

Bootstrap Methods for Finance: Review and Analysis*

Philippe Cogneau[†] and Valeri Zakamouline[‡]

This revision: May 27, 2010

Abstract

In finance one often needs to estimate the risk and reward of an asset over a long-run given a sample of observations over a short-run. Two common obstacles in these estimations are a lack of sufficient data and the uncertainty in the nature of the data generating process. To overcome the first obstacle the researches rely on statistical bootstrap methods. A practical realization of a bootstrap method depends crucially on whether the data are assumed to be serially dependent or not. In this paper we review some popular bootstrap methods and argue that the application of these methods is not well understood. This especially concerns the application of a moving block bootstrap method, which is used if there is a serial dependency in data. Namely, the estimates provided by a moving block bootstrap method are generally biased. We demonstrate the estimation bias and propose a method of bias adjustment. Moreover, in this paper we also analyze the precisions of estimations provided by bootstrap methods. We show that the precision of estimation provided by the moving block bootstrap methods is rather poor, which means that one should be aware that the estimation risk is a big issue.

Key words: time-series data, parameter estimation, bootstrap, block bootstrap.

JEL classification: C13, C14, C15, G11.

*This is the first draft. All comments are very welcome!

[†]University of Liege, HEC Management School, Rue Louvrex 14, Bat. N1, B-4000 Liege, Belgium, philippe.cogneau@student.ulg.ac.be

[‡]Corresponding author. University of Agder, Faculty of Economics, Service Box 422, 4604 Kristiansand, Norway, Tel.: (+47) 38 14 10 39, Valeri.Zakamouline@uia.no

1 Introduction

An important issue in finance is the relationship between the optimal portfolio asset allocation and the length of the investment horizon. Despite the abundance of academic models that describe the investor's preferences, in theory the construction of the optimal investor's portfolio is pretty simple: make up a portfolio that either maximizes the investor's expected utility or the investor's performance measure over a specific investment horizon. Yet the practical implementation of a theoretical model requires the estimation of the model's input parameters. This requires, among other things, estimating risk and reward of each asset in the portfolio. Financial advisors and mutual fund sales literature commonly recommend that the investment horizon should be relatively long in order to benefit from the "time diversification". In finance, the problem of estimating risk and reward over a relatively long horizon comes from a lack of sufficient data and the uncertainty in the nature of the return generating process.

In particular, the majority of historical databases used in financial research have a span of about 80 years of monthly observations. The estimation of risk and reward of an asset on horizon of, say, 5 years using only 16 available non-overlapping 5-year return observations is subject to big errors, to say nothing of errors of estimations on horizons of 15-20 years. To tackle this problem the researches rely on statistical bootstrap methods. A bootstrap is a computer-intensive method of estimation of parameters and distributions by resampling the original data. Yet the practical realization of a bootstrap method depends crucially on whether the returns are assumed to be serially dependent or not. If the returns are assumed to be independent, then one employs the standard bootstrap method, suggested by Efron (1979), which is pretty straightforward in implementation.

However, most of the practitioners believe that the stock returns are mean-reverting. This means that the stocks become less risky as the investment horizon lengthens. There is a large strand of academic literature which supports the practitioners' point of view. Among other things, this literature documents long-term reversal in stock prices. If this is really so, then the standard bootstrap method cannot be used, because this method destroys any serial dependency in data. In order to capture the dependence structure in the return series while estimating the parameter of interest of the return distribution on long horizons, many researches rely on the moving block bootstrap methods that were introduced by Hall (1985) and

Künch (1989). These methods use overlapping blocks of data instead of individual observations to estimate the parameters and distributions.

In this paper we review some popular bootstrap methods and argue that the application of the moving block bootstrap methods to the financial time-series data and problems is not well understood. First of all, the statistical bootstrap methods were not supposed, in the first place, to estimate the long-run parameters of processes sampled over short-run. Second, the estimates provided by a moving block bootstrap method are generally biased and rather imprecise. In particular, it is well known that the usage of overlapping blocks leads to a potentially serious estimation bias (often called “small sample bias”), see, among others, Orcutt and Irwin (1948), Marriott and Pope (1954), Müller (1993), Hall, Horowitz, and Jing (1995), Härdle, Horowitz, and Kreiss (2003), and Sun, Nelken, Han, and Guo (2009). Therefore, the estimates obtained by a moving block bootstrap method must be adjusted for the bias. Such a bias adjustment requires estimating the bias itself. A more elaborate bias adjustment method relies on choosing a proper block length in a moving block bootstrap method (see, for example, Hall et al. (1995)). Moreover, it is also known that the precision of estimation (as measured by, for example, the standard error of estimation) of parameters of serially dependent data provided by a moving block bootstrap method is much poorer as compared to the precision of estimation of parameters of independent data provided by the standard bootstrap method (see, for example, Härdle et al. (2003)).

Taking into account the growth popularity of moving block bootstrap methods in financial applications, this paper aims to demonstrate the estimation bias and proposes a method of bias adjustment based on choosing the optimal block length which minimizes the bias. Moreover, in this paper we also analyze the precisions of estimations provided by bootstrap methods. All this is done by using a widely accepted flexible parametric model for the stock price process. In this model the stock log return is composed of two components: a permanent component which is a random walk, and a transitory component which is a strictly stationary process. Such a mixture of two components is able to generate short-term momentum and long-term reversal in stock prices documented by numerous empirical studies.

The paper’s organization is as follows. In Section 2 we review the statistical bootstrap methods, both for independent and serially dependent data. In Section 3 we review the financial problem which is supposed to be solved using a bootstrap method. Section 4 describes the

implementation of the standard bootstrap method and points on a possible pitfall in implementation and on convergence to normality as the length of horizon increases. Section 5 describes a possible implementation of a moving block bootstrap method. However, the main goal of this section is to present the model for the stock price process and, based on some feasible parametric assumptions in this model, to demonstrate the estimation bias and how this bias can be mitigated by a proper choice of the block length. In Section 6 we analyze the precisions of estimations provided by bootstrap methods. Here we also show how big is the improvement in the precision of estimation provided by a bootstrap method as compared with the precision of estimation provided by a method which is based on using non-overlapping blocks of data. In the same section we try to answer the question whether the bootstrap methods can help to discriminate if the empirical stock market returns follow a pure random walk or a mixture of a random walk and a stationary process. Section 7 concludes the paper.

2 Review of Bootstrap Methods

2.1 Setup

Let $X = (x_1, x_2, \dots, x_n)$ be an observed stretch from a strictly stationary time series $\{x_t\}$ (note that x_t can also be a vector). The assumption of stationarity implies that the joint probability law of $(x_t, x_{t+1}, \dots, x_{t+k})$ does not depend on t for any $k \geq 0$. The probability distribution of x_t is denoted by F (we say that $x_t \sim F$) and is completely unspecified. Let $\theta(F)$ be some parameter of interest such as the mean, median, correlation, or standard deviation of F . Let $\hat{\theta}(X)$ be an estimator of $\theta(F)$ computed using observations X .

Since the estimator $\hat{\theta}(X)$ is a random variable, for the purpose of statistical inference one needs to know the probability distribution of $\hat{\theta}(X)$. This is necessary in order to be able to estimate standard errors and confidence intervals for $\theta(F)$. This is also necessary for constructing hypothesis tests about the value of $\theta(F)$. The bootstrap is a method for estimating the distribution of estimator or test statistic by resampling the data. Bootstrapping is becoming the most popular general purpose approach to statistical inference because it does not require any parametric assumption on F , and because it can be effectively utilized with smaller sample sizes ($n < 30$). Bootstrapping is also often used as an alternative to inference based on parametric assumptions when theoretical calculation of the statistics of interest is

not possible.

The practical realization of a bootstrap method depends crucially on whether the observations of $\{x_t\}$ are assumed to be independent or dependent. We will refer to the bootstrap method for independent data as to the standard bootstrap. The most popular bootstrap methods for dependent data are block bootstraps. The standard bootstrap and some often used block bootstrap methods are reviewed below. The readers should be aware, however, that the bootstrap methods for dependent data is a lively research area. For a more comprehensive review of different bootstrap methods, the interested readers can consult Gentle, Hardle, and Mori (2004).

2.2 Standard Bootstrap

Efron (1979) was the first to introduce this bootstrap method. The method is implemented by sampling the data randomly *with replacement*. More formally, this method consists in drawing random resamples $X^* = (x_1^*, x_2^*, \dots, x_n^*)$ from (x_1, x_2, \dots, x_n) . Note that one does not get a permutation since the values of X^* are selected with replacement. Note also that the number of data points in a bootstrap resample is equal to the number of data points in the original sample. By doing this several times and computing for each resample $\hat{\theta}(X^*)$, one obtains an approximate probability distribution of the estimator $\hat{\theta}(X)$. The number of resamples is supposed to be as many as possible and is mainly limited by available computing power and time.

2.3 Block Bootstrap

If the data are dependent, in order not to break up the dependence while performing a bootstrap method, Hall (1985) suggested to resample the data using blocks of data instead of individual observations. There are basically two different ways of proceeding, depending on whether the blocks are overlapping or non-overlapping. Both approaches were introduced by Hall (1985) in the context of spatial data. Carlstein (1986) proposed non-overlapping blocks for univariate time series data, whereas Künch (1989) suggested overlapping blocks in the same setting. We will refer to the methods of Carlstein (1986) and Künch (1989) as to non-overlapping block bootstrap method and moving block bootstrap method respectively.

More formally, let b and l denote integers such that $n = bl$. In particular, b denotes

the number of blocks and l denotes a block length. In the non-overlapping block bootstrap method X is divided among b disjoint blocks, where the k th block is given by $B_k = (x_{(k-1)l+1}, x_{(k-1)l+2}, \dots, x_{kl})$ for $1 \leq k \leq b$. In the moving block bootstrap method the number of overlapping blocks is given by $n - l + 1$, where the k th block is given by $B_k = (x_k, x_{k+1}, \dots, x_{k+l-1})$ for $1 \leq k \leq n - l + 1$. The block bootstrapping method consists in choosing blocks $B_1^*, B_2^*, \dots, B_b^*$ by resampling randomly *with replacement* from among available blocks B_1, B_2, \dots, B_N where N is either b or $n - l + 1$ depending on whether block are non-overlapping or overlapping. If $B_i^* = (x_{i1}^*, x_{i2}^*, \dots, x_{il}^*)$, then the bootstrap version of X is

$$X^* = (B_1^*, B_2^*, \dots, B_b^*) = (x_{11}^*, x_{12}^*, \dots, x_{1l}^*, x_{21}^*, x_{22}^*, \dots, x_{2l}^*, x_{b1}^*, x_{b2}^*, \dots, x_{bl}^*).$$

Again, as in the standard bootstrap method, the number of resamples is supposed to be as many as possible and in practice is limited by available computing power and time.

By construction, in the moving block bootstrap method the bootstrapped time series has a nonstationary (conditional) distribution. The resample becomes stationary if the block length is random. This version of the block bootstrap is called the stationary bootstrap and was introduced in Politis and Romano (1994). In particular, unlike the moving block bootstrap method where the block length is fixed, in the stationary block bootstrap method the length of block k , l_k , is generated from a geometric distribution with probability p . Thus, the average block length equals $\frac{1}{p}$. The k th block begins from a random index i which is generated from the discrete uniform distribution on $\{1, \dots, n\}$. Since a generated block length is not limited from above, $l_k \in [1, \infty)$, and the block can begin with observation x_n , the stationary bootstrap method “wraps” the data around in a “circle”, so that x_1 follows x_n and so on.

Another method that can potentially improve the performance of the moving block bootstrap was suggested by Carlstein, Do, Hall, Hesterberg, and Künch (1998). This method consists in aligning with higher likelihood those blocks that match at their ends. This is achieved by resampling the blocks according to a Markov chain whose transition probabilities depend on data. This method is usually called either the “matched block bootstrap”, or the “Markov bootstrap” (“Markovian bootstrap”). Graflund (2001) and Sanfilippo (2003) implement a modification of the same idea. Instead of matching blocks at their ends, these authors match two subsequent blocks of data using their standard deviations.

The question of paramount importance in the implementation of the block bootstrap method is how to choose the optimal block length l . The paper by Hall et al. (1995) addresses this issue. The authors find that the optimal block length depends very much on context. In particular, the optimal asymptotic formula for block length is $l \sim Cn^{\frac{1}{h}}$, where C is a constant and $h = 3, 4, \text{ or } 5$. For computing block bootstrap estimators of bias or variance $h = 3$. For computing block bootstrap estimators of one-sided and two-sided distribution functions $h = 4$ and 5 respectively.

3 Financial Problem Supposed to Be Solved by Bootstrapping

Assume that the time interval $[t, T]$ is divided into n equally spaced subintervals of length Δt such that $n = \frac{T-t}{\Delta t}$. Assume further that at times $t_k = t + k\Delta t$, $k = 0, 1, \dots, n$, we observe the value of a risky asset which we denote by $P(t_k)$. Denote the one-period log return during the interval k , $1 \leq k \leq n$, by

$$x_k = \log \left(\frac{P(t_k)}{P(t_{k-1})} \right),$$

where $P(t_{k-1})$ and $P(t_k)$ are the values of the risky asset at the beginning and at the end, respectively, of interval k . Note that the number of observed returns is n such that our sample is $X = (x_1, x_2, \dots, x_n)$.

Unlike the existing and well-understood bootstrap methods in statistics (that have been reviewed in the preceding section), in finance one is not often interested in knowing the probability distribution of the estimator $\hat{\theta}(X)$. Instead, quite often one wants to estimate the parameter of interest $\theta(F_m)$, where F_m is the probability distribution of log return over m successive intervals of Δt , that is, the probability distribution of

$$x_{k,k+m} = \log \left(\frac{P(t_{k+m-1})}{P(t_{k-1})} \right) \sim F_m.$$

In other words, the problem can also be formulated as follows: given an empirical probability distribution of returns over a short-run, one is interested in computing the probability distribution of returns over a long-run. The most important obstacle in the estimation of F_m or $\theta(F_m)$ is that the number of non-overlapping intervals of length m is relatively small. For example, most often the researches use the databases supplied by either the Center for Research in Se-

curity Prices (CRSP) or the Ibbotson Associates. These databases start around 1926 which means that one has roughly 80 successive annual return observations. Suppose one wants to compute the mean and standard deviation of returns over a horizon of 20 years. In this case the number of non-overlapping intervals of 20 years is only 4 which makes the precision of estimation to be very low. It is assumed that a bootstrap method helps to increase substantially the precision of estimation of the parameter of interest. Depending on the assumption on whether the returns x_k are independent or not, researches rely on either standard bootstrap or moving block bootstrap method (note that non-overlapping block bootstrap is not used in this context due to very small number of available blocks). Yet apparently the well-known bootstrap methods reviewed in the preceding section were not supposed to handle this kind of problem in the first place. Moreover, recall that the implementation of the moving block bootstrap method requires choosing the optimal block length. How to do it in this setting remains unanswered and the researches seem to choose the block length completely arbitrarily. Finally, the researches that used bootstrap methods had never tried to find out how big are the standard errors of estimations of parameters of interest.

In the subsequent sections we will analyze the standard and moving block bootstrap methods as applied to the financial problem described here. For the purpose of illustrating results, we will employ the annual (nominal) returns on the broad U.S. stock market index for the period from 1927 to 2009 obtained from the data library of Kenneth French.¹ This index represents the value-weighted return on all NYSE, AMEX, and NASDAQ stocks. The total number of annual observations is 83. Our ultimate goal is to compute some parameters of the return distribution over periods of up to 20 years. The descriptive statistics of the log market returns are presented in Table 1. Note that the log market returns exhibit relatively large negative skewness and quite small positive excess kurtosis.² On the basis of the Jarque-Bera test, the assumption of normally distributed log market returns can be rejected at the 1% significance level. Figure 1 shows the autocorrelation and partial correlation functions of annual market returns. Apparently, there are no significant autocorrelations and partial correlations at any lag.

¹See http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

²We remind the readers that the excess kurtosis is computed by subtracting the value of 3 from the value of kurtosis.

Statistics	Value
Mean	0.091
Std. deviation	0.203
Skewness	-0.999
Kurtosis	4.080
Minimum	-0.586
Maximum	0.454

Table 1: Descriptive statistics of annual log market returns.

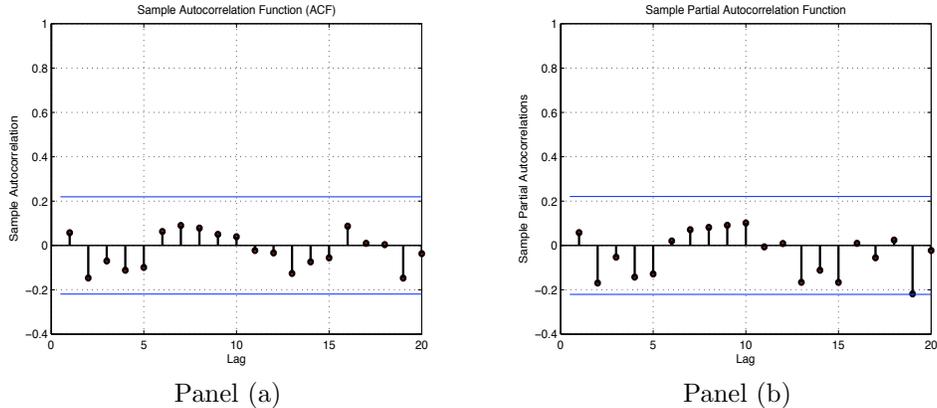


Figure 1: Panel (a) plots the autocorrelation function of the annual log market returns together with confidence bounds. Panel (b) plots the corresponding partial correlation function.

4 Analysis of the Standard Bootstrap Method

The standard bootstrap method is used when one assumes that the one-period returns are independent and identically distributed. This method has been used by many researchers, among others by Lloyd and Modani (1983), Lloyd and Haney (1985), Leibowitz and Langetieg (1989), Butler and Domian (1991), Hodges, Taylor, and Yoder (1997), Hickman, Hunter, Byrd, Beck, and Terpening (2001), Mukherji (2003), and Sinha and Sun (2005). More formally, this method consists in drawing random resamples $X_m^* = (x_1^*, x_2^*, \dots, x_m^*)$ from (x_1, x_2, \dots, x_n) with replacement. Note that unlike the genuine method of Efron (1979), here the number of data points in a bootstrap resample is less than the number of data points in the original sample. Then one compounds the returns to get m -period return

$$x_{t,t+m}^* = \sum_{k=1}^m x_k^*.$$

By doing this many times one obtains an approximate probability distribution of F_m that can be used to compute the parameter of interest $\theta(F_m)$. Again, the number of resamples is

supposed to be as many as possible and is mainly limited by available computing power and time.

It can be easily shown that on long horizons the bootstrapped (log) returns converge to the normally distributed returns. Indeed, the m -period return is just the sum of m one-period independent and identically distributed returns. By the Central Limit Theorem the distribution of this sum converges in the limit to the normal distribution as m becomes sufficiently large. To illustrate the Central Limit Theorem, denote the first four moments of probability distribution of x_t by $E[x_t]$ (mean), $Var[x_t]$ (variance), $Skew[x_t]$ (skewness), and $Kurt[x_t]$ (kurtosis). Then the first four moments of distribution of $x_{t,t+m}$ are given by (see, for example, Bakshi, Kapadia, and Madan (2003))

$$E[x_{t,t+m}] = mE[x_t], \quad Var[x_{t,t+m}] = mVar[x_t],$$

$$Skew[x_{t,t+m}] = \frac{1}{\sqrt{m}}Skew[x_t], \quad Kurt[x_{t,t+m}] - 3 = \frac{1}{m}(Kurt[x_t] - 3).$$

Note that as m increases the skewness and excess kurtosis of $x_{t,t+m}$ approach zero, while the mean and variance of $x_{t,t+m}$ increase linearly with m . Table 2 demonstrates the rate of converges of the skewness and kurtosis of the log market returns as m increases. Apparently, under the i.i.d. assumption and on the horizons up to 30 years the stock market returns deviate noticeable from the normality. However, the point is that if the probability distribution of a risky asset deviates insignificantly from normality in the short-run and one assumes that the returns are i.i.d, then for sufficiently long investment horizons one does not really need to bootstrap returns, but instead one can safely assume that the log returns are normal. In this case the risky asset price process can be modelled as a Geometric Brownian Motion which often has closed-form solutions for many types of reward and risk.

Horizon, years	Skewness	Kurtosis
1	-0.999	4.080
5	-0.447	3.216
10	-0.316	3.108
15	-0.258	3.072
20	-0.223	3.054
25	-0.200	3.043
30	-0.182	3.036

Table 2: Values of the skewness and excess kurtosis of the log market returns for different horizons under the i.i.d. assumption.

All in all, the implementation of the standard bootstrap method for financial problem considered in the preceding section is rather straightforward. However, at the end of this section we would like to point out on a potentially serious pitfall in the implementation of the standard bootstrap method. If one resamples *without the replacement* (as it was done in a couple of papers), then one gets a severely biased estimation of the parameter of interest. For the sake of motivation, suppose that one wants to estimate the variance of m -period returns given a sample of n one-period return observations. Then as $m \rightarrow n$ each resample becomes a permutation of the original sample. Consequently, as $m \rightarrow n$ the estimated variance converges towards zero. As a result of such a resampling, in the long-run the stocks appear to be much less risky than they are in reality.

5 Analysis of the Moving Block Bootstrap Method

5.1 Motivation and Review

There is a large strand of financial literature, which emerged during the 1980s, that advocates that the time series of stock returns are dependent. In particular, this literature documents excess volatility, short-term momentum and long-term reversal in stock prices, and the long-run predictability of stock returns. The most important earlier contributions to this literature include, among others, the papers by Shiller (1981), Summers (1986), Fama and French (1988), Poterba and Summers (1988), Lo and MacKinlay (1988), and Campbell and Shiller (1998) (see also Cochrane (2008) and references therein). In addition, it is well known that the hedge fund returns may exhibit a strong degree of serial correlation (see, for example, Lo (2002)).

Due to dependence between the return observations, the standard bootstrap method cannot be used. This is because the standard bootstrap destroys serial dependency in data. In order to capture the dependence structure in the return series while estimating the parameter of interest of the return distribution on long horizons, many researches rely on the moving block bootstrap method (see, for example, Hansson and Persson (2000), Graffund (2001), Lin and Chou (2003), Sanfilippo (2003), Beach (2007), and Jan and Wu (2008)).

A possible realization of the moving block bootstrap method consists in the following. Suppose one wants to simulate b times the returns over m -period horizon. Thus, one needs to simulate a sample of mb one-period returns. Given the block length l we assume for simplicity

that there is an integer d such that $dl = bm$ where d is the required number of blocks to make up the sample consisting of mb one-period observations. Since the original data sample length is n , the number of available blocks of length l is given by $n - l + 1$, where the k th block is given by $B_k = (x_k, x_{k+1}, \dots, x_{k+l-1})$ for $1 \leq k \leq n - l + 1$. The block-bootstrapping method consists in choosing blocks $B_1^*, B_2^*, \dots, B_d^*$ by resampling randomly *with replacement* from among available blocks. If $B_i^* = (x_{i1}^*, x_{i2}^*, \dots, x_{il}^*)$, then the bootstrapped sample is

$$(B_1^*, B_2^*, \dots, B_d^*) = (x_{11}^*, x_{12}^*, \dots, x_{1l}^*, x_{21}^*, x_{22}^*, \dots, x_{2l}^*, \dots, x_{d1}^*, x_{d2}^*, \dots, x_{dl}^*).$$

Finally the long bootstrapped sample of one-period returns is converted to m -period returns by compounding each disjoint sequence of m one-period returns. Whereas Hansson and Persson (2000), Lin and Chou (2003), Beach (2007), and Jan and Wu (2008) exploit the original moving block bootstrap method of Künch (1989), Graflund (2001) and Sanfilippo (2003) implement a modification of the Markov moving block bootstrap method of Carlstein et al. (1998). Unfortunately, the choice of the block length has never been properly justified and its length varies from paper to paper. The researches employ either a fixed block length or a variable block length. In the former case the block length does not depend on the horizon length and varies from 6 months to 5 years. In the latter case the block length equals the horizon length.

5.2 The Choice of the Optimal Block Length

Our goal in this subsection is to determine the optimal block length within the framework of the financial problem described in Section 3. Our method³ of finding the optimal block length is as follows. First, we construct an approximate but flexible parametric model for the log market returns. Then we estimate its parameters. Finally, by simulations, we choose the block length that is optimal for the model under specific parameters. The criterion of optimality for the choice of the block length is the minimization of the estimation bias of the parameter of interest (this idea is exploited in, for example, Hall et al. (1995)).

After having studied a number of possible parametric specifications, we decided to employ a widely accepted model for the stock or stock portfolio price. This model was for the first time considered by Muth (1960). Later on this model was used in many papers and books,

³In fact, this method appeared as a result of our discussions with Peter Hall.

among others by Summers (1986), Campbell and Mankiw (1987), Fama and French (1988), Lo and MacKinlay (1988), Poterba and Summers (1988), and Campbell, Lo, and MacKinlay (1997). As before, denote by P_t the value of a stock market index at time t . Denote by p_t the log of the index, that is, $p_t = \log(P_t)$. In this model the index value is composed of two components: a permanent component which is a random walk, and a transitory component which is a stationary process

$$p_t = q_t + z_t. \quad (1)$$

In particular, the process q_t is a random walk

$$q_t = q_{t-1} + \mu + e_t, \quad (2)$$

where μ is expected drift and e_t is white noise. z_t is any zero-mean stationary process. The common interpretation is that q_t is the “fundamental” component that reflects the efficient market value, and z_t reflects a deviation from the efficient market value q_t , implying the presence of “fads” or “bubbles”.

We assume that z_t is an ARMA(1,1) process

$$z_t = \phi z_{t-1} + w_t + \theta w_{t-1}, \quad (3)$$

where w_t is white noise, and ϕ and θ are some constants. The ARMA(1,1) process becomes AR(1) process when we set $\theta = 0$, and MA(1) process when we set $\phi = 1$. The one-period return (say, the return from t to $t + 1$) is given by $x_t = p_{t+1} - p_t$. The compounded return from t to $t + m$ is

$$x_{t,t+m} = p_{t+m} - p_t = (q_{t+m} - q_t) + (z_{t+m} - z_t).$$

If in this model ϕ is close to 1.0 and θ is close to 0.0, then it is very difficult to detect the presence of the autoregressive and moving-average components by studying the time series of one-period returns. These components become visible only in the time series of multi-period returns. To detect the presence of the autoregressive and moving-average components, one usually computes either the variance ratio or the first-order autocorrelation function of the returns compounded over several periods. In particular, the variance ratio is given by (this

ratio was used by, for example, Lo and MacKinlay (1988))

$$VR(m) = \frac{Var(x_{t,t+m})}{mVar(x_t)}.$$

The first-order autocorrelation function of returns compounded over m periods (this autocorrelation was used by, for example, Fama and French (1988))

$$AC1(m) = \frac{Cov(x_{t,t+m}, x_{t-m,t})}{\sqrt{Var(x_{t,t+m})Var(x_{t-m,t})}}.$$

The fact is that without the presence of the stationary component z_t , the (theoretical) variance ratio is 1.0 and the first-order autocorrelation function is 0.0 irrespective of the horizon length m . In other words, without the presence of the stationary component, the variance of m -period returns equals m times the variance of one-period returns and there is no correlation between returns compounded over m periods.

Using the standard logic, the derivation of the expressions for $VR(m)$ and $AC1(m)$ for our model given by (1)-(3) yields (see Appendix for the details of the derivation)

$$VR(m) = \frac{m \left(\frac{\sigma_e}{\sigma_w}\right)^2 + 2 \frac{(1+2\phi\theta+\theta^2)-\phi^{m-1}(\phi+\theta)(1+\phi\theta)}{1-\phi^2}}{m \left(\frac{\sigma_e}{\sigma_w}\right)^2 + 2m \frac{(1+2\phi\theta+\theta^2)-(\phi+\theta)(1+\phi\theta)}{1-\phi^2}}, \quad (4)$$

$$AC1(m) = \frac{\phi^{m-1}(2-\phi^m)(\phi+\theta)(1+\phi\theta) - (1+2\phi\theta+\theta^2)}{m(1-\phi^2) \left(\frac{\sigma_e}{\sigma_w}\right)^2 + 2((1+2\phi\theta+\theta^2) - \phi^{m-1}(\phi+\theta)(1+\phi\theta))}, \quad (5)$$

where σ_e and σ_w are the standard deviations of e_t and w_t respectively. The reasonable value of ϕ should be close to but less than 1.0, and the value of θ should be close to 0.0. Otherwise, the presence of ARMA(1,1) components could be easily detected by looking at the plots of sample autocorrelation and partial correlation functions of x_t . We believe that the ranges of feasible values for ϕ and θ are $\phi \in [0.85, 1.00]$ and $\theta \in (0.00, 0.15]$.

Figure 2 demonstrates possible shapes of the $VR(m)$ and $AC1(m)$ for $\phi = 0.9$, $\theta = 0.1$, and different values of the ratio $\frac{\sigma_e}{\sigma_w}$. Observe that whereas the random walk component produces white noise in returns, the stationary component causes negative autocorrelation in returns on longer horizons. Fama and French (1988) show that this result does not depend on any particular parameterization of z_t . The presence of the stationary component also causes a

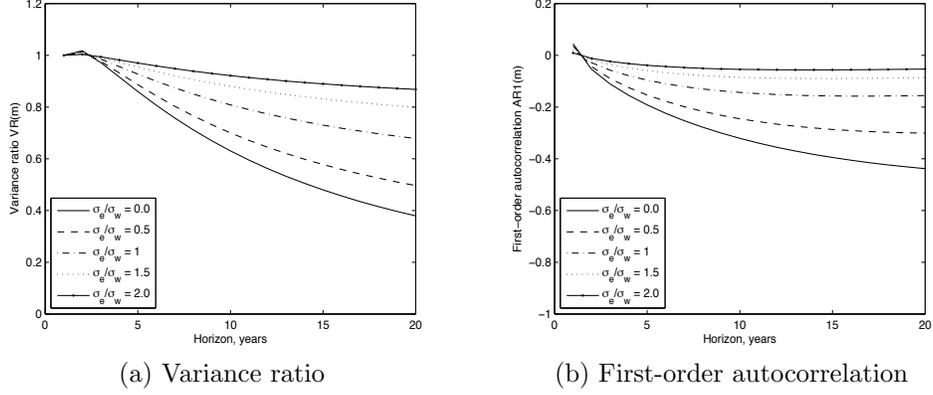


Figure 2: Variance ratio and first-order autocorrelation for $\phi = 0.9$, $\theta = 0.1$, and different values of the ratio $\frac{\sigma_e}{\sigma_w}$.

decrease in the variance ratio on longer horizons. On shorter horizons the presence of the moving-average component (when $\theta > 0$) causes positive autocorrelation in returns and an increase in the variance ratio. Thus, the presence of ARMA(1,1) components in the stock returns can generate short-term momentum and long-term reversal in stock prices documented by numerous empirical studies. Finally note that the smaller the value of the ratio $\frac{\sigma_e}{\sigma_w}$, the more pronounced the effects of the ARMA(1,1) components.

Before proceeding to the simulation method to determine the optimal block length, we need to estimate the value of the ratio $\frac{\sigma_e}{\sigma_w}$. Fortunately, this estimate is readily available in the paper by Poterba and Summers (1988). For the value-weighted U.S. stock market index, the range of possible values for this ratio is $\frac{\sigma_e}{\sigma_w} \in [0.77, 1.12]$. That is, according to the estimation method proposed by Poterba and Summers (1988), the variance of the permanent component in the stock returns is of about the same magnitude as the variance of the transitory component, $\sigma_e \approx \sigma_w$. Yet, in the seminal paper by Shiller (1981), the author argues that the stock prices are too volatile, that is, the movements in stock price indexes could not realistically be attributed to changes in fundamental values. In his view, $\sigma_e \ll \sigma_w$, putting it into words this means that the variance of the fundamental (permanent) component is much less than the variance of the transitory component. Using a simple efficient market model, Shiller shows that $\frac{\sigma_e}{\sigma_w} \in [0.1, 0.2]$. The big discrepancy between the estimates provided by Poterba and Summers from one side, and Shiller from the other side, stems from the use of different estimation methods and the need of making some assumptions in the process of estimation. Since the researches do not agree on the value of the ratio $\frac{\sigma_e}{\sigma_w}$, in our study we consider a wider range of possible values of

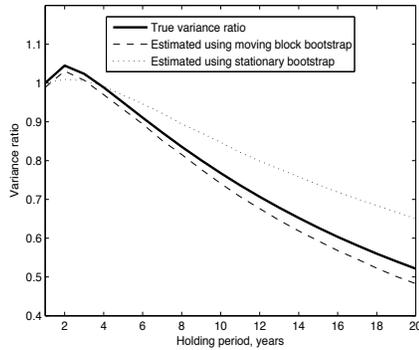
this ratio, in particular we assume that $\frac{\sigma_e}{\sigma_w} \in [0.0, 2.0]$.

Observe that in this model for the stock market returns the ARMA(1,1) components are transitory, but their effects are highly persistent through time. That is, the effects of the ARMA(1,1) components do not disappear after some time, but persist on all lags and horizons. Given this knowledge it is tempting to conclude that in this model the optimal block length in the moving block bootstrap method should equal the horizon length. However, this is not so because the estimations of the model parameters using overlapping blocks of data are subject to so-called “small sample bias”. The sources of this bias in the estimation of autocorrelation are described in details by Orcutt and Irwin (1948) and Marriott and Pope (1954). In particular, these authors show that an estimate of autocorrelation obtained using overlapping blocks is downward biased. Similarly, an estimate of variance obtained using overlapping blocks is also downward biased. This was shown in, for example, Müller (1993) and Sun et al. (2009).

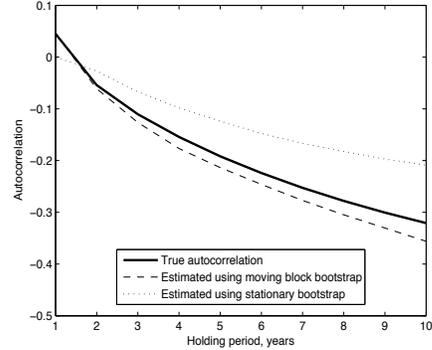
To demonstrate the bias in the estimation of the variance ratio and first-order autocorrelation we perform a simulation experiment. In particular, using our model we simulate 5,000 times a series of returns with the following model parameters: $n = 83$, $\phi = 0.9$, $\theta = 0.1$. The value of the ratio $\frac{\sigma_e}{\sigma_w}$ takes either 0.0 or 2.0. Note that the length of a series equals the number of annual observations in our empirical sample of log market returns. The mean and standard deviation of simulated annual returns are also fitted to the respective estimates of the market returns (strictly speaking, this is not necessary). After each simulation we estimate $VR(m)$ and $AC1(m)$ using three different block bootstrap methods: the moving block bootstrap method of Künch (1989); the matched block bootstrap method of Sanfilippo (2003); and the stationary block bootstrap method of Politis and Romano (1994). In the moving block bootstrap and matched block bootstrap methods we set the block length equal to the horizon length, $l = m$. In the stationary block bootstrap method we also set the average block length equal to the horizon length. At the end of the simulation experiment we compute the mean values of estimated $VR(m)$ and $AC1(m)$ and compare them to the true values given by expressions (4) and (5).

The results of our simulations and estimations are presented in Figure 3. Observe that for the moving block bootstrap method the estimation bias is negative for both parameters. In other words, the bias always decreases the estimates of variance and first-order autocorrelation as compared to their true values. The magnitude of the estimation bias depends on the sample

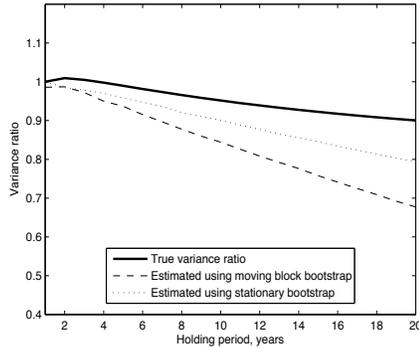
size, block length, and the model parameters, mainly on the value of the ratio $\frac{\sigma_e}{\sigma_w}$. In particular, the magnitude of the bias decreases when the value of the ratio $\frac{\sigma_e}{\sigma_w}$ decreases.⁴ That is, the larger the proportion of the variance of the transitory component in the total variance of the market returns, the smaller the estimation bias. Virtually similar results for the estimation biases are obtained for the matched block bootstrap method of Sanfilippo (2003). Yet, the matched block bootstrap method is much more time-consuming as one needs to compute a matching statistics and many generated blocks are rejected because of mismatching. For the stationary bootstrap method the estimation bias can be both negative and positive. In particular, the bias is positive when the value of the ratio $\frac{\sigma_e}{\sigma_w}$ is rather low and negative when the value of the ratio $\frac{\sigma_e}{\sigma_w}$ is rather high. Again, as compared to the moving block bootstrap method the stationary bootstrap method is more time-consuming as one needs to generate random block lengths.



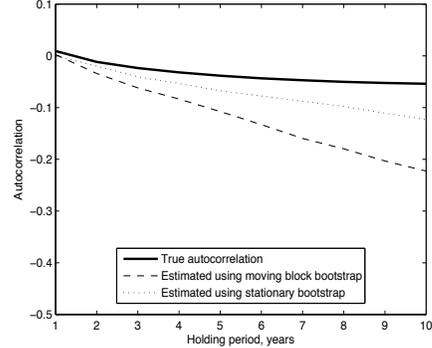
(a) Variance ratio, $\frac{\sigma_e}{\sigma_w} = 0.0$



(b) First-order autocorrelation, $\frac{\sigma_e}{\sigma_w} = 0.0$



(c) Variance ratio, $\frac{\sigma_e}{\sigma_w} = 2.0$



(d) First-order autocorrelation, $\frac{\sigma_e}{\sigma_w} = 2.0$

Figure 3: Demonstration of the bias in the estimation of the variance ratio and first-order autocorrelation using the *moving block bootstrap* and *stationary bootstrap* methods when the block length equals the horizon length. The model parameters are: $n = 83$, $\phi = 0.9$, and $\theta = 0.1$.

⁴See a similar note in Fama and French (1988), page 253.

In the econometrics estimations the method of the bias adjustment is based on the correction of the value of the estimated parameter by the estimated magnitude of the bias. For example, in Fama and French (1988) after estimation of the value of $AC1(m)$ the authors simulated the returns assuming they follow a random walk. Since the true first-order autocorrelation of m -periods returns is zero when returns follow a random walk, the estimation of the value of $AC1(m)$ for the simulated returns provides the bias estimation. Note that for the moving block bootstrap and matched block bootstrap this method of the bias correction is very conservative since the true bias can be much smaller than the bias implied by a pure random walk (when $\frac{\sigma_e}{\sigma_w}$ is rather low the bias is quite small). For the stationary block bootstrap this method of bias correction does not work at all since both the magnitude and the sign of the bias depend on the value of the ratio $\frac{\sigma_e}{\sigma_w}$.

The fact is that in many cases the bias adjustment can be performed by choosing appropriately the block length in the moving block bootstrap method. Intuitively, the bias occurs because the blocks are overlapped. The use of overlapping blocks of returns makes the multi-period returns to be (negatively) serially dependent even if there is no serial dependency in one-period returns. If the returns exhibit some degree of serial correlation, then the use of blocks induces additional negative serial dependence. Note that if the returns are dependent and we use the standard bootstrap method, then this method completely destroys any dependence in returns. Similarly, if the returns are serially dependent and we use relatively short blocks in the moving block bootstrap method, then such a method partially destroys the serial dependency in returns. The shorter the blocks, the bigger the destruction. Consequently, the idea behind a feasible bias adjustment method is to perform the moving block bootstrap simulation with blocks whose length is shorter than the horizon length in order to offset the artificially induced negative serial correlation. Note, however, that this method of bias adjustment works only when the true serial correlation is negative. This is because the bias always reduces the value of serial correlation. When the true serial correlation is positive, in this case we need to adjust the value of serial correlation in the positive direction from zero, whereas the use of shorter blocks can only adjust the value of serial correlation in the direction towards zero. Fortunately, the stock returns exhibits small positive serial correlation only on relatively short horizons up to one year length (see, for example, Lo and MacKinlay (1988) and Poterba and Summers (1988)). On longer horizons the stock returns exhibit negative serial correlation.

We use a simulation method to find the optimal block length in the moving block bootstrap method. The criterion of optimality is the minimization of the bias between the true value of a parameter of interest and the mean of the estimated value of the parameter after a series of simulations. For the sake of demonstration, let us suppose that we want to estimate the long-run variance of stock returns. In this case the criterion of optimality might be to choose the block length to minimize the bias in the estimation of the variance ratio. The simulations proceed as follows. Using our model we simulate 5,000 times a series of returns given a set of the model parameters: n , ϕ , θ , and $\frac{\sigma_e}{\sigma_w}$. After each simulation we estimate $VR(m)$ using the moving block bootstrap method where the block length varies from one to the horizon length, $l \in [1, \dots, m]$. At the end of the simulations we compute the mean values of estimated $VR(m)$ for each block length and find the block length which produces the value of estimated mean variance ratio closest to the theoretical value given by expression (4). The results of this exercise for $\phi = 0.9$ and $\theta = 0.1$ are reported in Table 3. These results suggest that the optimal block length tends to decrease with respect to the horizon length when either the sample length decreases, the horizon length increases, or the value of the ratio $\frac{\sigma_e}{\sigma_w}$ increases. The optimal block length is roughly equal to the horizon length only when $\sigma_e \ll \sigma_w$. Generally, the optimal block length is shorter than the horizon length.

Equipped with the knowledge of how the choice of the block length affects the estimation bias, we are able now to explain why the bias in the stationary bootstrap method can be both positive and negative (when the average block length equals the horizon length). First, note that when the value of the ratio $\frac{\sigma_e}{\sigma_w}$ is relatively small, the artificially induced (as a result of using overlapping blocks) serial dependence is also small. In contrast, when the value of the ratio $\frac{\sigma_e}{\sigma_w}$ is relatively high, the induced serial dependence is also high. Second, note that a random block length also partially destroys serial dependence in data. Yet, the degree of “destruction” is the same regardless the magnitude of the bias. Thus, when the value of the ratio $\frac{\sigma_e}{\sigma_w}$ is relatively small, the stationary bootstrap decreases the serial dependency more than necessary to reduce the real bias, so the resulting bias becomes positive. On the other hand, when the value of the ratio $\frac{\sigma_e}{\sigma_w}$ is relatively high, the stationary bootstrap decreases the serial dependency less than necessary to reduce the real bias, so the resulting bias remains negative.

Finally in this section we would like to point out that when the block length is equal to the horizon length, $l = m$, then the moving block bootstrapping is completely unnecessary.

Ratio σ_e/σ_w	Horizon, years									
	2	4	6	8	10	12	14	16	18	20
Sample length $n = 60$										
0.5	2	1	6	7	9	11	13	15	17	19
1.0	2	2	5	7	7	10	11	11	13	14
1.5	1	2	4	4	5	5	6	7	7	8
2.0	1	1	2	3	3	5	5	6	6	8
Sample length $n = 83$										
0.5	2	1	6	8	10	12	14	16	18	19
1.0	2	3	5	7	9	11	13	14	15	16
1.5	2	2	4	5	5	6	7	8	9	12
2.0	1	1	2	4	5	6	7	7	8	8
Sample length $n = 100$										
0.5	2	1	6	8	10	12	14	16	18	20
1.0	2	3	6	7	9	11	13	15	16	18
1.5	2	2	5	5	9	6	12	12	14	14
2.0	2	1	3	5	6	7	8	9	10	10
Sample length $n = 150$										
0.50	2	1	6	8	10	12	14	16	18	20
1.00	2	1	6	8	10	12	13	15	17	19
1.50	2	3	5	7	9	11	13	15	16	19
2.00	2	1	5	6	7	8	9	10	11	11

Table 3: Optimal block length (in years) for the estimation of the variance ratio given the length of the investment horizon and the value of the ratio $\frac{\sigma_e}{\sigma_w}$. Model parameters are $\phi = 0.9$ and $\theta = 0.1$. One-period mean and standard deviation of returns mimic the respective estimates of the market returns.

This is because in this case the probability distribution of F_m is implicitly approximated by the sample probability distribution \hat{F}_m where the number of observations equals the number of overlapped blocks, $n - m + 1$, and the probability mass of each block equals $\frac{1}{n-m+1}$. Thus, as the number of simulations goes to infinity, the estimation of a parameter of interest of F_m converges to the computation of the value of this parameter using the discrete probability distribution \hat{F}_m .

6 Benefits of Bootstrap Methods and Precisions of Estimations

Without a bootstrap, to estimate the parameters of interest of F_m the only alternative is to use all available non-overlapping m -period returns. In particular, suppose that b is the number of available non-overlapping blocks of length m such that $bm \leq n$, but $(b + 1)m > n$. That is, without using the bootstrap we are left with only b observations of m -period returns, where the k th observation is based on the compounding of $(x_{(k-1)m+1}, x_{(k-1)m+2}, \dots, x_{km})$. With a

bootstrap, the number of resampled blocks is much bigger than the number of available non-overlapping blocks. Thus, one can reasonably suppose that the precision of estimation (of a parameter of interest) delivered by a bootstrap should be better. The first question we would like to answer in this section is how big is the improvement in the precision of estimation.

To estimate the benefits provided by bootstrap methods, we perform the following simulation experiments based on realistic model parameters and the number of available observations. To evaluate the improvement in the precision of estimation provided by the standard bootstrap method, we simulate 83 observations of normally distributed independent returns whose mean and variance equal the estimated annual mean and variance of empirical log market returns. In contrast, to estimate the improvement in the precision of estimation provided by the moving block bootstrap method, we simulate 83 observations of serially dependent returns using our model with $\phi = 0.9$, $\theta = 0.1$, and $\frac{\sigma_e}{\sigma_w} = 1.0$. Again, the mean and standard deviation of simulated annual returns are fitted to the respective estimates of the market log returns. After each simulation we compute the standard deviation of multi-period returns by two distinct methods: the first one is based on using available non-overlapping intervals of length m ; the other one is based on 1,000 (properly) bootstrapped returns. The standard bootstrap method is implemented in accordance with the description given in Section 4, whereas the moving block bootstrap method is implemented in accordance with the description given in Section 5 with the optimally chosen block length (see Table 3). We repeat this exercise 5,000 times and compute the standard error of estimation and the 95% confidence intervals for the estimation of standard deviation for each distinct method. The results are presented in Table 4 and Figure 4. It is pretty evident that as m increases, the benefits of the bootstrap methods also increase. In particular, even though the standard error of estimation of standard deviation increases with horizon regardless the method used, the rate of increase in the standard error of estimation provided by a bootstrap method is substantially lower than that provided by the method that uses non-overlapping returns.

Yet, the amount of benefits provided by a bootstrap method depends on whether the returns are serially dependent or not. In particular, the benefits provided by a bootstrap method are substantially bigger when the returns are independent. For example, if the returns are i.i.d. and $m = 20$ then using the (standard) bootstrap method the standard error of the estimation of standard deviation is approximately 5 times smaller than using only 4 available

non-overlapping returns over 20 periods. In contrast, if the returns are serially dependent then at the same horizon $m = 20$ the standard error of the estimation of standard deviation is only approximately 2 times smaller than using 4 available non-overlapping returns over 20 periods. Apparently, the use of overlapping blocks instead of non-overlapping block provides relatively little new information about the value of the parameter of interest.

Horizon, years	Standard error of estimation			
	Independent returns		Dependent returns	
	Method 1	Method 2	Method 1	Method 2
1	0.016	0.016	0.016	0.016
2	0.031	0.023	0.033	0.029
4	0.066	0.033	0.067	0.041
6	0.101	0.041	0.102	0.060
8	0.133	0.046	0.135	0.077
10	0.169	0.051	0.163	0.092
12	0.221	0.057	0.207	0.108
14	0.259	0.062	0.247	0.123
16	0.276	0.065	0.256	0.137
18	0.340	0.069	0.314	0.152
20	0.356	0.074	0.317	0.162

Table 4: This table demonstrates the improvements in the precision of estimation provided by the bootstrap methods. In particular, the table reports the values of the standard errors of estimation of standard deviation of multi-period returns provided by each distinct method. Method 1 uses available non-overlapping multi-period returns, whereas Method 2 is a proper bootstrap method.

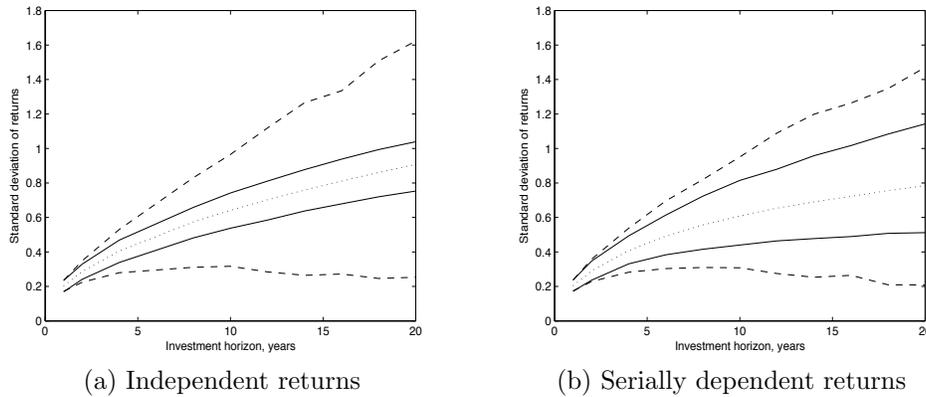


Figure 4: This figure demonstrates the improvements in the precision of estimation provided by the bootstrap methods. Dotted line shows the theoretical standard deviation over m -period horizon. Solid lines show the boundaries of the 95% confidence interval for the estimation of the standard deviation using the proper bootstrap method. Dashed lines depict the boundaries of the 95% confidence interval for the estimation of the standard deviation using non-overlapping periods of length m .

The second question we would like to answer in this section is whether the bootstrap methods can help to discriminate if the empirical stock market returns follow a pure random

walk or a mixture of a random walk and a stationary process. Recall that there are numerous studies that document that an empirically computed variance ratio of the stock market returns is decreasing with horizon. This is usually attributed to the evidence of mean-reversion in stock returns. However, despite the fact that the decrease in the empirical variance ratio appears to be economically significant, one also needs to know whether this decrease is statistically significant. To estimate the statistical significance of mean-reversion in the stock market returns, we perform a statistical test where the null hypothesis is that the returns follow a pure random walk. We can reject this hypothesis if the variance ratio of the empirical stock market returns lies outside of the confidence interval for the variance ratio of a pure random walk. To compute the boundaries of this confidence interval, we will again rely on simulations.

In these simulations we bootstrap the empirical stock market returns 10,000 times to get many possible return realizations of the same length as that of the original sample, $n = 83$. Here we use the *standard bootstrap* method which results in realizations that follow a pure random walk. After each simulation we compute the variance ratio for each horizon $m \in [1, \dots, 20]$ using the *moving block bootstrap method* implemented with the optimal block length for the following set of the values of the ratio $\frac{\sigma_e}{\sigma_w} \in \{0.5, 1.0, 1.5, 2.0\}$ (see Table 3). Note that in this manner we compute the probability distribution of $VR(m)$ when the underlying returns follow a pure random walk, but the computation is done under the assumption of mean-reversion. Finally for each value of $\frac{\sigma_e}{\sigma_w}$ we use the moving block bootstrap method with a proper block length to compute the variance ratio of the empirical stock market returns. The locations of the variance ratio of the empirical stock market returns with respect to the probability distributions of the variance ratio of a pure random walk are presented in Figure 5. Apparently, this figure suggests that the random walk behavior of the stock market returns cannot be rejected at conventional statistical levels (1% or 5%). However, on some horizons and the values of the ratio of $\frac{\sigma_e}{\sigma_w}$ the significance level becomes less than 20%. For example, when $\frac{\sigma_e}{\sigma_w} = 1.0$ or $\frac{\sigma_e}{\sigma_w} = 1.5$, we are from 60% to 80% confident that the stock market returns do not follow a random walk on horizons above 7 years.

It is worth noting that even though the empirical data favor the mean-reversion hypothesis and under the assumption of mean-reversion the estimate for the standard deviation of long-term stock market returns is lower than that under the assumption of a random walk, this may not actually imply that the stock market is less volatile in the long-run if its returns are

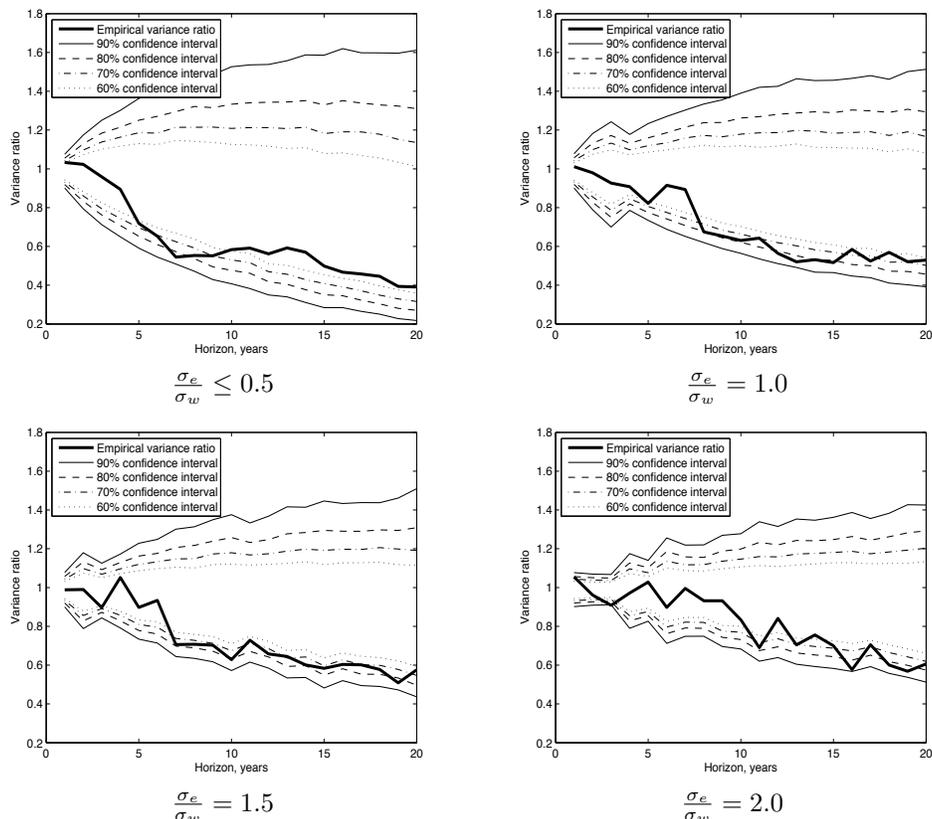


Figure 5: Each panel in this figure depicts the location of the variance ratio of the empirical stock market returns with respect to the probability distributions of the variance ratio of a pure random walk.

mean-reverting. The problem is that if the stock returns are mean-reverting, then the precision of the estimation of volatility is quite low. To illustrate this point, we estimate the standard deviation of multi-period log stock market returns under two competing assumptions: a pure random walk and a mixture of a random walk and a stationary process (here we assume that $\frac{\sigma_e}{\sigma_w} = 1.0$). When we assume that the returns are i.i.d., we use the standard bootstrap method to estimate the standard deviation. When the returns are assumed to be serially dependent, we use the moving block bootstrap method with the optimally chosen block length. The estimates for the standard deviation of the stock market returns together with the boundaries of the 95% confidence intervals are reported in Table 5. Observe that under the i.i.d. assumption the width of the confidence interval increases as \sqrt{m} , where m is the horizon length, whereas the width of the confidence interval under the mean-reversion assumption increases with a higher rate. In fact, the upper boundaries of both confidence intervals are virtually the same, whereas the lower boundary of the confidence interval under the mean-reversion assumption is much below

the respective boundary under the random walk assumption. Thus, under the mean-reversion the uncertainty in the long-run *annualized* stock volatility is much higher than the uncertainty in the short-run *annualized* stock volatility. Consequently, taking into account the parameter uncertainty, under the mean-reversion the stocks appear to be more volatile in the long-run. A similar point is made by Pastor and Stambaugh (2009).

Horizon, years	Estimates for standard deviation	
	Independent returns	Dependent returns
1	0.203 \pm 0.032	0.203 \pm 0.032
2	0.286 \pm 0.046	0.298 \pm 0.058
4	0.405 \pm 0.066	0.397 \pm 0.082
6	0.489 \pm 0.082	0.469 \pm 0.120
8	0.573 \pm 0.092	0.478 \pm 0.154
10	0.633 \pm 0.102	0.500 \pm 0.184
12	0.689 \pm 0.114	0.534 \pm 0.216
14	0.755 \pm 0.124	0.563 \pm 0.246
16	0.814 \pm 0.130	0.604 \pm 0.274
18	0.853 \pm 0.138	0.634 \pm 0.304
20	0.900 \pm 0.148	0.645 \pm 0.324

Table 5: The estimates for the standard deviation of the empirical stock market returns together with the boundaries of the 95% confidence intervals under two distinct assumptions.

7 Conclusions

It is generally accepted that time horizon plays a crucial role in the determination of the optimal portfolio asset allocation. Financial advisors and mutual fund sales literature commonly recommend that the investment horizon should be relatively long in order to benefit from the “time diversification”. In finance, the problem of estimating risk and reward over a relatively long horizon comes from a lack of sufficient data and the uncertainty in the nature of the return generating process. To overcome the first obstacle the researches rely on statistical bootstrap methods. However, a practical realization of a bootstrap method depends crucially on whether the data are assumed to be serially dependent or not. In this paper we reviewed some popular bootstrap methods and argued that the application of these methods is not well understood. This especially concerns the application of the moving block bootstrap methods, which are used if there is a serial dependency in data. Namely, the estimates provided by a moving block bootstrap method are generally biased and rather imprecise.

For the sake of being specific, we considered the well-established model for the equity price process where the return is composed of two components: a permanent component which is a random walk, and a transitory component which is a strictly stationary process. Based on some feasible parametric assumptions in this model, we demonstrated the estimation bias and how this bias can be mitigated by a proper choice of the block length in a moving block bootstrap method. We found that the optimal block length depends mainly on the investment horizon and the ratio of the standard deviation of the permanent component in the stock returns to the standard deviation of the temporary component. In particular, we found that when the value of this ratio is rather small, then the optimal block length is close to the horizon length. In other cases the optimal block length is shorter than the horizon length. We also analyzed the precision of estimation provided by bootstrap methods and demonstrated that the precision of estimation of parameters of serially dependent data provided by a moving block bootstrap method is much poorer as compared to the precision of estimation of parameters of independent data provided by the standard bootstrap method. Finally, we studied whether the bootstrap methods can help to discriminate if the empirical stock market returns follow a pure random walk or a mixture of a random walk and a stationary process. We found that the random walk behavior of the stock market returns cannot be rejected at conventional statistical levels. Nevertheless, the empirical data favors the mean-reverting hypothesis.

One can, of course, question the validity of the model that we have employed. Nevertheless, the main message of this paper remains valid regardless of the model used. Namely, the moving block bootstrap methods provide generally biased and poor estimations and the correct accounting for the bias is impossible without making some specific assumptions about the nature of the data generating process. After having made the required assumptions, the bias adjustment method remains completely the same: using simulations one needs to choose the block length that minimizes the estimation bias. In other words, our results imply that the choice of the optimal block length cannot be made arbitrarily and, generally, the optimal block length is model dependent. In addition, since the precision of estimation provided by the moving block bootstrap methods is rather poor, one should be aware that the estimation risk is a big issue. All in all we conclude that the bootstrap methods for finance are of considerable practical importance and merit further extensive research.

Having reviewed the papers where the researches use bootstrap methods, we also noted

a couple of pitfalls in the implementation of these methods. In particular, sometimes in the standard bootstrap method the resampling is done without the replacement which leads to risk underestimation. In addition, the simulations are completely unnecessary when the block length in the moving block bootstrap method equals the horizon length.

Appendix: Derivation of the Formulas for $VR(m)$ and $AR1(m)$

Note that z_t can be written as $MA(\infty)$

$$z_t = w_t + (\phi + \theta) \sum_{i=1}^{\infty} \phi^{i-1} w_{t-i}.$$

This allows us to compute the variance of z_t

$$Var(z_t) = (1 + (\phi + \theta)^2(1 + \phi^2 + \phi^4 + \phi^6 + \dots)) \sigma_w^2 = \frac{1 + 2\phi\theta + \theta^2}{1 - \phi^2} \sigma_w^2.$$

The variance ratio is given by

$$VR(m) = \frac{Var(p_t - p_{t-m})}{m Var(p_t - p_{t-1})} = \frac{Var(\sum_{i=1}^m e_{t-m+i}) + Var(z_t - z_{t-m})}{m (Var(e_t) + Var(z_t - z_{t-1}))},$$

where

$$p_t - p_{t-1} = \mu + e_t + (1 - \phi)z_{t-1} + w_t + \theta w_{t-1}.$$

Let us find $Var(z_t - z_{t-m}) = 2Var(z_t) - 2Cov(z_t, z_{t-m})$. We know that

$$\begin{aligned} z_t &= w_t + (\phi + \theta)w_{t-1} + \phi(\phi + \theta)w_{t-2} + \dots \\ &+ \phi^{m-1}(\phi + \theta)w_{t-m} + \phi^m(\phi + \theta)w_{t-m-1} + \phi^{m+1}(\phi + \theta)w_{t-m-2} + \phi^{m+2}(\phi + \theta)w_{t-m-3} + \dots, \\ z_{t-m} &= w_{t-m} + (\phi + \theta)w_{t-m-1} + \phi(\phi + \theta)w_{t-m-2} + \phi^2(\phi + \theta)w_{t-m-3} + \dots \end{aligned}$$

Therefore

$$\begin{aligned} Cov(z_t, z_{t-m}) &= \left(\phi^{m-1}(\phi + \theta) + (1 + \phi^2 + \phi^4 + \phi^6 + \dots)\phi^k(\phi + \theta)^2 \right) \sigma_w^2 \\ &= \left(\phi^{m-1}(\phi + \theta) + \frac{\phi^m(\phi + \theta)^2}{1 - \phi^2} \right) \sigma_w^2 = \frac{\phi^{m-1}(\phi + \theta)(1 + \phi\theta)\sigma_w^2}{1 - \phi^2}. \end{aligned} \quad (6)$$

Thus

$$\begin{aligned} \text{Var}(z_t - z_{t-m}) &= 2 \frac{1 + 2\phi\theta + \theta^2}{1 - \phi^2} \sigma_w^2 - 2 \frac{\phi^{m-1}(\phi + \theta)(1 + \phi\theta) \sigma_w^2}{1 - \phi^2} \\ &= 2 \frac{(1 + 2\phi\theta + \theta^2) - \phi^{m-1}(\phi + \theta)(1 + \phi\theta)}{1 - \phi^2} \sigma_w^2. \end{aligned} \quad (7)$$

Observe that formula (7) is valid for any $m \geq 1$. Therefore, the formula for the variance ratio becomes

$$VR(m) = \frac{m\sigma_e^2 + 2 \frac{(1+2\phi\theta+\theta^2) - \phi^{m-1}(\phi+\theta)(1+\phi\theta)}{1-\phi^2} \sigma_w^2}{m\sigma_e^2 + 2m \frac{(1+2\phi\theta+\theta^2) - \phi^{m-1}(\phi+\theta)(1+\phi\theta)}{1-\phi^2} \sigma_w^2}.$$

The division of both the numerator and the denominator in the formula above by σ_w^2 produces formula (4).

Now we turn to the derivation of the formula for

$$AC1(m) = \frac{\text{Cov}(p_{t+m} - p_t, p_t - p_{t-m})}{\text{Var}(p_t - p_{t-m})}.$$

We already know the expression for $\text{Var}(p_t - p_{t-m})$ in the denominator. Thus, it remains only to find the expression for the numerator. First observe that

$$\text{Cov}(p_{t+m} - p_t, p_t - p_{t-m}) = \text{Cov}(z_{t+m} - z_t, z_t - z_{t-m}),$$

because the terms e_{t+i} for all i are independent. Then

$$\text{Cov}(z_{t+m} - z_t, z_t - z_{t-m}) = -\text{Cov}(z_t, z_t) + \text{Cov}(z_{t+m}, z_t) + \text{Cov}(z_t, z_{t-m}) - \text{Cov}(z_{t+m}, z_{t-m}).$$

Now note that since z_t is stationary, then $\text{Cov}(z_{t+m}, z_t) = \text{Cov}(z_t, z_{t-m})$ and we have already derived the formula for this, see (6). Now note that by stationarity $\text{Cov}(z_{t+m}, z_{t-m}) = \text{Cov}(z_t, z_{t-2m})$. Hence, we can use the same formula as (6). Finally

$$AC1(m) = \frac{-\frac{1+2\phi\theta+\theta^2}{1-\phi^2} \sigma_w^2 + 2 \frac{\phi^{m-1}(\phi+\theta)(1+\phi\theta)}{1-\phi^2} \sigma_w^2 - \frac{\phi^{2m-1}(\phi+\theta)(1+\phi\theta)}{1-\phi^2} \sigma_w^2}{m\sigma_e^2 + 2 \frac{(1+2\phi\theta+\theta^2) - \phi^{m-1}(\phi+\theta)(1+\phi\theta)}{1-\phi^2} \sigma_w^2}.$$

After some re-arrangements and the division of both the numerator and the denominator by $\frac{\sigma_w^2}{1-\phi^2}$, we arrive to formula (5).

References

- Bakshi, G., Kapadia, N., and Madan, D. (2003). “Stock Return Characteristics, Skew Laws, and the Differential Pricing of Individual Equity Options”, *Review of Financial Studies*, 16(1), 101–143.
- Beach, S. L. (2007). “Semivariance in Asset Allocations: Longer Horizons Can Handle Riskier Holdings”, *Journal of Financial Planning*, 20(1), 60–69.
- Butler, K. C. and Domian, D. L. (1991). “Risk, Diversification, and the Investment Horizon”, *Journal of Portfolio Management*, 17(3), 41–47.
- Campbell, J. Y., Lo, A. W., and MacKinlay, A. C. (1997). *The Econometrics of Financial Markets*. Princeton University Press.
- Campbell, J. Y. and Mankiw, N. G. (1987). “Are Output Fluctuations Transitory?”, *Quarterly Journal of Economics*, 102(4), 857–880.
- Campbell, J. Y. and Shiller, R. J. (1998). “Valuation Ratios and the Long-Run Stock Market Outlook”, *Journal of Portfolio Management*, 24(4), 11–26.
- Carlstein, E. (1986). “The Use of Subseries Values for Estimating the Variance of a General Statistic from a Stationary Sequence”, *Annals of Statistics*, 14(3), 1171–1179.
- Carlstein, E., Do, K.-A., Hall, P., Hesterberg, T., and Künch, H. R. (1998). “Matched-Block Bootstrap for Dependent Data”, *Bernoulli*, 4(3), 305–328.
- Cochrane, J. H. (2008). “The Dog That Did Not Bark: A Defense of Return Predictability”, *Review of Financial Studies*, 21(4), 1533–1575.
- Efron, B. (1979). “Bootstrap Methods: Another Look at the Jackknife”, *Annals of Statistics*, 7(1), 1–26.
- Fama, E. F. and French, K. R. (1988). “Permanent and Temporary Components of Stock Prices”, *Journal of Political Economy*, 96(2), 246–273.
- Gentle, J. E., Hardle, W. K., and Mori, Y. (2004). *Handbook of Computational Statistics: Concepts and Methods*. Springer.
- Graflund, A. (2001). “Expected Real Returns from Swedish Stocks and Bonds”, Working Paper, University of Lund.
- Hall, P. (1985). “Resampling a Coverage Process”, *Stochastic Processes and Their Applications*, 20, 231–246.
- Hall, P., Horowitz, J. L., and Jing, B.-Y. (1995). “On Blocking Rules for the Bootstrap with Dependent Data”, *Biometrika*, 82(3), 561–574.

- Hansson, B. and Persson, M. (2000). “Time Diversification and Estimation Risk”, *Financial Analysts Journal*, 56(5), 55–62.
- Härdle, W., Horowitz, J., and Kreiss, J.-P. (2003). “Bootstrap Methods for Time Series”, *International Statistical Review*, 71(5), 435–459.
- Hickman, K., Hunter, H., Byrd, J., Beck, J., and Terpening, W. (2001). “Life Cycle Investing, Holding Periods, and Risk”, *Journal of Portfolio Management*, 27(2), 101–111.
- Hodges, C. W., Taylor, W. R., and Yoder, J. A. (1997). “Stock, Bonds, the Sharpe Ratio, and the Investment Horizon”, *Financial Analysts Journal*, 53(6), 74–80.
- Jan, Y.-C. and Wu, Y.-L. (2008). “Revisit the Debate of Time Diversification”, *Journal of Money, Investment and Banking*, 6, 27–33.
- Künch, H. R. (1989). “The Jackknife and the Bootstrap for General Stationary Observations”, *Annals of Statistics*, 17(3), 1217–1241.
- Leibowitz, M. L. and Langetieg, T. C. (1989). “Shortfall Risk and the Asset Allocation Decision: A Simulation. Analysis of Stock and Bond Risk Profiles”, *Journal of Portfolio Management*, 16(1), 61–68.
- Lin, M.-C. and Chou, P.-H. (2003). “The Pitfall of Using Sharpe Ratio”, *Finance Letters*, 1, 84–89.
- Lloyd, W. P. and Haney, R. L. (1985). “Time Diversification: Surest Route to Lower Risk”, *Journal of Portfolio Management*, 11(4), 24–26.
- Lloyd, W. P. and Modani, N. K. (1983). “Stocks, Bonds, Bills, and Time Diversification”, *Journal of Portfolio Management*, 9(3), 7–11.
- Lo, A. W. and MacKinlay, A. G. (1988). “Stock Market Prices do not Follow Random Walks: Evidence from a Simple Specification Test”, *Review of Financial Studies*, 1(1), 41–66.
- Lo, A. W. (2002). “The Statistics of Sharpe Ratios”, *Financial Analysts Journal*, 52(4), 36 – 52.
- Marriott, F. H. C. and Pope, J. A. (1954). “Bias in the Estimation of Autocorrelations”, *Biometrika*, 41(3/4), 390–402.
- Mukherji, S. (2003). “Optimal Portfolios for Different Holding Periods and Target Returns”, *Financial Services Review*, 12(1), 61–71.
- Müller, A. (1993). “Statistics of Variables Observed Over Overlapping Intervals”, Working Paper, Olsen and Associates.
- Muth, J. (1960). “Optimal Properties of Exponentially Weighted Forecasts”, *Journal of the American Statistical Association*, 55, 229–306.

- Orcutt, G. H. and Irwin, J. O. (1948). “A Study of the Autoregressive Nature of the Time Series Used for Tinbergen’s Model of the Economic System of the United States”, *Journal of the Royal Statistical Society. Series B*, 10(1), 1–53.
- Pastor, L. and Stambaugh, R. F. (2009). “Are Stocks Really Less Volatile in the Long Run?”, NBER Working Paper No. 14757.
- Politis, D. and Romano, J. (1994). “The Stationary Bootstrap”, *Journal of the American Statistical Association*, 89(428), 1303–1313.
- Poterba, J. M. and Summers, L. H. (1988). “Mean Reversion in Stock Prices: Evidence and Implications”, *Journal of Financial Economics*, 22, 27–59.
- Sanfilippo, G. (2003). “Stocks, Bonds and the Investment Horizon: A Test of Time Diversification on the French Market”, *Quantitative Finance*, 3(4), 345–351.
- Shiller, R. J. (1981). “Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?”, *American Economic Review*, 71(3), 421–436.
- Sinha, A. K. and Sun, M. Y. (2005). “Asset Choice And Time Diversification Benefits”, *Journal of Business and Economics Research*, 3(6), 23–34.
- Summers, L. H. (1986). “Does the Stock Market Rationally Reflect Fundamental Values?”, *Journal of Finance*, 41(3), 591–601.
- Sun, H., Nelken, I., Han, G., and Guo, J. (2009). “Error of VAR by Overlapping Intervals”, *Risk Magazine*, 2, 86–91.