.4015 (-28.3%)
15	0.6921 (8.2%)	0.5617 (-9.6%)	0.4653 (-18.4%)	0.3923 (-26.7%)	0.3396 (-33.7%)
20	0.6052 (1.1%)	0.4880 (-15.0%)	0.4022 (-23.3%)	0.3409 (-29.7%)	0.2979 (-35.3%)
25	0.5446 (-2.6%)	0.4401 (-16.6%)	0.3654 (-23.9%)	0.3130 (-29.1%)	0.2756 (-34.1%)
30	0.5031 (-4.4%)	0.4090 (-15.7%)	0.3426 (-22.4%)	0.2962 (-26.9%)	0.2627 (-31.7%)

Table 2: Calibrated swaption volatility surface of the 1-Factor Hull-White model with $\boldsymbol{\beta = 0.1}$. The table reports the calibrated swaption volatility surface of the 1-Factor Hull-White model with mean reversion parameter $\beta = 0.1$ for the given maturity tenor combinations (5,5), (5,10), …, (30, 25). Figures in parentheses denote the relative deviation from the target value.

[6] Due to the characteristics of the German life traditional saving products with retirement features and profit participation.
[7] More specifically, this means the Last Liquid Point is set at 30 years and the Ultimate Forward Rate (UFR) corresponds to 3.75%.

1-FACTOR HULL-WHITE
$\boldsymbol{\beta = 0.01,\ \sigma = 0.007}$

MATURITY / TENOR	5	10	15	20	25
	CALIBRATED SWAPTION VOLATILITY SURFACE				
5	0.6826 (9.7%)	0.6621 (4.8%)	0.6470 (7.7%)	0.6364 (7.9%)	0.6213 (8.0%)
10	0.6706 (0.5%)	0.6557 (0.0%)	0.6430 (5.9%)	0.6266 (7.7%)	0.6024 (7.6%)
15	0.6559 (2.5%)	0.6412 (3.2%)	0.6224 (9.2%)	0.5985 (11.8%)	0.5732 (12.0%)
20	0.6371 (6.5%)	0.6183 (7.7%)	0.5948 (13.4%)	0.5712 (17.7%)	0.5487 (19.2%)
25	0.6221 (11.2%)	0.6027 (14.2%)	0.5810 (20.9%)	0.5609 (27.1%)	0.5413 (29.4%)
30	0.6130 (16.5%)	0.5949 (22.6%)	0.5762 (30.5%)	0.5585 (37.9%)	0.5412 (40.7%)

Table 3: Calibrated swaption volatility surface of the 1-Factor Hull-White model with $\boldsymbol{\beta = 0.01}$. The table reports the calibrated swaption volatility surface of the 1-Factor Hull-White model with mean reversion parameter $\beta = 0.01$ for the given maturity tenor combinations (5,5), (5,10), …, (30, 25). Figures in parentheses denote the relative deviation from the target value.

G2++
$\boldsymbol{a = 0.025, b = 0.182,}$
$\boldsymbol{\sigma = 0.008, \eta = 0.001, \rho = 0.025}$

MATURITY / TENOR	5	10	15	20	25
	CALIBRATED SWAPTION VOLATILITY SURFACE				
5	0.7016 (12.7%)	0.6584 (4.2%)	0.6221 (3.5%)	0.5929 (0.5%)	0.5624 (-2.2%)
10	0.6664 (-0.1%)	0.6284 (-4.1%)	0.5967 (-1.7%)	0.5638 (-3.1%)	0.5285 (-5.6%)
15	0.6308 (-1.4%)	0.5952 (-4.2%)	0.5595 (-1.9%)	0.5225 (-2.4%)	0.4887 (-4.5%)
20	0.5941 (-0.7%)	0.5567 (-3.0%)	0.5192 (-1.0%)	0.4852 (-0.0%)	0.4556 (-1.0%)
25	0.5640 (0.8%)	0.5276 (-0.0%)	0.4940 (2.8%)	0.4645 (5.2%)	0.4389 (4.9%)
30	0.5417 (2.9%)	0.5086 (4.8%)	0.4787 (8.4%)	0.4524 (11.7%)	0.4297 (11.7%)

Table 4: Calibrated swaption volatility surface of the $G2++$ model. The table reports the calibrated swaption volatility surface of the $G2++$ model for the given maturity tenor combinations (5,5), (5,10), …, (30, 25). Figures in parentheses denote the relative deviation from the target value.

Starting with the results for the 1-Factor Hull-White model with benchmark parametrization $\beta = 0.1$ in Table 2, several findings are noteworthy. First, with the estimated volatility parameter $\sigma = 0.015$ the target volatility of the (10,10)-swaption is nearly perfectly replicated. Moreover, regarding the relative deviations of 1.1%, -2.6% and -4.4% also the volatilities of the maturity-tenor-combinations (20,5), (25,5) and (30,5) are good approximations concerning to the corresponding target values. But, for other volatilities, the calibration shows more significant deviations, in many cases the relative difference is even higher than 20%. The maximum deviation with a value of 55.4% can be observed for maturity-tenor-combination (5,5).

We now turn our attention to the results of the 1-Factor Hull-White model with the alternative mean reversion parametrization $\beta = 0.01$ presented in Table 3. Here, the optimization algorithm yields the estimated volatility parameter $\sigma = 0.015$ and, again, the target volatility of the (10,10)-swaption is replicated almost perfectly. Furthermore, the alternative modeling approach also shows an adequate fit to some swaption volatilities with tenor 5, but in contrast to the benchmark model this holds only for shorter maturities. To be precise, the relative deviations for the maturity-tenor-combinations (10,5) and (15,5) are very low with 0.5% and 2.5%, respectively. The maximum deviation here is 40.7% and corresponds to the maturity-tenor-combination (30,25). Overall, the results emphasize that the alternative parametrization of the 1-Factor Hull-White model in most cases reveals significant lower differences between the calibrated volatilities and the corresponding target values and thus clearly outperforms the benchmark model. Finally, and also quite remarkable is the fact that for this modeling approach, the volatilities are generally overestimated, i.e. the calibrated values are always greater than the target values.

The parameters of the $G2++$ model estimated by the calibration methodology presented in chapter 2.2. correspond to $a = 0.025, b = 0.182,\ \sigma = 0.008, \eta = 0.001$ and $\rho = 0.025$. Regarding the results of the calibrated swaption volatility surface presented in Table 4, it is not surprising that the model does not fit the (10,10)-swaption volatility perfectly like both 1-Factor Hull-White model approaches. This of course stems from the fact that the Hull-White model approaches are calibrated on exactly this target at-the-money volatility. But, however, the (10,10)-swaption is met very satisfactory, the relative deviation is only -4.1% and thus quite moderate. Besides this, several target volatilities are met within a one percent deviation window. This applies for the calibrated volatilities with maturity-tenor-combinations (5,20), (10,5), (20,5), (20,15), (20,20), (20,25), (25,5) and (25,10). The maximum relative difference for the $G2++$ model is 12.7%, which corresponds to the (5,5)-swaption. But, however, short tenor swaptions are usually less relevant for valuation purposes, so that calibration result is still quite satisfying. Thus, it becomes clearly evident that the $G2++$ model seems to adapt well the given volatility surface and significantly outperforms both 1-Factor Hull-White modeling approaches.

It is very important to have a good fit for all term/tenor combinations as it typically is not known, which of these are relevant in valuing life insurance liabilities. Indeed, if the liabilities, by some coincidence, would resemble a (15,15) swaption, using the 1-Factor Hull-White model with benchmark parametrization $\beta = 0.1$ would result in a valuation error of -18.4%, using the 1-Factor Hull-White model with alternative parametrization $\beta = 0.01$ would result in a valuation error of +9.2% and using the $G2++$ model would result in a valuation error of -1.9%. Using the $G2++$ model results in much more stable results for all possible cases.

To further substantiate the previous findings, we show the absolute value of the relative differences between target prices and calibrated prices for the three calibrated models in Figure 1.

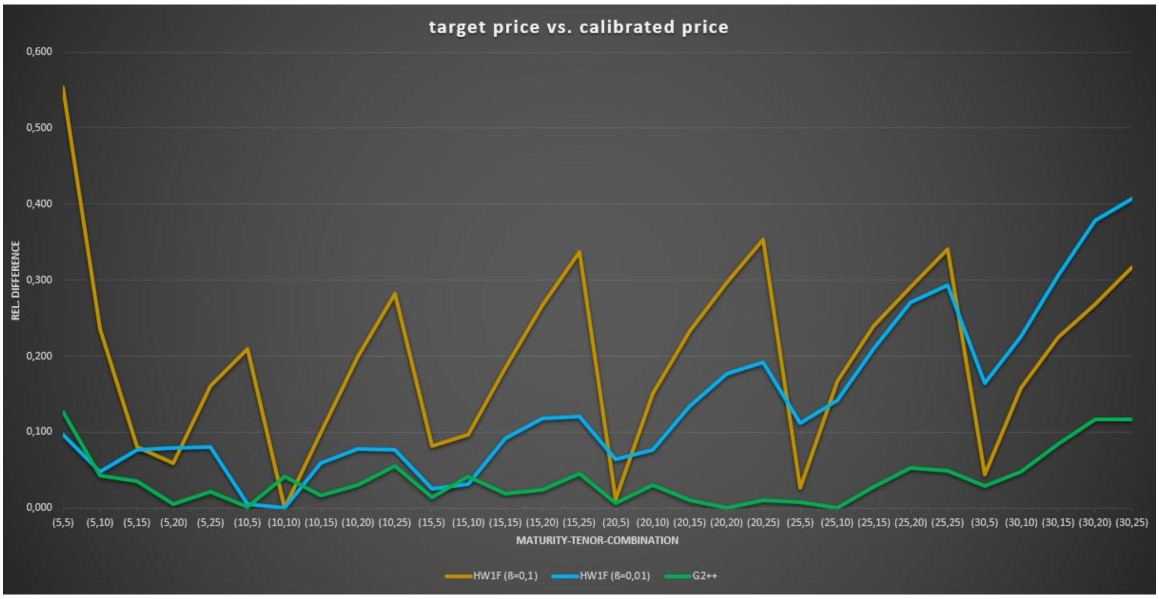


Figure 1: **Relative differences between target prices and calibrated prices.** The figure shows of the relative difference between the target price and the calibrated prices of the three different modeling approaches. The different term-tenor combinations are represented on the x-axis, while the relative differences are represented on the y-axis.
The relative difference of benchmark 1-Factor Hull-White model with mean reversion parameter $\beta = 0.1$ is represented by the orange line, the alternative modeling approach with parameter $\beta = 0.01$ is depicted by the blue line and the green line corresponds to the $G2++$ model.

From the plot the well-suited calibration of the $G2++$ model becomes apparent. Only at the left and at the right margin, respectively, the relative deviation for this model approach, depicted by the green line, exceeds 10%. Apart from this, the difference is quite low, which underlines the excellent adaptability of the model to capture the characteristics of the given volatility surface. Another interesting finding is the shape of the blue curve, which represents the relative difference between target and calibrated prices for the alternative 1-Factor Hull-White model ($\beta = 0.01$). Here, for short-term and mid-term maturities the calibrated model yields appropriate swaption prices, whereas for longer maturities the fit becomes worse and worse and, at the right margin the relative differences even exceed those ones stemming from the benchmark 1-Factor Hull-White model with mean reversion parameter $\beta = 0.1$. Except at the left margin, from the plot it becomes clear that the shape of the curve of the corresponding benchmark model equals the well-known sawtooth pattern. In other words, for a fixed maturity the relative difference tends to rise up with increasing tenor.

The results emphasize that both 1-Factor Hull-White model approaches yield excellent replications of the (10,10)-swaption volatility. However, as a whole it can be pointed out that both, the benchmark model as well as the alternative model approach, fit the volatility surface considerably inadequate. In contrast, the $G2++$ model does not replicate the (10,10)-swaption volatility perfectly, but provides a very appropriate fit throughout the complete volatility surface and thus seems to be a sensible approach to capture characteristics of the market. To deepen our analysis, we now turn our attention to the usage of the three modeling approaches within the scope of ESG and address the frequent problem in insurance applications of generating quite extreme and to many negative interest rate scenarios.

3.2 SIMULATION STUDY

The 1-Factor Hull-White model has often been criticized for the possibility of (unbounded) negative interest rates which, for German life insurance products, could have a major impact on the valuation of the liabilities and the TVOGs. The general behavior of the $G2++$ model regarding this point is much less clear. We therefore perform an extended simulation study with 100 simulation runs each generating 1.000 simulation paths for the three calibrated modeling approaches presented in section 3.1. In the analysis we turn our attention on the number and intensity of the produced negative interest rates with respect to the short rate and different maturities of the spot rate.

First of all, simulations of the short rate are presented in Figure 2. For increased transparency, we provide only one randomly chosen simulation plot for each modeling approach, which are representatives for all 100 generated simulation runs.

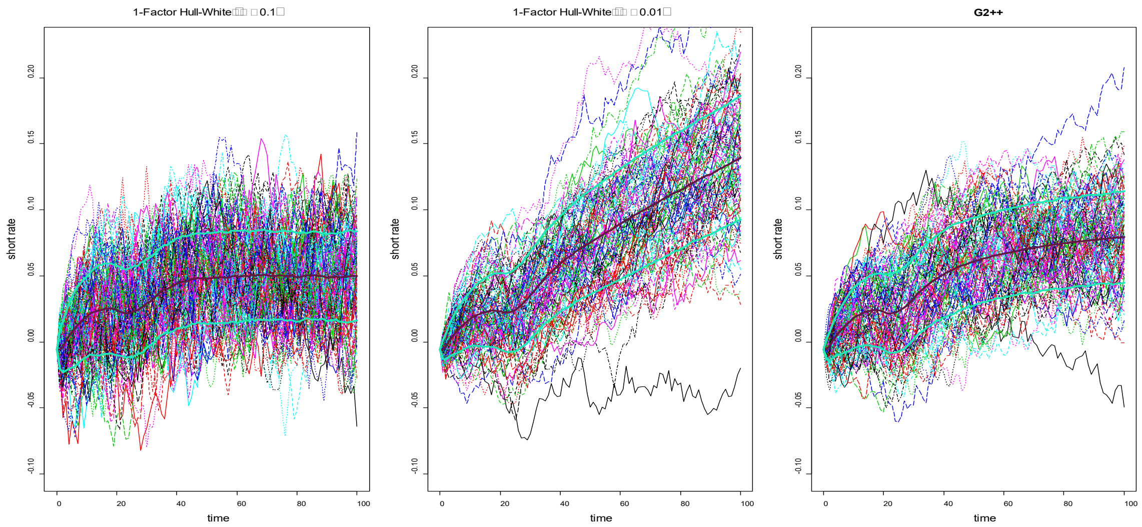


***Figure 2*: Short rate simulations.** 1000 simulated future paths for the instantaneous short rate for the calibrated 1-Factor Hull-White model with mean-reversion parameter $\beta = 0.1$ (left panel) and mean-reversion parameter $\beta = 0.01$ (mid panel), respectively, and the $G2++$ model (right panel), each of them along with the mean trajectory and the standard deviation around the mean. While time is represented on the horizontal axis, the short rates are represented on the vertical axis.

The plots highlight several interesting characteristics of the different modeling approaches. First, it is obvious that all three models produce negative interest rates. But while the benchmark 1-Factor Hull-White modeling approach with mean-reversion parameter $\beta = 0.1$ in the left panel depicts the highest number and the most negative level of interest rates, the corresponding model parametrization with $\beta = 0.01$, presented in the mid panel, leads to significantly larger short rates towards the end of the projection. Here, a considerable number of simulation paths end up at 20%. These findings are also underlined by the course of the mean trajectories in both plots, which remains at a relatively constant level of 5% from point of time 40 onwards for the model parametrization $\beta = 0.1$ and sharply rises nearly up to 15% for the alternative modeling approach with mean-reversion parameter $\beta = 0.01$, respectively. Regarding the right panel, with respect to the number and the level of interest rates the simulations of the $G2++$ model seem to move between the both 1-Factor Hull-White modeling approaches. Moreover, the mean trajectory increases slightly up to 7.5% approximately. Note, that these findings hold throughout the whole simulation sample.

In the next step, we turn our attention to the characteristics of simulated 10y spot rates. The 10y spot rate is generally an important benchmark and plays a distinct role in the German regulatory framework. Since 2011, life insurers have been subject to an Additional Interest Provision, the so-called “Zinszusatzreserve” (ZZR), which forces them to increase the reserve requirements for those cohorts of contracts which have a guaranteed rate of return that is higher than the reference interest rate. In this context the reference interest rate corresponds to the ten year moving average of the ten year Euro Swap rate.

Therefore we now investigate the characteristics of the 10y spot rate throughout the whole projection period and present plots of simulations of the 10y spot rate in the following Figure 3. Again, for increased transparency we provide only one simulation plot for each modeling approach.

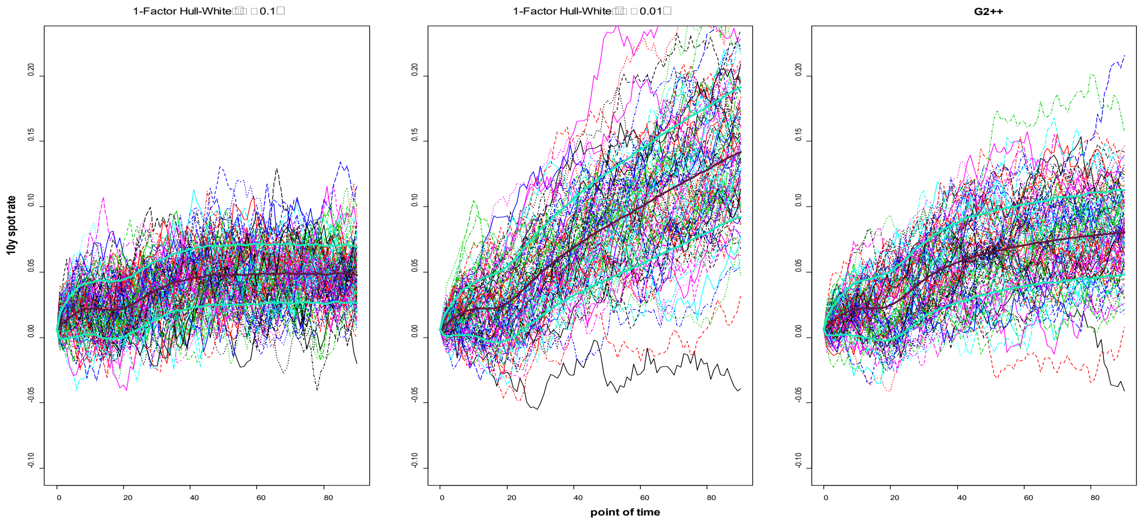


Figure 3: **Simulations of 10y spot rate.** 1000 simulated future paths for the 10y spot rate for the calibrated 1-Factor Hull-White model with mean-reversion parameter $\beta = 0.1$ (left panel) and with mean-reversion parameter $\beta = 0.01$ (mid panel), respectively, and the $G2++$ model (right panel), each of them along with the mean trajectory and the standard deviation around the mean. While time is represented on the horizontal axis, the 10y spot rates are represented on the vertical axis.

The plots in the left and mid Panel in Figure 3 show the different characteristics of the 1-Factor Hull-White model emerging from distinct parametrizations. Regarding simulations of the benchmark 1-Factor Hull-White model on the left hand side and in the mid panel, respectively, the impact of mean-reversion parameter becomes apparent. The relatively high mean-reversion speed parameter of $\beta = 0.1$ (left Panel) causes a stabilization around the mean, i.e. spikes tend to revert very quickly. On the other hand, the alternative parametrization of $\beta = 0.01$ indicates a rather slow mean-reversion and therefore reveals some extremely high spot rate scenarios (mid panel). On the contrary, in the simulations of the $G2++$ model (right Panel) extremely low and high simulation paths occur only in a small number of cases.

To analyze extreme interest rate levels more deeply, we now explore characteristics of the 10y spot rate throughout the whole projection period. To be precise, we specify a threshold level of -2.5% and depict the average number of exceedances (i.e. interest rates below -2.5%) of the 10y spot rate of the 100 simulation runs in Figure 4.

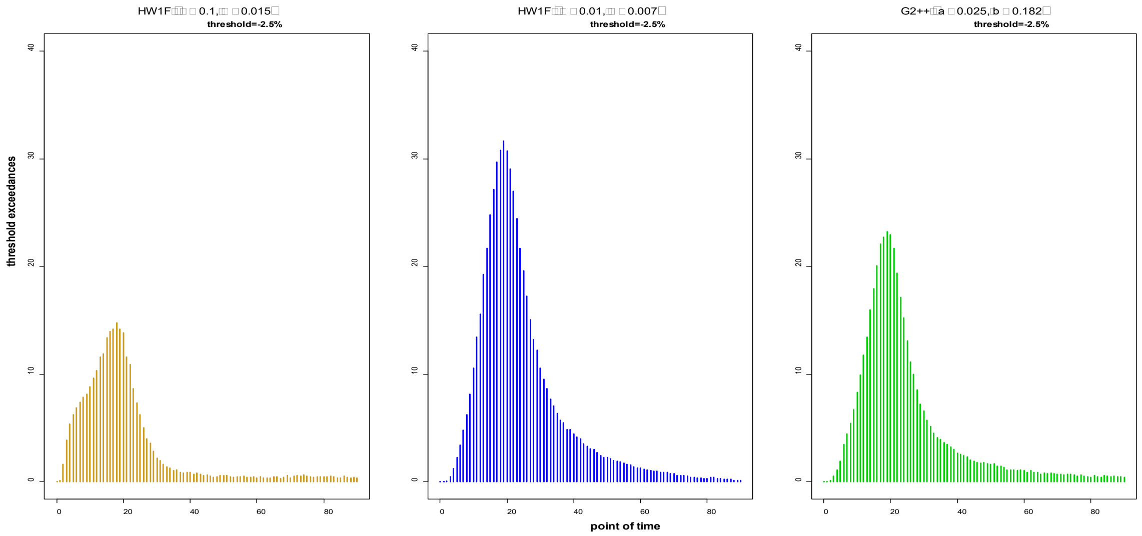


***Figure 4*: Comparison of the average number of threshold exceedances.** The plots depict the average number of exceedances of the 10y spot rate across the 100 simulation runs subject to the point of time. The left Panel and the mid Panel show the average number of exceedances resulting from simulations of the 1-Factor Hull-White model with mean-reversion parameter $\beta = 0.1$ and $\beta = 0.01$, respectively. In the right Panel the corresponding average number of exceedances of the $G2++$ model is illustrated. While time is represented on the horizontal axis, the average number of exceedances is represented on the vertical axis.

The plots in the left and mid Panels in Figure 4 show the different characteristics of the 1-Factor Hull-White model emerging from distinct parametrizations. On the left hand side, at the beginning the plot of the benchmark 1-Factor Hull-White model with mean-reversion parameter $\beta = 0.1$ shows a hump with a maximum average number of exceedances after around 20 years (peak value ~ 15) and a sharp decline subsequently. In contrast, the alternative 1-Factor Hull-White model parametrization ($\beta = 0.01$) in the mid panel shows considerably more threshold exceedances for almost every point of time. In particular, after the maximum peak (value ~ 32) the decrease is rather moderate. In comparison, the pattern of threshold exceedances of the G2++ model in the right panel depicts a remarkable similarity to the alternative 1-Factor Hull-White model parametrization, but, however, the average number of exceedances is significantly lower (maximum peak value ~ 24). Finally, it is worth mentioning the three plots underline our first impression from Figure 2 that the $G2++$ model seem to move between the both 1-Factor Hull-White modeling approaches.

The analysis so far shows that the $G2++$ model outperforms both 1-Factor Hull-White modeling approaches with respect to the accuracy of fit of the volatility surface. But, the $G2++$ model reveals a considerably higher number of average threshold exceedances at -2.5% regarding 10y spot rate simulations from the 1-Factor Hull-White model with mean-reversion parameter $\beta = 0.1$. At his point it is natural to ask, whether or not the number of threshold hits crucially depends on the specified interest rate level and spot rate maturity, respectively. To further assess these dependencies, we calculate the relative frequency of threshold exceedances over the whole projection period for various spot rate maturities and threshold levels and present the results in Table 5.

RELATIVE FREQUENCY OF THRESHOLD EXCEEDANCES

		THRESHOLD LEVEL						
Model	Spot Rate	-4%	-3.5%	-3%	-2.5%	-2%	-1.5%	-1%
$HW1F\ (\beta = 0.1)$	1y	0.0087	0.0133	0.0199	0.0291	0.0415	0.0577	0.0782
$HW1F\ (\beta = 0.01)$	1y	0.0031	0.0049	0.0077	0.0119	0.0182	0.0278	0.0420
$G2++$	1y	0.0027	0.0045	0.0073	0.0117	0.0185	0.0288	0.0440
$HW1F\ (\beta = 0.1)$	5y	0.0022	0.0040	0.0070	0.0116	0.0188	0.0294	0.0446
$HW1F\ (\beta = 0.01)$	5y	0.0022	0.0036	0.0058	0.0091	0.0141	0.0214	0.0322
$G2++$	5y	0.0016	0.0028	0.0047	0.0078	0.0127	0.0202	0.0315
$HW1F\ (\beta = 0.1)$	10y	0.0003	0.0007	0.0015	0.0031	0.0061	0.0113	0.0202
$HW1F\ (\beta = 0.01)$	10y	0.0014	0.0024	0.0041	0.0066	0.0105	0.0164	0.0250

$G2++$	10y	0.0008	0.0015	0.0027	0.0048	0.0082	0.0137	0.0222
$HW1F\ (\beta = 0.1)$	20y	0.0000	0.0000	0.0000	0.0001	0.0003	0.0010	0.0030
$HW1F\ (\beta = 0.01)$	20y	0.0005	0.0009	0.0017	0.0031	0.0054	0.0093	0.0155
$G2++$	20y	0.0001	0.0003	0.0007	0.0014	0.0029	0.0057	0.0108

Table 5: Comparison of the relative frequency of threshold exceedances. The table presents a comparison of the relative frequency of threshold exceedances with respect to various spot rate maturities and threshold levels, respectively, for the 1-Factor Hull-White model with mean reversion parameter $\beta = 0.1$ and $\beta = 0.01$ and the $G2++$ model. The relative frequency is given by the ratio of the number of threshold hits by the number of simulation paths times the number of time observations. All results are given in averages of 100 simulation runs.

The results shown in Table 5**Error! Reference source not found.** present several interesting insights concerning the characteristics of the three modeling approaches with respect to simulated negative interest rates. First, we can see from Table 5 that for the 1y spot rate the relative frequency of threshold exceedances of the benchmark 1-Factor Hull-White model with mean reversion parameter $\beta = 0.1$ is considerably larger than the alternative 1-Factor Hull-White parametrization and the $G2++$ model, respectively, for all considered threshold levels. Apart from this, the alternative 1-Factor Hull-White parametrization with $\beta = 0.01$ and the $G2++$ model show similar results. While the relative frequency of threshold hits of the $G2++$ model are slightly lower for the threshold levels -4%, -3.5%, -3% and -2.5%, the opposite holds for the levels -2%, -1.5% and -1%. Turning the attention to the 5y spot rate, it can be stated that simulations of the $G2++$ model reveal the least frequency of threshold exceedances in comparison with the both 1-Factor Hull-White modeling approaches. Regarding these, it is noteworthy that both models generate the same number of extreme negative simulation paths, but for lower negative threshold levels the benchmark parametrization of $\beta = 0.1$ exhibits considerably more threshold hits. The results of the 10y and 20y spot rate, respectively, show a clear order. The alternative 1-Factor Hull-White parametrization with $\beta = 0.01$ has the highest relative frequency of thresholds hits, followed by the $G2++$ model and benchmark 1-Factor Hull-White model parametrization with $\beta = 0.1$. Finally, from these findings it can be concluded that with increasing maturity the relative frequency of threshold exceedances decreases. Moreover, the results emphasize that the $G2++$ model reveals the lowest relative frequency of threshold hits for short-term interest rates and the benchmark 1-Factor Hull-White model the lowest for the long-term interest rates.

4 Summary and Conclusion

In this paper we provide a comparative study of two the widely used 1-Factor Hull-White and $G2++$ short rate models regarding their calibration methodology provided by DAV and FINMA. From the results of the model calibrations it becomes apparent that the $G2++$ model adapts well the given volatility surface and significantly outperforms both 1-Factor Hull-White modeling approaches. The results show that the alternative parametrization of the 1-Factor Hull-White $(\beta = 0.01)$ model reveals significant lower differences between the calibrated volatilities and the corresponding target values than the benchmark model $(\beta = 0.1)$.

In the simulation study, we explore the number and intensity of the generated negative interest rates with respect to the short rate and different maturities of the spot rate. The results show that the $G2++$ model reveals the lowest relative frequency of threshold exceedances for short-term interest rates and, on the other hand, the benchmark 1-Factor Hull-White model the lowest for the long-term interest rates. Also, for the long-term maturities the alternative 1-Factor Hull-White modeling approach by far displays the highest relative frequency of thresholds hits.

Our main result can be summarized as follows: Compared to the 1-Factor Hull-White model the $G2++$ model can be implemented without considerable additional efforts while at the same time reflecting the given volatility surface much better. This is extremely important as it is typically not clear which swaptions represent the liabilities best and thus, to avoid model error, the overall fit to the calibration surface has to be sufficiently good.

Moreover, for all examined spot rate maturities and a threshold level of -2.5% the $G2++$ model generates less negative interest rate scenarios than the alternative 1-Factor Hull-White modeling approach which avoids pathological valuation results.

5 Appendix

TERM STRUCTURE OF INTEREST RATES					
Maturity (year)	1	2	3	4	5
Zero coupon rates (percent p.a.)	-0.558	-0.452	-0.291	-0.131	0.014
	6	7	8	9	10
	0.146	0.268	0.385	0.496	0.601
	11	12	13	14	15
	0.697	0.782	0.859	0.928	0.989
	16	17	18	19	20
	1.043	1.088	1.125	1.153	1.171
	21	22	23	24	25
	1.181	1.185	1.187	1.188	1.190
	26	27	28	29	30
	1.195	1.203	1.215	1.232	1.254
	31	32	33	34	35
	1.281	1.312	1.346	1.383	1.421
	36	37	38	39	40
	1.460	1.499	1.539	1.579	1.619
	41	42	43	44	45
	1.658	1.697	1.735	1.773	1.809
	46	47	48	49	50
	1.845	1.881	1.915	1.948	1.981

Table 6: Structure of Interest Rates. The table presents the term structure of (nominal) interest rates for the euro zone at 31/12/2017 based on Bloomberg data as prescribed for the SST 2018. See *(FINMA, Swiss Solvency Test (SST), 2018)* for the specifications of the SST 2018.

References

Brigo, D., & Mercurio, F. (2006). *Interest Rate Models - Theory and Practice.* Berlin : Springer-Verlag .

DAV - Ausschuss Invest. (2015). *Zwischenbericht zur Kalibrierung und Validierung spezieller ESG unter Solvency II.* Von https://aktuar.de/unsere-themen/fachgrundsaetze-oeffentlich/2015-11-09_DAV-Ergebnisbericht_Kalibrierung%20und%20Validierung%20spezieller%20ESG_Update.pdf abgerufen

DAV - Ausschuss Invest. (2018). *Beispielhafte Kalibrierung und Validierung des ESG im BSM zum 31.12.2017.* Von https://aktuar.de/unsere-themen/fachgrundsaetze-oeffentlich/2018-01-22_Ergebnisbericht-Kalibrierung-Validierung-ESG-Q4-2017_final.pdf abgerufen

FINMA. (2018). *SST-Standardmodell für Lebensversicherungen.* Von https://www.finma.ch/de/ueberwachung/versicherungen/spartenuebergreifende-instrumente/schweizer-solvenztest-sst/ abgerufen

FINMA. (2018). *Swiss Solvency Test (SST).* Von https://www.finma.ch/de/~/media/finma/dokumente/dokumentencenter/myfinma/2ueberwachung/sst/sst-inputdaten.xlsx?la=de abgerufen

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

Marcus Scheffer
marcus.scheffer@milliman.com

Mario Zacharias
mario.zacharias@milliman.com

© 2018 Milliman, Inc. All Rights Reserved. The materials in this document represent the opinion of the authors and are not representative of the views of Milliman, Inc. Milliman does not certify the information, nor does it guarantee the accuracy and completeness of such information. Use of such information is voluntary and should not be relied upon unless an independent review of its accuracy and completeness has been performed. Materials may not be reproduced without the express consent of Milliman.