



## Stock Returns and Dividend Yields Revisited: A New Way to Look at an Old Problem

Michael Wolf

*Journal of Business & Economic Statistics*, Vol. 18, No. 1. (Jan., 2000), pp. 18-30.

Stable URL:

<http://links.jstor.org/sici?sici=0735-0015%28200001%2918%3A1%3C18%3ASRADYR%3E2.0.CO%3B2-4>

*Journal of Business & Economic Statistics* is currently published by American Statistical Association.

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/astata.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

---

The JSTOR Archive is a trusted digital repository providing for long-term preservation and access to leading academic journals and scholarly literature from around the world. The Archive is supported by libraries, scholarly societies, publishers, and foundations. It is an initiative of JSTOR, a not-for-profit organization with a mission to help the scholarly community take advantage of advances in technology. For more information regarding JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

# Stock Returns and Dividend Yields Revisited: A New Way to Look at an Old Problem

Michael WOLF

Departamento de Estadística y Econometría, Universidad Carlos III de Madrid, 28903 Getafe, Spain  
(mwolf@est-econ.uc3m.es)

The problem of whether stock returns can be predicted from dividend yields is discussed. I apply a new statistical method for finding reliable confidence intervals for regression parameters in the context of dependent and possibly heteroscedastic data, called subsampling. The method works under very weak conditions and avoids the pitfalls of having to choose a structural model to fit to observed data. Appropriate simulation studies suggest that it has better small-sample properties than the generalized method of moments, which is also model free and works under weak conditions. Applying the subsampling method to three datasets, I do not find convincing evidence for the predictability of stock returns.

KEY WORDS: Constant expected returns; Dividend yield regressions; Subsampling.

There has been considerable debate in the recent finance literature over the predictability of stock returns. Several studies appear to provide empirical support for the use of the current dividend–price ratio, or dividend yield, as a measure of expected stock returns. See, for example, Rozeff (1984), Campbell and Shiller (1988a), Fama and French (1988), Hodrick (1992), and Nelson and Kim (1993). The problem with such studies is that stock-return regressions face several kinds of statistical problems, among them strong dependency structures and biases in the estimation of regression coefficients. These problems tend to make findings against the no-predictability hypothesis appear more significant than they really are.

Having recognized this, Goetzmann and Jorion (1993) argued that previous findings might be spurious and largely due to the poor small-sample performance of commonly used inference methods. They employed a bootstrap approach and concluded that there is no strong evidence indicating that dividend yields can be used to forecast stock returns. Note, however, that their special approach is not shown to be backed up by theoretical properties. Moreover, it requires a lot of custom tailoring to the specific situation at hand. For other scenarios, a different tailoring would be needed.

I intend to help in resolving some of the disagreement by applying a new technique, called subsampling. It has been shown to give correct results under very weak conditions, including dependency and heteroscedasticity. Moreover, it makes use of the observed data in a very intuitive and simple way and does not require any modifications to be applicable in different scenarios. The article is organized as follows. In Section 1, I give a brief description of the stock-returns regression problem, as well as a summary of previously used approaches and corresponding findings. Section 2 introduces the proposed subsampling method. Section 3 contains some practical details concerning the actual implementation. I use a simulation study to evaluate small-sample properties concerning stock-return regressions in Section 4. In Section 5, I apply the subsampling method to three datasets and present the results. Section 6 provides

some additional insight dealing with a reorganization of long-horizon return regressions and a joint test for multiple horizons. The article ends with some concluding remarks in Section 7.

## 1. BACKGROUND AND DEFINITIONS

I shall now describe the stock-return problem in a formal way and look at some of the previous studies in more detail. Most of the empirical studies use monthly data. Define the one-period real total return as

$$R_{t+1} = (P_{t+1} + d_{t+1})/P_t, \quad (1)$$

where  $P_t$  is the end-of-month real stock price and  $d_t$  is the real dividends paid during month  $t$ . The total return can be decomposed into capital and income return:

$$R_{t+1} = R_{t+1}^C + R_{t+1}^I \equiv P_{t+1}/P_t + d_{t+1}/P_t. \quad (2)$$

In computing the dividend yield, I follow the approach of Hodrick (1992). Because dividend payments are highly seasonal, a monthly annualized dividend series  $D_t$  is computed from compounding 12 monthly dividends at the one-month treasury-bill rate  $r_t$ :

$$D_t = d_t + (1 + r_t)d_{t-1} + (1 + r_t)(1 + r_{t-1})d_{t-2} + \dots + (1 + r_t)(1 + r_{t-1}) \dots (1 + r_{t-10})d_{t-11}.$$

Then, annual dividend yield is defined as  $Y_t = D_t/P_t$ . The historic random-walk model specifies that the returns  $R_t$  are iid (independent and identically distributed) according to some distribution. The distribution is often assumed to be lognormal, implying that log returns are normal. One implication of this particular model, but also of other models implying constant expected returns, is that future returns are unpredictable. Especially, a linear regression model like

$$\ln(R_{t+k,t}) = \alpha_k + \beta_k(D_t/P_t) + \varepsilon_{t+k,k} \quad (3)$$

would have a true  $\beta_k$  coefficient of 0. Here,  $\ln(R_{t+k,t}) = \ln(R_{t+1}) + \dots + \ln(R_{t+k})$  is the continuously compounded  $k$ -period return. All of the aforementioned studies are concerned with testing the null hypothesis  $H_0: \beta_k = 0$ . Typically, several return horizons  $k$  are considered because for some theoretical reasons (e.g., present-value model), predictability might be suspected to increase with the return horizon. Most studies are able to reject the null hypothesis at conventional significance levels for all horizons considered, suggesting that future returns can be partially forecasted using present dividend yields. The empirical evidence typically increases with the return horizon.

It is clear that under the null hypothesis the stochastic behavior of the error variables  $\varepsilon_{t+k,k}$  in (3) is completely determined by the stochastic behavior of the  $R_t$  process. Even under the random-walk model—which is stronger than the null hypothesis of  $\beta_k = 0$ —the  $\varepsilon_{t+k,k}$  are uncorrelated only for  $k = 1$ . For  $k > 1$ , the errors will always exhibit serial correlation due to the resulting overlap. For example, under the random-walk model they follow an MA( $k - 1$ ) process. In case the log returns are correlated, or under the alternative hypothesis, the  $\varepsilon_{t+k,k}$  can be arbitrarily serially correlated for all values of  $k$ . The estimation of  $\beta_k$  can be easily done by ordinary least squares (OLS), but testing the null hypothesis is nontrivial for several reasons.

First, in the case of correlated residuals, as in the case of long-horizon regressions, the usual OLS standard errors are not valid. Second, the independent variable in the regression (3) is predetermined but not exogenous. That is to say that  $D_t/P_t$  is uncorrelated with the current error term  $\varepsilon_{t+k,k}$  but generally not with past error terms  $\varepsilon_{t+k-j,k}$ ,  $j > 1$ . This is because  $\varepsilon_{t+k-j,k} = \ln(R_{t+k-j,t}) - \alpha_k - \beta_k(D_{t-j}/P_{t-j})$  and the dividend yield series  $D_t/P_t$  is highly persistent at monthly intervals. It is well known that regressions with predetermined independent variables can lead to biased, although consistent, estimates; a standard reference is Stambaugh (1986). In the case of stock-return horizons, the OLS estimator  $\hat{\beta}_k$  is typically upward biased. Third, there is evidence that the finite-sampling distribution of  $\hat{\beta}_k$  is skewed to the right; see, for example, Goetzmann and Jorion (1993).

In the remainder of the article, I shall discuss various inference methods for  $\beta_k$  according to two criteria, asymptotic consistency and small-sample properties.

### 1.1 The GMM Approach

A very common approach for making inference on  $\beta_k$  in the context of dependent and possibly heteroscedastic observations is to correct the standard errors of regression coefficient estimates for serial correlation according to the generalized method of moments (GMM) of Hansen and Hodrick (1980) and Hansen (1982). Most of the literature that follows this idea bases the correction on the additional hypothesis that log returns are uncorrelated, in which case the residuals of a  $k$ -horizon regression follow a simple MA( $k - 1$ ) process. The GMM fares well in terms of asymptotic consistency. It has been shown to converge to the right answer under weak and very general conditions (see the preceding references).

Small-sample properties, on the other hand, might pose a problem. Because GMM uses asymptotic normality, centered at the true  $\beta_k$ , it accounts neither for the finite-sample bias of  $\hat{\beta}_k$  nor for its skewness to the right. In addition, there is evidence that in the context of serial correlation the GMM corrections of the standard errors are often insufficient in finite samples. For example, see Richardson and Stock (1989), Ferson and Foerster (1994), and Politis, Romano, and Wolf (1997). I therefore expect the GMM approximation to the true sampling distribution of  $\hat{\beta}_k$  to be centered at too small a value and to have a right tail that is too short. The consequence is that observed (positive) values of  $\hat{\beta}_k$  will be judged as overly significant, and hence tests for  $\beta_k$  will be biased toward false rejection of  $H_0$ .

Two examples of studies employing the GMM were given by Fama and French (1988) and chapter 7 of Campbell, Lo, and MacKinlay (1997). Both studies reject the null hypothesis  $\beta_k = 0$  at conventional significance levels, at least for return horizons of one year and beyond.

### 1.2 The VAR Approach

An alternative approach is to estimate the finite-sampling distribution of  $\hat{\beta}_k$  under the null hypothesis and to use this estimated distribution to attach a  $P$  value to the observed value of  $\hat{\beta}_k$ . To this end, some data-generating mechanism that imposes the null has to be specified.

Campbell and Shiller (1988b), Hodrick (1992), Nelson and Kim (1993), and Goetzmann and Jorion (1995), among others, considered a first-order vector autoregression (VAR) in at least two variables, log return and dividend yield. Sometimes, additional variables are included. For example, Campbell and Shiller (1988b) included a term corresponding to earnings price ratio. Hodrick (1992) included the one-month treasury-bill return relative to its previous 12-month moving average, which is denoted  $rb_t$ . To describe his model, say, let

$$Z_t \equiv [\ln(R_t) - E(\ln(R_t)), D_t/P_t - E(D_t/P_t), rb_t - E(rb_t)]'.$$

Then a first-order VAR, or VAR(1), is given by

$$Z_{t+1} = AZ_t + u_{t+1}, \quad (4)$$

where  $A$  is a  $3 \times 3$  matrix and  $u_t$  is a three-dimensional white-noise innovation sequence.

Hodrick (1992) fitted this model to the observed data and then set the first row of the estimated VAR(1) matrix equal to 0 and the constant term corresponding to log returns equal to the unconditional mean implied by the original VAR. Of course, specifying the VAR parameters is not sufficient because an innovation sequence  $u_t$  has to be fed to the VAR model. Because there is strong empirical evidence for return data to exhibit (conditional) heteroscedasticity, Hodrick fitted a generalized autoregressive conditionally heteroscedastic (GARCH) model to the fitted innovations. He then generated artificial innovation sequences according to the estimated GARCH process, in which the innovations have a conditional normal distribution. Using this approach, Hodrick also found evidence of predictability in stock returns, both for short and long horizons.

Nelson and Kim (1993) employed a similar method, simulating from a VAR model under the null hypothesis. They randomized fitted innovations for the artificial innovation sequences, however, to better match the dispersion of true marginal distribution of the innovations. The disadvantage of this method is that it destroys any potential dependence in the innovation sequence. The study reports that the simulated distributions of the regular  $t$  statistics are upward biased and that these biases should be taken into account when making inference. Even after a bias correction, however, the authors found some evidence for predictability, especially when looking at postwar data.

Unlike the GMM, the VAR approach tries to capture the finite-sampling distribution of  $\hat{\beta}_k$  by generating artificial data having the sample size as the observed data. It succeeds in correcting for both upward biases and skewness to some extent, as demonstrated by Nelson and Kim (1993) and Goetzmann and Jorion (1993). For many financial data, however, using GARCH innovations with a conditional normal distribution tends to underestimate the tails of the true sampling distribution; see also Remark 4.1. Underestimating the tails will result in overstating the significance of observed  $\hat{\beta}_k$  values again. This might explain why the findings of Nelson and Kim (1993), who randomized fitted innovations, are not as significant as those of Hodrick (1992). On the other hand, the small-sample effect of destroying the correlation in the second moments of the innovations is not clear.

The obvious shortcoming of the VAR approach is the use of a structural model. Asymptotic consistency will only be assured if VAR(1) is the true model. This is doubtful. Of course, how big the asymptotic mistake is depends on how far the true mechanism is away from VAR(1). The problem is magnified if a parametric model for the innovations, such as GARCH(1, 1), is used. In addition, it is noteworthy that the VAR model is estimated from monthly, nonoverlapping data. Small mistakes for  $k = 1$  will therefore be magnified for long horizons, such as  $k = 48$ , via adding up  $k$  one-month returns to construct a  $k$ -month return. Another shortcoming of the VAR approach, as pointed out by Goetzmann and Jorion (1993), for example, is that it only indirectly models the serial dependence from the lagged price effect in dividend yield regressions: The variable  $P_t$  appears both on the right side and on the left side of Equation (3). This motivated Goetzmann and Jorion (1993) to develop a bootstrap approach, designed to fix this shortcoming, as will be discussed later.

Note that it is very awkward to judge the small-sample properties of the VAR method via simulation studies. Hodrick (1992) presented a simulation study that paints a very favorable picture. The problem is that he used VAR(1) with GARCH(1, 1) innovations as the data-generating mechanism in the study; that is, he pretended to know what the true mechanism is. Such a study is bound to be overly optimistic.

### 1.3 A Bootstrap Approach

As an alternative to the VAR method, Goetzmann and

Jorion (1993) used a bootstrap approach to generate artificial data sequences under the null hypothesis. The motivation is that a model-free method such as the bootstrap should avoid any mistakes due to a potentially misspecified structural model. Their particular bootstrap works as follows:

1. Form the empirical distribution of monthly total stock returns  $R_t$  and their associated income returns  $R_t^I$ , as defined in (2), from the observed data.
2. Generate a pseudo return sequence  $R_t^*$  iid according to the empirical distribution of the observed total returns  $R_1 \cdots R_n$ .
3. Subtract the contemporaneous income returns  $R_t^{I,*}$  to create a pseudo capital-return series  $R_t^{C,*}$ :  $R_t^{C,*} = R_t^* - R_t^{I,*}$ . Compound these to create a pseudo price series  $P_t^*$ .
4. Create a pseudo dividend yield sequence  $D_t/P_t^*$ , in which the  $D_t$  are the actual annual dividend flows.

It is obvious that some custom tailoring is employed here in the attempt to capture the relationship between price levels and dividends. The key problem with this approach is seen in the fact that total returns are resampled at random, implying that returns are iid according to some unknown distribution, while dividend flows remain fixed, implying that dividend payments are completely nonstochastic. By implicitly imposing these two assumptions this bootstrap is not really model free anymore. Although the first assumption is slightly troublesome—the null hypothesis of a random walk is stronger than the null hypothesis of no predictability—the second one seems unrealistic. For example, in the bootstrap world, dividend payments are completely independent of prices. For this reason, the asymptotic consistency of this bootstrap approach is doubtful. Goetzmann and Jorion (1993) did not discuss the asymptotic properties of their method.

Note that Goetzmann and Jorion came to basically the opposite conclusion of all previous studies. They did not find strong statistical evidence in favor of predictability of stock returns.  $P$  values of the observed  $\hat{\beta}_k$  values are typically slightly above 10%, even for long-term horizons.

## 2. THE SUBSAMPLING APPROACH

Politis and Romano (1994) introduced the subsampling approach as a new, general-inference method. The idea is to recompute an estimator on all subsamples or blocks of the observed sequence, where the block size is smaller than the original sample size. The motivation is as follows. Each block, as a part of the original series, was generated by the true underlying probability mechanism. It then seems reasonable to hope that one can gain information about the sampling distribution of a statistic—such as the least squares estimator  $\hat{\beta}_k$ —by evaluating it on all subsamples, or “subsamples.” Note that this approach is completely model free but computer intensive. An attractive feature of the subsampling method is that it has been shown to be asymptotically consistent under weaker assumptions than bootstrap methods.

I shall proceed to give a brief description of the subsampling method applied to simple (or straight-line) linear regression here to make this article self-contained. For a general theory of subsampling and many other applications, the reader is referred to Politis and Romano (1994) and Politis et al. (1997). The latter will be abbreviated by PRW (1997) in the remainder of this article.

Consider the simple linear model  $\mathbf{y} = X(\alpha, \beta)' + \varepsilon$ , where  $\mathbf{y}$  and  $\varepsilon$  are  $n \times 1$  vectors,  $(\alpha, \beta)'$  is a  $2 \times 1$  vector, and  $X$  is an  $n \times 2$  matrix, with the first column being equal to a vector of ones and the second column being equal to a vector  $\mathbf{x} = (x_1, \dots, x_n)'$ . In my application,  $\mathbf{y}$  corresponds to log returns and  $\mathbf{x}$  corresponds to dividend yields. The goal is to draw inference on  $\beta$ . To be able to apply the subsampling method, I need to define subvectors and submatrices:  $\mathbf{y}_{b,a} \equiv (y_a, \dots, y_{a+b-1})'$ ,  $\varepsilon_{b,a} \equiv (\varepsilon_a, \dots, \varepsilon_{a+b-1})'$  and

$$X_{b,a} \equiv \begin{pmatrix} 1 & x_a \\ \vdots & \vdots \\ 1 & x_{a+b-1} \end{pmatrix}, \quad \text{where } X = \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}.$$

Here,  $b$  denotes the block size and  $a$  the starting index of the block. The estimator of  $\beta$  based on the entire sample is the OLS estimator, denoted by  $\hat{\beta}$ . The idea of the subsampling method is to approximate the unknown sampling distribution of an estimator by recomputing it on subsamples of size  $b < n$  of the data and using the empirical distribution of those subsample estimates, after an appropriate normalization. The estimator of  $\beta$  based on  $X_{b,a}$  and  $\mathbf{y}_{b,a}$  then is given by the OLS estimator  $\hat{\beta}_{b,a} \equiv (X_{b,a}' X_{b,a})^{-1} X_{b,a}' \mathbf{y}_{b,a}$ . The subsampling approximation of

$$\Pr\{n^{1/2}(\hat{\beta} - \beta) \leq x\} \quad (5)$$

is defined by

$$L_n(x) = \frac{1}{n-b+1} \sum_{a=1}^{n-b+1} 1\{b^{1/2}(\hat{\beta}_{b,a} - \hat{\beta}) \leq x\}, \quad (6)$$

which is simply the empirical distribution of the normalized subsample statistics. Note the use of the normalizing constants  $n^{1/2}$  and  $b^{1/2}$  in Equations (5) and (6), respectively. These are necessary because the subsample statistics are computed on blocks of size  $b < n$  and therefore exhibit a greater variation than the estimator based on the entire sample. In general, the normalizing constants are chosen to ensure a proper limiting distribution of the normalized estimators. In regular cases, this is just the square root of the respective sample size (in irregular cases, the normalizing constant could be another power of the sample size, for example).

Using the approximation (6) to the unknown sampling distribution (5) allows me to construct one-sided confidence intervals for  $\beta$  in the obvious way. For example, a one-sided lower  $(1 - \alpha)$  interval is given by  $I_{\text{LOW}} = [\hat{\beta} - n^{-1/2} c_n(1 - \alpha), \infty)$ , where  $c_n(1 - \alpha)$  denotes a  $1 - \alpha$  quantile of the subsampling distribution  $L_n$  defined in (6). Two-sided confidence intervals can be constructed as the intersection of two one-sided intervals. Such intervals are

called *equal-tailed* because they have approximately equal probability in each tail. As an alternative approach, two-sided *symmetric* confidence intervals can be constructed. Their name stems from the fact that they extend equally far to the left as to the right of the estimate  $\hat{\beta}_n$ , just as normal intervals do. The common way to construct symmetric confidence intervals is to estimate the two-sided cumulative distribution function

$$\Pr\{n^{1/2}|\hat{\beta} - \beta| \leq x\}. \quad (7)$$

The corresponding subsampling approximation is defined as

$$L_{n,|\cdot|}(x) = \frac{1}{n-b+1} \sum_{a=1}^{n-b+1} 1\{b^{1/2}|\hat{\beta}_{b,a} - \hat{\beta}| \leq x\}. \quad (8)$$

A two-sided symmetric  $(1 - \alpha)$  confidence interval is then given by  $I_{\text{SYM}} = [\hat{\beta} - n^{-1/2} c_{n,|\cdot|}(1 - \alpha), \hat{\beta} + n^{-1/2} c_{n,|\cdot|}(1 - \alpha)]$ , where  $c_{n,|\cdot|}(1 - \alpha)$  denotes a  $1 - \alpha$  quantile of the subsampling distribution  $L_{n,|\cdot|}$  defined in (8). Why is it useful to distinguish between equal-tailed and symmetric intervals? It is known that symmetric intervals often enjoy enhanced coverage properties and, even in asymmetric circumstances, can be shorter than equal-tailed intervals (e.g., Hall 1988). I shall use symmetric subsampling intervals for the remainder of this article.

It can be shown that, under weak conditions, the subsampling method will yield confidence intervals for  $\beta$  with asymptotically correct coverage probability. The following is a set of sufficient conditions, allowing for dependence and considerable local heteroscedasticity. Assume  $b \rightarrow \infty$  and  $b/n \rightarrow 0$  as  $n \rightarrow \infty$ . In addition assume that, for some  $\delta > 0$ ,

- $E(y_t \varepsilon_t) = 0$  for all  $t$ ,
- $E|y_t|^{4+\delta} \leq \Delta_1$  for all  $t$ ,
- $E|\varepsilon_t|^{2+\delta} \leq \Delta_2$  for all  $t$ ,
- $E|y_t \varepsilon_t|^{2+\delta} \leq \Delta_3$  for all  $t$ ,
- $\text{cov}(b^{-1/2} \sum_{t=1}^{a+b-1} (1, y_t)' \varepsilon_t) \rightarrow V > 0$  as  $b \rightarrow \infty$  uniformly in  $a$ ,
- $E(X'_{b,a} X_{b,a}/b) \rightarrow M > 0$  as  $b \rightarrow \infty$  uniformly in  $a$ , and
- $\sum_{k=1}^{\infty} (k+1)^2 \alpha^{\delta/(4+\delta)}(k) \leq \Delta_4$ .

In the last line, the  $\alpha(k)$  are the strong mixing coefficients corresponding to the process  $\{(y_t, \varepsilon_t)'\}$ . The  $\Delta_i$  are finite constants. For a formal proof in the more general setup of multivariate OLS regression, see PRW (1997).

*Remark 2.1.* An important advantage of the subsampling method is that it is enough to know about the existence of a limiting distribution. It does not have to be known exactly or estimated in practice. Numerous examples exist in which the limiting distribution depends in a complicated way on the underlying data-generating mechanism, making inference very difficult or even impossible if explicit estimation of this distribution is necessary. One example is the area-of-variance ratio tests, in which the estimation of the limiting variance of the test statistic is usually done under simplifying assumptions (e.g., Lo and MacKinlay 1988). Another

example is given in Subsection 6.2, in which the goal is to make joint inference for several return horizons.

*Remark 2.2.* Another feature of the subsampling method is that it can be used as a simple and model-free tool to describe the sampling distribution of an estimator. In my application, the sampling distribution (5) is approximated by the subsampling distribution (6). Therefore, the distribution of  $\hat{\beta} - \beta$  is approximated by the empirical distribution of the scaled subsample values  $(b/n)^{1/2}(\hat{\beta}_{b,a} - \hat{\beta})$ ,  $a = 1, \dots, n - b + 1$ . Plotting these values as a histogram, I get a visual picture of the approximating distribution. In particular, the histogram allows one to judge properties such as bias, skewness, departure from normality, and so on. Previous studies have reported that the sampling distribution of  $\hat{\beta}_k$  is upward biased and skewed to the right. These studies, however, rely on some parametric model such as the VAR(1) model (e.g., Hodrick 1992; Nelson and Kim 1993) or on some custom-tailored bootstrap (Goetzmann and Jorion 1993). Because all subsample values come from the true probability mechanism, the corresponding histograms provide a truly model-free alternative. Some results are shown in Section 5. Note that subsample histograms could be smoothed by any of the conventional techniques to get a smoother estimate of the sampling distribution. The use of subsampling for the description of a sampling distribution, although not its use for the construction of confidence intervals, was also recognized by Shermann and Carlstein (1996).

*Remark 2.3.* A frequent concern in stock-return regressions is the high persistence of dividend yield, at least when sampled at monthly intervals. From a statistical viewpoint and the given sample sizes, a unit-root process often cannot be discarded. It has been noted in the literature that standard asymptotical methods based on mean-reverting regressors, such as GMM, can give poor small-sample approximations when a regressor is nearly integrated; for example, see Elliot and Stock (1994) and Cavanagh, Elliot, and Stock (1995).

An alternative approach, designed to yield improved small-sample performance, makes use of *local-to-unity asymptotics*; see, for example, again the two preceding references. Roughly speaking, this assumes that the largest root of the regressor process is in a  $1/n$  neighborhood of 1, allowing for a nearly integrated process. Viceira (1997) derived the asymptotic distribution for  $\hat{\beta}_1$  under local-to-unity assumptions. Applying his method to four datasets, he did not find evidence for predictability at the one-month horizon. Note that his particular method could not be applied to long-horizon regression because the theory requires the regression residuals to be a martingale difference sequence. Even if this is true for  $k = 1$ , it cannot be true for bigger  $k$  due to moving-average-like behavior of the residuals. It seems possible, however, to extend the theory to serially correlated error terms.

The subsampling theory presented in this article also allows for nearly integrated regressors—in the sense of having a *fixed* root near 1—and is intended to be an improvement over standard methods in terms of small-sample performance. It would not work, on the other hand, for exactly integrated regressors. I do not feel that this is a very seri-

ous restriction because economic theory speaks against this case. Moreover, if dividend yield were an integrated process, the same would have to be true for log returns in case  $\beta_k$  differs from 0, by cointegration theory. Because there is strong evidence against a unit root in log returns, this implies that either  $\beta_k$  equals 0 or that dividend yield is not exactly integrated. Because this article will not make a case for predictability, I feel justified in focusing on the latter scenario.

### 3. CALIBRATION

A practical problem in applying the subsampling method lies in choosing the block size  $b$ . To ensure the asymptotic properties of the method it is only necessary that the block size  $b$  tend to infinity with the sample size  $n$ , but at a smaller rate:  $b \rightarrow \infty$  and  $b/n \rightarrow 0$ , as  $n \rightarrow \infty$ . Of course, this rule gives us very little practical guidance. To deal with the problem of choosing the block size  $b$  for the subsampling method, I suggest a calibration technique that in a sense avoids having to find the “best” block size.

One can think of the accuracy of an approximate or asymptotic confidence procedure—such as normal, bootstrap, or subsampling methods—in terms of its *calibration* (Loh 1987). Suppose I use the procedure to construct a confidence interval with nominal confidence level  $1 - \lambda$ . I can denote the actual confidence level by  $1 - \alpha$ .  $\lambda$  is known to us,  $\alpha$  typically is not. An asymptotic method only ensures that  $1 - \lambda$  will tend to  $1 - \alpha$  as the sample size tends to infinity. For a finite sample size, the two levels might not be the same. If I knew the calibration function  $h: 1 - \lambda \rightarrow 1 - \alpha$ , I could construct a confidence region with exactly the desired coverage by selecting the value of  $\lambda$  that satisfies  $h(1 - \lambda) = 1 - \alpha$ . For example, if  $h(.98) = .95$ , then a confidence interval with nominal level 98% would be an actual 95% confidence interval.

Fortunately, the calibration function  $h(\cdot)$  can be estimated using bootstrap methods. This is achieved by generating artificial sequences from a bootstrap distribution  $P_n^*$ , then constructing a confidence interval from each generated pseudo sequence and observing how frequently the parameter  $\hat{\beta}$  is contained in those intervals. In the context of dependent data, I need to employ a bootstrap suitable for time series. The moving-blocks bootstrap (Künsch 1989) lends itself to the task. It generates pseudo sequences  $X_1^*, \dots, X_n^*$  by resampling entire blocks from the original data and joining these together rather than using single data points. Formally, let  $Y_{b,a}$  be the block of size  $b$  of the data  $\{X_a, \dots, X_{a+b-1}\}$ . For simplicity, I assume that  $n = lb$ , for some integer  $l$ . Moreover, let  $P_n^*$  denote the empirical distribution of the blocks  $Y_{b,1}, Y_{b,2}, \dots, Y_{b,n-b+1}$ . Then a pseudo sequence is constructed by choosing  $Y_{b,1}^*, \dots, Y_{b,l}^*$  iid from  $P_n^*$  and concatenating them. In case  $n$  is not a multiple of  $b$ , I use the same algorithm with the smallest  $l$  for which  $n < lb$  and truncate the so-obtained sequence at  $n$ .

In case I want to apply the calibration scheme to the subsampling method, I can do it conditional on a *reasonable* block size. This means that I fix a sensible block size and calibrate the subsampling intervals using that particu-

lar block size. This eliminates the problem of finding the “best” block size. In some scenarios, I will have a pretty good idea what a reasonable block size will be, either from prior experience or related simulation studies. Otherwise, see Remark 3.1. To describe the calibration technique more formally, I can use the following algorithm.

*Algorithm 3.1 (Calibration by Adjusting the Confidence Level).*

1. Generate  $K$  pseudo sequences  $X_1^{*k}, \dots, X_n^{*k}$ , according to a moving-blocks bootstrap distribution  $P_n^*$ .
  - a. For each sequence,  $k = 1, \dots, K$ , compute a  $1 - \lambda$  level confidence interval  $CI_{1-\lambda}^k$ , for a grid of values of  $\lambda$  in the neighborhood of  $\alpha$ .
2. For each  $\lambda$ , compute  $\hat{h}(1 - \lambda) = \#\{\hat{\beta} \in CI_{1-\lambda}^k\}/K$ .
3. Interpolate  $\hat{h}(\cdot)$  between the grid values.
4. Find the value of  $\lambda$  satisfying  $\hat{h}(1 - \lambda) = 1 - \alpha$ .
5. Construct a confidence interval with nominal level  $1 - \lambda$ .

*Remark 3.1.*

1. The moving-blocks bootstrap in Step 1 of the preceding algorithm requires its own block size  $b_{MB}$ . The choice of this block size has a second-order effect and is therefore not very important. If an automatic selection method is preferred, however, a “nested bootstrap” can be used. That means that I would use the moving-blocks bootstrap in both Steps 1 and 1a of the preceding algorithm with the same block size  $b_{MB}$ , limiting the grid of  $\lambda$  values to  $\lambda = \alpha$ . Repeating this algorithm for several  $b_{MB}$  values, I then would select the value  $b_{MB}$  that yields estimated coverage closest to  $1 - \alpha$ .

2. If I use the calibration scheme to calibrate the subsampling method, I need to start out with a reasonable block size  $b$ . In situations in which I do not know what a reasonable block size is, I can use the following idea. In the same way as the actual confidence level can be regarded as function of the nominal confidence level (conditional on a fixed block size), it can be considered as a function of the block size (conditional on a fixed nominal level). Fixing the nominal level at the desired level—that is, choosing  $\lambda = \alpha$ —I can therefore estimate the block calibration function  $g: b \rightarrow 1 - \alpha$ , using an analogous calibration algorithm:

*Algorithm 3.2 (Calibration by Adjusting the Block Size).*

1. Generate  $K$  pseudo sequences  $X_1^{*k}, \dots, X_n^{*k}$ , according to a moving-blocks bootstrap distribution  $P_n^*$ .
  - a. For each sequence,  $k = 1, \dots, K$ , compute a  $1 - \alpha$ -level confidence interval  $CI_b^k$ , for a selection of block sizes  $b$ .
2. For each  $b$ , compute  $\hat{g}(b) = \#\{\hat{\beta} \in CI_b^k\}/K$ . A reasonable block size will then satisfy  $\hat{g}(b) \simeq 1 - \alpha$ .
3. Two-sided equal-tailed intervals should always be computed as the intersection of two separately calibrated one-sided intervals. Particularly if the sampling distribution of  $\hat{\theta}_n$  is asymmetric, the amount of calibration needed in the lower tail can be different from the one needed in the upper tail. Of course, for symmetric intervals only one calibration is necessary.

As an illustration of how I would use the calibration method, see Figure 1 for an artificial example. Suppose I want to construct a 95% confidence interval, which corresponds to  $\alpha = .05$ . I estimate the calibration function  $h(\cdot)$  at the discrete points  $.90, .91, \dots, .99$  and linearly interpolate in between. My estimate tells me that I should construct a confidence region using a nominal level around  $.978$ .

To be fair, calibration techniques can potentially be used to enhance any asymptotic method. An obvious idea in the context of stock-return regressions would be to use a similar calibration method as the preceding algorithm for GMM confidence intervals. It is one of the purposes of this article, however, to compare my new results with previous results in the literature. Therefore, I use the “simple” GMM, as employed by Fama and French (1988) and Campbell et al. (1997) in my simulation studies.

#### 4. TWO SMALL-SAMPLE COMPARISONS

As noted before, I use two criteria to judge inference methods for  $\beta_k$ , asymptotic consistency and small-sample properties. I have already mentioned that both GMM and subsampling give asymptotically correct results under reasonable assumptions, whereas VAR and the Goetzmann and Jorion (1993) bootstrap only work under restrictive conditions. In this section, I compare the small-sample properties of the GMM and subsampling via simulation studies.

For my simulations, I need a data-generating mechanism that jointly models log returns and dividend yields. Although the true mechanism that yielded the observed data will always be unknown, I aim for a reasonable approxima-

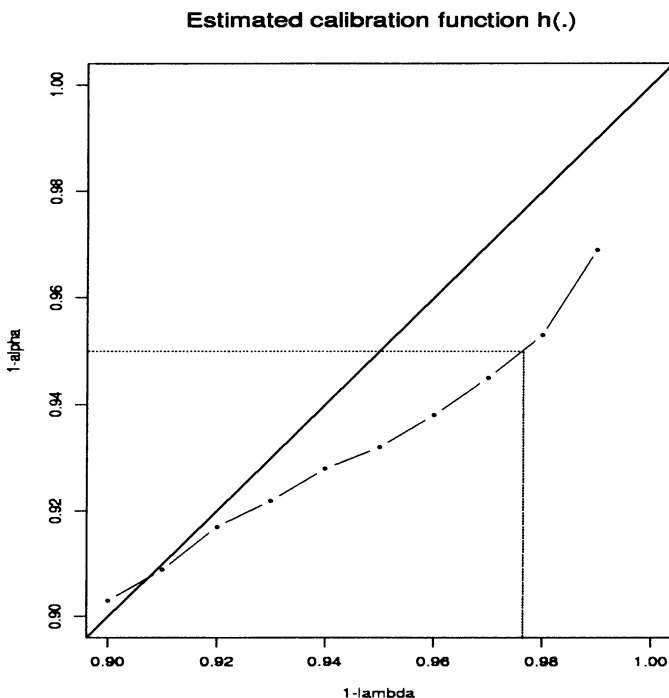


Figure 1. An Artificial Example: Estimates of  $h(1 - \lambda)$  for a grid of  $\lambda$  Values  $\{.01, .02, \dots, .1\}$ . For a confidence interval with desired confidence level  $.95$ , one should construct an interval with nominal level  $.978$ .

tion that includes at least two important features—on the one hand, the bias of  $\hat{\beta}_k$  due to the predetermined predictor, and on the other hand, the increasing autocorrelation of the residuals with the return horizon  $k$ . Both features are captured by the VAR model and the Goetzmann and Jorion (1993) bootstrap. Note that it is not a contradiction to employ models for a *simulation study*, which I criticized earlier when used for making *inference*. By definition, I need some model to generate data in a simulation study. It does not have to be the true model; it only has to capture its important features to be useful for comparing model-free inference methods, such as GMM and subsampling. On the other hand, if used for making inference, I need to be sure that I approximate the true model with a high degree of accuracy. For example, getting the marginal distributions right is crucial for making inference, although it is of relatively minor importance for a simulation study.

4.1 Simulating VAR Data

I use a first-order VAR model as my data-generating mechanism, jointly modeling log return and dividend yield as the vector  $X_t \equiv (\ln(R_t), D_t/P_t)$ . Let  $Z_t \equiv [\ln(R_t) - E(\ln(R_t)), D_t/P_t - E(D_t/P_t)]'$ . Then, the VAR(1) is given by

$$Z_{t+1} = AZ_t + u_{t+1}, \tag{9}$$

where  $A$  is a  $2 \times 2$  matrix and  $u_t$  is a white-noise innovation process. I fit this model to the observed data by least squares. The null hypothesis can be enforced by setting the first row of  $A$  equal to 0.

Because I am concerned with a simulation study only, I do not have to worry about the overall mean and can set it equal to 0 without loss of generality. I look at three different datasets, the New York Stock Exchange (NYSE) equal-weighted and value-weighted indexes and the Standard & Poor (S&P) 500 index, all starting in December 1947. Both of the NYSE datasets consist of 480 basic observations (12/1947 to 12/1986), and the S&P 500 dataset consists of 577 observations (12/1947 to 01/1995). The

fitted VAR parameters for the datasets under consideration are presented in Table 1.

To generate artificial  $X_t^*$  sequences, I use GARCH(1, 1) vector innovation sequences  $u_t^*$ . Let  $H_t = E_t(u_{t+1}u_{t+1}')$  be the conditional covariance matrix of the first-order VAR in (9) with typical element  $h_{ij,t}$ . The conditional variances and covariance follow ARMA(1, 1) processes:

$$h_{ij,t} = \omega_{ij} + \beta_{ij}h_{ij,t-1} + \alpha_{ij}u_{i,t}u_{j,t}, \quad i = 1, 2. \tag{10}$$

Equation (10) is known as the *diagonal vech* model proposed by Bollerslev, Engle, and Wooldridge (1988); of course,  $\omega_{12} = \omega_{21}$  and so on. It is not the most general multivariate GARCH(1, 1) model because it assumes that the conditional covariance of variables  $u_1$  and  $u_2$  depends on past realizations of the product  $u_1u_2$  only. Nonetheless, it is popular because it is a reasonable way to reduce the number of free parameters to a manageable size. In particular, note that it is more general than the constant correlation model by Bollerslev (1990), which assumes that the conditional correlation is constant over time; this model was used by Hodrick (1992), for example.

The nine parameters of my Model (10) are estimated by maximum likelihood, assuming conditional normality. The parameter estimates for my three different datasets are reported in Table 2. To judge the size of the  $\omega$  parameters, it should be mentioned that the models were fitted on the percentage scale; that is, a typical monthly return was on the order of .5 to .8 rather than .005 to .008. Of course, the  $\alpha$  and  $\beta$  parameters do not depend on the choice of scale. Artificial innovation sequences  $u_t^*$  as input to the VAR model (9) are generated by computing the Cholesky decomposition of the conditional covariance matrix,  $C_t' C_t = H_t$ , and setting  $u_{t+1}^* = C_t' \varepsilon_{t+1}$ , where  $\varepsilon_t$  is a sequence of independent bivariate standard normal random variables. When actually generating those sequences, I discard the first 100 observations to avoid start-up effects. Long-horizon return data  $X_t^*$  can be created by feeding the artificial innovation sequences  $u_t^*$  into the fitted VAR models, after imposing the null hypothesis by setting the first

Table 1. Parameter Estimates for VAR Matrix

Dependent variable	Coefficients on regressors	
	$\ln(R_t^0)$	$D_t^0/P_t^0$
NYSE equal-weighted, 12/1947 to 12/1986		
$\ln(R_{t+1}^0)$	.154	.325
$D_{t+1}^0/P_{t+1}^0$	-.004	.985
NYSE value-weighted, 12/1947 to 12/1986		
$\ln(R_{t+1}^0)$	.062	.423
$D_{t+1}^0/P_{t+1}^0$	-.002	.984
S&P 500, 12/1947 to 01/1995		
$\ln(R_{t+1}^0)$	.023	.450
$D_{t+1}^0/P_{t+1}^0$	.000	.985

NOTE: This table presents least squares estimates for the VAR matrix  $A$  of the following first-order VAR model:  $Z_{t+1} = AZ_t + u_{t+1}$ . Here,  $Z_t$  is the joint vector of log return and dividend yield having their respective means subtracted; that is,  $Z_t = [\ln(R_t) - E(\ln(R_t)), D_t/P_t - E(D_t/P_t)]'$ , and  $u_t$  is white noise. To simplify the notation within the table, I denote the mean zero variables by  $\ln(R_t^0)$  and  $D_t^0/P_t^0$ , respectively.

Table 2. Parameter Estimates for GARCH Model

Element	$\omega_{ij}$	$\alpha_{ij}$	$\beta_{ij}$
NYSE equal-weighted, 12/1947 to 12/1986			
$h_{11,t}$	.398	.084	.894
$h_{12,t}$	-.017	.057	.913
$h_{22,t}$	.0007	.044	.903
NYSE value-weighted, 12/1947 to 12/1986			
$h_{11,t}$	.283	.036	.937
$h_{12,t}$	-.012	.036	.932
$h_{22,t}$	.0005	.040	.927
S&P 500, 12/1947 to 01/1995			
$h_{11,t}$	.257	.047	.926
$h_{12,t}$	-.012	.058	.908
$h_{22,t}$	.0005	.072	.891

NOTE: This table presents parameter estimates for the GARCH(1, 1) model for the white-noise innovation sequence  $u_t$  of the VAR. Let  $H_t = E_t(u_{t+1}u_{t+1}')$  be the conditional covariance matrix of the VAR(1) in (9) with typical element  $h_{ij,t}$ . The conditional variances and covariance follow ARMA(1, 1) processes:  $h_{ij,t} = \omega_{ij} + \beta_{ij}h_{ij,t-1} + \alpha_{ij}u_{i,t}u_{j,t}$ ,  $i = 1, 2$ . All parameters are estimated simultaneously via maximum likelihood, assuming conditional normality.



row of the VAR matrix equal to 0. I also discard the first 100 observations in this step. Finally, the artificial long-horizons returns are compounded according to the formula  $\ln(R_{t+k,t}^*) = \ln(R_{t+1}^*) + \dots + \ln(R_{t+k}^*)$ . When generating the artificial data, I obviously need to match the original sample sizes, which are bigger for the S&P 500 data.

For every scenario, I generate 1,000 artificial sequences and compute a 95% calibrated subsampling interval for  $\beta_k$  for each sequence. For comparison, I also compute confidence intervals using the GMM employing the quadratic spectral (QS) kernel. This kernel was found to have some optimality properties by Andrews (1991). The bandwidth for the kernel was chosen according to the automatic selection procedure of Andrews (1991). I report the percentage of intervals that contain the true parameter 0 in Table 3. Note that I carried out the same simulations using conditional innovations having a (scaled)  $t$  distribution with 4 df. The results were essentially the same and are therefore not reported.

One can see that the GMM intervals undercover consistently. In other words, the GMM is biased toward falsely rejecting the null hypothesis. The undercoverage is already around 5% at the one-month horizon, and it increases in roughly linear fashion to about 30% to 35% at the 48-month horizon! The subsampling intervals perform much better, although they too undercover significantly at long horizons. The improvement of subsampling over GMM intervals is around 5% at the one-month horizon, and it increases to about 20% at the 48-month horizon.

*Remark 4.1.* In Section 1, I commented on the danger of simulating from a VAR model using GARCH innovation sequences to compute a  $P$  value for an *observed* statistic such as  $\hat{\beta}_k$ . Even in case the fitted VAR model is a good approximation, if the tails of the artificial GARCH sequences

are too light, then one overestimates the significance of observed statistics. It is therefore of interest to test whether the estimated innovations are consistent with the corresponding fitted GARCH distribution. My test statistic is the empirical .95 quantile of the estimated innovations. To compute the  $P$  value, I generate 1,000 innovation sequences of the corresponding GARCH(1, 1) model and compute the empirical .95 quantile for each of them. I then calculate the percentage of GARCH quantiles greater than or equal to the test statistic and the percentage of GARCH quantiles less than or equal to the test statistic. The  $P$  value is two times the smaller of these two percentages. To provide some additional information, I also characterize the sampling distribution of the GARCH quantiles by computing the mean, the median, the .01 quantile, and the .99 quantile of the 1,000 numbers. The results are presented in Table 4. Except for the log return innovations of the model for the equal-weighted NYSE data, all empirical .95 quantiles of the fitted innovations are too big to be compatible with the corresponding GARCH distribution. Four of the two-sided  $P$  values are equal to 0; the other one is around .05.

#### 4.2 Simulating Bootstrap Data

Although the VAR model captures some of the bias of  $\hat{\beta}_k$  due to the predetermined predictor, it fails to directly model the lagged priced effect in dividend yield regressions: The variable  $P_t$  appears both on the right side and on the left side of Equation (3). The Goetzmann and Jorion (1993) bootstrap, on the other hand, explicitly models this effect. Therefore, I use it as a second data-generating mechanism for my simulation studies. By design, it imposes the null hypothesis. This approach requires more knowledge than just the two-dimensional series of log returns and dividend yields—namely, also the split of total returns in capital returns and income returns; see Subsection 1.3. This information I only have for the S&P 500 data, so I restrict the bootstrap simulations to this dataset. Given the reasonably high correlation between S&P 500 and NYSE data, this does not seem a very serious restriction.

Table 3. Estimated Coverage Probabilities Under VAR Model

Horizon	GMM	Subsampling	Target
<i>NYSE equal-weighted, 12/1947 to 12/1986</i>			
$k = 1$	.90	.95	.95
$k = 12$	.83	.93	.95
$k = 24$	.75	.91	.95
$k = 36$	.69	.86	.95
$k = 48$	.64	.81	.95
<i>NYSE value-weighted, 12/1947 to 12/1986</i>			
$k = 1$	.90	.97	.95
$k = 12$	.81	.93	.95
$k = 24$	.72	.90	.95
$k = 36$	.66	.85	.95
$k = 48$	.59	.80	.95
<i>S&amp;P 500, 12/1947 to 01/1995</i>			
$k = 1$	.91	.96	.95
$k = 12$	.83	.93	.95
$k = 24$	.77	.91	.95
$k = 36$	.72	.87	.95
$k = 48$	.67	.84	.95

NOTE: This table presents estimated coverage probabilities of nominal 95% confidence intervals. The data-generating process is a VAR(1) with GARCH(1, 1) innovations. The null hypothesis of no predictability is enforced by setting the first row of the VAR matrix equal to 0. Two types of confidence intervals are considered, GMM intervals and calibrated symmetric subsampling intervals. The GMM uses the QS kernel with the automatic bandwidth selection procedure of Andrews (1991). Estimated coverage probabilities are based on 1,000 simulations for each scenario.

Table 4. VAR Innovation .95 Quantile: GARCH Model Versus Observed Quantile

	.01 quantile	Mean	Median	.99 quantile	Empirical	$P$ value
<i>NYSE equal-weighted, 12/1947 to 12/1986</i>						
Log return	4.909	6.751	6.628	10.156	5.554	.179
Dividend yield	.176	.211	.211	.252	.339	0
<i>NYSE value-weighted, 12/1947 to 12/1986</i>						
Log return	4.279	5.211	5.179	6.356	7.764	0
Dividend yield	.170	.209	.208	.255	.355	0
<i>S&amp;P 500, 12/1947 to 01/1995</i>						
Log return	4.073	5.033	5.010	6.324	5.982	.054
Dividend yield	.155	.193	.191	.247	.333	0

NOTE: This table compares the empirical .95 quantile of the estimated VAR innovations with the sampling distribution of the empirical .95 quantile from the corresponding GARCH(1, 1) model [based on 1,000 simulated GARCH(1, 1) innovation sequences]. The GARCH(1, 1) model was obtained via maximum likelihood from the estimated innovations. The sampling distribution is characterized by the .01 quantile, the mean, the median, and the .99 quantile. The two-sided  $P$  value tests the null hypothesis that the GARCH(1, 1) model gave rise to the estimated innovations.

The method is analogous to the previous one. I generate 1,000 sequences for each return horizon, compute a subsampling and a GMM 95% confidence interval for each sequence, and check how frequently the true parameter 0 is contained in the intervals. The results are reported in Table 5 and are very different from the ones of the VAR simulations. GMM intervals exhibit horrendous coverage, as high as .77 for  $k = 1$  and as low as .46 for  $k = 48$ . Subsampling intervals, on the contrary, have almost constant coverage between .97 and .98. These numbers are very surprising indeed and do not conform with my intuition that, as the return horizon increases, due to the increasing correlation of the residuals the performance of both GMM and subsampling should get worse. This behavior was noticed in the VAR simulations and also in some related simulations of PRW (1997). It may be that the unrealistic relationship of prices and dividend payments in the bootstrap world is the underlying cause.

## 5. A NEW LOOK AT RETURN REGRESSIONS

In this section, I apply the subsampling methodology to real-life stock-return regression data. There is strong consensus in the literature that the time series properties of stock data differ significantly in the prewar and postwar periods. In particular, predictability seems to be mostly a postwar phenomenon (e.g., Hodrick 1992; Nelson and Kim 1993). I therefore only provide postwar results. Another reason to focus on subperiods is that recent work in finance suggests that conditioning on samples from surviving markets renders the *ex post* observed process nonergodic. For example, Goetzmann and Jorion (1995) showed how this problem can bias stock-return regressions toward false rejection of the null hypothesis.

I use three different datasets that have been previously analyzed in the literature. Fama and French (1988) and Nelson and Kim (1993) reported regressions of log returns for value-weighted and equally weighted stock portfolios based on the Center for Research in Security Prices files for NYSE stocks. Goetzmann and Jorion (1993) used monthly data on the S&P 500 index. In accordance with the majority of the literature, I consider return horizons of 1, 12, 24, 36, and 48 months. Both of the NYSE postwar datasets consist of 480 basic observations (12/1947 to 12/1986), the S&P 500 dataset consists of 577 observations (12/1947 to 01/1995).

Table 5. Estimated Coverage Probabilities Under Bootstrap Model

Horizon	GMM	Subsampling	Target
<i>S&amp;P 500, 12/1947 to 01/1995</i>			
$k = 1$	.77	.97	.95
$k = 12$	.60	.98	.95
$k = 24$	.54	.98	.95
$k = 36$	.50	.97	.95
$k = 48$	.46	.97	.95

NOTE: This table presents estimated coverage probabilities of nominal 95% confidence intervals. The data-generating process is the Goetzmann and Jorion (1993) bootstrap, as described in Subsection 1.3. Estimated coverage probabilities are based on 1,000 simulations for each scenario.

Our strategy is to construct 95% confidence intervals for the regression parameter  $\beta_k$  and to check whether 0 is contained in the intervals or not. I use two-sided symmetric confidence intervals (see end of Sec. 2) in conjunction with the calibration technique described in Section 3. For the reader interested in the details of the implementation, some remarks are in order. They refer to Algorithm 3.1.

1. I used the moving-blocks bootstrap with block size  $b_{MB} = 100$  to generate the pseudo sequences in Step 1.

2. To find reasonable block sizes for the subsampling method, I used Algorithm 3.2. The so-chosen block sizes were between  $b = 40$  and  $b = 120$ , with the great majority of them between  $b = 60$  and  $b = 100$ .

3. A practical issue is the number  $K$  of bootstrap samples that I generate to estimate the calibration function  $h(\cdot)$ . I chose  $K = 1,000$ .

The resulting confidence intervals are listed in Table 6. There is no evidence for predictability for horizons of 1, 12, and 24 months because 0 is contained in all corresponding confidence intervals. For the horizon of 36 and 48 months, the findings are inconclusive. The equal-weighted NYSE index intervals contain 0, but the other two do not. For comparison, I also compute GMM intervals, using the QS kernel; again see Table 6. Note that they are much shorter and that only three of them contain 0. Employing GMM would suggest a strong case for predictability.

At this point, it is natural to ask two questions. First, the simulation study in Subsection 4.1 suggests some undercoverage of subsampling intervals at long horizons due to the very strong correlation of the residuals. How much does this evidence take away from the (weak) case for predictability that could be made at the three- and four-year horizons? Second, I always look at five return horizons simultaneously; namely,  $k = 1, 12, 24, 36$ , and 48. If I am

Table 6. 95% Confidence Intervals for  $\beta_k$

Horizon	$\hat{\beta}_k$	GMM	Subsampling
<i>NYSE equal-weighted, 12/1947 to 12/1986</i>			
$k = 1$	.28	[-.05, .61]	[-1.11, 1.67]
$k = 12$	4.54	[-.32, 9.40]	[-11.99, 21.07]
$k = 24$	8.70	[-.41, 17.81]	[-24.55, 41.94]
$k = 36$	11.32	[.72, 21.92]	[-17.90, 40.55]
$k = 48$	13.24	[2.59, 23.88]	[-10.55, 37.02]
<i>NYSE value-weighted, 12/1947 to 12/1986</i>			
$k = 1$	.41	[.11, .72]	[-.45, 1.28]
$k = 12$	5.67	[2.39, 8.95]	[-5.22, 16.57]
$k = 24$	10.40	[3.72, 17.07]	[-5.54, 26.34]
$k = 36$	13.81	[6.62, 20.99]	[.08, 27.53]
$k = 48$	17.37	[10.07, 24.67]	[7.50, 27.24]
<i>S&amp;P 500, 12/1947 to 01/1995</i>			
$k = 1$	.45	[.18, .72]	[-.49, 1.39]
$k = 12$	5.47	[3.10, 7.84]	[-5.11, 16.05]
$k = 24$	9.41	[5.21, 13.60]	[-2.27, 21.09]
$k = 36$	12.38	[7.85, 16.90]	[2.66, 22.10]
$k = 48$	15.38	[10.82, 19.95]	[6.88, 23.89]

NOTE: This table presents 95% confidence intervals for the return regression coefficient  $\beta_k$ , together with the estimated coefficient  $\hat{\beta}_k$ . I use monthly data, and various return horizons  $k$  are considered. The confidence intervals are GMM intervals using the QS kernel and calibrated symmetric subsampling intervals.

more interested in the overall null hypothesis of no predictability than in individual hypotheses concerning particular horizons, it seems preferable to derive a test for the joint null hypothesis of  $\beta_1 = \beta_{12} = \beta_{24} = \beta_{36} = \beta_{48} = 0$ . This avoids the usual pitfalls of multiple testing. I shall deal with both questions in Section 6.

In Remark 2.2, I discussed how to use subsampling to describe the sampling distribution of  $\hat{\beta}_k - \beta_k$ . Figure 2 displays the corresponding histograms for the return horizons  $k = 1$  and  $k = 24$  and all three datasets. Note that the histograms are somewhat ragged because the approximations are based on  $n - b + 1$  values, which is on the order of 500 here. Two well-known features of the sampling distribution of  $\hat{\beta}_k$ ; namely, an upward bias and skewness to the right are clearly visible. I also indicate the corresponding point estimates  $\hat{\beta}_k$  by vertical lines. Because the lines are well within the histograms for all six scenarios,  $\beta_k = 0$  is a reasonable value. This is in accordance with the confidence intervals of Table 6.

## 6. ADDITIONAL LOOKS AT RETURN REGRESSIONS

### 6.1 A Reorganization of Long-Horizon Regressions

Because the compound  $k$ -period return is simply the sum of  $k$  one-period returns, the numerator of the regression coefficient  $\beta_k$  in Equation (3) is the same as

$$\text{cov}[\ln(R_{t+1}) + \dots + \ln(R_{t+k}), (D_t/P_t)]. \quad (11)$$

Under the assumption of stationarity, the covariance (11) is identical to

$$\text{cov}[\ln(R_{t+1}), (D_t/P_t) + \dots + (D_{t-k+1}/P_{t-k+1})], \quad (12)$$

which is the numerator of  $\tilde{\beta}_k^*$  in the following, reorganized regression:

$$\ln(R_{t+1}) = \alpha_k^* + \beta_k^* [(D_t/P_t) + \dots + (D_{t-k+1}/P_{t-k+1})] + u_{t+1}. \quad (13)$$

The test  $H_0: \beta_k = 0$  is therefore equivalent to the test  $H_0: \beta_k^* = 0$ . This fact was recognized by Hodrick (1992), among others. The advantage of the latter test is that, under the null hypothesis, the stochastic behavior of the error terms  $u_{t+1}$  in (12) is determined by the behavior of the one-period returns  $\ln(R_{t+1})$  only, regardless of the horizon  $k$ . Hence, the problem of increasing correlation in the error terms due to an increasing return horizon is eliminated.

Hodrick (1992) carried out this alternative test, using critical values obtained by simulating from a VAR model that imposes the null hypothesis. He still found evidence for predictability at horizons of one year and beyond. The problem is that the critical values might be too small because the conditionally normal GARCH innovations of the VAR model tend to underestimate the tails of the true sampling distribution.

To provide an alternative viewpoint, I apply the subsampling method. The method of inference about  $\beta_k^*$  is analogous to the method of inference about  $\beta_k$ . The results are reported in Table 7. Notice that all subsampling confidence intervals contain 0, and therefore not even at the four-year horizon could a case for predictability be made. On the other hand, all but two of the GMM intervals exclude 0.

*Remark 6.1.* An alternative approach to dealing with long return horizons was developed by Richardson and Stock (1989). Instead of thinking of the return horizon  $k$  fixed while the sample size  $n$  tends to infinity, they derived asymptotic theory under the assumption of  $k/n$  tending to a positive constant less than 1. The idea is to get a better small-sample approximation when the return horizon is large compared to the sample size. Their theory only applies to autocorrelations, however, not to regressions of log returns on another variable such as dividend yield.

### 6.2 A Joint Test for Multiple Return Horizons

I now turn to the problem of making joint inference

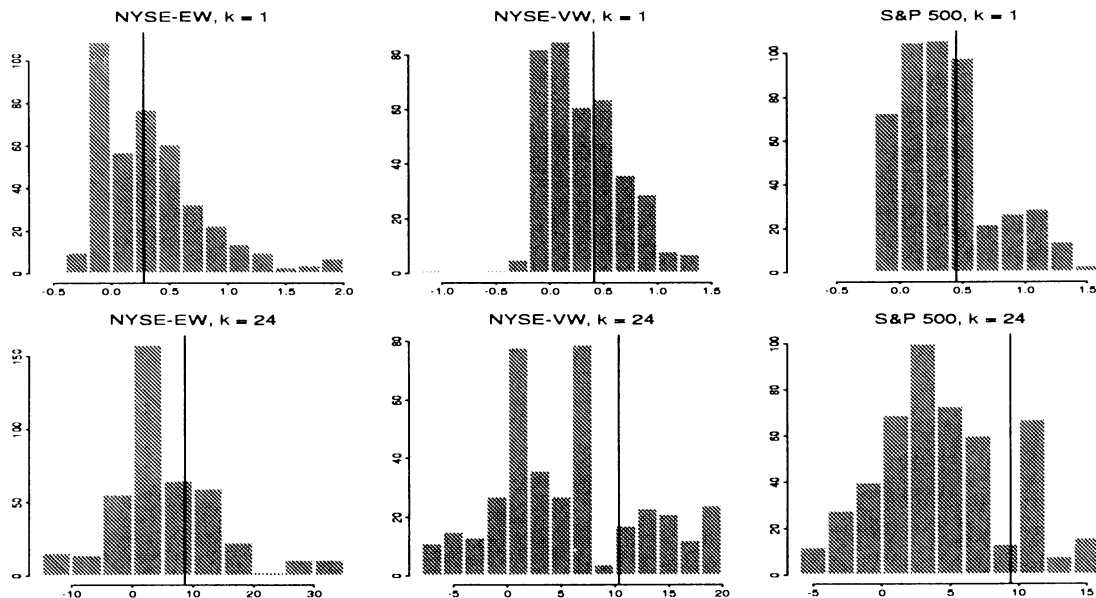


Figure 2. Histograms of the Scaled Subsampling Values  $(b/n)^{1/2}(\hat{\beta}_{k,b,a} - \hat{\beta}_k)$ , Approximating the Sampling Distribution of  $\hat{\beta}_k - \beta_k$ . The block sizes  $b$  were chosen by Algorithm 3.2. The corresponding point estimates  $\hat{\beta}_k$  are indicated by vertical lines.

Table 7. 95% Confidence Intervals for  $\beta_k^*$

Horizon	$\hat{\beta}_k^*$	GMM	Subsampling
<i>NYSE equal-weighted, 12/1947 to 12/1986</i>			
$k = 1$	.281	[-.052, .614]	[-1.109, 1.671]
$k = 12$	.035	[.006, .064]	[-.195, .266]
$k = 24$	.019	[.004, .035]	[-.237, .276]
$k = 36$	.011	[.001, .022]	[-.125, .147]
$k = 48$	.007	[-.002, .016]	[-.065, .078]
<i>NYSE value-weighted, 12/1947 to 12/1986</i>			
$k = 1$	.412	[.107, .718]	[-.453, 1.278]
$k = 12$	.043	[.017, .069]	[-.150, .235]
$k = 24$	.022	[.008, .036]	[-.104, .147]
$k = 36$	.013	[.003, .023]	[-.134, .160]
$k = 48$	.009	[.001, .017]	[-.036, .054]
<i>S&amp;P 500, 12/1947 to 01/1995</i>			
$k = 1$	.450	[.187, .713]	[-.485, 1.385]
$k = 12$	.040	[.017, .062]	[-.069, .148]
$k = 24$	.019	[.007, .030]	[-.057, .094]
$k = 36$	.011	[.003, .019]	[-.053, .075]
$k = 48$	.008	[.001, .014]	[-.031, .046]

NOTE: This table presents 95% confidence intervals for the return regression coefficient  $\beta_k^*$  of the reorganized regression (13), which avoids additional correlation in the residuals for long return horizons. The estimated coefficient  $\hat{\beta}_k^*$  is also presented. I use monthly data, and various return horizons  $k$  are considered. The confidence intervals are GMM intervals using the QS kernel and calibrated symmetric subsampling intervals.

for the collection of all  $\beta_k$  considered; that is, the vector  $\beta = (\beta_1, \beta_{12}, \beta_{24}, \beta_{36}, \beta_{48})'$ . The null hypothesis of interest is that  $\beta = \mathbf{0} \equiv (0, 0, 0, 0, 0)'$ . The joint estimation is easily done by combining the individual estimates for each horizon into a vector. The joint inference is more complicated, however. Under reasonable conditions, the vector  $\hat{\beta} = (\hat{\beta}_1, \hat{\beta}_{12}, \hat{\beta}_{24}, \hat{\beta}_{36}, \hat{\beta}_{48})'$  will have a limiting normal distribution, centered at  $\beta$ . Obviously, the limiting covariance matrix is not diagonal and therefore cannot be estimated by simply combining the individual variance estimates. To make matters worse, the explicit estimation of the  $5 \times 5$  covariance matrix of  $\hat{\beta}$  requires along the way the estimation of the  $10 \times 10$  covariance matrix of  $(\hat{\alpha}_1, \hat{\beta}_1, \dots, \hat{\alpha}_{48}, \hat{\beta}_{48})'$ . This does not appear to be a promising endeavor with sample sizes on the order of 500. Hodrick (1992) ran into this problem when he tried to estimate the limiting covariance matrix by GMM but found that "simultaneous estimation of the five equations ... results in failure of the GMM matrix to be positive definite" (p. 374). Because, in this instance, the limiting covariance matrix cannot be estimated, Hodrick was unable to test the null hypothesis of  $\beta = \mathbf{0}$ .

Fortunately, the subsampling methodology can handle multivariate parameters without much difficulty and avoids the problem of *explicit* estimation of the limiting distribution; for a general theorem, see PRW (1997). Here, I shall briefly outline how to handle the case  $\hat{\beta}$ . The unknown (multivariate) sampling distribution of  $n^{1/2}(\hat{\beta} - \beta)$  is estimated by the empirical distribution of the  $n - b + 1$  subsample statistics  $b^{1/2}(\hat{\beta}_{b,a} - \hat{\beta})$ ,  $a = 1, \dots, n - b + 1$ . With the help of a norm  $\|\cdot\|$  on  $\mathbb{R}^p$ , a confidence region for  $\beta$  can be found quite easily. Suppose a  $(1 - \alpha)$  confidence region for  $\beta$  is desired. An asymptotically correct region is given

by the collection of all vectors  $\beta^\dagger$  that satisfy

$$n^{1/2} \|\hat{\beta} - \beta^\dagger\| \leq c_{n,b,\|\cdot\|}(1 - \alpha). \quad (14)$$

Here,  $c_{n,b,\|\cdot\|}(1 - \alpha)$  denotes a  $(1 - \alpha)$  quantile of the univariate "normed" subsampling distribution  $L_{n,b,\|\cdot\|}$  having distribution function

$$L_{n,b,\|\cdot\|}(x) = \frac{1}{n - b + 1} \sum_{a=1}^{n-b+1} 1\{b^{1/2} \|\hat{\beta}_{b,a} - \hat{\beta}\| \leq x\}. \quad (15)$$

Although it would be cumbersome to explicitly exhibit such a confidence region, it is trivial to check whether a specific vector  $\beta^\dagger$  is contained in the region via examining condition (14).

The problem of choosing the block size  $b$  is analogous to the univariate case. I can use the same remedy, the calibration technique described in Section 3. The modifications of the algorithm outlined there should be obvious. Note that to do Step 2 the explicit computation of the confidence region in Step 1a is not really needed. It is only necessary to check whether  $\hat{\beta}$  is contained in the region.

When applying this method to three datasets under consideration, one has to be concerned with the magnitudes of the individual coefficients. Note that  $\hat{\beta}_k$  naturally will increase with the return horizon  $k$  because a  $k$ -horizon compounded return is predicted. It therefore seems sensible to standardize by dividing each estimated regression coefficient by its respective return horizon. Thus, the following modified Euclidean norm is employed:

$$\|(\beta_1, \beta_{12}, \dots, \beta_{48})'\|_{\text{mod}} = \sqrt{\beta_1^2 + (\beta_{12}/12)^2 + \dots + (\beta_{48}/48)^2}. \quad (16)$$

The results are reported in Table 8. For all three datasets, a block size of  $b = 80$  was used. Because the reader might wish to know some more details rather than simply whether the vector  $(0, 0, 0, 0, 0)'$  is contained in the corresponding confidence region, I provide the following information. The observed norm gives the numerical value of  $\|(\hat{\beta}_1 - 0,$

Table 8. Joint Test for  $\beta_1 = \beta_{12} = \dots = \beta_{48} = 0$

Observed norm	Observed P value	Cut-off point for .05 test
<i>NYSE equal-weighted, 12/1947 to 12/1986</i>		
.727	.518	.027
<i>NYSE value-weighted, 12/1947 to 12/1986</i>		
.927	.385	.060
<i>S&amp;P 500, 12/1947 to 01/1995</i>		
.850	.268	.081

NOTE: This table presents results for the joint test of all individual regression coefficients being equal to 0. The observed norm gives the numerical value of  $\|(\hat{\beta}_1 - 0, \hat{\beta}_{12} - 0, \dots, \hat{\beta}_{48} - 0)\|_{\text{mod}}$ , with  $\|\cdot\|_{\text{mod}}$  as defined in (16). The observed P value reports the percentage of subsample statistics  $b^{1/2} \|\hat{\beta}_{b,a} - \hat{\beta}\|_{\text{mod}}$  exceeding the scaled observed norm  $n^{1/2} \|\hat{\beta} - 0\|_{\text{mod}}$ . Finally, the cut-off point says how small the observed P value has to be to be deemed significant at the 5% level by the calibration technique of Section 3. In other words, if the observed P value is bigger than the cut-off point, then the vector  $(0, 0, 0, 0, 0)'$  is contained in the 95% confidence region.

$\hat{\beta}_{12} - 0, \dots, \hat{\beta}_{48} - 0)'_{\text{mod}}$ , with  $\|\cdot\|_{\text{mod}}$  as defined in (16). The observed  $P$  value reports the percentage of subsample statistics  $b^{1/2}\|\hat{\beta}_{b,a} - \hat{\beta}\|_{\text{mod}}$  exceeding the scaled observed norm  $n^{1/2}\|\hat{\beta} - 0\|_{\text{mod}}$ . Finally, the cut-off point says how small the observed  $P$  value has to be to be deemed significant at the 5% level by the calibration technique of Section 3. In other words, if the observed  $P$  value is bigger than the cut-off point, then the vector  $(0, 0, 0, 0, 0)'$  is contained in the 95% confidence region.

For all three datasets, the observed  $P$  value is substantially bigger than the cut-off point at the 5% level. Therefore, in none of the three cases can the null hypothesis be rejected.

*Remark 6.2.* A related usage of a multiple-horizon test was discussed by Richardson (1993). Instead of regressing stock returns on dividend yields, he considered autocorrelations of stock returns. Note that autocorrelations can be thought of regressing stock returns on past stock returns. Similarly to our findings, when Richardson compared the evidence for individual return horizons to the joint evidence for all horizons together, the latter turned out to be weaker. To carry out his test, Richardson used a Wald test statistic with which he was able to compute the limiting covariance matrix theoretically under the null hypothesis. Thereby he avoided the problem of having to explicitly estimate the matrix. Note that this approach would not work for regressions involving other variables such as dividend yield.

*Remark 6.3.* In my analysis, I have used U.S. postwar data. Goetzmann and Jorion (1995) demonstrated that regression statistics based on a sample drawn solely from surviving markets can bias the results toward finding predictability. My subsampling approach, since conditioning on samples from a surviving market, cannot address the size of this bias; see Goetzmann and Jorion (1995) for some simulations that allow markets to disappear with positive probability. This bias does not, however, weaken my results because they do not point toward predictability.

## 7. CONCLUSIONS

In this article, I presented a new statistical tool to make inference in the context of dependent and possibly nonstationary observations, as needed when examining the predictability of stock returns from dividend yields. The gist of the new method, called subsampling, is to recompute the statistic of interest on smaller blocks of the entire data sequence to approximate the sampling distribution of the estimator based on the complete data. This enables me to construct asymptotically correct confidence regions for unknown parameters under very weak conditions.

When comparing the subsampling method with previous approaches for testing the predictability of stock returns, I found it more trustworthy than the VAR approach and Goetzmann and Jorion's (1993) bootstrap on grounds of asymptotic consistency. A simulation study suggested that subsampling has better small-sample properties than GMM, which is a valid competitor in terms of asymptotic properties.

I applied the subsampling method to three different postwar datasets, the NYSE equal- and value-weighted indexes and the S&P 500 index, and included five return horizons ranging between one month and four years. I did not find any evidence for predictability for short and medium horizons, but findings at the long horizons appeared significant. Some potential undercoverage of subsampling confidence intervals for long horizons due to very strong dependencies in the residuals, as well as the issue of multiple testing, cast some doubt on this evidence.

A reorganization of long-horizon returns, avoiding increasing correlation in the residuals by means of summing dividend yields rather than returns, resulted in insignificant outcomes for all horizons. Moreover, a joint test for all five return horizons also failed to find any evidence. I therefore conclude that no convincing case for the predictability of stock returns from dividend yields can be made.

## ACKNOWLEDGMENTS

I am grateful to Ibbotson and Associates for providing me with data on the S&P 500 index and to Charles Nelson and Myung Kim for providing me with data on the NYSE index. I would like to thank Geert Bekaert, William Goetzmann, and Robert Hodrick for helpful comments and suggestions. I also got valuable feedback from seminar participants at Stanford University, UCLA, UCSB, UCSD, and USC. The final version of this article has benefited from insightful reports of two anonymous referees, an associate editor, and Mark Watson, the coeditor. Last, I am grateful to Olivier Ledoit for providing Matlab routines for fitting GARCH models. Any remaining errors are my own.

[Received April 1997. Revised April 1999.]

## REFERENCES

- Andrews, D. W. K. (1991), "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59, 817-858.
- Bollerslev, T. (1990), "Modeling the Coherence on Short Run Nominal Exchange Rates: A Multivariate GARCH Model," *Review of Economics and Statistics*, 72, 498-505.
- Bollerslev, T., Engle, R. F., and Wooldridge, J. M. (1988), "A Capital Asset Pricing Model With Time-Varying Covariances," *Journal of Political Economy*, 96, 116-131.
- Campbell, J. Y., Lo, A. W., and MacKinlay, A. C. (1997), *The Econometrics of Financial Markets*, Princeton, NJ: Princeton University Press.
- Campbell, J. Y., and Shiller, R. J. (1988a), "The Dividend Ratio Model and Small Sample Bias: A Monte Carlo Study," *Economics Letters*, 29, 325-331.
- (1988b), "Stock Prices, Earnings, and Expected Dividends," *Journal of Finance*, 43, 661-676.
- Cavanagh, C. L., Elliot, G., and Stock, J. H. (1995), "Inference in Models With Nearly Integrated Regressors," *Econometric Theory*, 11, 1131-1147.
- Elliot, G., and Stock, J. H. (1994), "Inference in Time Series Regression When the Order of Integration of a Regressor Is Unknown," *Econometric Theory*, 10, 672-700.
- Fama, E., and French, K. (1988), "Dividend Yields and Expected Stock Returns," *Journal of Financial Economics*, 22, 3-25.
- Ferson, W. E., and Foerster, S. (1994), "Finite Sample Properties of the Generalized Method of Moments in Tests of Conditional Asset Pricing Models," *Journal of Financial Economics*, 36, 29-55.
- Goetzmann, W. N., and Jorion, P. (1993), "Testing the Predictive Power of Dividend Yields," *Journal of Finance*, 48, 663-679.

- (1995), "A Longer Look at Dividend Yields," *Journal of Business*, 68, 483–508.
- Hall, P. (1988), "On Symmetric Bootstrap Confidence Intervals," *Journal of the Royal Statistical Society, Ser. B*, 50, 35–45.
- Hansen, L. (1982), "Large Sample Properties of Generalized Method of Moments Estimation," *Econometrica*, 50, 1029–1054.
- Hansen, L., and Hodrick, R. (1980), "Forward Exchange Rates as Optimal Predictors of Future Spot Rates," *Journal of Political Economy*, 88, 829–853.
- Hodrick, R. J. (1992), "Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement," *Review of Financial Studies*, 5, 357–386.
- Künsch, H. R. (1989), "The Jackknife and the Bootstrap for General Stationary Observations," *The Annals of Statistics*, 17, 1217–1241.
- Lo, A. W., and MacKinlay, A. C. (1988), "Stock Market Prices Do Not Follow Random Walks: Evidence From a Simple Specification Test," *Review of Financial Studies*, 1, 41–66.
- Loh, W.-Y. (1987), "Calibrating Confidence Coefficients," *Journal of the American Statistical Association*, 82, 155–162.
- Nelson, C. R., and Kim, M. J. (1993), "Predictable Stock Returns: The Role of Small Sample Bias," *Journal of Finance*, 48, 641–661.
- Politis, D. N., and Romano, J. P. (1994), "Large Sample Confidence Regions Based on Subsamples Under Minimal Assumptions," *The Annals of Statistics*, 22, 2031–2050.
- Politis, D. N., Romano, J. P., and Wolf, M. (1997), "Subsampling for Heteroskedastic Time Series," *Journal of Econometrics*, 81, 281–317.
- Richardson, M. (1993), "Temporary Components of Stock Prices: A Skeptic's View," *Journal of Business & Economic Statistics*, 11, 199–207.
- Richardson, M., and Stock, J. H. (1989), "Drawing Inferences From Statistics Based on Multiyear Asset Returns," *Journal of Financial Economics*, 25, 323–348.
- Rozeff, M. (1984), "Dividend Yields Are Equity Risk Premium," *Journal of Portfolio Management*, 11, 68–75.
- Sherman, M., and Carlstein, E. (1996), "Replicate Histograms," *Journal of the American Statistical Association*, 91, 566–576.
- Stambaugh, R. (1986), "Biases in Regressions With Lagged Stochastic Regressors," Working Paper 156, The University of Chicago, Graduate School of Business Administration.
- Viceira, L. M. (1997), "Testing for Structural Change in the Predictability of Asset Returns," preprint, Harvard University, Graduate School of Business Administration.

## LINKED CITATIONS

- Page 1 of 4 -



You have printed the following article:

### **Stock Returns and Dividend Yields Revisited: A New Way to Look at an Old Problem**

Michael Wolf

*Journal of Business & Economic Statistics*, Vol. 18, No. 1. (Jan., 2000), pp. 18-30.

Stable URL:

<http://links.jstor.org/sici?sici=0735-0015%28200001%2918%3A1%3C18%3ASRADYR%3E2.0.CO%3B2-4>

---

*This article references the following linked citations. If you are trying to access articles from an off-campus location, you may be required to first logon via your library web site to access JSTOR. Please visit your library's website or contact a librarian to learn about options for remote access to JSTOR.*

## References

### **Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation**

Donald W. K. Andrews

*Econometrica*, Vol. 59, No. 3. (May, 1991), pp. 817-858.

Stable URL:

<http://links.jstor.org/sici?sici=0012-9682%28199105%2959%3A3%3C817%3AHAACCM%3E2.0.CO%3B2-P>

### **Modelling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized Arch Model**

Tim Bollerslev

*The Review of Economics and Statistics*, Vol. 72, No. 3. (Aug., 1990), pp. 498-505.

Stable URL:

<http://links.jstor.org/sici?sici=0034-6535%28199008%2972%3A3%3C498%3AMTCISN%3E2.0.CO%3B2-N>

### **A Capital Asset Pricing Model with Time-Varying Covariances**

Tim Bollerslev; Robert F. Engle; Jeffrey M. Wooldridge

*The Journal of Political Economy*, Vol. 96, No. 1. (Feb., 1988), pp. 116-131.

Stable URL:

<http://links.jstor.org/sici?sici=0022-3808%28198802%2996%3A1%3C116%3AACAPMW%3E2.0.CO%3B2-7>

## LINKED CITATIONS

- Page 2 of 4 -



### **Stock Prices, Earnings, and Expected Dividends**

John Y. Campbell; Robert J. Shiller

*The Journal of Finance*, Vol. 43, No. 3, Papers and Proceedings of the Forty-Seventh Annual Meeting of the American Finance Association, Chicago, Illinois, December 28-30, 1987. (Jul., 1988), pp. 661-676.

Stable URL:

<http://links.jstor.org/sici?sici=0022-1082%28198807%2943%3A3%3C661%3ASPEAED%3E2.0.CO%3B2-1>

### **Testing the Predictive Power of Dividend Yields**

William N. Goetzmann; Philippe Jorion

*The Journal of Finance*, Vol. 48, No. 2. (Jun., 1993), pp. 663-679.

Stable URL:

<http://links.jstor.org/sici?sici=0022-1082%28199306%2948%3A2%3C663%3ATTPPOD%3E2.0.CO%3B2-7>

### **A Longer Look at Dividend Yields**

William N. Goetzmann; Philippe Jorion

*The Journal of Business*, Vol. 68, No. 4. (Oct., 1995), pp. 483-508.

Stable URL:

<http://links.jstor.org/sici?sici=0021-9398%28199510%2968%3A4%3C483%3AALLADY%3E2.0.CO%3B2-D>

### **On Symmetric Bootstrap Confidence Intervals**

Peter Hall

*Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 50, No. 1. (1988), pp. 35-45.

Stable URL:

<http://links.jstor.org/sici?sici=0035-9246%281988%2950%3A1%3C35%3AOSBCI%3E2.0.CO%3B2-2>

### **Large Sample Properties of Generalized Method of Moments Estimators**

Lars Peter Hansen

*Econometrica*, Vol. 50, No. 4. (Jul., 1982), pp. 1029-1054.

Stable URL:

<http://links.jstor.org/sici?sici=0012-9682%28198207%2950%3A4%3C1029%3ALSPOGM%3E2.0.CO%3B2-O>



## LINKED CITATIONS

- Page 3 of 4 -



### **Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis**

Lars Peter Hansen; Robert J. Hodrick

*The Journal of Political Economy*, Vol. 88, No. 5. (Oct., 1980), pp. 829-853.

Stable URL:

<http://links.jstor.org/sici?sici=0022-3808%28198010%2988%3A5%3C829%3AFERAOP%3E2.0.CO%3B2-J>

### **Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement**

Robert J. Hodrick

*The Review of Financial Studies*, Vol. 5, No. 3. (1992), pp. 357-386.

Stable URL:

<http://links.jstor.org/sici?sici=0893-9454%281992%295%3A3%3C357%3ADYAESR%3E2.0.CO%3B2-C>

### **The Jackknife and the Bootstrap for General Stationary Observations**

Hans R. Kunsch

*The Annals of Statistics*, Vol. 17, No. 3. (Sep., 1989), pp. 1217-1241.

Stable URL:

<http://links.jstor.org/sici?sici=0090-5364%28198909%2917%3A3%3C1217%3ATJATBF%3E2.0.CO%3B2-%23>

### **Stock Market Prices do not Follow Random Walks: Evidence from a Simple Specification Test**

Andrew W. Lo; A. Craig MacKinlay

*The Review of Financial Studies*, Vol. 1, No. 1. (Spring, 1988), pp. 41-66.

Stable URL:

<http://links.jstor.org/sici?sici=0893-9454%28198821%291%3A1%3C41%3ASMPDNF%3E2.0.CO%3B2-U>

### **Calibrating Confidence Coefficients**

Wei-Yin Loh

*Journal of the American Statistical Association*, Vol. 82, No. 397. (Mar., 1987), pp. 155-162.

Stable URL:

<http://links.jstor.org/sici?sici=0162-1459%28198703%2982%3A397%3C155%3ACCC%3E2.0.CO%3B2-C>

### **Predictable Stock Returns: The Role of Small Sample Bias**

Charles R. Nelson; Myung J. Kim

*The Journal of Finance*, Vol. 48, No. 2. (Jun., 1993), pp. 641-661.

Stable URL:

<http://links.jstor.org/sici?sici=0022-1082%28199306%2948%3A2%3C641%3APSRTRO%3E2.0.CO%3B2-9>

## LINKED CITATIONS

- Page 4 of 4 -



### **Large Sample Confidence Regions Based on Subsamples under Minimal Assumptions**

Dimitris N. Politis; Joseph P. Romano

*The Annals of Statistics*, Vol. 22, No. 4. (Dec., 1994), pp. 2031-2050.

Stable URL:

<http://links.jstor.org/sici?sici=0090-5364%28199412%2922%3A4%3C2031%3ALSCRBO%3E2.0.CO%3B2-A>

### **Temporary Components of Stock Prices: A Skeptic's View**

Matthew Richardson

*Journal of Business & Economic Statistics*, Vol. 11, No. 2. (Apr., 1993), pp. 199-207.

Stable URL:

<http://links.jstor.org/sici?sici=0735-0015%28199304%2911%3A2%3C199%3ATCOSP%3E2.0.CO%3B2-N>

### **Replicate Histograms**

Michael Sherman; Edward Carlstein

*Journal of the American Statistical Association*, Vol. 91, No. 434. (Jun., 1996), pp. 566-576.

Stable URL:

<http://links.jstor.org/sici?sici=0162-1459%28199606%2991%3A434%3C566%3ARH%3E2.0.CO%3B2-H>