Bootstrapping the economy – a non-parametric method of generating consistent future scenarios

Müller, Ulrich A, Bürgi, Roland and Dacorogna, Michel M
SCOR Switzerland

30. October 2004

Bootstrapping the economy – a non-parametric method of generating consistent future scenarios

Ulrich A. Müller [1], *Roland Bürgi* [2] *and Michel M. Dacorogna* [3]

October 29, 2004

Abstract

The fortune and the risk of a business venture depends on the future course of the economy. There is a strong demand for economic forecasts and scenarios that can be applied to planning and modeling. While there is an ongoing debate on modeling economic scenarios, the bootstrapping (or resampling) approach presented here has several advantages. As a non-parametric method, it directly relies on past market behaviors rather than debatable assumptions on models and parameters. Simultaneous dependencies between economic variables are automatically captured. Some aspects of the bootstrapping method require additional modeling: choice and transformation of the economic variables, arbitrage-free consistency, heavy tails of distributions, serial dependence, trends and mean reversion. Results of a complete economic scenario generator are presented, tested and discussed.

Keywords: economic scenario generator (ESG), asset-liability management (ALM), bootstrapping, resampling, simulation, Monte-Carlo simulation, non-parametric model, yield curve model

[1]SCOR Switzerland Ltd, General-Guisan Quai 26, 8022 Zürich, Switzerland, email: UMueller@scor.com
[2]SCOR Switzerland Ltd, General-Guisan Quai 26, 8022 Zürich, Switzerland
[3]SCOR Switzerland Ltd, General-Guisan Quai 26, 8022 Zürich, Switzerland

Contents

1 Introduction and motivation

In this paper, a concept to model and simulate major parts of the world economy is presented. The economy is represented by a few key variables such as interest rates (yield curves, risk-free and with credit risk), inflation, Gross Domestic Product (GDP) and indices for equity, hedge funds and real estate investments, all of these for several currency zones, plus the foreign exchange (FX) rates between these zones. The goal is to generate scenarios that, in their entirety, represent the space of likely future developments. These scenarios can be used for simulating anything that depends on the economy.

Our main application is asset-liability management (ALM). ALM (see e.g. [Ziemba and Mulvey, 1998]) and Dynamic Financial Analysis (DFA, see [Casualty Actuarial Society, 1998] or [Blum and Dacorogna, 2003]) require models for all the assets and liabilities of a firm and thus a comprehensive, dynamic model for all economic variables that determine asset and liability values. Our Economic Scenario Generator (ESG) has been developed to fulfill this requirement. Partial models for a restricted set of economic variables cannot do this, no matter how sophisticated they are, because of the complex dependencies between the variables.

The goal is ambitious. Our method is bootstrapping, also called resampling. Initially, bootstrapping was a non-parametric method for limited tasks such as assessing confidence limits of models estimated on finite data samples [Efron and Tibshirani, 1993]. [Barone-Adesi et al., 1999] then applied bootstrapping to portfolio risk assessment, followed by [Zenti and Pallotta, 2000], [Barone-Adesi et al., 2002] and [Marsala et al., 2004]. The historical returns of certain assets became objects of resampling in simulations.

In this paper, bootstrapping constitutes the core of the model rather than being an additional tool. Our basis is a set of historical time series of economic key variables. The returns or *innovations* of all economic variables as observed in a randomly selected historical time interval are taken and used for the simulation of future time intervals. While there is an ongoing debate on modeling economic scenarios, the bootstrapping approach has several advantages. It can be implemented in a straightforward way and relies on past behaviors of real markets rather than debatable assumptions on models and parameters. Empirical distribution functions and simultaneous dependencies between economic variables are automatically captured. Bootstrapping belongs to the family of non-parametric methods. Like other "non-parametric" models, our method still needs some parameters in order to define the method in a useful way, which ultimately makes the model semi-parametric. Another advantage of bootstrapping is flexibility. We can easily add more economic variables, which typically leads to large, comprehensive models.

Bootstrapping also has some disadvantages. Random trends may be continued to the future with no limitation, serial correlations are disrupted by the random selection of past intervals, and the statistical variety of behaviors may be too small in the historical time series, which implies that the probability of extreme events may be underestimated. These problems are solved by adding some preprocessing algorithms to the bootstrapping method. The following aspects have to be considered: choice and transformation of variables, data frequency, dependence (serial and between variables), fat tails of distributions, the treatment of trends and mean reversion, and an arbitrage-free consistency of the resulting scenarios. Some of these refinements are not new. [Barone-Adesi et al., 1999] already found that some variables should preferably be resampled in a mapped rather than raw form, so they developed the method of filtered bootstrapping [Zenti and Pallotta, 2000, Barone-Adesi et al., 2002, Marsala et al., 2004]. This paper offers a wide set of bootstrapping refinements, based on economic principles and facts. These refinements eliminate the major pitfalls of bootstrapping and turn this technique into a reliable generator of realistic scenarios.

A major difficulty for any parametric or non-parametric simulation model is to determine reasonable expectations for economic variables such as inflation or the growth of GDP or equity indices. Empirical means based on available samples of, say, 10 years have stochastic errors or are biased by long-term economic "cycles". Long samples are not available in some cases (e.g. for the Euro, where synthetic data are needed). If they are available, they may be useless, such as foreign exchange data from before 1973, when currencies were under a different regime. Our ESG is based on historical data and naturally takes empirical means as expectations, but these can be modified on the basis of expert opinion or special long-term studies such as [Dimson et al., 2003].

The quality of economic scenarios and forecasts based on bootstrapping has to be measured. Typical time horizons of economic scenarios are measured in years and quarters, so we have a limited number of historical observations that can be used for backtesting. An out-of-sample backtesting study based on a Probability Integral Transform (PIT) confirms the validity of our approach.

The document is organized as follows. After a general introduction of the bootstrapping method, some generic steps of the method are presented in Section 2. The implementation of these general bootstrapping steps demands a lot of specific treatment of individual economic variables in Section 3, where the subsections 3.2 - 3.6 deal with the particularly complex case of interest rates. Some resulting scenarios and out-of-sample backtesting results are shown and discussed in Section 4. Section 5 concludes.

2 Bootstrapping - the method and its refinements

2.1 The idea of bootstrapping

Our concept of bootstrapping is presented in Figure 1 in a schematic, simplified form. Before introducing methodological details or economic variables, we discuss the bootstrapping method by means of a simple example.

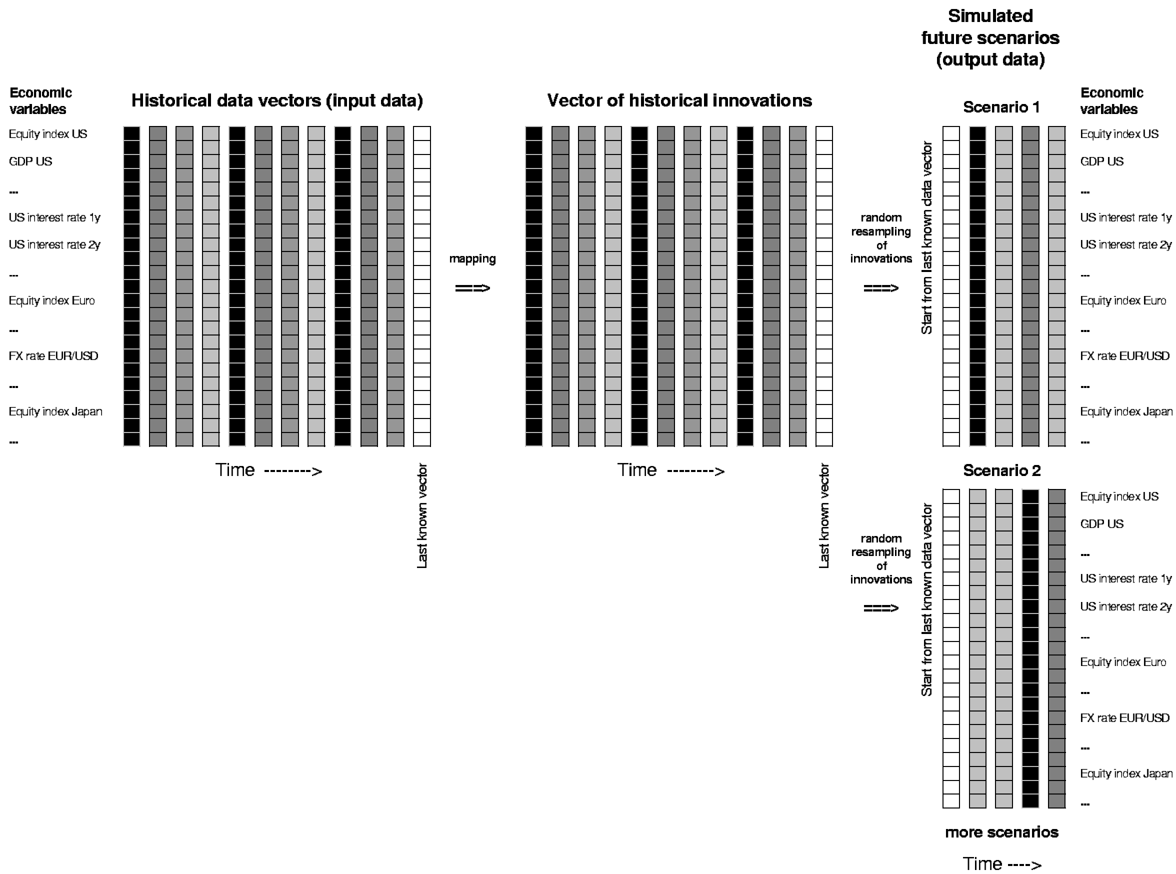
We start from a sample of historical data, that is a set of time series with historical observations over a certain time period. There is a regular[4] time sequence t_i with time steps of size Δt:

$$t_i \; = \; i \, \Delta t \tag{2.1}$$

The corresponding time series values are $X_i = X(t_i)$ (e.g. an equity index) and $Y_i = Y(t_i)$ (e.g. the GDP figures of the same country). The observations of all the series are synchronous and cover the same historical period (e.g. the last 10 years).

The last available values ("the values now") are X_n and Y_n. Our task is to simulate future values at times $t > t_n$: the vectors (X_{n+1}, Y_{n+1}), (X_{n+2}, Y_{n+2}), $\ldots$, where the future values are in the same regular sequence, i.e. $t_{n+k} = (n+k)\Delta t$. The basic idea of resampling is randomly picking an old time t_i of the sample and assuming the same set of observations for a future time of a scenario, e.g. for t_{n+1}.

[4] While the method relies on regular historical input data for bootstrapping, an algorithmic enhancement allows for starting a simulation from an irregular time point. We do not have to wait for the end of a quarter to produce up-to-date scenarios based on quarterly data.



This simplified diagram shows the bootstrapping method. We start with a historical series of data vectors containing different economic variables. Then we compute the innovations ($\approx$ returns) of the (mapped) economic variables and store them in a series of historical innovation vectors. The simulated scenarios start from the last available data vector and continue by adding innovations, which are are taken from randomly resampled innovation vectors.

Figure 1: The bootstrapping method

This is bootstrapping in its raw form, which will be modified in several respects. If we applied direct bootstrapping to the observations X_i and Y_i, the simulated values would never leave the range given by historical values. A GDP figure could never grow to a yet unobserved value. Therefore, our main concept is to bootstrap *innovations* in economic variables rather than the variables themselves. These innovations will be resampled and added to old variable values at each simulation step in a cumulative way.

A simple definition of innovations might be first differences of variables. When cumulating randomly resampled first differences, the simulated variable may become negative, which is not appropriate for positive definite economic variables. Returns are usually better than first differences. Logarithmic returns are an obvious choice. We can first transform the economic variable by taking the logarithm and then take first differences. In the general case, we first introduce a variable transformation,

$$x(t_i) \;=\; x_i \;=\; F(X_i, \mathcal{I}_i) \;=\; F[X(t_i), \mathcal{I}(t_i)] \tag{2.2}$$

where F can be a logarithm or a more complex function, which may depend not only on X_i but also some simultaneous values of other economic variables such as Y_i or, in general, the information set $\mathcal{I}_i$ available at time t_i, which includes earlier values of the considered variables. The function F should be invertible to determine X from x; its choice will be discussed for different variables. The innovation is defined in terms of x_i rather than X_i, for example as the first difference $x_i - x_{i-1}$. Most suitably, the innovation is defined as the deviation of x_i from its *expectation*[5] $E_{i-1}[x_i]$ that the market had at the previous time point t_{i-1}:

$$I_i = x_i - E_{i-1}[x_i] \tag{2.3}$$

The innovation I_i can be negative as well as positive. It constitutes the unanticipated element of surprise in a new value x_i and is thus unrelated to the market conditions at t_{i-1}. In case of the martingale hypothesis, if the expectation of x_i made at t_{i-1} was x_{i-1}, I_i would indeed be the first difference of x_i. In reality, the market often has a slightly different expectation $E_{i-1}[x_i]$ at t_{i-1}, so the innovation somewhat differs from the first difference. The market expectation $E_{i-1}[x_i]$ depends on the economic variable. For FX and interest rates, it is a forward rate. We agree with [James and Webber, 2000], Section 1.4.1, that forward rates are not particularly good predictors of spot rates, because the innovations I_i are large and unanticipated. Yet, an appropriate definition of $E_{i-1}[x_i]$ matters for long-term simulations, where seemingly weak modifications sum up to substantial effects. In Section 3, there are formulas for different economic variables, sometimes including some weak mean-reversion effects in $E_{i-1}[x_i]$.

The bootstrapping method will produce realistic results only if the I_i values are independent over time and identically distributed (i.i.d.) with zero mean. It should be impossible to reject the i.i.d. hypothesis, given the empirical sample of historical innovations. Then the expectation of I_i^2 is independent of current market conditions, in sufficient approximation. The mapping function F of Equation 2.2 has to be chosen accordingly. There is however the empirical phenomenon of volatility clustering which violates the independence of I_i^2: a large I_{i-1}^2 tends to be followed by a large I_i^2 with increased probability. In Section 2.9 this problem is solved.

In the course of simulation, the resampled innovations are used to modify the simulated, future x values. For a future time t_j, we randomly pick a historical index i and the innovation I_i of t_i to obtain the new simulated value[6]:

$$x_j = E_{j-1}[x_j] + I_i \tag{2.4}$$

This is an iteration. The next simulation time t_{j+1} will be treated the same way, picking a new historical index i' and re-using Equation 2.4 to obtain x_{j+1}. After a few iterative simulation steps, the resulting x value will contain an accumulation of many resampled innovations I_i. The variable x can drift to any value and will not observe any range constraints. Most original economic variables X_i, on the other hand, are positive definite. The logarithmic function transforms a positive definite variable to an unlimited real variable and is thus a standard choice for the mapping function F of Equation 2.2.

A main strength of the bootstrapping method is preservation of dependencies and correlations between variables. If the innovations $I_i[x]$ and $I_i[y]$ (the corresponding innovation of the variable Y_i) exhibit some dependence in the historical sample, the simulated variables x_j and y_j will be characterized by the same dependence structure. This is due to the fact

[5] $E_{i-1}[x_i]$ is used as a shortcut for the correct notation $E[x(t_i) \,|\, \mathcal{I}_{i-1}]$.

[6] The first simulation step starts at the last regular time $t_{j-1} = t_n$ and leads to x_j at time t_j. Sometimes, there is information available at an irregular time t_{irreg} after the last regular historical time t_n. In order to include this information in the first simulation step, the resampled innovation I_i can be modified to $I_{\mathrm{modified}} = I_{\mathrm{irreg}} + [(t_j - t_{\mathrm{irreg}})/\Delta t]^{1/2} I_i$, where I_{irreg} is the historical innovation from t_{j-1} to t_{irreg}.

that the resampled innovations $I_i[x]$ and $I_i[y]$ are always taken from the same historical time t_i within a simulation step.

The simulated mapped values x_j can be transformed back to standard values X_j by applying the function $F^{-1}(., \mathcal{I}_j)$ inverse to $F(., \mathcal{I}_j)$, see Equation 2.2.

2.2 General overview of the bootstrapping method

An economically meaningful bootstrapping procedure requires a set of well-thought steps in addition to the simple bootstrapping principle as outlined in the previous section. The general sequence of analysis steps is as follows:

- Start from a complete and representative *sample* of historical economic time series for several economic variables, regularly updated to the newest values.
- *Transform* the economic variables (see Equation 2.2, sometimes with deseasonalization, see Section 2.5), in order to attain unlimited additivity of innovations.
- Compute the market's *expectations* of variables at each time t_{i-1} for time t_i (e.g. a forward rate as market predictor for a foreign exchange spot rate), including some weak, long-term mean-reversion effects.
- Compute the *innovations* of variables as the differences between the current variable values and their previous market expectations, see Equation 2.3.
- *Remove stochastic trends* by forcing a zero mean of innovations, to avoid arbitrary trends in later simulations, see Section 2.6.
- Treat autoregressive conditional heteroskedasticity (clusters of volatility) of innovations by fitting a GARCH process, leading to *GARCH-corrected innovations*, see Section 2.9.

After this preparation, we are able to simulate future scenarios. We start by initializing all the variables (including auxiliary ones) to the latest historical values (the "values now"). The following sequence of steps describes one time step into the simulated future, which can be iteratively repeated.

- Do the central *bootstrapping* step, taking a vector of past GARCH-corrected innovations, all from the same randomly picked historical time interval.
- Multiply all these innovations by a random *tail correction factor*, thus injecting some rare shocks or stress scenarios that are not present in the initial data sample.
- *Re-transform* the GARCH-corrected innovations to the actual innovations to be used, and update the GARCH volatility equation.
- Compute the *simulated variable values* as sums of previous market expectations and innovations, see Equation 2.4.
- Compute the *market expectations* of variables for the next simulation step.
- Compute the simulated values of *economic variables in their original definitions* by doing transformations inverse to Equation 2.2 (reestablishing seasonality, if needed).

Notice that this sequence of steps mirrors the initially taken analysis steps in reverse order.

This elaborated methodology applies to all economic variables, but the details of each step may look different for them. More details and problems of the method are described below. The special treatment of different economic variables follows in Section 3.

2.3 Time steps: Using high-frequency observations?

The size of the bootstrapping time steps depends on the application. Commercial simulation and planning tools may have yearly time steps, but the generation of the underlying economic scenarios should be done in shorter steps. We can take quarterly steps and only use every fourth set of variable values, resulting in a yearly simulation.

When using a past of 10 years, we have a basis of only 40 quarterly time intervals. This is better than 10 (the number for using yearly steps), but still rather low. A resampled year will consist of four quarters, each having the randomly selected innovations of a historical quarter. Thus there will be a wide variety in the behavior of simulated years: 40^4 (more than 2 million) possible sequences of quarterly innovations.

Of course, we can further increase the number of historical intervals by taking monthly, weekly or daily time steps. For some variables such as GDP, high-frequency observations are not available. The clustering of volatility has been found to be stronger for high-frequency data in the literature, so the GARCH analysis (see Section 2.9) becomes more important. In our actually implemented economic scenario generator, we are always using quarterly time steps.

2.4 Noise

For some economic variables, the available data exhibit some noise. Here we mean mean-reverting short-term noise rather than the natural volatility of economic variables. Noise affects the innovation values computed by Equation 2.3 and leads to an increased variance of innovations. This increase is spurious because it reflects mean-reverting movements rather than true drifts, so it may lead to a too high volatility of results simulated over several time steps.

When using reliable data sources, this phenomenon is restricted to those variables whose definition is sensitive to such noise. In practice, this means the innovations of inflation and quarterly forward interest rates. The noise in consumer price index (CPI) figures is reinforced when computing inflation (a kind of first difference of the logarithmic CPI) and a second time when computing inflation innovations (which are similar to second differences of the logarithmic CPI). Quarterly forward interest rates have some almost inevitable noise due to small interpolation errors in the rather coarse grid of maturities supported by the yield curve data.

In these cases of noise, some smoothing techniques such as averaging are recommended in order to avoid spurious volatility in simulation results.

2.5 Seasonality

A variable recorded in a time series is called seasonal if its values or its first differences (or returns) have a seasonal pattern. This means that averages sampled at certain regular time intervals (e.g. second quarters of each year) significantly deviate from averages sampled at shifted intervals (e.g. third quarters).

Prices in liquid markets such as FX, fixed income or equity hardly exhibit any significant seasonality, as empirical studies have shown. (Otherwise, these could be exploited in a systematic way.) Other financial variables such as inflation and GDP may be seasonal as there is no investment strategy to exploit seasonality. Quarterly inflation rates (first differences of logarithms of the CPI) indeed exhibit some seasonality aside from the noise discussed in Section 2.4.

In order to use a seasonal variable for bootstrapping, we have to deseasonalize its historical observations before computing the innovations. The simulation results will be reseasonalized at the end. This is further discussed in Section 3.7.

2.6 Detrending

In our simulations, we use innovations according to Equation 2.4 in an iterative way, thereby cumulating the innovations. Innovations are defined as deviations from prior market forecasts. If the market forecasts are reasonable[7], we expect a mixture of positive and negative innovations in the long run, but the empirical mean of innovations within a historical sample may slightly (stochastically) deviate from zero. In that case, we risk introducing a trend into the simulated future.

Generating such a trend is not justified even if it existed as a random phenomenon in the historical data. Therefore we force the innovations to have a zero mean:

$$I_i \;=\; \sqrt{\frac{n}{n-1}}\;\left(I_{\mathrm{raw},i} - \frac{1}{n}\;\sum_{j=1}^{n} I_{\mathrm{raw},j}\right) \tag{2.5}$$

Each raw innovation $I_{\mathrm{raw},i}$ is corrected by subtracting the sample mean. When doing so, we implicitly minimize the variance of I_i about zero by using one degree of freedom. Therefore we need the correction factor $\sqrt{n/(n-1)}$ to restore the expected variance of innovations. Equation 2.5 is used for the correction of all innovations of the algorithm.

2.7 Mean reversion effects

When cumulating our detrended innovations, we obtain a stochastic random walk of the resulting variable, similar to a Brownian motion. Such motions do not exhibit any mean reversion. For most variables such as equity indices, this behavior conforms to theory and empirical findings. For other variables such as interest rates, however, there is a weak mean-reverting force which makes sure that interest rate levels do not drift to arbitrarily high (or low) values, even after decades and centuries. Another law with mean-reverting character is purchasing-power parity (PPP). FX and inflation rates observe this law only hesitantly, with time lags of several years (see the article by Cheung in [Chan et al., 2000]).

In our bootstrapping algorithm, a natural place to implement the small mean-reverting correction is the market forecast $E_{i-1}[x_i]$ of Equation 2.3. Mean reversion is a known phenomenon rather than an innovative surprise, so it belongs to the market forecast in the form of a small correction of the purely technical market forecast. Although such corrections are small, they may persist over years and exert a decisive force in real markets as well as in our simulations.

[7] This does not mean free of trends. Some economic variables such as equity indices or some FX rates ("Peso effect") have a natural trend that we have to model in the market forecast.

The mathematical nature of the small correction differs between economic variables and will be discussed in Section 3. In many cases of mean reversion, we use varying target values rather than constant means. Some of these are moving averages. This can be implemented in the form of an exponentially weighted moving average (EMA), which has the advantage of a very simple iteration formula for updating:

$$EMA_i[x; \Delta t_{\text{range}}] \;=\; \mu \; EMA_{i-1}[x; \Delta t_{\text{range}}] + (1-\mu)\; x_i \tag{2.6}$$

with

$$\mu \;=\; e^{-\Delta t / T} \tag{2.7}$$

where x_i stands for any variable to be averaged over time, and the time constant Δt_{range} is the range (= center of gravity of the weighting kernel) of the EMA. There are more complex moving averages and iterations, see [Dacorogna et al., 2001], but the simple mechanism of Equations 2.6 and 2.7 is certainly good enough to describe the behavior of means that are only used to make weak corrections. At the beginning, each EMA has to be initialized, using a sufficiently large sample of x_i . We use the best estimate for the EMA at the very beginning of the historical sample and iteratively use Equation 2.6 through the whole historical sample.

Mean-reversion effects often involve several variables with different volatility levels. In this case, we often prefer applying the mean-reversion correction to the high-volatility variable, where the low-volatility variable acts as a sort of dragging anchor. In the example of purchasing power parity (PPP), the high-volatility FX rate is weakly anchored by the low-volatility consumer price indices (CPI) of the two currency zones.

2.8 Dependence

Simultaneous innovations in different time series often depend on each other. Equity indices in different countries, for example, rarely move in different directions. The bootstrapping method captures these dependencies very well, as all innovations of a simulation step are resampled from the same historical time interval. Contemporaneous dependencies found in historical data are thus preserved.

Other forms of economic dependency pose some problems. Dependencies do not only exist for innovations, but also for the original and simulated variables. This can often be described as a mean-reversion effect such as purchasing-power parity (PPP) and has already been discussed in Section 2.7.

Serial dependence of innovations would mean that new innovations are partially anticipated by older ones. This is not the case here, since we define innovations as unpredictable in Equation 2.3.

If the serial dependence is in the volatility of an economic variable rather than the variable itself, we talk about autoregressive conditional heteroskedasticity. This is treated in Section 2.9.

2.9 Heteroskedasticity modeled by GARCH

Heteroskedasticity means a variation in the volatility of a variable over time. This is only a useful concept if a model for this volatility can be formulated. One way to model future volatility would be using implied volatility from option markets. For market variables such as FX, IR and equity indices, this is feasible as long as such volatility data are available (which is not the case for the long time horizons of certain ALM studies). We do not pursue this idea here.

Autoregressive conditional heteroskedasticity (clustering of volatility) is a well-known phenomenon [Engle, 1982, Bollerslev, 1986, Baillie et al., 1996, Dacorogna et al., 2001] in finance. For economic high-frequency data such as daily market prices, this is very significant. However, the effect is weaker for the data frequency of economic scenarios (such as quarterly or yearly). We model the effect by approximately assuming a GARCH(1,1) process [Bollerslev, 1986]. We fit a GARCH(1,1) process to the innovations as computed from the raw data.

The GARCH(1,1) process for the observed innovations I_i is

$$\begin{aligned} I_i &= \sigma_i \, \varepsilon_i \, , \\ \sigma_i^2 &= \alpha_0 + \alpha_1 \, I_{i-1}^2 + \beta_1 \, \sigma_{i-1}^2 \end{aligned} \tag{2.8}$$

with three positive parameters α_0, α_1 and β_1. The variable ε_i is identically and independently distributed (i.i.d.), with mean 0 and variance 1. The GARCH process is stationary with finite variance if $\alpha_1 + \beta_1 < 1$.

Calibrating the parameters of a GARCH process to the innovations I_i is not a routine task, although its feasibility with the help of commercial software may lead a user to that assumption. The usual quasi-maximum-likelihood method poses some problems in practice, such as convergence to non-stationary solutions (especially if the GARCH process is misspecified), local maximums of the likelihood function and other convergence problems. In some cases, [Zumbach, 2000] finds a maximum of the likelihood function for a GARCH process whose unconditional variance is about ten times the empirical variance of the data sample. The reasons are misspecification and limited sample size. [Zumbach, 2000] even finds such effects for data generated by a GARCH(1,1) process. Finite sample sizes pose problems for GARCH fitting that tend to be underestimated in the literature.

Our historical innovations based on low-frequency data definitely constitute small samples. Yet we need a reliable, robust GARCH calibration algorithm for the repeated analysis of dozens of economic variables without any human intervention. Standard GARCH fitting methods or software packages requiring human review and intervention are not sufficient. Our robust, automated GARCH calibration is described in Appendix 6 and follows [Zumbach, 2000] with additional emphasis on avoiding local optima and a careful build-up procedure. ARCH(1) and white noise are embedded as special cases of GARCH(1,1). The "white noise" solution means that there is no GARCH correction and the original innovations I_i are kept.

We apply GARCH corrections to all innovations, with one exception. The forward interest rate innovations of all maturities have a common GARCH model which we calibrate for a weighted sum of these innovations. This sum has positive weights and approximately stands for a first principal component. [Ballocchi et al., 1999] have shown that the first principal component of the term structure of forward interest rates is the only component with significant autoregressive conditional heteroskedasticity. This finding supports our GARCH modeling approach for interest rates.

After calibrating the GARCH process, we assume all the innovations I_i to be products of the volatility σ_i as resulting from the GARCH process and "normalized" innovations J_i which can be seen as the historically determined GARCH residuals, that is the ε_i values of Equation 2.8. We obtain J_i by dividing the original innovations I_i by σ_i. The normalized innovations J_i are the final results of our preprocessing of economic variables. We can also call them “GARCH-filtered” innovations, using an analogy to the “filtered bootstrap” by [Barone-Adesi et al., 1999].

In the simulation of the future time t_j, we resample a J_i and compute the volatility σ_j by using Equation 2.8, initially starting at the last computed historical GARCH variance σ_n^2. The newly constructed innovations $I_j = \sigma_j J_i$ will be used. The sequence of these I_j will have the desired property of volatility clustering, unlike the randomly resampled J_i values.

2.10 Fat tails of distribution functions

All the economic variables and their innovations have their empirical distributions as determined by the historical data. When using quarterly observations over ten years, we have 40 innovations. This is a small sample size for detailed statistics.

From the literature [Dacorogna et al., 2001, Embrechts et al., 1997], we know that many financial variables exhibit fat tails in their distribution functions, if studied with enough data, using high frequency or very long samples. Typical tail indices of high-frequency foreign-exchange data are around $\alpha = 3.5$.

Our economic scenarios are made for studying risk as well as average behaviors. We use tail-based risk measures such as value at risk (VaR) and, more importantly, the expected shortfall, see [Artzner et al., 1997]. The simulation of extreme events (such as the “1 in 100” event) should be realistic. How is this possible based on only 40 quarterly innovations for bootstrapping? Pure bootstrapping will underestimate risks, except for the unlikely case that the most extreme historical observation substantially exceeds the quantile that can reasonably be expected for the maximum in a small sample.

Some risk and ALM specialists rely on a few arbitrary ”stress scenarios”, that is some stylized extreme events. Here we propose a more consequent way to include a rich variety of many possible stress scenarios. When doing the simulations, we add some stochastic variation to the resampled innovations to attain a more elaborated tail behavior. We do not really change the tail behavior, we just add some small stochastic variability on both sides, increasing or decreasing an original innovation. Technically, this can be done without increasing the overall variance. The stochastic variation of historically observed innovations is small and harmless, except for very rare, extreme tail events. We explain the method for an economic innovation I_i. If the GARCH analysis of Section 2.9 is made, we apply the tail correction to the GARCH-corrected innovations, so I_i actually stands for the normalized innovation J_i.

I_i has an unknown distribution function with mean 0. We assume a tail index $\alpha > 2$ in both tails. A suitable value for many economic variables might be $\alpha = 4$. Now we define an auxiliary, Pareto-distributed random variable η to modify the original, resampled innovations I_i in a multiplicative way:

$$I'_i \;=\; \eta\, I_i \tag{2.9}$$

The new variable η is defined to have the same tail index α:

$$\eta \;=\; A + B\,(1-u)^{-1/\alpha} \tag{2.10}$$

where u is a uniformly distributed random variable in the range between 0 and 1. Thus η is confined:

$$\eta \;\geq\; \eta_{\min} \;=\; A + B \tag{2.11}$$

This minimum corresponds to $u = 0$. We always choose $A+B > 0$, so η is positive definite. The inverse form is

$$u \;=\; 1 - \left(\frac{\eta - A}{B}\right)^{-\alpha} \tag{2.12}$$

This is the cumulative probability distribution of η, where u is the probability of η being below the specific η value inserted in Equation 2.12. This is indeed a Pareto distribution with tail index α.

We choose the parameters A and B in a way that η is normally close to 1, so the modified variable $I'_i = \eta I_i$ is similar to the original, resampled value I_i, and the overall character of the bootstrapping method is maintained. However, the modified innovation I'_i based on the random variable η and the independently chosen resampling index i will exhibit a fat tail in simulations. The larger the number of simulations, the denser the coverage of this fat tail will be. Tail observations of I'_i will occur if two unlikely events coincide: very large values of both $|I_i|$ and η.

The resulting tail index[8] of I' is α, as assumed for I. Thus we do not make the tail fatter than it should be, we just introduce enough variation in the tail for realistic simulations.

The parameters A and B must be defined in a suitable way. We have to keep the original variance of innovations unchanged. This is important when using the GARCH correction of Section 2.9. GARCH is a variance model, so we should not modify the unconditional variance in our simulations here. The condition is

$$\mathrm{E}[I'^2_i] \;=\; \mathrm{E}[I^2_i] \tag{2.13}$$

Considering Equation 2.9 and the independence of η, this implies the condition

$$\mathrm{E}[\eta^2] \;=\; A^2 + \frac{2\,\alpha}{\alpha - 1}\,A\,B + \frac{\alpha}{\alpha - 2}\,B^2 \;=\; 1 \tag{2.14}$$

which is the result of an integration over the distribution of η, using Equation 2.10. In order to keep the variance $\mathrm{E}[\eta^2]$ finite, we need $\alpha > 2$, which turns out to be well satisfied by empirical economic data. The second equation to determine A and B is given by Equation 2.11: $A + B = \eta_{\min}$. Solving this equation together with Equation 2.14, we obtain

$$B \;=\; \frac{1}{2}\left[\sqrt{\eta^2_{\min}\,(\alpha-2)^2 + 2\,(\alpha-1)\,(\alpha-2)\,(1-\eta^2_{\min})} - \eta_{\min}\,(\alpha-2)\right] \tag{2.15}$$

and

$$A \;=\; \eta_{\min} - B \tag{2.16}$$

We still need to choose the minimum $\eta_{\min}$ of the correction factor η. We argue that the tail correction should neither be too timid nor too strong (which would mean to destroy the character of the bootstrapping method). We allow it to be just strong enough to fill the gap between the largest and the second largest historical innovation. In reality, the empirical values of these innovations are subject to wide stochastic variations. Just for the sake of a reasonable definition of $\eta_{\min}$, we assume them to be regular quantiles here. We locate the largest observation of I_i, called $I_{\max}$, at a cumulative probability between

[8] A closer tail analysis shows that η should be based on a tail index infinitesimally larger than α, otherwise the resulting tail index of x' is infinitesimally less than α. This theoretical consideration does not matter in practice.

$1-1/n$ and 1, in fact in the middle of this range, at $1-1/(2n)$. Assuming a Pareto behavior at the tail around I_{max} with tail index α, we obtain the heuristic approximation

$$I_{\text{max}} \;\approx\; (2\; c\; n)^{\frac{1}{\alpha}} \tag{2.17}$$

where the constant c stays undetermined. Following the same logic, the second largest value of I_i can be associated to the cumulative probability range between $1-2/n$ and $1-1/n$. The probability value $1-1/n$ separates the expected domain of I_{max} from the domain of the second largest value. The I value corresponding to this separating limit is

$$I_{\text{limit}} \;\approx\; (c\; n)^{\frac{1}{\alpha}} \tag{2.18}$$

By applying the tail correction of Equation 2.9, the largest observation can be reduced to $\eta_{\text{min}} I_{\text{max}}$, but not more. We identify this reduced value with the limit I_{limit}:

$$\eta_{\text{min}}\; I_{\text{max}} \;\approx\; (c\; n)^{\frac{1}{\alpha}} \tag{2.19}$$

Equations 2.17 and 2.19 can be solved for η_{min}. The unknown constant c cancels out. We obtain the following recommended choice:

$$\eta_{\text{min}} \;=\; 2^{-\frac{1}{\alpha}} \tag{2.20}$$

This result is independent of n and always < 1. For an α of 4, we obtain $\eta_{\text{min}} \approx 0.841$, which is rather close to 1. Our definition of η is complete now and consists of Equations 2.10, 2.15, 2.16 and 2.20.

Eventually, the tail correction will be made for all resampled innovations, not only for one variable I_i. When doing it for all innovations in a multi-dimensional setting, two issues have to be addressed:

- Do we use the same tail index α for all economic variables? This is not necessary. Detailed statistical studies of all variables may lead to specific α values. In a simpler approach, we can use a general assumption such as taking $\alpha = 4$ for all economic variables.
- Do we use the same random variable u for all economic variables? In the case that we also take the same α (which is not necessary, see above), this implies using the same η for all variables. Using different u values for different variables adds some noise and blurs the dependence in the tails. Using the same u or η leads to an emphasis on the dependence in the extreme tails of all those variables that simultaneously have extreme observations. Some findings [Dacorogna et al., 2001] indeed indicate that dependencies between variables are larger in the tails than under less extreme circumstances. In a parametric model, this effect could be modeled through copulas. In our bootstrapping approach, we obtain a conservative, risk-conscious effect by assuming the same u for all variables. At the same time, this reduces the number of computations per simulation step.

Using the proposed method, we can successfully reconcile the bootstrapping method with the requirement of realistic tail simulations. There is some room for human judgement. If conservative users have reasons to believe that future behaviors will be more extreme than historical behaviors, they can decrease the assumed tail index α.

3 Bootstrapping of different economic variables

3.1 Choice of economic variables

The set of economic variables to be modeled depends on the availability of raw data and the needs of the model user. There are interactions between economic variables (e.g. weak mean reversion effects) that can only be modeled if a sufficiently large set of variables is chosen.

The following economic variables are included in a reasonable implementation of an economic scenario generator based on bootstrapping:

- Interest rates (IRs). These have different maturities. We have to deal with whole yield curves. The interest rate model is the heart of any comprehensive economic model.
- Foreign Exchange (FX) rates between the supported currencies of the generator.
- Equity indices. It is possible to include several indices per currency zone, e.g. different sector indices, real-estate fund indices or hedge fund indices. We prefer total-return indices which include reinvested dividends, because these indices are directly related to investment performance. However, the bootstrapping technique also works for price indices.
- Inflation, in the form of a Consumer Price Index (CPI). It is possible to add other indices, e.g. wage inflation or medical inflation.
- Gross Domestic Product (GDP).

The variables have different levels of volatility. We can roughly sort them, from low to high volatility: real gross domestic product (GDP), consumer price index (CPI), interest rates, inflation (which is a temporal derivative of the CPI), FX rates, equity indices.

All the variables are modeled for several major currency zones. Major currencies should be included as well as those minor currencies that are relevant for an application. We are using the currencies USD, EUR, JPY, GBP, CHF and AUD.

The lists of variables and currencies can be varied. One of the advantages of the bootstrapping method is that adding or removing an economic variable from the model is technically easy. As an example, we may include rating-dependent credit spreads as a new variable to simulate the behavior of corporate bonds.

Other economic variables such as the values of certain bonds, including mortgage-backed securities with their special behavior, can be derived from the simulated values of primary variables such as interest rates in sufficiently good approximation.

In the following sections, the treatment of different variables is discussed in detail. For each of them, the steps of the bootstrapping method as outlined in Sections 2.1 and 2.2 take different forms.

3.2 Interest rate forwards and futures

When modeling interest rates, we refer to “risk-free” market interest rates as extracted from different, liquid financial instruments, which are issued by governments or institutions of the highest ratings. Such interest rates, for different maturities, can be summarized in the form of a zero-coupon yield curve, or just yield curve, such as the "fair market" yield curves composed by Bloomberg.

An interest rate (IR) as quoted in a yield curve has a complex dynamic behavior. Interest rates for different maturities are available at the same time, with a complicated dependence structure. Long-term interest rates have maturity periods of many years, over which the economic conditions can be expected to change. The dynamic behavior of an IR with constant maturity period is characterized by the fact that this period is continuously moving over time. The IR thus refers to a moving target.

A way to disentangle the complex dynamics and dependencies of interest rates – both in market practice and in modeling – is using forward interest rate or IR futures. Using IR futures is the most consequent solution, as these future contracts always focus on the same future time interval, for example from 15 March 2007 to 15 June 2007. For such a fixed, well-defined period, the price-finding process in the market is more efficient than for large, heterogeneous, moving time intervals. This fact helped to make IR futures the most liquid financial instrument in the IR market for maturities from 3 months to about 2 years. We shall see that IR futures have similar advantages[9] in modeling, too. A major advantage is arbitrage-free consistency. If all IR-based financial instruments are constructed from the same forward IRs and thus the same market prices of IR futures, there is no way to generate riskless profits, no matter how sophisticated the IR portfolio composition.

There is a rich literature on interest rate modeling; we use [James and Webber, 2000] as a main reference. The basics of yield curve mathematics can be found in Section 3.1 of that book. We transform the information contained in a yield curve and package it as an equivalent set of forward interest rates. The yield curve consists of annualized interest rates $r(T)$ as a function of the time interval to maturity, T. We use interest rates R in logarithmic form,

$$R(T) \;=\; \log\left(1 + \frac{r(T)}{100\%}\right) \tag{3.1}$$

This has the advantage of transforming the multiplicative compounding of interest rates to simple additive compounding. Now we regard the forward interest rate $\varrho(T_1, T_2)$ for the interval between the future time points $T_2 > T_1$. From elementary interest compounding rules, we derive

$$T_2 \; R(T_2) \;=\; T_1 \; R(T_1) + (T_2 - T_1) \; \varrho(T_1, T_2) \tag{3.2}$$

which is additive due to the logarithmic transformation of Equation 3.1. We solve for $\varrho(T_1, T_2)$:

$$\varrho(T_1, T_2) \;=\; \frac{T_2 \; R(T_2) - T_1 \; R(T_1)}{T_2 - T1} \tag{3.3}$$

When starting from a yield curve, this equation serves as a definition and computation formula for the empirical forward rates $\varrho(T_1, T_2)$, where T_1 and T_2 are neighboring maturities of the yield curve. In practice, $R(T_1)$ and $R(T_2)$ are often interpolated values from a more coarsely defined yield curve. We need a good yield curve interpolation formula, but even an excellent formula may lead to small interpolation errors which are reinforced by building the difference of Equation 3.3. This problem requires an additional smoothing procedure later in the bootstrapping algorithm.

[9] There is also a disadvantage when using futures. IR futures markets require a collateral margin account which leads to a small deviation between the values of forward rate agreements and futures, called the convexity adjustment (see Section 5.5 of [James and Webber, 2000]). We assume that our basic curves are "fair market" yield curves where the convexity adjustment is accounted for when they are constructed from futures prices.

The forward rate of an infinitesimally small maturity interval, from T to $T+dT$, is denoted by $\varrho(T)$. The logarithmic interest rate $R(T)$ can be written as

$$R(T) \;=\; \frac{1}{T}\int_0^T \varrho(T')\; dT' \tag{3.4}$$

In fact, $R(T)$ is the average forward IR as measured over the whole maturity axis from 0 to T.

At the end of a simulation step, the resulting set of forward interest rates can be retransformed to a yield curve, following the notion of Equation 3.4.

3.3 The innovations of forward interest rates

Setting up a satisfactory bootstrapping algorithm for forward interest rates is a complex task. For the sake of completeness, we formulate in this section an intuitive direct approach to resampling forward rates. However, this approach leads to problems, so we shall need a more sophisticated method as described in Sections 3.4 and 3.5. At the end of Section 3.5, the steps for bootstrapping interest rates are summarized.

First we add the dimension of time t, using the regular time points of Equation 2.1. We write $\varrho_i(T)$ for the forward rate at time t_i, named $\varrho(T, T+\Delta t)$ in Section 3.2. For bootstrapping, we are only interested in rates with a forward period of the size of the basic time step Δt (= 3 months for quarterly steps) and a time to maturity T that is an integer multiple of Δt. For the corresponding spot rate with maturity Δt, we write R_i ($= \varrho_i(0)$). How do forward rates $\varrho_i(T)$ evolve over time? At first glance, we might consider the behavior of the forward rate $\varrho_i(T)$ for a fixed maturity period T. However, the time $t_i + T$ of the maturity would move in parallel with time t_i. The value of $\varrho_i(T)$ would therefore refer to changing time points with changing market conditions, which makes the assessment difficult.

Instead, we focus on the forward IR for a fixed time interval in the future. This is exactly the point of view of IR futures markets. The price of an IR future reflects the current market consensus forecast $\varrho_i(T)$ of the underlying interest rate. When the futures contract reaches maturity, at time $t_i + T$, we can directly read the value $R_{i+T/\Delta t}$ of this interest rate from the yield curve. In other words, $\varrho_i(T)$ is the market's forecast[10] of $R_{i+T/\Delta t}$. There is a stream of unanticipated news that leads to innovations in this forecast. At the earlier time t_{i-1}, the market forecast for $R_{i+T/\Delta t}$ was $\varrho_{i-1}(T+\Delta t)$; at t_i it is $\varrho_i(T)$. We observe the following innovation from t_{i-1} to t_i:

$$I_i[\varrho(T)] \;=\; \varrho_i(T) - \varrho_{i-1}(T+\Delta t) \tag{3.5}$$

This is Equation 2.3 applied to forward interest rates. The innovation $I_i[\varrho(T)]$ can be resampled and cumulated in our bootstrapping method. However, such a direct procedure may lead to negative interest rates in the simulation and some other shortcomings as shown below. We need a deeper analysis of $\varrho_i(T)$ and a more sophisticated method.

3.4 Mapping and the asymmetry of interest rates

Three problems arise when directly using $I_i[\varrho(T)]$ from Equation 3.5 for resampling:

[10] This statement will be qualified twice: first in Section 3.6, due to the asymmetry in interest rates, then in Section 3.6, where a small mean-reverting correction term is added.

1. Innovations can be negative as well as positive. When cumulating $I_i[\varrho(T)]$ values from randomly resampled historical time points t, the resulting ϱ values in some scenarios may drift to a value less than zero after some simulation steps. Such a behavior cannot be accepted as it violates an economic principle which states that no increase of profit can be reached at zero risk. As soon as an IR (or forward IR) is negative, a risk-free profit can be made by storing money physically instead of investing it in a deposit. In historical data, we hardly find any negative interest rates.

2. Interest rates are more volatile on a high level than on a low level close to zero. The same innovation value $I_i[\varrho(T)]$ may produce high volatility in the context of low ϱ values and low volatility when resampled in a high-interest regime. This is against the bootstrapping principle. A resampled innovation should always model approximately the same force on the market, regardless of the current economic condition.

3. The empirical forward rate $\varrho_i(T)$ as determined by the market is a forecast with uncertainty rather than a simple quantity. Market participants know that the distribution is skewed: negative values of $R_{i+T/\Delta t}$ are unlikely while the positive part of the distribution is unlimited. Under normal conditions, they will thus agree on a forward rate $\varrho_i(T)$ exceeding the expected median of $R_{i+T/\Delta t}$ by an amount that is related to the "term premium".

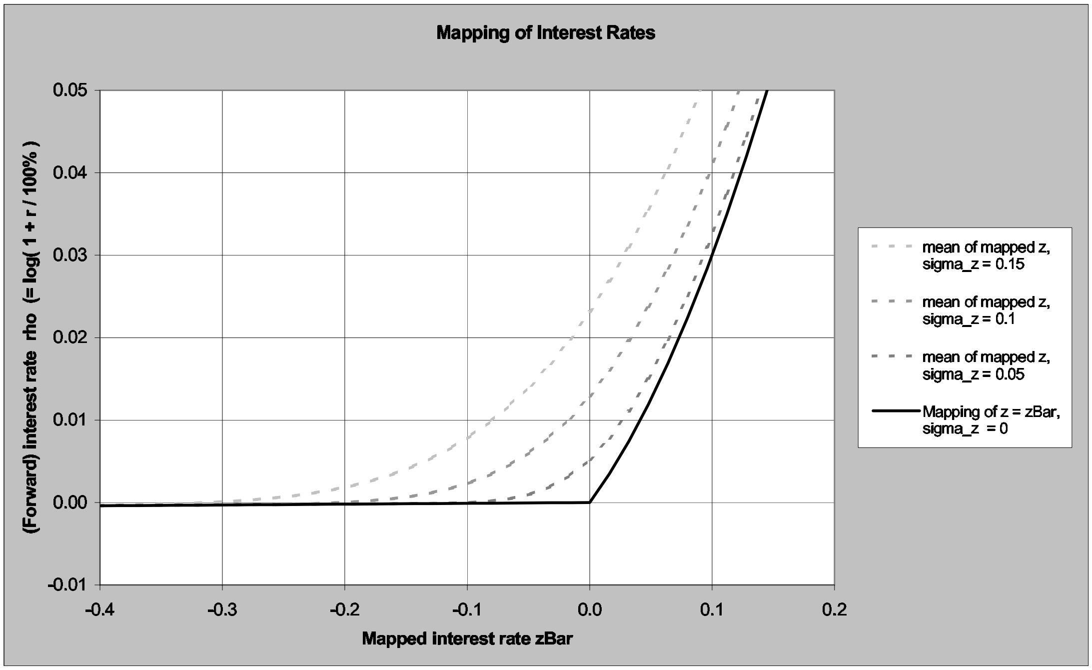
All these problems are related to the asymmetry or skewness of interest rate distributions. There is a mathematical method that solves all of them at the same time: non-linear mapping of short-term interest rates R_i, for which we simply write R here. We define a mapped variable z:

$$z \;=\; z(R) \;=\; \begin{cases} \sqrt{R+\varepsilon} - \sqrt{\varepsilon} & \text{for } R \;\geq\; 0 \\ A\; R & \text{for } R \;<\; 0 \end{cases} \tag{3.6}$$

with a small offset $\varepsilon \approx 0.01$ and a large factor $A \approx 1000$. The idea behind the mapping of Equation 3.6 is to eliminate the asymmetry of interest rates. At time t_i, the distribution expected for the rate $R_{i+T/\Delta t}$ at maturity time $t_i + T$ is asymmetric with a variance depending on the value of $\varrho_i(T)$. In contrast, we define z in a way to fulfill two working hypotheses: (1) the distribution of the z value expected for time $t_i + T$ is symmetric around a mean $\bar{z}$; (2) the variance σ_z^2 of z is independent of the $\bar{z}$ level. If the parameters of Equation 3.6 are chosen appropriately, both hypotheses should hold in sufficiently good approximation. The working hypotheses are no firm claims, they are just used to motivate and introduce our method of bootstrapping interest rates.

Historical interest rates are rarely negative. In simulations, the large parameter A will cause a sort of soft boundary for interest rates below zero. This boundary is not as absolute as in Section 6.4.3 of [James and Webber, 2000]. The function $z(R)$ is continuous and has a pronounced kink at $R = z = 0$, which is natural for a quantity for which the limit $R = 0$ plays an important role.

In the upper part ($R \geq 0$), z approximately grows with the square root of R. This is in agreement with the Cox-Ingersoll-Ross (CIR) model of interest rates (which is very different in other aspects, see Section 3.3.2 of [James and Webber, 2000]). The CIR model assumes the volatility of interest rates to be proportional to the square root of the current IR value. Our mapping implies a similar behavior by assuming a fixed distribution of z and translating the behavior of z back to the behavior of interest rates R. The square-root



The solid curve shows the mapping of interest rates in the inverse form of Equation 3.7, $R = R(z)$. In the region of $\bar{z} < 0$, the curve is not horizontal, but has a tiny positive slope of size $1/A$. The dotted curves show forward interest rates ϱ as functions of the bootstrapping variable $\bar{z}$, following Equation 3.10 and assuming different variances σ_z^2 of z about $\bar{z}$. The values σ_z = 0.05, 0.1 and 0.15 approximately represent three maturity periods: 1/2 year, 2 years and the long-term limit. In the case of the solid line, the maturity and the variance σ_z^2 are zero, $z = \bar{z}$, and ϱ stands for the spot interest rate R.

Figure 2: Interest rate mapping

law is modified by adding a constant ε to R in Equation 3.6. This makes the volatility at very low interest rates less aberrant and more similar to that of higher IR levels, a behavior we have observed for Japanese interest rates. The very low Japanese rates since the late 1990s have given us some useful hints on how to model low levels realistically. Our model based on Equation 3.6 is robust for a wide range of different IR levels, using the term "robustness" as in Section 1.5.2 of [James and Webber, 2000] and relating to the discussion of Section 6.4 of the same book.

The function $z(R)$ is strictly monotonic and can thus be inverted:

$$R \;=\; R(z) \;=\; \begin{cases} (z + \sqrt{\varepsilon})^2 - \varepsilon & \text{for } z \;\geq\; 0 \\ \frac{z}{A} & \text{for } z \;<\; 0 \end{cases} \tag{3.7}$$

A is a very large parameter, so R will be very close to zero even if z is distinctly negative, as shown in Figure 2. This is a first reason why the simulation will never produce strongly negative interest rates. If it ever produces negative IRs, these are so close to zero that they can be rounded to zero in most practical applications.

Equation 3.6 relates the new variable z to the short-term interest rate R. In order to use z in practice, we need to define its relation to observable forward rates ϱ. This follows from the distribution function of z which we approximately assume to be normal[11] with mean $\bar{z}$ and variance σ_z^2:

$$z_{i+T/\Delta t} \;\sim\; \mathcal{N}[\bar{z}_i(T), \sigma_z^2(T)] \tag{3.8}$$

This ensures mathematical tractability. Now we express $\varrho_i(T)$ as the expectation value of $R_{i+T/\Delta t}$. Taking the expectation value is justified if the values of simple IR-based portfolios at time $t_i + T$ are linear functions of $R_{i+T/\Delta t}$ and risk aversion effects are negligible. In good approximation, this is the case for efficient markets with low to moderate rate levels, where risk aversions of large lenders and borrowers are low, act in opposite directions and approximately cancel out. Using Equation 3.8, the expectation value of $R_{i+T/\Delta t}$ is

$$\varrho_i(T) \;=\; \frac{1}{\sqrt{2\,\pi}\;\sigma_z} \int_{-\infty}^{\infty} R(z)\; \mathrm{e}^{-\frac{(z-\bar{z}_i)^2}{2\;\sigma_z^2}}\, \mathrm{d}z \tag{3.9}$$

where $R(z)$ is defined by Equation 3.7. This means averaging R with a Gaussian weighting kernel. The integral can be solved:

$$\varrho_i(T) \;=\; \bar{\varrho}(\bar{z}_i, \sigma_z^2) \;=\; P\left(-\frac{\bar{z}_i}{\sigma_z}\right)\; \left[\frac{\bar{z}_i}{A} - (\bar{z} + \sqrt{\varepsilon})^2 + \varepsilon - \sigma_z^2\right] + \tag{3.10}$$

$$\frac{\sigma_z}{\sqrt{2\;\pi}}\; \mathrm{e}^{-\frac{\bar{z}_i^2}{2\;\sigma_z^2}}\; \left(\bar{z}_i + 2\;\sqrt{\varepsilon} - \frac{1}{A}\right) + (z + \sqrt{\varepsilon})^2 - \varepsilon + \sigma_z^2$$

where $P(.)$ is the cumulative standard normal distribution function. Whenever a simulation produces a value of $\bar{z}_i(T)$, Equation 3.10 is used to transform it to a forward rate $\varrho_i(T)$ which then can be used to construct a simulated yield curve. Figure 2 shows forward rates ϱ as functions of $\bar{z}$ for several values of σ_z^2 according to Equation 3.10. What happens if $\bar{z}$ is drifting in the negative region in a simulation? The corresponding ϱ values will stay close to zero for quite some time. This can be a real behavior, as we have seen for Japanese rates over the last few years.

The variance σ_z^2 has to be known in order to fully establish the link between ϱ and $\bar{z}$. In our model[12], Our σ_z^2 only depends on T and is independent of the current $\bar{z}$ level; this was one of the goals when we introduced the variable z. When z is normally distributed and innovations in $\bar{z}$ are assumed to be unexpected (caused by news) and independent, we can describe its dynamics in terms of a Brownian motion of $\bar{z}$. At maturity ($T = 0$), $\sigma_z^2 = 0$, as no uncertainty on the outcome remains. The longer the maturity period, the more unexpected news may increase the uncertainty. For a Brownian motion, we obtain $\sigma_z^2 \propto T$. However, σ does not grow to infinity with increasing T. Historical interest rate plots over several decades or even centuries (e.g. Figures 1.1, 1.2 and 17.2 of [James and Webber, 2000]) show that interest rate levels hardly drift to very extreme values (such as -0.5% or 40%) and never stay extreme for a long time. We rather observe a weak mean reversion[13] of IR levels that brings these levels back to a certain range in

[11] We cannot investigate z empirically here, because $\bar{z}$ is not yet defined as a function of observable forward rates.

[12] Again, we cannot use an empirical variance of z here, because we are still in the process of defining $\bar{z}$ as a function of observable variables. As soon as the model is complete, we can verify and calibrate it. In the further course of the algorithm, we are using a GARCH model for the variance of innovations of $\bar{z}$, see Section 2.9. That sophisticated volatility model should not be confused with the simple σ_z^2 model introduced here for the sole purpose of defining a suitable variable $\bar{z}$.

[13] Mean reversion effects in the long run are explicitly discussed in Section 3.8. At the moment, we are only interested in the behavior of σ_z^2.

the long run. Thus our σ_z^2 will not infinitely grow but rather converge to a finite value at very long maturities T. The variance behaves as follows, approximately:

$$\sigma_z^2 \;=\; \sigma_z^2(T) \;=\; b\,\frac{T}{T_{\text{rev}} + T} \tag{3.11}$$

This is just one possible function to model σ_z^2. The proposed function has two interesting properties. First, we look at short maturities and obtain

$$\sigma_z^2 \;\approx\; b\,\frac{T}{T_{\text{rev}}} \quad \text{for} \quad T \ll T_{\text{rev}} \tag{3.12}$$

This is indeed proportional to T. A reasonable choice of the constant T_{rev} is around 3 years. Now we also look at very long maturities:

$$\sigma_z^2 \;\approx\; b \quad \text{for} \quad T \gg T_{\text{rev}} \tag{3.13}$$

The constant b is the asymptotic value which defines the maximum volatility. Values roughly around $b \approx 0.02$ lead to realistic models.

3.5 The innovation of mapped forward interest rates

Now we are finally able to define $\bar{z}$ as a function of the observed forward interest rate ϱ. The variable $\bar{z}$ is the variable that satisfies Equation 3.10. This is the definition:

$$\bar{z}_i(T) \;=\; Z[\varrho_i(T), \sigma_z^2(T)] \tag{3.14}$$

where $Z(.,.)$ is the inverse function of $\bar{\varrho}(.,.)$, with

$$Z[\bar{\varrho}(\bar{z}, \sigma_z^2), \sigma_z^2] \;=\; \bar{z} \tag{3.15}$$

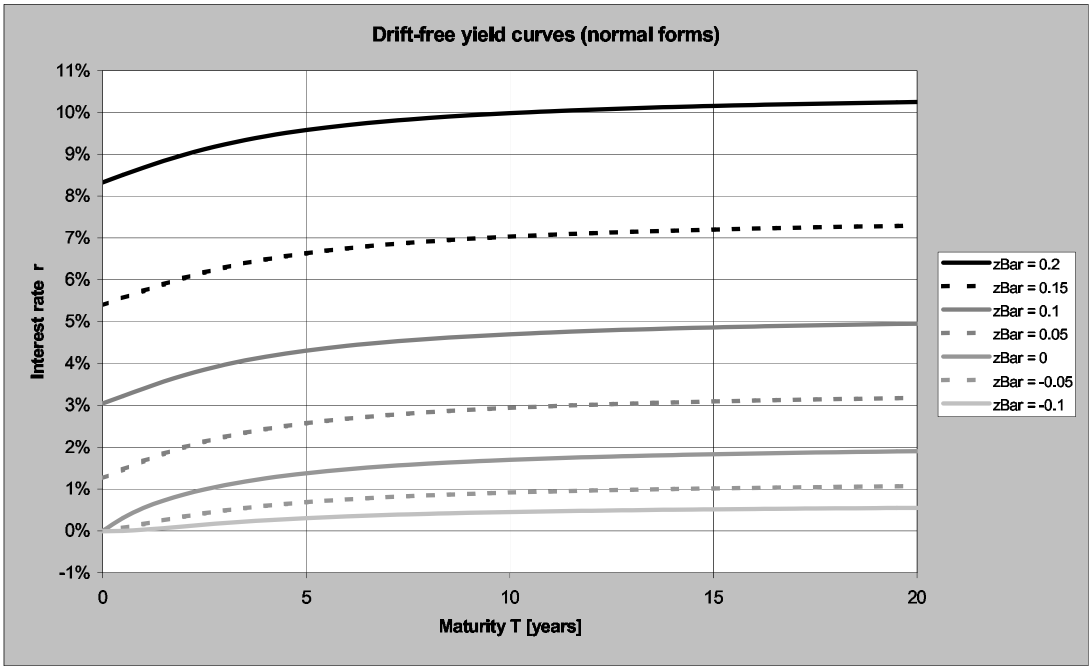
and $\sigma_z^2(T)$ is the result of Equation 3.11. There is no analytical formula for the function $Z(.,.)$, so we have to invert Equation 3.10 numerically. This is not a large problem as $\bar{\varrho}(\bar{z}, .)$ and $Z(\varrho, .)$ are monotonic functions for a constant σ_z^2. There is always exactly one finite solution of each function, given a finite argument.

For our bootstrapping algorithm, we shall use $\bar{z}_i(T)$ from Equation 3.14 instead of the unsuitable variable $\varrho_i(T)$. Now we can define the innovations in the sense of Equation 2.3:

$$I_i[\bar{z}(T)] \;=\; \bar{z}_i(T) - \bar{z}_{i-1}(T + \Delta t) \tag{3.16}$$

where both $\bar{z}_i(T)$ and $\bar{z}_{i-1}(T+\Delta t)$ result from Equation 3.14. This replaces the insufficient definition of Equation 3.5. In Section 3.8, this definition of innovations will be slightly modified as we correct the expectation $\bar{z}_{i-1}(T + \Delta t)$ of $\bar{z}_i(T)$ by a small mean-reversion term. The historically determined innovations $I_i[\bar{z}(T)]$ will not necessarily be normally distributed. We include the fat-tail correction of Section 2.10 and the GARCH correction of Section 2.9 in simulations, which implies a further deviation from the normal distribution. The empirical distribution of $z_{i+T/\Delta t} - \bar{z}_i(T)$ may also deviate from a theoretical normal distribution as assumed by Equation 3.8. The deviations should however be limited, in general and also under specific market conditions. This is a guideline when calibrating the four parameters of the $\bar{z}$ definition: ε, A, T_{rev} and b. Another useful study is to test if the historical innovations $I_i[\bar{z}(T)]$ are serially independent, as they should be for resampling.

The bootstrapping of interest rates is rather complex. This is why we summarize the steps in the following list, which is just a specific implementation of the general list of Section 2.2:



These stylized yield curves are based on fixed, constant values of the mapped interest rate $\bar{z}$ and represent standard forms. Simulated yield curves resulting from the bootstrapping method fluctuate around these standard forms and exhibit a wider variety of different forms.

Figure 3: Drift-free yield curves based on constant $\bar{z}$

- Compute all historical forward rates $\varrho_i(T)$ for a wide range of maturities T from 0 to 30 years (or more) as integer multiples of the basic time step Δt, using Equation 3.3 with an interpolation formula for interest rates R on the yield curve.
- Map all forward rates $\varrho_i(T)$ to $\bar{z}_i(T)$ by using Equations 3.14 and 3.11.
- Compute all innovations $I_i[\bar{z}(T)]$ through Equation 3.16, including the small mean-reverting correction of Section 2.7.
- Apply detrending and a GARCH analysis of innovations as specified in Sections 2.6 and 2.9.

A simulation step, which can be repeated arbitrarily often, is done as follows:

- Resample the innovations of a randomly selected historical index i, apply standard tail and GARCH corrections.
- Add the resampled innovations $I_i[\bar{z}(T)]$ to $\bar{z}_{j-1}(T+\Delta t)$ (or a more complex expectation of $\bar{z}_j(T)$, including the small mean-reverting correction of Section 3.8) to obtain $\bar{z}_j(T)$.
- Convert all $\bar{z}_j(T)$ to the simulated forward rates $\varrho_j(T)$, using Equation 3.10.
- Construct the simulated yield curve by averaging the obtained forward rates $\varrho_j(T)$ in the sense of Equation 3.4.

3.6 Interest rate mapping and the form of the yield curve

In this small section, we demonstrate that our asymmetric mapping of forward interest rates is closely related with the typical form of yield curves.

Let us imagine a static yield curve, free of drift and surprise, with constant values of $z = \bar{z}$ over the whole maturity axis, unchanged over time. The corresponding yield curve is not flat, in spite of the constant value of $\bar{z}$. This is due to the variance σ_z^2 which grows with increasing maturity T, following Equation 3.11. We can compute the forward rates as functions of T by inserting that equation in Equation 3.10 and average them to build the conventional yield curve of interest rates r $(= [\exp(R) - 1] \cdot 100\%)$ as functions of T. This has been made in Figure 3.

The yield curves of Figure 3 indeed look familiar, they have a "standard" or "classical" form, see e.g. [James and Webber, 2000], Section 1.3.1. A curve starts at a low value for very short maturities, has a positive slope, becomes flatter with increasing maturity and reaches an asymptotic value at long maturities $T \gg T_{\mathrm{rev}}$. Values at longer maturities exceed the short rates by a positive amount which can be called the "term premium". The curves with $\bar{z} < 0$ look slightly different with an almost horizontal tangent at $T = 0$. This is similar to the current form of the Japanese yield curve.

The fact that the drift-free yield curves have familiar forms confirms the suitability of our interest mapping. In reality and in our simulations, $\bar{z}$ is not constant across the maturity axis, and its values are affected by varying innovations, so the resulting yield curves will not be drift-free. Real yield curves make complex movements centered around the normal forms of Figure 3 and exhibit many different forms, as in Figures 1.3 and 1.4 of [James and Webber, 2000]: sometimes with steeper slopes, sometimes flatter, sometimes with inverted (negative) slopes, sometimes with humped forms. The variety of simulated yield curve forms will be shown in Section 4.2, Figure 4.

3.7 Processing inflation

Historical values of the Consumer Price Index (CPI) are the basis of our inflation processing. The CPI is the nominal value of a representative basket of consumer goods and results from rather sophisticated statistical methods. There is a debate on these methods, and there are alternative price indices that might be used instead of or in addition to the CPI.

We take logarithms of the CPI and define inflation as a first difference:

$$x''_i[\mathit{Infl}] \;=\; \log CPI_i - \log CPI_{i-1} \tag{3.17}$$

For quarterly data, we obtain quarterly inflation figures that are not annualized. Annualized inflation in the usual sense can be computed as $[\exp(4x''_i[\mathit{Infl}]) - 1] \cdot 100\%$.

Inflation exhibits serial correlation and behaves more like Brownian motion than like white noise. It is highly correlated with the IR level. Therefore, we resample innovations of inflation, which are correlated with IR innovations, rather than innovations of the CPI.

Inflation figures $x''_i[\mathit{Infl}]$ as computed by Equation 3.17 are not yet suitable for bootstrapping. Inflation exhibits seasonality: values in winter are typically higher than in summer. We should not resample winter innovations to simulate summers directly. Our solution is to deseasonalize inflation values. A simple way of deseasonalizing quarterly inflation figures is to subtract the mean of all historical inflation figures that occurred in the same quarter, for example the second quarter of a year. Let us denote the deseasonalized inflation by $x'_i[\mathit{Infl}]$.

Inflation is a difference in Equation 3.17 and is highly sensitive against small changes in the underlying CPI data[14]. Inflation innovations are like second differences of log(CPI) and are thus more affected by small changes than inflation itself. In order to prevent spurious volatility of these innovations due to noise, we need some smoothing of the deseasonalized $x'_i[\mathit{Infl}]$ values. An obvious way of doing this is to take a short-term moving average of $x'_i[\mathit{Infl}]$ in the sense of Equation 2.6:

$$x_i[\mathit{Infl}] \;=\; EMA[x'_i[\mathit{Infl}]; \Delta t_{\mathrm{smooth}}] \tag{3.18}$$

We choose a moderately short smoothing time constant $\Delta t_{\mathrm{smooth}}$. There is a trade-off: a large $\Delta t_{\mathrm{smooth}}$ will cause better smoothing, but then $x_i[\mathit{Infl}]$ will no longer be up to date.

The innovations of the deseasonalized and smoothed inflation are computed by Equation 2.3. This requires a formula for the expectation of inflation. The formula is given in Section 3.8, where the interaction of inflation and interest rates is described.

The bootstrapping algorithm leads to simulated values of the deseasonalized and smoothed inflation $x[\mathit{Infl}]$. We have to reseasonalize $x[\mathit{Infl}]$ to obtain simulated inflation values $x''[\mathit{Infl}]$ and cumulate these results to obtain simulated values of log(CPI). Undoing the smoothing of Equation 3.18 is also possible by adding some artificial noise, but this is probably useless in practice.

3.8 Interaction and expectation of interest rates and inflation

In the long run, neither interest rates nor inflation are freely drifting out of their usual range. This fact is implemented in the form of weak mean-reverting forces. We use constant target values, where high precision is not required, as the mean reversion is a very weak force. For inflation, we choose a target $x_{\mathrm{target}}[\mathit{Infl}]$, for mapped forward interest rates $\bar{z}$ a target $\bar{z}_{\mathrm{target}}$, both based on long-term historical experience.

Empirical research shows that a slightly stronger force holds IR levels and inflation together. In other words, real interest rates ($\approx$ IR minus inflation) have a lower volatility and a stronger mean reversion than IRs or inflation alone. This effect affects IRs of all maturities, but we choose a a rather short-term reference maturity $T = m\Delta t$ with a low integer number m to model it. We define an adjusted forward IR,

$$\varrho_{\mathrm{adj},i} \;=\; \varrho_i(m\Delta t) - x_i[\mathit{Infl}] \tag{3.19}$$

where $x_i[\mathit{Infl}]$ results from Equation 3.18. This adjusted rate is similar to a real interest rate, except for the timing: $\varrho_i(m\Delta t)$ refers to a time interval after t_i whereas $x_i[\mathit{Infl}]$ is the inflation of the interval before t_i. It has a sample mean $\bar{\varrho}_{\mathrm{adj}}$ which we estimate from historical data and possibly modify on the basis of expert opinion or a specialized study. At time t_{i-i}, the adjusted rate $\varrho_{\mathrm{adj},i-1}$ probably deviates from the mean $\bar{\varrho}_{\mathrm{adj}}$, so we model a force reverting to that mean. We obtain a corresponding target for inflation:

$$x'_i[\mathit{Infl}] \;=\; x_i[\mathit{Infl}] - \mu\;[\bar{\varrho}_{\mathrm{adj}} - \varrho_{\mathrm{adj},i}] \tag{3.20}$$

with a positive constant $\mu \approx 0.4$ which is discussed below. The target for $\varrho_i(m\Delta t)$ is

$$\varrho'_i(m\Delta t) \;=\; \varrho_i(m\Delta t) + (1-\mu)\;[\bar{\varrho}_{\mathrm{adj}} - \varrho_{\mathrm{adj},i}] \tag{3.21}$$

[14]Due to the complex computation procedure with debatable assumptions, CPI figures have natural uncertainties, are computed with a delay and may be modified by posterior corrections. Sometimes, we are forced to extrapolate the most recent CPI value in order to have a complete set of historical data.

where $1 - \mu \approx 0.6$ is also positive. The mean-reverting force due to the interaction of interest rates and inflation acts on both variables, but there are empirical indications for a slight lead-lag effect. Inflation affects IR levels slightly more than the other way around. This can be modeled by choosing $\mu < 1 - \mu$. The difference of the two target values of Equations 3.21 and 3.20 is the mean adjusted rate $\bar{\varrho}_{\text{adj}}$, as it should be.

For inflation, we arrive at the following formula for the expectation of $x_i[\mathit{Infl}]$:

$$E_{i-1}[\mathit{Infl}_i] \;=\; x_{i-1}[\mathit{Infl}] + \varepsilon_{\text{infl}}\, \{x_{\text{target}}[\mathit{Infl}] - x_{i-1}[\mathit{Infl}]\} + \varepsilon_{\text{adj}}\, \{x'_{i-1}[\mathit{Infl}] - x_{i-1}[\mathit{Infl}]\} \tag{3.22}$$

For mapped forward interest rates, we obtain the following modified version of the expectation of $\bar{z}_i(T)$:

$$\begin{aligned} E_{i-1}[\bar{z}_i(T)] \;&=\; \bar{z}_{i-1}(T + \Delta t) + \varepsilon_{\text{IR}}\, \{\bar{z}_{\text{target}} - \bar{z}_{i-1}(T + \Delta t)\} \\ &+\varepsilon_{\text{adj}} \sqrt{\frac{m\Delta t}{T}}\, \{Z(\varrho'_{i-1}(m\Delta t), \sigma_z^2(m\Delta t)) - \bar{z}_{i-1}(T + \Delta t)\} \end{aligned} \tag{3.23}$$

where the function $Z(.,.)$ of Equation 3.15 is used to convert a mean-reversion target from an unmapped rate to mapped one. The small factor ε_{adj} determines the mean-reversion effect due to the adjusted IR and is slightly larger than the tiny constants $\varepsilon_{\text{infl}}$ and ε_{IR}. The factor $\sqrt{m\Delta t/T}$ is used to modify the corrections for maturities T other than $m\Delta t$. The choice of this function as well as the diverse ε parameters and μ should be made on the basis of a study of the behavior of IRs and inflation. The resulting expectation $E_{i-1}[\bar{z}_i(T)]$ is used to compute the innovations of mapped forward interest rates:

$$I_i[\bar{z}(T)] \;=\; \bar{z}_i(T) - E_{i-1}[\bar{z}_i(T)] \tag{3.24}$$

This is the corrected version of Equation 3.16 to be used in the bootstrapping algorithm.

3.9 Foreign exchange rates

Foreign exchange (FX) rates can be treated in a simpler way than interest rates. We take the logarithm as our mapping function in the sense of Equation 2.2. An FX rate such as EUR/USD is defined as the value of a unit of an exchanged currency ("exch", here EUR) expressed in another currency ("expr", here USD). We take the logarithm as our mapping function in the sense of Equation 2.2, which is the only function that leads to an equivalent treatment of inverted FX rates (USD/EUR in our example):

$$x_i[FX] \;=\; \log FX_i \tag{3.25}$$

The market forecast of a spot FX rate $x_i[FX]$ at time t_i is the forward FX rate at time t_{i-1}. A forward FX rate depends on the difference of interest rates of the two involved currency zones, see Equation 2.2 of [Dacorogna et al., 2001]:

$$E'_{i-1}[FX_i] \;=\; x_{i-1}[FX] + (R_{\text{expr},i-1} - R_{\text{exch},i-1})\, \frac{\Delta t}{1\ \mathit{year}} \tag{3.26}$$

where R_{i-1} is a logarithmic interest rate at time t_{i-i} for the maturity period Δt as defined by Equation 3.1, for the exchanged currency ("exch") and the currency in which the FX rate is expressed ("expr"). Here we express all FX rates in USD, so the index "expr" always refers to the US market. Equation 3.26 is also known as covered interest parity.

There is no static mean reversion for FX rates, but there is purchasing-power parity (PPP). If the values of consumer good baskets (see section 3.7) in two different currency zones strongly deviate, there is a reverting market force which should be added to the forecast $E'_{i-1}[FX_i]$. This force is weak (see the article by Cheung in [Chan et al., 2000]) and only matters for long-term simulations. PPP can be modeled as a force towards a target level of the FX rate:

$$x_{\mathrm{ppp},i} \;=\; EMA_i[x[FX] - \log CPI_{\mathrm{expr}} + \log CPI_{\mathrm{exch}}; \Delta t_{\mathrm{ppp}}] + \log CPI_{\mathrm{expr},i} - \log CPI_{\mathrm{exch},i} \tag{3.27}$$

where the exponential moving average of Equation 2.6 is used. PPP is a slow effect, and the time constant Δt_{ppp} should be long enough to cover more than one PPP cycle (many years). An extremely large Δt_{ppp} would not be appropriate, because the consumer price indices (CPI), as values of baskets of varying composition, may not be entirely consistent in two countries over many decades.

The technical forward rate $E'_{i-1}[FX_i]$ is now modified by a weak force towards $x_{\mathrm{ppp},i-1}$ which is proportional to the distance between the two values:

$$E_{i-1}[FX_i] \;=\; E'_{i-1}[FX_i] + \varepsilon_{\mathrm{ppp}}\; [x_{\mathrm{ppp},i-1} - E'_{i-1}[FX_i]] \tag{3.28}$$

The force is weak, so $\varepsilon_{\mathrm{ppp}}$ is a small constant. The innovations of FX rates can now be computed by Equation 2.3, using the expectation $E_{i-1}[FX_i]$.

3.10 GDP

The modeling of the Gross Domestic Product (GDP) is rather simple. We use real (= deflated or inflation-corrected) GDP figures[15] rather than the nominal GDP, as this is less volatile. There are reliable, yearly OECD data that can be used to verify the quarterly data. The mapped real GDP is logarithmic:

$$x_i[GDP] \;=\; \log GDP_i \tag{3.29}$$

The volatility of the real GDP is modest as compared to the fluctuations of other variables such as equity indices or FX rates. Thus, normal applications (with no particular emphasis on GDP) do not need a sophisticated model for the market expectation of GDP growth. We simply take the sample mean of historical growth:

$$E_{i-1}[GDP_i] \;=\; x_{i-1} + \frac{1}{n} \sum_{j=1}^{n} x_j[GDP] \tag{3.30}$$

This is our expectation of x_i made at time t_{i-1} for any i, in all historical and simulated cases, independent of the market situation. This may be improved if an elaborated GDP model is available. Following Equation 2.3, the GDP innovation is

$$I_i[GDP] \;=\; x_i[GDP] - E_{i-1}[GDP_i] \tag{3.31}$$

The further steps of the algorithm follow the standard procedure as described in Sections 2.1 and 2.2. If we need a nominal GDP figure historically or in a simulation, we can always compute it as

$$NominalGDP \;=\; c\; GDP \cdot CPI \;=\; c\; \mathrm{e}^{x[GDP]+x[CPI]} \tag{3.32}$$

where a constant c determines the basis of the nominal GDP.

[15] Due to the complex computation procedure, GDP figures are computed with a delay and may be modified by posterior corrections. Sometimes, we are forced to extrapolate the most recent GDP value in order to have a complete set of historical data.

3.11 Equity indices

Many applications focus on the performance of equity investments. Thus we prefer modeling total-return (or "gross") indices that are based on the assumption that dividends are immediately and fully reinvested. An obvious choice is to take standard indices such as the MSCI gross indices which are available for all major currency zones. It is possible to have several equity indices per currency zone in the same scenario generator. We can technically treat price indices (excluding dividends), sector indices, hedge-fund indices or real-estate indices like the main total-return index.

Equity indices have a high volatility which may lead to extreme drifts in long-term simulations based on pure resampling. Is there a mean-reverting force to prevent this? In the long run, the "stock exchange economy" cannot arbitrarily drift away from the "real economy" as expressed by the GDP. Indeed, the "equity-to-GDP ratio" does not show a large trend over time. In the 1920s, before the "Black Friday", the ratio of the Standard & Poors 500 index to the US GDP reached similar levels as in the late 1990s. This notion leads us to choosing the equity-GDP ratio for resampling. The mapping is

$$x_i[\mathit{Equity}] \;=\; \log \mathit{EquityIndex}_i - (\log \mathit{GDP}_i + \log \mathit{CPI}_i) \tag{3.33}$$

The equity index has no inflation correction. Therefore we need the nominal GDP for our equity-GDP ratio, using the CPI as in Equation 3.32.

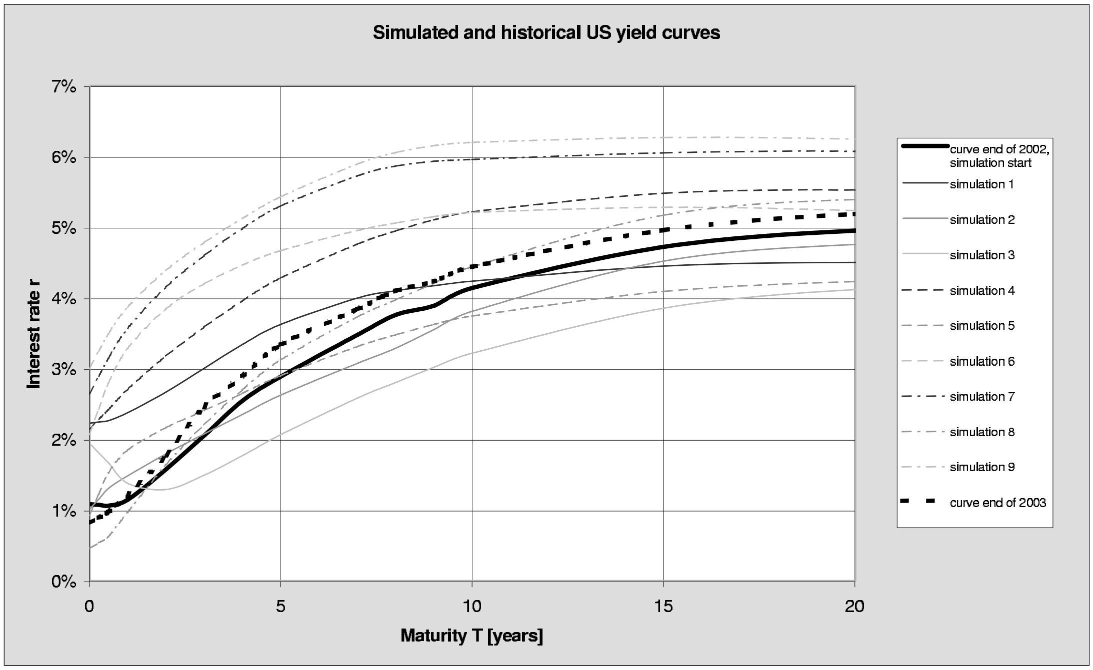
Now we compute the market expectation[16] of $x_i[\mathit{Equity}]$. The simplest model is taking the mean growth within the historical sample as a constant expectation, similar to Equation 3.34:

$$E_{i-1}[\mathit{Equity}_i] \;=\; x_{i-1}[\mathit{Equity}] + \frac{1}{n}\sum_{j=1}^{n} x_j[\mathit{Equity}] \tag{3.34}$$

A first refinement is to add a bias to this market expectation of growth. If an external model gives rise to a certain assumption on future growth, this can be incorporated in the formulation of the market expectation. We are doing this in our scenario generator. For long-term simulations, we need a mean reversion of our equity-GDP ratio. The value of $x_i[\mathit{Equity}]$ slowly reverts to a static long-term mean (a good assumption for price indices) or to a slow growth path (for total-return indices). This mean reversion can be implemented as in the case of FX rates, similar to Equation 3.28.

Equity returns have been found to have a negatively skewed distribution. Rare crashes imply few strongly negative returns, whereas long boom phases generate many moderately positive returns. This skewness of historical stock returns will be reflected by the resampled innovations and thus maintained by the bootstrapping method, at least when looking at one-step returns, i.e. quarterly returns in our implementation. The historical correlation between stock and bond returns is also respected. Returns of long-term bond portfolios are mainly driven by changes in IR levels (rather than the IR levels themselves). Since IR innovations values are resampled along with equity index innovations, a realistic correlation between bond and equity returns will be reproduced by the simulation. This is important for the application to asset allocation, where most portfolios contain bonds as well as equity.

[16] We do not use market-implied forecasts extracted from market prices of derivatives because these are generally not available for the time horizons we are interested in (especially in the case of long-term simulations).



As an example for simulation results of the Economic Scenario Generator (ESG), US yield curves are shown. The bold, solid US yield curve of the end of 2002 marks the starting point of the simulation. The thin yield curves are simulated for the end of 2003, based on the information set available at the end of 2002, representing 9 different simulation scenarios. The bold, dotted curve is the yield curve actually observed at the end of 2003. We can compare this curve to the set of simulated curves. Such comparisons are a basis for backtesting. In usual simulation studies and in backtests, we use thousands of scenarios rather than just 9 of them.

Figure 4: Simulation of US yield curves

4 Results and testing

4.1 Calibration

As a non-parametric method, the pure bootstrapping algorithm has a natural advantage over parametric models: there is no need for any calibration. Historical behaviors are reproduced in the simulated future. If past behaviors provide a reliable guidance for future behaviors (which is not always an accurate assumption), the simulated results will automatically be in line.

On a subordinated level, our refined bootstrapping method has some parametric elements such as some mapping formulas, GARCH and small corrections of market expectations, mainly the weak mean-reversion effects of some variables. The models of these refinements rely on a few parameters that were calibrated by some special studies during the development phase. Some calibrations are done as a preparatory step at the beginning of each new scenario generation, based on the same historical data that are also used by the bootstrapping method. This is the case for the GARCH calibrations as presented in the Appendix 6.

At the end of our development process, some experienced finance specialists looked at many graphs with simulation results. Given the complexity of setting up a comprehensive set of calibration and testing procedures, the value of judgments by human experts should not be underestimated.

Although a calibration procedure may lead to an in-sample quality measure of a model, it does not provide an independent, reliable quality test. At the end, an out-of-sample test is necessary to finally assess a scenario generator and its results.

4.2 Example of results: yield curve simulation

An economic scenario generator produces a wealth of results: many complete scenarios for many economic variables, simulated over many time steps. In Figure 4, we show a tiny part of these results for the example of simulated US yield curves. The historical curve for the end of 2002 marks the starting point of the simulations. The 9 simulated curves are for the end of 2003, based on four quarterly simulation steps. Each curve belongs to a complete, consistent scenario for all economic variables, including interest rates with times to maturity up to 30 years, yield curves for all currencies and many other variables. Although the simulation period is just one year, we see a variety of yield curve forms on different levels. Simulation 3 has an inversion for maturities up to 2 years, simulation 6 has a more humped form.

The true US yield curve at the end of 2003 is also plotted in Figure 4 as a bold, dotted curve. It lies slightly above the curve of 2002. The simulated curves show a wider range of deviations. This indicates a considerable IR risk at start of 2003; the actual move in 2003 was rather small in comparison. The overall IR level increased, which is in line with the majority of simulated curves. The very low short rate at the end of 2003 is below most of the simulated values, indicating a particular low-rate policy of the US Federal Reserve.

4.3 Out-of-sample backtesting

Our quality testing method is backtesting, see e.g. [Christoffersen, 2003]. Comparisons between historical variable values and their prior scenario forecasts, as in Figure 4, are a basis for backtesting. In Chapter 8 of [Christoffersen, 2003], backtests are proposed for three different types of forecasts: (1) point forecasts for the value of a variable, (2) probability range forecasts (e.g. the value at risk (VaR) which is the projected quantile at a certain probability, often 1%) and (3) forecasts of the complete probability distribution. Such distribution forecasts are the most comprehensive type as they imply range forecasts and point forecasts (using the mean or median of the distribution, for example).

Scenarios produced by a scenario generator are no forecasts in the usual sense. In typical studies, we produce many thousands of scenarios. Each of these scenarios has its own forecast value for a certain variable at a certain future time. All the scenario values together define an empirical distribution for the variable. Hence we have distribution forecasts rather than just point or range forecasts.

Our task is comprehensive out-of-sample backtesting of distribution forecasts. Even the limited task of testing specialized models such as an interest rate model is difficult, as discussed in Section 1.5.2 of [James and Webber, 2000]. Here we propose a methodology based on the Probability Integral Transform (PIT). [Diebold et al., 1998, Diebold et al., 1999] have introduced the PIT (also known as Lévy or Rosenblatt Transform) as a method for testing distribution forecasts in finance. The whole test is described in detail in [Blum, 2004]. This is a summary of the steps:

1. We define an in-sample period for building the bootstrapping method with its innovation vectors and parameter calibrations (e.g. for the GARCH model). The out-of-sample period starts at the end of the in-sample period. Starting at each regular time point out-of-sample, we run a large number of simulation scenarios and observe the scenario forecasts[17] for each of the many variables of the model.

2. The scenario forecasts of a variable x at time t_i, sorted in ascending order, constitute an empirical distribution forecast. In the asymptotic limit of very many scenarios, this distribution converges to the marginal cumulative probability distribution $\Phi_i(x) = P(x_i < x|\mathcal{I}_{i-m})$ that we want to test, conditional to the information $\mathcal{I}_{i-m}$ available up to the time t_{i-m} of the simulation start. In the case of a one-step forecast, $m = 1$. The empirical distribution $\hat{\Phi}_i(x)$ slightly deviates from this. The discrepancy $\Phi_i(x) - \hat{\Phi}_i(x)$ can be quantified by using a formula given by [Blum, 2004]. Its absolute value is less than 0.019 with a confidence of 95% when choosing 5000 scenarios, for any value of x and any tested variable. This is accurate enough, given the limitations due to the rather low number of historical observations.

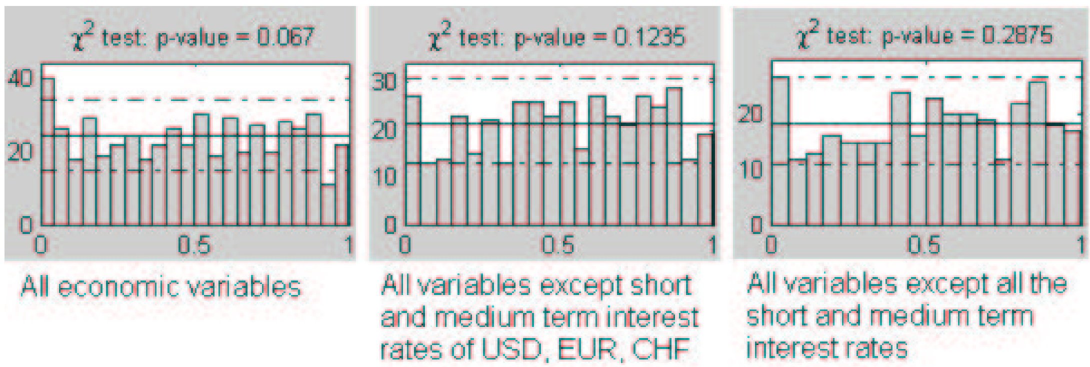
3. For a set of out-of-sample time points t_i, we now have a distribution forecast $\hat{\Phi}_i(x)$ as well as a historically observed value x_i. The cumulative distribution $\hat{\Phi}_i(x)$ is used for the following Probability Integral Transform (PIT): $Z_i = \hat{\Phi}_i(x_i)$. The probabilities Z_i, which are confined between 0 and 1 by definition, are used in the further course of the test. A proposition proved by [Diebold et al., 1998] states that the Z_i are i.i.d. with a uniform distribution $U(0, 1)$ if the conditional distribution forecast $\Phi_i(x)$ coincides with the true process by which the historical data have been generated. The proof is extended to the multivariate case in [Diebold et al., 1999]. If the series of Z_i significantly deviates from either the $U(0, 1)$ distribution or the i.i.d. property, the model does not pass the out-of-sample test.

Testing the hypotheses of $U(0, 1)$ and i.i.d. can now be done by using any suitable method from statistics. We pursue two approaches here:

1. An approach which we call *non-parametric* is suggested by [Diebold et al., 1998, Diebold et al., 1999]. It consists of considering histograms in order to detect deviations from the $U(0, 1)$ property, and correlograms of the Z_i's and their low integer powers to detect deviations from the independence property. We complement these graphical evaluations by the usual χ^2 test for uniformity, and by Kendall-Stuart bounds for the significance of the autocorrelations.

2. [Chen and Fan, 2004] suggest another approach, which we call *parametric*. It relies on the assumption that the Z_i's form a Markov chain with stationary distribution $G^*(.)$ and copula $C^*(.,.)$ for the dependence structure of (Z_i, Z_{i-1}). One can then select some model for $G^*(.)$ which contains $U(0, 1)$ as a special case, and some model for $C^*(.,.)$ which contains the independence copula as a special case. The joint null hypothesis of independence and uniformity can then be tested by standard likelihood ratio or Wald procedures. In this study, we specifically use the Farlie-Gumbel-Morgenstern copula as a model for dependence structure and the β-distribution as a model for the marginal distribution. In a *semi-parametric* variant of this procedure,

[17] Our main test is for one-step forecasts where the simulation is for one time step (a quarter in our case). Multi-step forecasts can be tested using the same methodology, but the number of available independent observations with non-overlapping forecast intervals will be distinctly smaller, given the same out-of-sample period.



The frequency of empirically found probabilities Z_i (results of the Probability Integral Transform, PIT) is plotted. A model is rejected if such a histogram significantly deviates from a uniform distribution, corresponding to a low p-value of the χ^2 test ($p < 0.05$). The left histogram is based on all economic variables, whereas some short and medium term interest rates are excluded from the computation of the other histograms. The dashed lines indicate a 95% confidence range for the individual frequencies.

Figure 5: Out-of-sample backtesting: uniform distribution of PIT-transformed variables

no model for $G^*(.)$ is chosen, but the empirical distribution of the Z_i's is plugged in instead. This allows to test for the isolated hypothesis of independence, irrespective of the marginal distribution.

A rejection by one of these tests does not necessarily mean that a model is valueless. It means that the model does not live up to the full predictive potential indicated by the data or that there is a structural difference between the in-sample and out-of-sample periods.

When applying the tests to our ESG results, the limited number of historical observations poses a problem. For a few economic variables, we have decades of historical data, but we are restricted to the period after September 1993 when constructing our comprehensive ESG with many variables and many currencies. This leaves very little space for defining a reasonable out-of-sample period. In order to increase this space, we cut the in-sample period (which normally covers 10 years) to 8 years, from end of September 1993 to September 2001. We obtain an ESG with only 32 quarterly innovations, which implies a less stable behavior than the production version with 40 innovations. This reduced ESG is tested out of sample.

The out-of-sample period starts at the end of September 2001 and ends in June 2004, which allows for testing 11 one-step forecasts, i.e. 11 observations of PIT-transformed values Z_i per economic variable. This is a low number for any statistical test. However, we obtain an sizable total number of Z_i observation if we consider all the economic variables for all the currencies. Our tested variables are equity index (MSCI gross), FX rate[18]against the USD, CPI and GDP. We add four interest rates to this set of variables, namely the extremes on our maturity scale, the 3-month and the 30-year rates, and two intermediate rates with times to maturity of 2 years and 10 years. Thus we obtain 8 variables for each of the 6 currency zones (USD, EUR, JPY, GBP, CHF, AUD). We link the small Z_i series of all variables together to obtain a set of 528 ($= 11 \cdot 8 \cdot 6$) observations of Z_i.

[18]For the currency USD, choosing the FX rate against itself makes no sense. Instead, we add a US hedge-fund index to the US-based variables to be tested.

The χ^2 test of the 528 Z_i observations and the underlying histogram is shown on the left-hand side of Figure 5. The p-value of 0.0607 exceeds the confidence limit of 0.05. The ESG forecasts are not rejected, but the low p-value does not instill wholehearted confidence. An autocorrelation analysis reveals a marginally significant first-lag autocorrelation between the Z_i. The semi-parametric evaluation has a high p-value and does not reject the ESG forecasts. The likelihood ratio test of the parametric evaluation, which is the most powerful test, significantly rejects the null hypothesis of i.i.d. $U(0,1)$ with a p-value of only 0.00021, which is far below a confidence limit of 0.05.

We have to accept the fact that the ESG forecasting method is rejected by our most powerful test. Fortunately, the testing methods also inform us on what exactly is rejected, and why. We need a closer look at the investigated out-of-sample period. In some respects, our out-of-sample period is characterized by a fundamental difference from the in-sample period. It covers an economic situation after a marked decline of equity markets. The worsening economic situation caused low demand, low inflation and low interest rates. Most importantly, the US Federal Reserve chose a distinct policy which kept short-term interest rates low and the US yield curve artificially steep. This policy is specific to the years 2001 -2004 and distinctly different from the policies of the in-sample period and the 1980s. It led to low values of low and medium term interest rates, much lower than the market forecasts based on forward interest rates indicated[19]. The example of Figure 4 can be seen as an illustration of the unexpectedly low short-term interest rates caused by this policy. In the first histogram of Figure 5, the low rates materialize in the significantly high frequency of Z_i values in the leftmost bar.

Our hypothesis is that the unusual low-interest policy is the reason for the rejection of the forecasts. We test this hypothesis by excluding the 3-month, 2-year and 10-year interest rates, so the 30-year rate is the only interest rate in the test. In an analysis called study B, we do this only for the currencies USD (directly affected by the US Feral Reserve policy) and EUR and CHF, where the central banks followed similar, if less pronounced policies. Thus the currencies JPY, GBP and AUD still have a full coverage of interest rates. Study B has a sample of 429 Z_i observations. In study C, we exclude short and medium term interest rates for all currencies and arrive at a sample of 330 observations. In both studies, B and C, the ESG forecasts are no longer rejected by any test. The χ^2 tests have p-values of 0.1235 (B) and 0.2875 (C), both on the good side of the confidence limit of 0.05, see the middle and right histograms of Figure 5. The strongest test, the parametric evaluation, confirms this with p-values of 0.2313 (B) and 0.6017 (C). We conclude that the ESG forecasts are rejected only in the case of low and medium term interest rates of USD, EUR and CHF. Thus we report a qualified success of our ESG forecasts.

Is there a way to improve the method in order to give optimal forecasts for all variables? This is only possible if factors such as the policy of the US Federal Reserve or, more generally, economic cycles can be predicted. Neither the bootstrapping method nor any of its algorithmic modifications are able to do this, to our knowledge. Long data samples covering many decades and many economic cycles would help, but we are restricted to shorter samples for most of the modeled economic variables. Shifts in policies, economic cycles and market structures make future developments less predictable. In our bootstrapping method, a way to accommodate this would be to augment the resampled innovations by a factor. Technically, this can be done the same way as the tail correction of Section 2.10, using Equation 2.9 with an increased "cycle uncertainty multiplier".

[19]Some traders made a bet on the persistence of this anomaly and made profitable "carry trades". They financed long-term deposits by rolling short-term debts forward.

Although the tests based on PIT are powerful, they cannot test all possible aspects of model quality. Several competing models or simulation methods might pass a PIT-based test at the same time, but one model might still be better than another[20]. Some properties stay untested in our PIT-based method, most notably the dependence between returns of different variables in the simulated scenarios. We have added a study comparing correlations of simulated returns to those of actual returns, with good results. This is expected for a bootstrapping method which preserves dependencies in the innovations by design.

5 Conclusion

Refined bootstrapping is our method to generate realistic scenarios of the future behavior of the global economy as represented by a set of key variables. We have presented many details that need to be observed in order to arrive at a realistic behavior of many different economic variables such as interest rates, foreign exchange rates, equity indices, inflation and GDP for several currency zones. A careful treatment of these modeling details, which include some subordinated parametric elements, is vital for the success of the bootstrapping method.

The following advantages of the bootstrapping method have been found:

- wide coverage of economic variables, modularity and flexibility when extending the set of covered economic variables;
- automatic preservation of distributions and simultaneous dependencies between the innovations of different economic variables;
- exact reproduction of initial conditions at simulation start (no fitting of a model needed for that);
- feasibility of long-term simulations (over decades), due to mean-reversion elements in expectations of variables;
- natural transition from the short-term behavior at start to the long-term behavior;
- easy ways to introduce modifications based on special studies or expert opinion (e.g. assuming expected equity returns lower than the mean of the historical sample);
- good coverage of extreme risks, relying on the tail correction of Section 2.10 and large numbers of simulations;
- no large calibration problems because the method is essentially non-parametric.

Out-of-sample tests have confirmed the validity of the approach. A certain problem arises from the behavior of short and medium term interest rates of some currencies, reflecting an unusual low-interest policy of central banks during the out-of-sample period. We have discussed this behavior and possible solutions.

The final goal of our project has always been the application of the method in practice. We have implemented the refined bootstrapping method in our Economic Scenario Generator (ESG). The results are regularly applied to Asset-Liability Management (ALM)

[20]Example: Some variables might follow a complex nonlinear process that is captured by model A, whereas model B sees the same behavior as random noise. While none of the models is rejected in a PIT-based test, the nonlinear model A is "better" as it predicts narrower distributions.

studies that are part of the strategic decision making of the analyzed companies. We plan to include corporate yield spreads and possibly other economic quantities to the set of bootstrapped variables in order to add new asset classes such as corporate bonds to ESG-based asset allocation studies.

6 Appendix: Robust calibration of a GARCH process

In Equation 2.8, a GARCH(1,1) process is defined. In our application, we need an especially robust calibration procedure. Following [Zumbach, 2000], we do not directly calibrate the three parameters α_0, α_1 and β_1. We rather reformulate the equation for the conditional variance as follows:

$$\sigma_1^2 \;=\; \sigma^2 + \mu_{\mathrm{corr}}\,[\mu_{\mathrm{ema}}\,\sigma_{i-1}^2 + (1-\mu_{\mathrm{ema}}\,r_{i-1}^2) - \sigma^2]\;, \tag{6.1}$$

$$\mu_{\mathrm{corr}} \;=\; \alpha_1 + \beta_1\;,\quad \mu_{\mathrm{ema}} \;=\; \frac{\beta_1}{\mu_{\mathrm{corr}}}\;,\quad \sigma^2 \;=\; \frac{\alpha_0}{1-\mu_{\mathrm{corr}}}$$

The parameters μ_{corr} and μ_{ema} have values between 0 and (less than) 1. While μ_{corr} describes the decay of the memory in conditional volatility, μ_{ema} determines the depth of averaging in the formation of the volatility memory.

The unconditional variance σ^2 is no longer regarded as a model parameter to be optimized through maximum likelihood. Instead, we directly take the empirical variance of the raw innovations as the "moment estimator" for σ^2. Thus we make sure that the unconditional variance of the process equals the empirical variance even if the GARCH process is misspecified or finite-sample problems lead to difficult behavior.

The two parameters μ_{corr} and μ_{ema} remain to be calibrated. The resulting GARCH(1,1) embeds two other processes[21]: ARCH(1) if $\mu_{\mathrm{ema}} = 0$ and a Gaussian random walk (Brownian motion, white noise) if $\mu_{\mathrm{corr}} = 0$. In the latter case, the value of μ_{ema} becomes irrelevant.

The GARCH equation is evaluated iteratively at each time series point with index i. Therefore all μ parameters correspond to an exponential decay with time constant τ:

$$\mu_{\mathrm{corr}} \;=\; \mathrm{e}^{-1/\tau_{\mathrm{corr}}}\;,\quad \mu_{\mathrm{ema}} \;=\; \mathrm{e}^{-1/\tau_{\mathrm{ema}}}\;,\quad \tau_{\mathrm{corr}} \;=\; -\frac{1}{\log\mu_{\mathrm{corr}}}\;,\quad \tau_{\mathrm{ema}} \;=\; -\frac{1}{\log\mu_{\mathrm{ema}}} \tag{6.2}$$

where the τ values are in units of the time step of the time series.

If the maximum-likelihood procedure leads to a μ very close to 1, the time constants τ may reach extremely high values. Reason demands that we do not choose a time constant exceeding the sample size. This is why our robust method sets an upper limit for τ:

$$\tau_{\max} \;=\; f\;n \tag{6.3}$$

where n is size of the sample used for fitting and f a constant factor; we usually take $f = 0.5$. If we use a 10-year sample, for example, we do not accept decay models with time constants longer than 5 years. At the limit, there are only two 5-year volatility clusters within the 10-year sample, at maximum. Two observations are not a large amount in statistics. This fact may lead to an intuitive understanding of why we are not willing

[21] [Zumbach, 2000] transforms μ_{corr} and μ_{ema} to other fitting variables by further mapping. We do not follow that approach as it pushes ARCH(1) and white noise (which are perfectly acceptable solutions) far away to the asymptotic limits of the parameter space.

to accept even longer clusters with even lower significance in our robust GARCH fitting procedure. Our condition is

$$0 \leq \tau_{\mathrm{corr}} \leq \tau_{\max} \; , \quad 0 \leq \tau_{\mathrm{ema}} \leq \tau_{\max} \; , \quad \mu_{\mathrm{corr}} \leq \mathrm{e}^{-1/\tau_{\max}} \; , \quad \mu_{\mathrm{ema}} \leq \mathrm{e}^{-1/\tau_{\max}} \tag{6.4}$$

where the conditions for μ are derived from Equation 6.2. The unconstrained solution of most practical fitting cases anyway obeys Equation 6.4. However, in some misspecified or small-sample cases, the maximum likelihood may lie outside those conditions, and we prefer the robust solutions ensured by Equation 6.4. The stationarity condition, $\mu_{\mathrm{corr}} = \alpha_1 + \beta_1 < 1$, is always fulfilled by the slightly stronger μ_{corr} condition of Equation 6.4. Our solutions not only observe the stationarity limit condition but also keep a safe distance from that limit.

The logarithm of the likelihood function[22] is

$$l(\mu_{\mathrm{corr}}, \mu_{\mathrm{ema}}) \; = \; -\frac{1}{2\,n} \sum_{i=m+1}^{n+m} \left[\ln 2\pi + \ln \sigma_i^2 + \frac{r_i^2}{\sigma_i^2} \right] \tag{6.5}$$

with a total number of $n + m$ observations in the sample. We reserve a considerable number m of initial observations for the build-up of σ_i^2. At start ($i = 1$), we use the initial value

$$\sigma_0^2 \; = \; \sigma^2 \tag{6.6}$$

which has an initial error that exponentially declines over the GARCH iterations, Equation 6.1, from $i = 1$ to m. The larger m, the smaller is the remaining error of σ_i^2. However, the remaining sample of size n also becomes smaller, given a limited total size $n+m$. This is a trade-off. In our low-frequency case with quarterly data, this trade-off is almost desperate. A 10-year sample has 40 quarterly observations – a modest number. We need these 40 observations for the likelihood function in order to produce meaningful results. Reserving 20 observations for build-up and using the remaining, meager 20 observations for GARCH fitting does not seem to be a reasonable approach. For some economic variables, we have past data older than 10 years that we can use for the build-up. For some other time series, this is not available. As a numerical trick, we can recycle the scarce available data to build up an initial σ_i^2 through a "zig-zag" method. We create a synthetic past. The real data are $r_{2n+1} \ldots r_{3n}$, so $m = 2n$; the synthetic past consists of $r_1 \ldots r_n$ with $r_i = r_{2n+i}$ and $r_{n+1} \ldots r_{2n}$ with $r_i = r_{4n+1-i}$. This is justified as the innovations r_i are already detrended and their temporal coherence, which is important for GARCH, is respected, though partially in reverse order. We claim that the thus obtained σ_{2n}^2 value is a better approximation of the true value than a simple initialization $\sigma_{2n}^2 = \sigma^2$. Of course, this claim should be substantiated through a theoretical or statistical study.

Now we have to determine the maximum of the log-likelihood of Equation 6.5 by varying the parameters μ_{corr} and μ_{ema} under the constraints of Equation 6.4. This cannot be done analytically. The solution of this non-linear optimization problem can be done with the help of any appropriate method[23]. All numerical methods lead to a local optimum depending on the initial guess of parameter values. In order to obtain the global optimum, it is important to run the optimization from different initial parameters and to take the best among the obtained solutions. For some GARCH(1,1) processes with large values of

[22] This formulation assumes that ε_i is generated by a Gaussian random process, which is an acceptable assumption for our low-frequency data and the limited role of GARCH within the whole bootstrapping algorithm. Assuming another distribution of ε_i leads to a different log-likelihood function.

[23] Contact the first author to learn about the set of methods actually used in the implemented scenario generator.

both μ_{corr} and μ_{ema}, the white noise solution ($\mu_{\text{corr}} = 0$) appears as a local optimum that is dominated by the true, global optimum. Therefore one should always start optimizations from at least two initial points: (1) white noise and (2) close-to-maximum values of μ_{corr} and μ_{ema}.

References

[Artzner et al., 1997] **Artzner P., Delbaen F., Eber J., and Heath D.**, 1997, *Thinking coherently*, Risk Magazine, **10**(11), 68–71.

[Baillie et al., 1996] **Baillie R. T., Bollerslev T., and Mikkelsen H.-O.**, 1996, *Fractionally integrated generalized autoregressive conditional heteroskedasticity*, Journal of Econometrics, **74**(1), 3–30.

[Ballocchi et al., 1999] **Ballocchi G., Dacorogna M. M., Hopman C. M., Müller U. A., and Olsen R. B.**, 1999, *The intraday multivatiate structure of the Eurofutures markets*, Journal of Empirical Finance, **6**, 479–513.

[Barone-Adesi et al., 1999] **Barone-Adesi G., Giannopoulos K., and Vosper L.**, 1999, *VaR without correlations for portfolios of derivative securities*, Journal of Futures Markets.

[Barone-Adesi et al., 2002] **Barone-Adesi G., Giannopoulos K., and Vosper L.**, 2002, *Backtesting derivativeportfolios with filtered historical simulation*, European Financial Management, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=308955, **8**, 31–58.

[Blum, 2004] **Blum P.**, 2004, *On some mathematical aspects of dynamic financial analysis*, preprint of Ph. D. thesis, ETH Zürich (Swiss Federal Institute of Technology, Zurich), to be published in 2005, 1–183.

[Blum and Dacorogna, 2003] **Blum P. and Dacorogna M. M.**, 2003, *Dynamic financial analysis*, In J. Teugels and B. Sundt (Eds.) Encyclopedia of Actuarial Science, **John Wiley & Sons (forthcoming 2003)**, 1–21.

[Bollerslev, 1986] **Bollerslev T.**, 1986, *Generalized autoregressive conditional heteroskedasticity*, Journal of Econometrics, **31**, 307–327.

[Casualty Actuarial Society, 1998] **Casualty Actuarial Society**, 1998, *DFA Research Handbook*, Casualty Actuarial Society, www.casact.org/research/dfa.

[Chan et al., 2000] Chan W. S., Li W. K., and Tong H., editors, 2000, *Statistics and Finance: An Interface*, Imperial College Press, London.

[Chen and Fan, 2004] **Chen X. and Fan Y.**, 2004, *Evaluating density forecasts via the copula approach*, Finance Research Letters, **1**(1), 74–84.

[Christoffersen, 2003] **Christoffersen P. F.**, 2003, *Elements of Financial Risk Management*, Academic Press, San Diego.

[Dacorogna et al., 2001] **Dacorogna M. M., Gençay R., Müller U. A., Olsen R. B., and Pictet O. V.**, 2001, *An Introduction to High Frequency Finance*, Academic Press, San Diego,CA.

[Diebold et al., 1998] **Diebold F., Gunther T., and Tay A.**, 1998, *Evaluating density forecasts with applications to financial risk management*, International Economic Review, **39**(4), 863–883.

[Diebold et al., 1999] **Diebold F., Hahn J., and Tay A.**, 1999, *Multivariate density forecasts and calibration in financial risk management: High-frequency returns on foreign exchange*, The Review of Economics and Statistics, **39**(4), 863–883.

[Dimson et al., 2003] **Dimson E., March P., and Staunton M.**, 2003, *Global evidence on the equity risk premium*, Journal of Applied Corporate Finance, **15**(4).

[Efron and Tibshirani, 1993] **Efron B. and Tibshirani R.**, 1993, *An introduction to the bootstrap*, Chapman & Hall.

[Embrechts et al., 1997] **Embrechts P., Klüppelberg C., and Mikosch T.**, 1997, *Modelling Extremal Events*, volume 33 of *Applications of Mathematics Stochastic Modelling and Applied Probability*, Springer, Berlin.

[Engle, 1982] **Engle R. F.**, 1982, *Autoregressive conditional heteroskedasticity with estimates of the variance of U. K. inflation*, Econometrica, **50**, 987–1008.

[James and Webber, 2000] **James J. and Webber N.**, 2000, *Interest Rate Modelling*, John Wiley & Sons, Chichester.

[Marsala et al., 2004] **Marsala C., Pallotta M., and Zenti R.**, 2004, *Integrated risk management with a filtered bootstrap approach*, Presented at the V Workshop di Finanza Quantitativa, Siena, 29-30 Jan 2004, www.econ.pol.unisi.it/wfq/PaperSiena/Marsala-Pallotta-Zenti.pdf, 1–16.

[Zenti and Pallotta, 2000] **Zenti R. and Pallotta M.**, 2000, *Risk analysis for asset managers: Historical simulation, the bootstrap approach and value at risk calculation*, EFMA 2001, Lugano, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=251669, 1–44.

[Ziemba and Mulvey, 1998] Ziemba W. T. and Mulvey J. M., editors, 1998, *Worldwide Asset and Liability Modelling*, Cambridge University Press, Cambridge.

[Zumbach, 2000] **Zumbach G.**, 2000, *The pitfalls in fitting GARCH processes*, Advances in Quantitative Asset Management, edited by Christian Dunis. Kluwer Academic Publisher.