SIMULATING COINTEGRATED TIME SERIES

Alexander Galenko

PENSON Financial Services
6805 North Capital of Texas Highway, Suite 320
Austin, TX, 78731, U.S.A.

David Morton

Operations Research and Industrial Engineering
The University of Texas at Austin
Austin, TX 78712, U.S.A.

Elmira Popova

Operations Research and Industrial Engineering
The University of Texas at Austin
Austin, TX 78712, U.S.A.

Ivilina Popova

Department of Finance and Economics
Texas State University
San Marcos, TX, U.S.A

ABSTRACT

When one models dependence solely via correlations, portfolio allocation models can perform poorly. This motivates considering dependence measures other than correlation. Cointegration is one such measure that captures long-term dependence. In this paper we present a new method to simulate cointegrated sample paths using the vector auto-regressive-to-anything (VARTA) algorithm. Our approach relies on new properties of cointegrated time series of financial asset prices and allows for marginal distributions from the Johnson system. The method is illustrated on two data sets, one real and one artificial.

1 INTRODUCTION

The recent financial crisis and the collapse of the hedge fund Long Term Capital Management in 1998 have shown the importance of correctly modeling the dependence between financial time series. When a large systematic shock affects the entire financial system, correlation becomes an inferior tool for risk management. The standard approach used in financial portfolio construction and management is based on the mean-variance methodology developed by Markowitz (1952). One of its important inputs is the covariance matrix (and as a consequence, correlation) of the financial returns. There are a variety of approaches designed to yield better estimates of these parameters. Another approach is based on maximizing the investor's utility function. Popova et al. (2007) used a combination of simulation and optimization to obtain solutions of such a stochastic optimization model for a family of utility functions. One important step in their approach is simulation of the possible future scenarios that are used in the optimization model. Given the current events in the financial markets, this paper focuses on creating a model for scenario simulation that is rooted in preserving long-term trends in financial time series.

A measure of long-term dependencies in financial time series is cointegration. Since the seminal work of Engle and Granger (1987, 1991), a number of researchers in finance and economics have used cointegration to model dependencies between security prices. Several statistical tests for cointegration have been developed (Engle and Granger 1991). Tsay (2005) presented numerous examples and a detailed justification for using cointegration in modeling financial time series. Alexander et al. (2002) used cointegration to construct an index-tracking portfolio. They showed that the optimal index-tracking portfolio has stationary tracking errors, and that efficient long-short hedge strategies can be achieved with relatively few stocks and less turnover. Alexander (2001) discussed in detail relevant published work in finance that used different properties of cointegration for portfolio optimization or the construction of trading strategies. In summary, published work on the use

483

of cointegration in portfolio construction—specifically on constructing time-dynamic trading strategies—is limited mainly to those aforementioned.

To better understand the relationship between cointegration and asset allocation, we need to first describe two methods for asset allocation. The process of selecting a *target asset allocation* or *index* is called *strategic asset allocation*. The variation from the target is called *tactical asset allocation*. One can think of strategic asset allocation as a process of selecting the appropriate benchmark for a portfolio. For example, pension plans regularly go through such a benchmark-selection process in order to establish their investment policies. Tactical asset allocation is usually identified with active portfolio management. For example, how one should maintain the allocation of sixty percent stocks and forty percent bonds over time comes into question.

Available research investigating the relationship between cointegration and asset allocation shows that cointegration affects strategic asset allocation. In other words, the decision about optimal portfolio mix (or setting the appropriate benchmark) is influenced by the common stochastic trend between assets. Lucas (1997) presented a model where a portfolio manager maximized expected utility of total earnings over a finite time period. The associated time-series model captured cointegrating relations among the assets. He showed that cointegration affects strategic asset allocation while error-correction mainly affects tactical asset allocation.

Using cointegration for portfolio allocation implies the existence of a long-term stochastic trend. In general, this contradicts one of the commonly-used hypotheses in finance, i.e., that stock-price returns follow a random walk. Lo and MacKinley (1988) tested the random-walk hypothesis with weekly stock market returns, by comparing variance estimators derived from data sampled at different frequencies. If the stock returns followed a random walk, then the variance should have grown with the square root of time. They rejected the random-walk hypothesis and showed that the rejection was due largely to the behavior of small stocks. Additionally, they showed that the autocorrelations of individual securities were generally negative, and the autocorrelations of equally and positively weighted CRSP (Center for Research in Security Prices) indices were positive.

In this paper we create a special input for the vector autoregressive-to-anything (VARTA) model of Biller and Nelson (2003), in order to simulate sample paths that preserve the cointegration property of financial time series. This new tool can be used to construct optimal portfolios using a variety of utility functions as well as to design trading strategies for index tracking. The focus of this paper is on simulating cointegrated time series via VARTA, and such applications will be addressed in future research.

2 COINTEGRATED FINANCIAL TIME SERIES

Portfolio allocation models, like the mean-variance model introduced by Markowitz (1952), and models that use value-at-risk introduced by Jorion (2001), include a measure of the risk associated with the set of assets being considered. The standard measure is based on the corresponding covariance (or equivalently, the correlation) matrix. The covariance between two random vectors X and Y is given by

$$Cov(X,Y) = E\left\{(X - E[X])(Y - E[Y])^T\right\},$$

where $E[X]$ stands for the expected value of the random vector X, and T denotes the transpose operator. It is well documented that the covariance (or correlation) is a measure of short-term linear dependencies; see, e.g., Theorem 4.5.7 in Casella and Berger (2002).

In contrast to correlation, cointegration is a measure of long-term dependencies, see Engle and Granger (1991). We will briefly outline the notion of cointegration. In order to do so, we need to first define stationary and integrated time series.

A stochastic process Y_t is stationary if its first and second moments are time invariant: in particular if $E[Y_t] = \mu$, $\forall t$ and $E[(Y_t - \mu)(Y_{t-h} - \mu)^T] = \Gamma_Y(h) = \Gamma_Y(-h)^T$, $\forall t, h = 0, 1, 2, \ldots$, where μ is a vector of finite mean terms, and $\Gamma_Y(h)$ is a matrix of finite covariances. Such a process is known as *integrated of order 0* and denoted by $I(0)$. A univariate process is called *integrated of order d*, $I(d)$, if in its original form it is non-stationary but becomes stationary after differencing d times. If all elements of the vector X_t are $I(1)$, and there exists a vector b such that $b^T X_t$ is $I(0)$, then the vector process X_t is said to be cointegrated, and b is called the cointegrating vector. For example, two univariate time series X and Y are cointegrated if X, Y are $I(1)$, and there exists a scalar b such that $Z = X - bY$ is $I(0)$.

Assume that we have $N \geq 2$ cointegrated financial assets, and their log-prices are $I(1)$ (integrated of order 1) processes. It is widely assumed that stock returns are integrated of order 0, whereas stock prices are integrated of order 1 (Alexander et al. 2002).

Denote the vector of the asset prices at time t by $P_t = (P_t^1, \ldots, P_t^N)$. $P_t^i = P_0^i e^{\sum_{j=0}^t r_j^i}$, $i = 1, \ldots, N$, where $r = (r_t^1, \ldots, r_t^N)$ denotes the continuously compounded asset returns, and $P_0^1, \ldots, P_0^N$ are the initial prices. Without loss of generality, we assume that $P_0^1 = \cdots = P_0^N = 1$ so that the log-prices can be written as $\ln P_t^i = \sum_{j=0}^t r_j^i, i = 1, \ldots, N$. Denote the corresponding cointegrating vector $b = (b^1, \ldots, b^N)$. By the definition of cointegration, the resulting time series $Y_t = \sum_{i=1}^N b^i \ln P_t^i$ is stationary and integrated of order 0. Galenko (2008) proved the following results.

Proposition 1. *Assume that the log prices of $N \geq 2$ assets, $\ln P^i, i = 1, \ldots, N$, are cointegrated with cointegrating vector b. Let $Y_t = \sum_{i=1}^N b^i \ln P_t^i$ be the corresponding stationary series, and $(r_t^1, \ldots, r_t^N)$ be the continuously compounded asset returns at time $t > 0$. Define $Z_t = Y_t - Y_{t-1} = \sum_{i=1}^N b^i r_t^i$. If $\lim_{p \to \infty} Cov[Y_t, Y_{t-p}] = 0$, then $\sum_{p=1}^\infty p Cov[Z_t, Z_{t-p}] = -Var Y_t$, where $Cov[Z_t, Z_{t-p}] = \sum_{i=1}^N \sum_{j=1}^N b^i b^j Cov[r_t^i, r_{t-p}^j]$.*

Proposition 2. *Assume $\ln P_t^i, i = 1, \ldots, N$, are the log-prices of N assets, and $(r_t^1, \ldots, r_t^N)$ denotes the continuously compounded asset returns at time $t > 0$. For some finite vector b, the process $Y_t = \sum_{i=1}^N b^i \ln P_t^i$, given $\lim_{p \to \infty} Cov[Y_t, Y_{t-p}] = 0$, is stationary, and therefore the time series of the assets' log-prices are cointegrated, if and only if the process $Z_t = \sum_{i=1}^N b^i r_t^i$ has the following three properties:*

(i) $EZ_t = 0$,
(ii) $Var Z_t = -2 \sum_{p=1}^\infty Cov[Z_t, Z_{t-p}]$, *and*
(iii) $\sum_{p=1}^\infty p Cov[Z_t, Z_{t-p}] < \infty$.

In what follows, we use the results of Propositions 1 and 2 to create special input for VARTA in order to form simulated sample paths that preserve the cointegrated property of the underlying time series. We first give a brief introduction to the VARTA algorithm.

3   VECTOR AUTOREGRESSIVE TO ANYTHING (VARTA)

VARTA simulates stationary multivariate time-series processes with marginal distributions from the Johnson translation system (Johnson 1949) and a specified autocorrelation structure. Such a multivariate process can be simulated by generating a base Gaussian vector autoregressive process whose autocorrelation structure is adjusted in such a way that after inverse transformation of the base process we obtain a time series with the specified Johnson marginal distribution and the desired autocorrelation structure.

In general, the Johnson translation system for a random variable Z is defined by a cumulative distribution function that has the form

$$F_Z(z) = \Phi(\gamma + \sigma f[(z - \xi)/\lambda]), \tag{1}$$

where γ and σ are shape parameters, ξ is a location parameter, λ is a scale parameter, Φ is the normal distribution function, and $f(\cdot)$ is one of the following transformations:

$$f(y) = \begin{cases} \log(y) & \text{for the lognormal family,} \\ \log(y + \sqrt{y^2 + 1}) & \text{for the unbounded family,} \\ \log(\frac{y}{1-y}) & \text{for the bounded family,} \\ y & \text{for the normal family.} \end{cases}$$

In this paper we use the unbounded Johnson system family to model financial asset returns.

Below is the VARTA algorithm applied to our problem:

- step 1: Estimate the parameters of the Johnson distribution for each component of the original process $\ln P_t$ and the specified autocorrelation lag p.
- step 2: Estimate the parameters of the base Gaussian process $Z_t = \sum_{j=1}^p \alpha_j Z_{t-j} + u_t$, where Z_t is the vector of observations at time t ($\ln P_t$ in our case), $\alpha_j, j = 1, \ldots, p$ are the autoregressive coefficients, and u_t is white noise.
- step 3: Simulate Z_t.

- step 4: Transform the generated base process into the original process with specified marginals by taking the inverse Johnson cumulative distribution function of the simulated values.

In step 1, we estimate the Johnson system parameters, γ, σ, ξ and λ, via weighted least squares (Swain et al. 1988), although there are several estimation methods that one can use; see Hill et al. (1976), Storer et al. (1988) and Slifker and Shapiro (1980). Step 2 of the VARTA algorithm is equivalent to fitting a normal copula. The parameters of the estimated copula form a system of linear equations. Its solution consists of the parameters of the base process Z_t, for details see Biller and Nelson (2003). Biller and Nelson (2005) show that combining step 1 and step 2 produce better estimators for both the original and base processes. Steps 3 and 4 are straightforward.

4 SIMULATION OF COINTEGRATED TIME SERIES

Proposition 2 says that cointegration, under certain assumptions, can be mapped to a function of the first two moments of the underlying time series. We then take the view that in simulating cointegrated time series one should employ an algorithm that exploits this special property. Proposition 2 also shows that the cointegration effect spreads over an infinite number of lags. However, VARTA can simulate time series preserving the given expected values and autocovariances up to a finite lag. Hence, in order to make time series generated by VARTA cointegrated with cointegration vector b, we need to assure that conditions (i) and (ii) of Proposition 2 are satisfied, and after prespecified lag K the simulated time series have zero autocovariances. Based on this idea we formulate the following optimization model to adjust autocovariances. The model finds an autocovariance matrix satisfying the conditions described above. Before formulating the model we summarize, and present additional, notation.

Model Parameters:

- N - number of time series (assets)
- P - number of autocovariance lags in the underlying Gaussian base process
- $K < P$ - number of autocovariance lags that can be changed
- b - cointegration vector of dimension N
- $\hat{\Sigma}_p, p = 1, \ldots, P$ - estimated autocovariance matrix of lag p and dimension $N \times N$
- $\hat{\Sigma}$ - estimated "big" autocovariance matrix of dimension $NP \times NP$ and of a form

$$\hat{\Sigma} = \begin{bmatrix} \hat{\Sigma}_0 & \hat{\Sigma}_1 & \hat{\Sigma}_2 & \hat{\Sigma}_3 & \ldots & \hat{\Sigma}_{p-1} & \hat{\Sigma}_p \\ \hat{\Sigma}_1 & \hat{\Sigma}_0 & \hat{\Sigma}_1 & \hat{\Sigma}_2 & \ldots & \hat{\Sigma}_{p-2} & \hat{\Sigma}_{p-1} \\ \hat{\Sigma}_2 & \hat{\Sigma}_1 & \hat{\Sigma}_0 & \hat{\Sigma}_1 & \ldots & \hat{\Sigma}_{p-3} & \hat{\Sigma}_{p-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \hat{\Sigma}_{p-1} & \hat{\Sigma}_{p-2} & \ldots & \hat{\Sigma}_2 & \hat{\Sigma}_1 & \hat{\Sigma}_0 & \hat{\Sigma}_1 \\ \hat{\Sigma}_p & \hat{\Sigma}_{p-1} & \ldots & \hat{\Sigma}_3 & \hat{\Sigma}_2 & \hat{\Sigma}_1 & \hat{\Sigma}_0 \end{bmatrix}$$

- χ - class of matrices that have the same block structure as $\hat{\Sigma}$

Decision Variables:

- X - $NP \times NP$ matrix of decision variables associated with $\hat{\Sigma}$
- X_p, $p = 1, \ldots, P$ - $N \times N$ submatrix of X

We want to find a matrix $X = \tilde{\Sigma}$ with the following properties:

- $\tilde{\Sigma}$ is positive semidefinite (PSD)
- Blocks $\tilde{\Sigma}_0, \tilde{\Sigma}_1, \ldots, \tilde{\Sigma}_p$ of $\tilde{\Sigma}$ satisfy $b^T[\tilde{\Sigma}_0 + 2\sum_{i=1}^{p}\tilde{\Sigma}_i]b = 0$ - condition (i) from Proposition 2
- After a given lag K, $\tilde{\Sigma}_p = 0$ for $\forall p > K$ - artificial way to force the autocovariances after lag p to be 0

$$\min_{X} \quad \| X - \hat{\Sigma} \|^2 \tag{2a}$$

$$\text{s.t.} \quad X \in \chi \tag{2b}$$

$$b^T[X_0 + 2\sum_{p=1}^{P} X_p]b = 0 \tag{2c}$$

$$X_p = 0, \quad p = K+1, \ldots, P \tag{2d}$$

$$X \succeq 0. \tag{2e}$$

The objective function (2a) minimizes the distance, in the sense of the two-norm, between the selected autocovariance matrix, $X = \tilde{\Sigma}$, and the estimated covariance matrix, $\hat{\Sigma}$. Constraints (2b)-(2d) can all be captured via finite linear constraints, but constraint (2e) cannot, as it requires X to be PSD. Constraint (2e) may be expressed as an infinite number of linear constraints, i.e., $q^T X q \geq 0, \forall q$. As a result, model (2) is a semidefinite program (SDP). We note that an SDP naturally arises for different reasons (Ghosh and Henderson 2002) in the context of the normal-to-anything (NORTA) procedure of Cario and Nelson (1997).

But for constraint (2e), model (2) is a standard linearly-constrained least-squares optimization model. Specialized algorithms have been developed for the SDP that arises with the addition of (2e); see, for example, Boyd and Xiao (2005) and Sivaramakrishnan (2002). We now show how constraint (2e) can be handled by solving a sequence of relaxations of model (2) in which constraint (2e) is enforced at a finite set of q vectors. We call these linear-programming relaxations *master problems*.

- step 0: Let the initial master problem be model (2) without constraint (2e).
- step 1: Solve the master problem and obtain candidate solution $\hat{X}$.
- step 2: Find the minimum eigenvalue $\lambda_{\min}$, and the corresponding eigenvector $q_{\min}$, of $\hat{X}$. If $\lambda_{\min} \geq 0$ then stop and output $\hat{X}$.
- step 3: Add $q_{\min}^T X q_{\min} \geq 0$ to the master problem. Go to step 1.

The solution $\hat{X}$ to the master problem is symmetric by construction and hence its eigenvalues are real. Step 2 allows us to check whether $\hat{X}$ violates constraint (2e). If $\lambda_{\min}$ is negative, $\hat{X}$ is not PSD, and we have $q_{\min}^T \hat{X} q_{\min} < 0$. So, we require future solutions to the master problem satisfy $q_{\min}^T X q_{\min} \geq 0$.

The resulting algorithm for simulation of cointegrated time series is as follows:

- Estimate the autocovariances up to lag P, the parameters of Johnson distribution and the cointegration vector b.
- Using the estimated autocovariances and the cointegration vector solve the fitting model (2) to obtain the adjusted autocovariances.
- Run VARTA using the estimated parameters of the Johnson distribution and the adjusted autocovariances as input.

5 SIMULATION RESULTS UNDER KNOWN COINTEGRATION VECTOR

Here we execute the described strategy for a bivariate time series for which we can compute the exact cointegration vector. Consider the following bivariate model describing the log-returns of two assets r_t^1 and r_t^2:

$$
\begin{aligned}
r_t^1 &= r_t^2 + c + (\gamma - 1)Y_{t-1} + \xi_t^1 \\
r_t^2 &= \mu + \xi_t^2 \\
Y_t &= c + \gamma Y_{t-1} + \xi_t^1,
\end{aligned}
$$

where

- μ - real-valued constant
- (ξ_t^1, ξ_t^2) - bivariate normal random variable with mean vector of 0, variances σ_1^2, σ_2^2, and covariance ρ
- γ - parameter of AR(1) process Y_t, a constant between 0 and 1

- c - parameter of AR(1) process Y_t, a constant.

Give their structure the log-prices of these two assets are cointegrated of order one with cointegration vector $(1, -1)$. We can now compute the moments:

$$
\begin{aligned}
E[r_t^1] &= \mu \\
E[r_t^2] &= \mu \\
\text{Var}[r_t^1] &= \sigma_2^2 + 2\rho + 2\sigma_1^2/(1+\gamma) \\
\text{Var}[r_t^2] &= \sigma_2^2 \\
\text{Cov}[r_t^1, r_t^2] &= \sigma_2^2 + \rho \\
\text{Cov}[r_t^1, r_{t-p}^1] &= -(1-\gamma)\sigma_1^2\gamma^{p-1}/(1+\gamma) + (\gamma-1)\rho\gamma^{p-1}, \quad p = 1, 2, \ldots \\
\text{Cov}[r_t^1, r_{t-p}^2] &= (\gamma-1)\rho\gamma^{p-1}, \quad p = 1, 2, \ldots \\
\text{Cov}[r_t^2, r_{t-p}^1] &= 0, \quad p = 1, 2, \ldots \\
\text{Cov}[r_t^2, r_{t-p}^2] &= 0, \quad p = 1, 2, \ldots
\end{aligned}
$$

To run our procedure we use the following specific values of these parameters: $\mu = 0$, $\rho = 0$, $\gamma = 0.75$ and $\sigma_1^2 = \sigma_2^2 = 0.25$. For $p = 30$ and $k = 7$, the population autocovariance matrices are:

$$
\Sigma_0 = \begin{bmatrix} 0.5357 & 0.25 \\ 0.25 & 0.25 \end{bmatrix}, \quad
\Sigma_1 = \begin{bmatrix} -0.0357 & 0 \\ 0 & 0 \end{bmatrix},
$$

$$
\Sigma_2 = \begin{bmatrix} -0.0268 & 0 \\ 0 & 0 \end{bmatrix}, \quad
\Sigma_3 = \begin{bmatrix} -0.0201 & 0 \\ 0 & 0 \end{bmatrix}
$$

$$
\Sigma_4 = \begin{bmatrix} -0.0151 & 0 \\ 0 & 0 \end{bmatrix}, \quad
\Sigma_5 = \begin{bmatrix} -0.0113 & 0 \\ 0 & 0 \end{bmatrix},
$$

$$
\Sigma_6 = \begin{bmatrix} -0.0085 & 0 \\ 0 & 0 \end{bmatrix}, \quad
\Sigma_7 = \begin{bmatrix} -0.0064 & 0 \\ 0 & 0 \end{bmatrix}
$$

$$
\Sigma_p = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad p = 8, \ldots, 30.
$$

The algorithm was coded in C++, the optimization was carried out by CPLEX, and the eigenvalue calculations by LAPACK. The resulting adjusted autocovariance matrix after solving the adjusting model (2) is:

$$
\tilde{\Sigma}_0 = \begin{bmatrix} 0.5357 & 0.2506 \\ 0.2506 & 0.25 \end{bmatrix}, \quad
\tilde{\Sigma}_1 = \begin{bmatrix} -0.0363 & 0.0006 \\ 0.0006 & -0.0006 \end{bmatrix},
$$

$$
\tilde{\Sigma}_2 = \begin{bmatrix} -0.0274 & 0.0006 \\ 0.0006 & -0.0006 \end{bmatrix}, \quad
\tilde{\Sigma}_3 = \begin{bmatrix} -0.0207 & 0.0006 \\ 0.0006 & -0.0006 \end{bmatrix},
$$

$$
\tilde{\Sigma}_4 = \begin{bmatrix} -0.0157 & 0.0007 \\ 0.0007 & -0.0007 \end{bmatrix}, \quad
\tilde{\Sigma}_5 = \begin{bmatrix} -0.012 & 0.0007 \\ 0.0007 & -0.0007 \end{bmatrix},
$$

$$
\tilde{\Sigma}_6 = \begin{bmatrix} -0.00919 & 0.0007 \\ 0.0007 & -0.0007 \end{bmatrix}, \quad
\tilde{\Sigma}_7 = \begin{bmatrix} -0.0071 & 0.0007 \\ 0.0007 & -0.0007 \end{bmatrix},
$$

$$\tilde{\Sigma}_p = \left[\begin{array}{cc} 0 & 0 \\ 0 & 0 \end{array} \right], \ \ p = 8, \ldots, 30.$$

We run VARTA first using the true autocovariance and second using the adjusted autocovariance. The length of the generated time series is 6000. We performed the Johansen cointegration rank test using SAS for both simulations. The cointegration test for time series when we used true autocovariance up to lag 7 failed. For the second case the same test indicated that the time series generated by VARTA using adjusted autocovariance are cointegrated with probability 0.99 with cointegration vector (1,-1.06). The cointegration rank test results for the adjusted and true covariance matrices are in Tables 1 and 2.

Table 1: Johansen cointegration rank test results for simulated time series using true autocovariance

Hypothesized order of cointegration	Eigenvalue	Trace Stat	99%CV	95% CV	Max Stat	99% CV	95% CV
H(0) (no cointegration)	0.0015	8.7263	12.53	16.31	8.7254	11.44	15.69
H(1) (cointegration of order 1)	0	0.0009	3.84	6.51	0.0009	3.84	6.51

Table 2: Johansen cointegration rank test results for simulated time series using adjusted autocovariance

Hypothesized order of cointegration	Eigenvalue	Trace Stat	99%CV	95% CV	Max Stat	99% CV	95% CV
H(0) (no cointegration)	0.0089	53.6002	12.53	16.31	53.5957	11.44	15.69
H(1) (cointegration of order 1)	0	0.0045	3.84	6.51	0.0045	3.84	6.51

The first row in both tables corresponds to a null hypothesis of no cointegration whereas the second row corresponds to testing cointegration of order 1. There are two test statistics: Trace and Maximum. To decide whether to reject the corresponding null hypothesis one needs to compare either test statistic to its critical value (coded as CV). We present 99% and 95% critical values. If the test statistic is less than the critical value then we cannot reject the null hypothesis, if it is greater then we reject it. Row 1 of Table 1 corresponds to the first situation since both the trace and maximum test statistics are less than their critical values. Table 2 shows the second scenario, where we reject the no-cointegration null hypothesis and conclude that the series are cointegrated of order 1. Figures 1 and 2 show the simulated time series using the original and adjusted covariances.
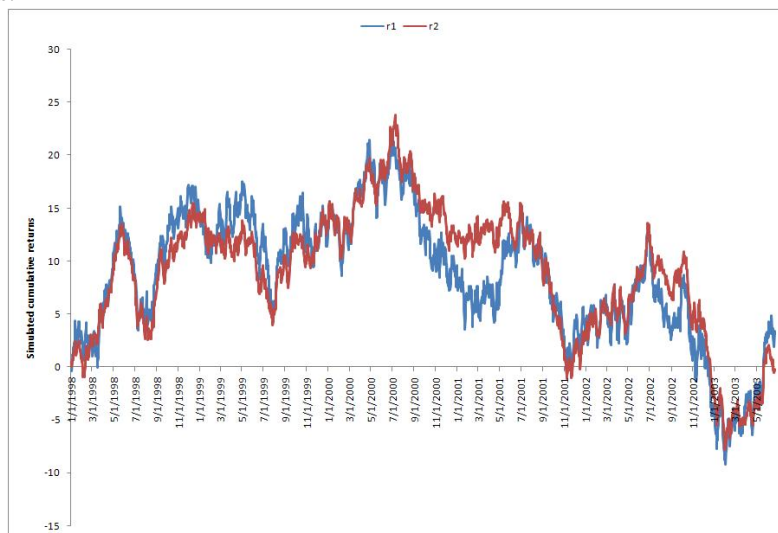

Figure 1: Constructed example: simulated time series using VARTA with true autocovariance with $p = 30$ and $k = 7$
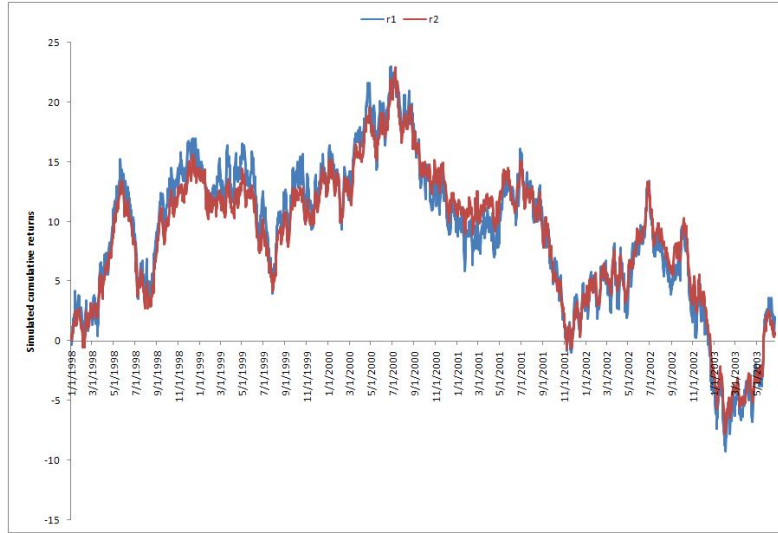
Figure 2: Constructed example: simulated time series using VARTA with adjusted autocovariance with $p = 30$ and $k = 7$

6   SIMULATION RESULTS FOR REAL FINANCIAL TIME SERIES

We used historical data for four equity world indices: AEX (The best-known index of Euronext Amsterdam, the AEX index is made up of the 25 most active securities in the Netherlands. This index provides a fair representation of the Dutch economy.), DAX (Deutsche Aktien Xchange 30 is a blue chip stock market index consisting of the 30 major German companies trading on the Frankfurt Stock Exchange.), CAC (The CAC 40, which takes its name from Paris Bourse's early automation system Cotation Assistée en Continu (Continuous Assisted Quotation), is a French stock market index. The index represents a capitalization-weighted measure of the 40 most significant values among the 100 highest market caps on the Paris Bourse.), and FTSE (The FTSE 100 Index is a share index of the 100 most highly capitalized companies listed on the London Stock Exchange.) These time series are cointegrated (based on the Johansen cointegration test with 99% significance) and the estimated cointegration vector can be found in Table 3.

Table 3: Real Example: estimated cointegration vector

	AEX	CAC	DAX	FTSE
Cointegration vector	3.685894	-4.65918	13.56943	-22.3922
Normalized coint. vector	1	-1.26406	3.681448	-6.0751

We ran the VARTA algorithm with $p = 15$ and $k = 8$. Increasing p much beyond this led to instability in VARTA. We simulated 6000 data points.

Table 4: Real example: Johansen cointegration rank test results for simulated time series using nonadjusted autocovariances

	Eigenvalue	Trace Stat	95% CV	99%CV	Max Stat	95% CV	99% CV
H(0)(no cointegration)	0.0034	40.7659	47.21	54.46	20.3613	27.07	32.24
H(1) (cointegration of order 1)	0.0026	20.4046	29.68	35.65	15.5304	20.97	25.52
H(2) (cointegration of order 2)	0.0008	4.8743	15.41	20.04	4.802	14.07	18.63
H(3) (cointegration of order 3)	0	0.0723	3.76	6.65	0.0723	3.76	6.65

Cointegration tests for adjusted and non-adjusted cases are presented in Tables 4 and 5. From Table 4 both trace and maximum test statistics are less than the corresponding critical values and therefore the *raw* time series are not cointegrated. However the adjusted ones are cointegrated of order 1 as shown in the first row of Table 5. Figures 3 and 4 show the simulated time series using the original and adjusted autocovariances.
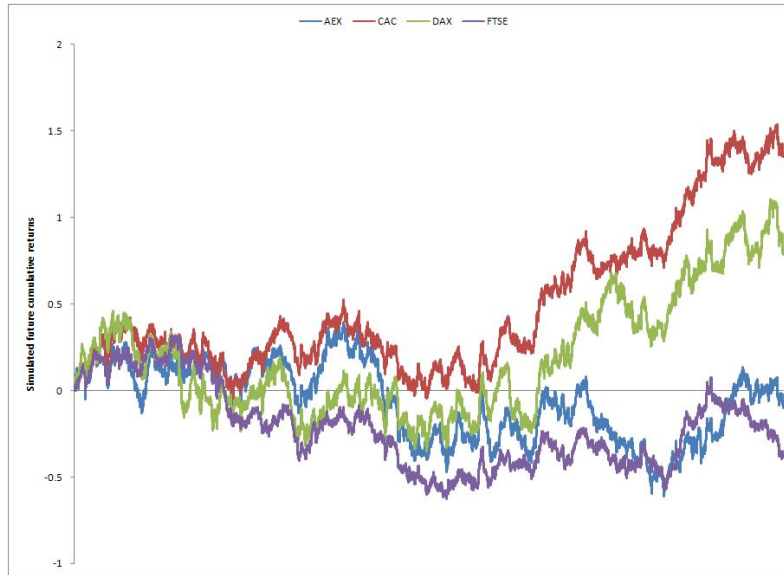
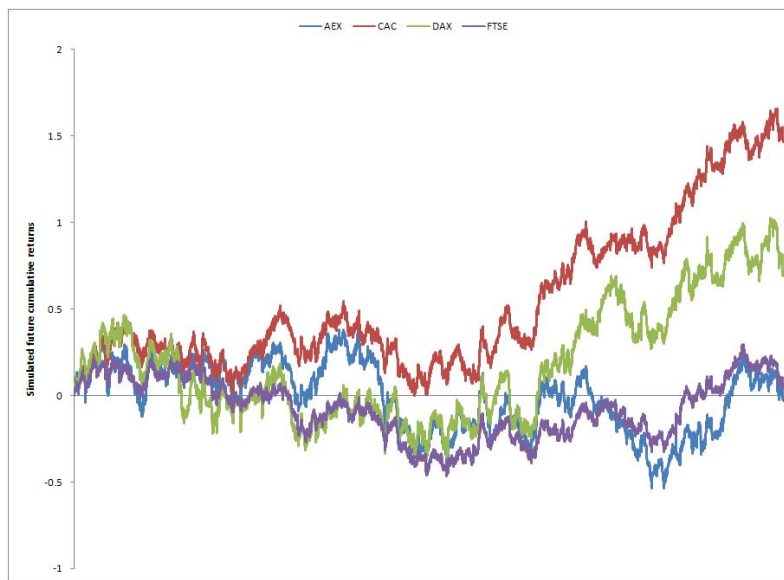Figure 3: Real example: simulated time series using VARTA with true autocovariance with p=15 and k=8

Figure 4: Real example: simulated time series using VARTA with adjusted autocovariance with p=15 and k=8

Table 5: Real example: Johansen cointegration rank test results for simulated time series using adjusted autocovariances

	Eigenvalue	Trace Stat	95% CV	99%CV	Max Stat	95% CV	99% CV
H(0) (no cointegration)	0.0061	55.26	47.21	54.46	36.4144	27.07	32.24
H(1) (cointegration of order 1)	0.0021	18.8456	29.68	35.65	12.5605	20.97	25.52
H(2) (cointegration of order 2)	0.001	6.2851	15.41	20.04	6.2816	14.07	18.63
H(3) (cointegration of order 3)	0	0.0035	3.76	6.65	0.0035	3.76	6.65

7   CONCLUSION

In this paper we have presented new theoretical results for cointegrated time series that allow us to simulate future cointegrated sample paths using VARTA. Our approach creates specialized input for VARTA in terms of optimally truncated autocovariance matrices that are obtained via solving a semidefinite optimization model. In future work, we will use the resulting sample paths in a stochastic portfolio optimization problem.

ACKNOWLEDGEMENTS

REFERENCES

Alexander, C. 2001. *Market models: A guide to financial data analysis*. John Wiley and Sons, Chichester.

Alexander, C., I. Giblin, and W. Weddington. 2002. Cointegration and asset allocation: A new active hedge fund strategy. ISMA Centre Discussion Papers in Finance Series.

Biller, B., and B. L. Nelson. 2003. Modeling and generating multivariate time-series input processes using a vector autoregressive technique. *ACM Transactions on Modeling and Computer Simulation (TOMACS)* 13 (3): 211–237.

Biller, B., and B. L. Nelson. 2005. Fitting time series input processes for simulation. *Operations Research* 53:549–559.

Boyd, S., and L. Xiao. 2005. Least-squares covariance matrix adjustment. *SIAM Journal on Matrix Analysis and Applications* 27 (2): 532–546.

Cario, M. C., and B. L. Nelson. 1997. Modeling and generating random vectors with arbitrary marginal distributions and correlation matrix. Technical report, Department of Industrial Engineering and Management Sciences, Northwestern University.

Casella, G., and R. L. Berger. 2002. *Statistical inference*. Duxbury Press, Belmont, CA.

Engle, R. F., and C. W. J. Granger. 1987. Portfolio selection. *Econometrica* 55 (2): 251–276.

Engle, R. F., and C. W. J. Granger. 1991. *Long-run economic relationships, readings in cointegration*. Oxford University Press, Oxford.

Galenko, A. 2008. *Modeling stochastic dependencies and their impact on optimal decision making*. The University of Texas at Austin. Ph.D. Dissertation.

Ghosh, S., and S. G. Henderson. 2002. Chessboard distributions and random vectors with specified marginals and covariance matrix. *Operations Research* 50 (5): 820–834.

Hill, I. D., R. Hill, and R. L. Holder. 1976. Fitting Johnson curves by moments. *Applied Statistics* 25 (2): 180–189.

Johnson, N. L. 1949. Systems of frequency curves generated by methods of translation. *Biometrika* 36:149–176.

Jorion, P. 2001. *Value at risk: The new benchmark for managing financial risk*. Second ed. McGraw-Hill, New York.

Lo, A. W., and A. C. MacKinley. 1988. Stock market prices do not follow random walks: evidence from a simple specification test. *Review of Financial Studies* 1:41–66.

Lucas, A. 1997. Strategic and tactical asset allocation and the effect of long-run equilibrium relations. Serie Research Memoranda 0042, Free University Amsterdam, Faculty of Economics, Business Administration and Econometrics, De Boelelaan 1105, 1081 HV Amsterdam.

Markowitz, H. M. 1952. Portfolio selection. *Journal of Finance* 7:77–91.

Popova, I., D. P. Morton, E. Popova, and J. Yau. 2007. Optimizing benchmark-based portfolios with hedge funds. *Journal of Alternative Investments* 10 (1): 35–55.

Sivaramakrishnan, K. K. 2002. *Linear programming approaches to semidefinite programming problems*. Rensselaer Polytechnic Institute. Ph.D. Dissertation.

Slifker, J. F., and S. S. Shapiro. 1980. The Johnson system: Selection and parameter estimation. *Technometrics* 22 (2): 239–246.

Storer, R. H., J. J. Swain, S. Venkatraman, and J. R. Wilson. 1988. Comparison of methods for fitting data using Johnson translation distributions. In *Proceedings of the 1988 Winter Simulation Conference*, ed. M. Abrams, P. Hughes, and J. Comfort, 476–481. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Swain, J. J., S. Venkatraman, and J. R. Wilson. 1988. Least-squares estimation of distribution functions in Johnson's translation system. *Journal of Statistical Computation and Simulation* 29:271–297.

Tsay, R. S. 2005. *Analysis of financial time series*. Second ed. John Wiley and Sons, New York.

AUTHOR BIOGRAPHIES

ALEXANDER GALENKO graduated from the Advanced Physical-Mathematical School, Novosibirsk, Russia in 1999. He received BS degree in Mathematical Economics from the Novosibirsk State University in June, 2003, and MS and PhD in Operations Research and Industrial Engineering from The University of Texas at Austin in 2008. Currently he works as Quantitative Research Analyst at Penson Financial Services. His email address for these proceedings is <agalenko@gmail.com>.

DAVID P. MORTON is Engineering Foundation Professor in the Graduate Program in Operations Research in the Mechanical Engineering Department at The University of Texas at Austin. His research interests include computational stochastic programming, including simulation-based approximations in stochastic programming. His email address is <morton@mail.utexas.edu>, and his web page is <http://www.me.utexas.edu/ orie/Morton.html>.

ELMIRA POPOVA is an Associate Professor in the Graduate Program in Operations Research in the Mechanical Engineering Department at The University of Texas at Austin. Her research interests are in uncertainty modeling of complex systems, advanced Bayesian analysis, Markov Chain Monte Carlo methods, and stochastic optimization. Her e-mail address is <elmira@mail.utexas.edu>.

IVILINA POPOVA is an Associate Professor of Finance at Texas State University. She taught at Purdue University and Seattle University prior to joining McCoy College of Business. Before that she worked for Deutsche Asset Management in New York and Koch Capital Markets in Houston. Her email address is <ip12@txstate.edu>.