City Research Online

City St George's, University of London

Citation: Mammen, E., Nielsen, J. P., Scholz, M. & Sperlich, S. (2019). Conditional variance forecasts for long-term stock returns. Risks, 7(4), 113. doi: 10.3390/risks7040113

This is the published version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/23329/

Link to published version: https://doi.org/10.3390/risks7040113

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to City Research Online policy.

City Research Online: http://openaccess.city.ac.uk/ publications@citystgeorges.ac.uk

risks

MDPI

Article

Conditional Variance Forecasts for Long-Term Stock Returns

Enno Mammen [1], Jens Perch Nielsen [2], Michael Scholz [3,*] and Stefan Sperlich [4]

[1] Institute for Applied Mathematics, Heidelberg University, Im Neuenheimer Feld 205, 69120 Heidelberg, Germany; mammen@math.uni-heidelberg.de
[2] Faculty of Actuarial Science and Insurance, Cass Business School, 106 Bunhill Row, London EC1Y 8TZ, UK; jens.nielsen.1@city.ac.uk
[3] Department of Economics, University of Graz, Universitätsstraße 15/F4, 8010 Graz, Austria
[4] Geneva School of Economics and Management, Université de Genève, Bd du Pont d'Arve 40, 1211 Genève, Switzerland; stefan.sperlich@unige.ch
* Correspondence: michael.scholz@uni-graz.at; Tel.: +43-316-380-7112

Received: 20 August 2019; Accepted: 29 October 2019; Published: 5 November 2019

Abstract: In this paper, we apply machine learning to forecast the conditional variance of long-term stock returns measured in excess of different benchmarks, considering the short- and long-term interest rate, the earnings-by-price ratio, and the inflation rate. In particular, we apply in a two-step procedure a fully nonparametric local-linear smoother and choose the set of covariates as well as the smoothing parameters via cross-validation. We find that volatility forecastability is much less important at longer horizons regardless of the chosen model and that the homoscedastic historical average of the squared return prediction errors gives an adequate approximation of the unobserved realised conditional variance for both the one-year and five-year horizon.

Keywords: benchmark; cross-validation; prediction; stock return volatility; long-term forecasts; overlapping returns; autocorrelation

JEL Classification: C14; C53; C58; G17; G22

1. Introduction

The volatility of financial assets has important implications for the theory and practice of asset pricing, portfolio selection, risk management, and market-timing strategies. Therefore, it is of fundamental interest to measure ex ante, or forecast successfully, the conditional variance of returns. Of course, the evaluation of the latter and the forecasting itself have been complicated by the unobservability of the realised conditional variance (Galbraith and Kisinbay 2005). An extensive amount of research is engaged in analysing the distributional and dynamic properties of stock market volatility; see, for example, Andersen et al. (2001) and citations therein. The standard approaches applied include parametric (G)ARCH-type or stochastic volatility models and estimate the underlying returns based on specific distributional assumptions. Alternatives, especially for data of higher frequency, are based on constructing model-free estimates of ex-post *realized* volatilities by adding up the squares and cross-products of intraday high-frequency returns (Andersen et al. 2001).

The present paper instead uses annual U.S. stock market data to construct excess stock returns at the one-year and five-year horizon and to examine their model-based variance forecasts. Note that the risk depends on the investment horizon considered and that different horizons are relevant for different applications (Christoffersen and Diebold 2000). Little is known about the forecastability of variance at horizons beyond a year. Here, we take the long-term actuarial view

and extend the work of Kyriakou et al. (2019a, 2019b). In a two-step procedure, we first apply machine learning (ML) to predict stock returns in excess of different benchmarks, considering the short- and long-term interest rate, the earnings-by-price ratio, and the inflation rate. Second, the squared residuals are used to analyse model-based volatility forecastability. Here, we compare these forecasts with the forecast implicit in the unconditional residual variance, as proposed, for example, by Galbraith and Kisinbay (2005). We find that volatility forecastability is much less important at longer horizons regardless of the chosen model and that the homoscedastic historical average of the squared return prediction errors gives an adequate approximation of the unobserved realised conditional variance for both the one-year and five-year horizon.

Our preferred ML technique applied in this paper is local-linear smoothing in combination with a leave-*k*-out cross-validation for the following reasons.[1] First, we are interested in longer-horizon stock returns based on annual observations and their volatility. Thus, we are not in the high-frequency context where the number of observations is huge and the set of possible predictive variable combinations is enormous (and, thus, dimension reduction or shrinkage are indispensable). Our data set is, instead, sparse and a careful imposition of structure to the statistical modelling process is much more promising, as shown, for example, by Nielsen and Sperlich (2003) and Scholz et al. (2015, 2016). Second, the evidence of stock return predictability is much stronger once one allows for nonlinear functions as documented, for example, in Lettau and Van Nieuwerburgh (2008), Chen and Hong (2010), or Cheng et al. (2019). Thus, the local-linear smoother is ideally suited as it can estimate a linear function—the classical benchmark in this context—without any bias. Finally, our procedures are analytically well studied, i.e., sound and rigorous, statistical tools which let us operate in a glasshouse, not in a black box—in contrast to other fancier but less clear ML methods.[2]

Note further that longer horizons are important to long-term investors, such as pension funds or market participants saving for distant payoffs. These investors are generally willing to take on more risk for higher rewards and, thus, volatility forecastability is for them of fundamental interest. Rapach and Zhou (2013) show that longer horizons tend to produce better estimates than shorter horizons, while Munk and Rangvid (2018) point out that major finance houses today use longer horizons—up to ten years—to stabilise and improve future predictions. In our paper, we exemplarily concentrate on the one-year and five-year view.[3] However, shorter horizons based on monthly, weekly, or even daily data do not seem to provide the pension saver with good information about future income as a pensioner. Therefore, these type of short-term predictions—sometimes called investment robots—are not suitable when a pensioner should define his or her risk appetite.

The remaining of this paper is organized as follows. Section 2 presents our framework for the purpose of conditional variance prediction. We define the underlying financial model, introduce our two-step procedure, and present our validation criterion for model selection. In addition, we review different ways of estimating the conditional variance and discuss bootstrap-tests for the null hypothesis of no predictability. In Section 3, we provide a description of our data set and of our empirical findings from different validated scenarios: (i) a single benchmarking approach that uses the dependent variable transformed with the benchmark, and (ii) the case where both the independent and dependent variables are transformed with the benchmark (full benchmarking approach). Finally, we take the

---

1 Our methodology of validating a fully nonparametric structure can be viewed as one of the simplest and therefore also most transparent version of machine learning; see Section 2 of Kyriakou et al. (2019a) for more details justifying the label machine learning for our approach.

2 Note that the use of a different ML method would come with the cost of losing interpretability, smoothness, or flexibility due to restrictions on the functional form. A comparison of different ML techniques in finding that one which gives the best predictions, wins an investment horse-race out-of-sample, or being the most robust method over different periods is out of the scope of our work.

3 The choice of the one-year horizon is related to the frequency of the data. In contrast, the five-year horizon is arbitrary but is intended to be a starting point for actuarial long-term models for real-income savings. Other horizons and related questions remain for future research.

long-term view and comment on real income pension prediction. Section 4 summarizes the key points of our analysis and concludes the paper.

2. A Framework for Conditional Variance Prediction

In this section, we focus on nonlinear predictive relationships between squared residuals of model-based predicted stock returns over the next T years in excess of a benchmark and a set of explanatory variables. Our aim is the investigation of different benchmark models and their volatility predictability over return horizons of one year and five years. We consider four different benchmarks: the short- and the long-term interest rate, the earnings-by-price ratio, and the inflation rate.

2.1. One-Year Predictions

Let P_t denote the (nominal) stock price at the end of year t and D_t the (nominal) dividends paid during year t. We investigate stock returns $S_t = (P_t + D_t)/P_{t-1}$ in excess (log-scale) of a given benchmark $B_{t-1}^{(A)}$:

$$Y_t^{(A)} = \ln \frac{S_t}{B_{t-1}^{(A)}}, \tag{1}$$

where $A \in \{R, L, E, C\}$ with, respectively,

$$B_t^{(R)} = 1 + \frac{R_t}{100}, \quad B_t^{(L)} = 1 + \frac{L_t}{100}, \quad B_t^{(E)} = 1 + \frac{E_t}{P_t}, \quad B_t^{(C)} = \frac{CPI_t}{CPI_{t-1}},$$

using the short-term interest rate, R_t, the long-term interest rate, L_t, the earnings accruing to the index in year t, E_t, and the consumer price index for year t, CPI_t. The predictive and fully nonparametric regression model for a one-year horizon is then given by the location-scale model

$$Y_t^{(A)} = m(X_{t-1}) + v(X_{t-1})^{1/2}\zeta_t, \tag{2}$$

where

$$m(x) = \mathbb{E}(Y^{(A)}|X = x) \text{ and } v(x) = Var(Y^{(A)}|X = x),\ x \in \mathbb{R}^q \tag{3}$$

are unknown smooth functions for the conditional mean and variance, resp., ζ_t are serially uncorrelated zero-conditional-mean random error terms, given the past, with the conditional variance of one, and X_{t-1} is a q-dimensional vector of available explanatory variables.[4]

Our aim is to forecast the conditional variance of excess stock returns $Y_t^{(A)}$ based on model (2) and popular explanatory variables with predictive power reported in the literature, for example, the dividend-by-price ratio, $d_{t-1} = D_{t-1}/P_{t-1}$, the earnings-by-price ratio, $e_{t-1} = E_{t-1}/P_{t-1}$, the short-term interest rate, $r_{t-1} = R_{t-1}/100$, the long-term interest rate, $l_{t-1} = L_{t-1}/100$, inflation, $\pi_{t-1} = (CPI_{t-1} - CPI_{t-2})/CPI_{t-2}$, the term spread, $s_{t-1} = l_{t-1} - r_{t-1}$, and lagged excess stock return, $Y_{t-1}^{(A)}$.

Based on (2), in a two-step procedure, we first estimate $\hat{Y}_t^{(A)} = \hat{m}(X_{t-1})$ as in Kyriakou et al. (2019b), and, in a second step, we estimate $\hat{v}(X_{t-1})$ from

$$v(x) = \mathbb{E}((Y^{(A)} - m(X))^2|X = x),\ x \in \mathbb{R}^q, \tag{4}$$

using the squared residuals $\hat{\varepsilon}_t^2 := (Y_t^{(A)} - \hat{m}(X_{t-1}))^2$ as the dependent variable and a local-linear smoother in both steps. The estimates $\hat{m}$ and $\hat{v}$ depend on smoothing parameters (bandwidths) h and g,

[4] Note that the set of explanatory variables in (2) could be different or overlapping for the mean and variance function.

respectively. As we are interested in predictions, we take the values which minimize the out-of-sample prediction error using cross-validation. More details are provided in Section 2.4.[5]

2.2. Longer-Horizon Predictions

For longer horizons T, we consider the sum of annual continuously compounded returns:

$$Z_t^{(A)} = \sum_{i=0}^{T-1} Y_{t+i}^{(A)}.$$

Note that we use here overlapping returns $Z_t^{(A)}$, which require a careful econometric modelling. For illustrative purposes, assume a linear relationship in (2) between $Y_t^{(A)}$ and X_{t-1}, as well as the persistence of the forecasting variable (treating the variables as deviations from their means):

$$Y_t^{(A)} = \beta X_{t-1} + \xi_t \quad \text{and} \quad X_t = \gamma X_{t-1} + \eta_t,$$

with $\xi_t := \nu_\theta(X_{t-1})^{1/2}\zeta_t$ similar to the error term in (2) and a parametric specification for the conditional variance $\nu_\theta(\cdot)$, and η_t being white noise. The T-year regression problem that is implied by this pair of one-year regressions is now

$$\begin{aligned} Z_t^{(A)} &= Y_t^{(A)} + \ldots + Y_{t+T-1}^{(A)} = (\beta X_{t-1} + \xi_t) + \ldots + (\beta X_{t+T-2} + \xi_{t+T-1}) \\ &= \beta \sum_{i=0}^{T-1} \gamma^i X_{t-1} + \beta \sum_{i=0}^{T-1} \sum_{j=0}^{T-1-i} \gamma^j \eta_{t+i} + \sum_{i=0}^{T-1} \xi_{t+i} = \phi X_{t-1} + \psi_t, \end{aligned}$$

i.e., the excess stock return for the year t over the next T years can be decomposed in a predictive part depending on the variable X_{t-1} and an unpredictable error term ψ_t. In estimating the conditional mean and variance functions for the T-year returns $Z_t^{(A)}$, we use nonparametric models because they can capture possible misspecification due to violation of the linear models assumed above. Thus, we set up our predictive nonparametric regression model in the same fashion as in (2)

$$Z_t^{(A)} = m(X_{t-1}) + \nu(X_{t-1})^{1/2}\omega_t, \tag{5}$$

where

$$m(x) = \mathbb{E}(Z^{(A)}|X = x) \text{ and } \nu(x) = Var(Z^{(A)}|X = x),\ x \in \mathbb{R}^q \tag{6}$$

are the unknown smooth conditional mean- and variance-function. The predictive variables X under consideration are the same as for the one-year horizon. The important difference between Equations (2) and (5) is now that the error process $\psi_t := \nu(X_{t-1})^{1/2}\omega_t$ in Equation (5) will be serially correlated by

---

[5] For a description and statistical properties of the local-linear smoother, see, for example, Section 2.3 in Kyriakou et al. (2019b). Note further that the smoothing parameters h and g are separately chosen in each step.

construction.[6,7] For a discussion on asymptotic properties of our nonparametric estimators of model (5) and (6), see Section 2.3 in Kyriakou et al. (2019b).

Based on (5), our two-step procedure consists now of, first, estimating $\hat{Z}_t^{(A)} = \hat{m}(X_{t-1})$, and second, estimating $\hat{v}(X_{t-1})$ from

$$v(x) = \mathbb{E}((Z^{(A)} - m(X))^2 | X = x), \; x \in \mathbb{R}^q, \tag{7}$$

using the squared residuals $\hat{\varepsilon}_t^2 := (Z_t^{(A)} - \hat{m}(X_{t-1}))^2$ as the dependent variable and a local-linear smoother again in both steps.

2.3. Alternative Ways in Estimating the Conditional Variance Function

For the estimation of the conditional variance or volatility function of a response variable Y in a location-scale model similar to (2) or (5), four different approaches are mainly proposed in the literature: the direct, the residual-based, the likelihood-based, and the difference-sequence method.

(i) The direct method uses the variance expressed as the difference of the first two conditional moments (see, for example, Härdle and Tsybakov 1997): $Var(Y|X = x) = \mathbb{E}(Y^2|X = x) - \mathbb{E}(Y|X = x)^2$. Both parts of the right-hand side are separately estimated and, thus, the result is not necessarily nonnegative and also not fully adaptive to the mean function.[8]

(ii) The residual-based method consists of two stages—first, estimating the conditional mean function $m(\cdot)$ and calculating the squared residuals $\hat{\varepsilon}^2 = (Y - \hat{m}(X))^2$. Second, estimating the conditional variance function $v(\cdot)$ by regressing $\hat{\varepsilon}^2$ on a set of explanatory variables X. There exist different variants of residual based methods for the second step.[9]

(iii) The preferred estimators of Yu and Jones (2004) build on a localised normal likelihood and use a standard local-linear form for estimating the mean, a local log-linear form for estimating the variance, and allow for separating bandwidths for mean and variance estimation.

(iv) Finally, examples for the difference-sequence method in a fixed design can be found for the homoscedastic case in Wang and Yu (2017) and citations therein. Wang et al. (2008) analyse for the heteroscedastic case the effect of the unknown (smooth) mean function on the estimation of the variance function. They also compare the performance of the residual-based estimators to a first-order-difference-based estimator. Their results indicate that it is not desirable to estimate the variance function based on the residuals from an optimal estimator of the mean in case the mean function is not smooth. Wang et al. (2008) recommend instead an estimator for the mean with minimal bias.

---

[6] Our flexible location-scale model in (5), could be easily extended to time-lags of higher order. However, in the empirical application in Section 3, we see that, for example, for real-earnings—the main driver of real-returns—an AR1-type model is ideally suited. This is in line with findings from Kothari et al. (2006). Note further that one might expect risk and return to be somehow related (see, for example, Merton 1973). The parametric GARCH-in-Mean process captures this idea (Linton and Yan 2011). However, the inclusion of an interaction of mean and variance in a fully nonparametric fashion is out of the scope of this paper. To our knowledge, only semiparametric versions where either the mean or variance function is modeled parametrically can be found in the literature, see, for example, Linton and Perron (2003); Pagan and Hong (1991); Pagan and Ullah (1988).

[7] For possible solutions to the problem of autocorrelation, see, for example, Xiao et al. (2003), Su and Ullah (2006), Linton and Mammen (2008), or more recently Geller and Neumann (2018). The implementation and analysis of these techniques remain for future research. In our approach, we account for autocorrelation in the validation criterion with a leave-k-out strategy, where $k = 2T - 1$; see Section 2.4.

[8] It does not estimate the volatility function as efficiently as if the true mean were known.

[9] Examples of these variants are: (i) Applying a local-linear kernel smoother in both stages (Fan and Yao 1998). The result is again not necessarily nonnegative but asymptotically fully adaptive to the unknown mean function. (ii) Using the local exponential estimator to ensure nonnegativity (Ziegelmann 2002). (iii) Implementing a combined estimator (a multiplicative bias reduction technique), where a parametric guide captures some roughness features of the unknown variance function (Glad 1998; Mishra et al. 2010). (iv) Utilising a re-weighted local constant estimator maximising the empirical likelihood such that it becomes a bias-reducing moment restriction (Xu and Phillips 2011).

In the empirical part of this paper in Section 3, we show the results of the residual-based method applying a local-linear kernel smoother in both stages. As a robustness check, we have implemented in the second step the local-exponential estimator (Ziegelmann 2002) and the combined estimator (Mishra et al. 2010) getting almost always very similar results.[10] We do not consider: (i) the direct method, since it is not fully adaptive to the mean function, (ii) the re-weighted local constant estimator (Xu and Phillips 2011) due to its asymptotic similarity to the local-linear method, (iii) the method based on the assumption of normal error terms (Yu and Jones 2004), since skewness and excess kurtosis are common properties of stock returns, and (iv) the difference-sequence method, since it was not convincingly performing in a small sample study, the mean functions are rather smooth in our problem, and bias reduction is key due to sparsity.[11]

2.4. The Validation Criterion for the Choice of Smoothing Parameters and Model Selection

For the nonparametric technique applied in this study, we require an adequate measure of predictive power. In-sample measures, such as the classical R^2 or the adjusted R^2, are not appropriate because they either prefer the most complex model or need a degrees of freedom adjustment which is an unclear concept in nonparametric estimation. Furthermore, our focus lies on prediction. Thus, we are interested in the out-of-sample performance of a model and not in how well it explains the variation inside the sample. Therefore, our preferred measure estimates the prediction error directly.

For the purpose of model selection and optimal bandwidth choice, we use the validated R^2_V introduced in the actuarial literature by Nielsen and Sperlich (2003) and based on a leave-k-out cross-validation. Note that this criterion is very similar to the forecast content function of Galbraith (2003) and Galbraith and Kisinbay (2005) defined as the proportionate reduction in the mean square forecast error achievable relative to the unconditional mean forecast.

Our validation criteria for the first and second step are defined as

$$R^2_{V,m} = 1 - \frac{\sum_t (Z_t^{(A)} - \hat{m}_{-t})^2}{\sum_t (Z_t^{(A)} - \bar{Z}_{-t}^{(A)})^2} \quad \text{and} \quad R^2_{V,\nu} = 1 - \frac{\sum_t (\hat{\varepsilon}_t^2 - \hat{v}_{-t})^2}{\sum_t (\hat{\varepsilon}_t^2 - \overline{\hat{\varepsilon}^2_{-t}})^2}. \tag{8}$$

Note that leave-k-out estimators are used: $\hat{m}_{-t}$ and $\hat{v}_{-t}$ for the nonparametric functions m and v, resp., $\bar{Z}_{-t}^{(A)}$ and $\overline{\hat{\varepsilon}^2_{-t}}$ for the unconditional mean of $Z_t^{(A)}$ and $\hat{\varepsilon}_t^2$, resp. These are computed by removing $k = 2T - 1$ observations: $(T-1)$ before the tth time point, t itself, and $(T-1)$ after t. We need to exclude $k = 2T - 1$ data points due to the construction of the dependent variable over a horizon of T years, i.e., we use for the one-year horizon the classical leave-one-out estimator, while, for example, for the five-year horizon the leave-nine-out estimator. Note that the validated R^2_V measures the predictive power of a model in comparison to the predictive power of the cross-validated historical mean. Thus, positive values imply that the regression model based on explanatory variables outperforms the corresponding historical average over T years. Negative values in the first step of our approach suggest that the historical mean return should be preferred over a model-based approach, while negative values in the second step indicate a constant homoscedastic conditional variance forecast. Note further that the numerator in the ratio of $R^2_{V,m}$ and $R^2_{V,\nu}$ corresponds to the classical cross-validation criterion. Thus, choosing the bandwidth which minimizes this criterion for a given set of explanatory variables is equivalent in maximizing the validated R^2_V. This means that we can use the validated R^2_V as a single criterion for both purposes: model and bandwidth selection.[12]

---

[10] Those results are available upon request by the authors.

[11] There is also a lack of studies using the difference-sequence method in a random design and in multivariate problems as in our case.

[12] Model selection in the sense of composition of the set of explanatory variables.

It is well known from the literature that cross-validation often requires to omit more than one observation and, possibly, additional correction when the omitted fraction of data are considerable (see, for example, Burman et al. 1994). In addition, when serial correlation arises, as in our longer-horizon application, and the structure of the error terms is ignored, De Brabanter et al. (2011) show that automatic methods for the choice of smoothing parameters, such as cross-validation or plug-in, fail. The problem is that the chosen bandwidths become smaller for increasing correlations (Opsomer et al. 2001), and the corresponding model fits become progressively more under-smoothed. The bias of the predictor reduces this way and, as it contributes in a squared fashion to the prediction mean squared error—the numerator of the ratio in (8), R_V^2 increases (not because the fit is good but due to the ignored correlation structure). A misleading decision on the bandwidth or model specification, as well the set of preferred covariates is the consequence. To overcome those problems, Chu and Marron (1991) propose the use of bimodal kernel functions. Such functions are known to remove the correlation structure very effectively, but the estimator $\hat{m}$ suffers from increased mean squared error, as discussed in De Brabanter et al. (2011). They also propose correlation-corrected cross-validation that consists of, first, finding the amount of data k to be left out in the estimation process when a bimodal kernel function is used; and, second, applying the actual choice of the smoothing parameter using leave-k-out cross-validation with a unimodal kernel function. In our application, we can skip the first step because k is known by construction. For example, in the five-year case, we have $Z_t^{(A)} = Y_t^{(A)} + \ldots + Y_{t+4}^{(A)}$. Now, we want to exclude the complete information included at time t, i.e., skip all $Z_s^{(A)}$ that include any of $Y_t^{(A)}, \ldots, Y_{t+4}^{(A)}$; it is easy to see that this amounts to a leave-nine-out set of $Z_{t-4}^{(A)}, \ldots, Z_{t+4}^{(A)}$ (see, for example, Kyriakou et al. 2019b, Figure 1).

2.5. A Bootstrap-Test: No Predictability vs. Predictability of the Conditional Variance

We test the null of no predictability of the conditional variance applying the tests proposed by Kreiss et al. (2008) (hereafter KNY-test) and Scholz et al. (2015) (hereafter SNS-test). Formally, this is equivalent to say that, under the null, ν is a constant function, which essentially corresponds to the historical average of the squared residuals, i.e., constant volatility. In particular, let $\nu(\cdot)$ be the true volatility function as in (2) or (5) for some specified set of regressors X_t, i.e., (4) or (7) holds. Let $\overline{\hat{\varepsilon}^2}$ be the sample mean of the squared residuals from step one in our approach. The KNY-test is based on the distance

$$\int \left| \nu(x) - \overline{\hat{\varepsilon}^2} \right|^2 w(x)\, dx, \tag{9}$$

for some weighting function w, which has been studied by several authors and statistics have been derived from the above, for example, in Härdle and Mammen (1993) or Kreiss et al. (2008). We use the statistic derived in Equation 2.3 of Kreiss et al. (2008)

$$h^{q/2} T \int \left| \frac{1}{T} \sum_{t=1}^{T} K_h(x - X_t) \left(\hat{\varepsilon}_t^2 - \overline{\hat{\varepsilon}^2} \right) \right|^2 w(x)\, dx, \tag{10}$$

where $K_h(x)$ is a symmetric kernel smoother with bandwidth h. The bandwidth is selected using R_V^2 for the Nadaraya–Watson kernel estimator rather than a local-linear one. We choose w to be proportional to the uniform density with support in the range of the sample data and replace integration by the mean over uniform independent observations $X'_1, X'_2, \ldots, X'_N$ in the range of the data:

$$\tau := \frac{h^{q/2} T}{N} \sum_{i=1}^{N} \left| \frac{1}{T} \sum_{t=1}^{T} K_h(X'_i - X_t) \left(\hat{\varepsilon}_t^2 - \overline{\hat{\varepsilon}^2} \right) \right|^2. \tag{11}$$

Then, the error in the integral is $O\left(N^{-1/2}\right)$ (Geweke 1996). Under the null, the above test statistic τ is small. This choice could lead to a statistic whose power is lower than the one in

Härdle and Mammen (1993) due to some implicit over-smoothing resulting in the weight function w (see comment in Kreiss et al. 2008, just after their Equation 2.5). Power may also improve by using a local-linear smoother in the test. However, the theory for this has not been developed yet, so we refrain from such extension.

Critical values for τ are best derived via wild bootstrap (Härdle and Mammen 1993). For the bootstrap critical values to be consistent, the procedure needs to be independent of whether the null is true or not. Hence, in correspondence with Equation 2.10 in Kreiss et al. (2008), for $b = 1, \ldots, B$,

$$\tau^b := \frac{h^{q/2}T}{N} \sum_{i=1}^{N} \left| \frac{1}{T} \sum_{t=1}^{T} K_h \left(X_i' - X_t \right) \left[u_t^b \left(\hat{\varepsilon}_t^2 - \hat{v} \left(X_t \right) \right) \right] \right|^2, \tag{12}$$

where the u_t^b's are independent and identically distributed random variables with a mean of zero and a variance of one, for example, $u_t^b \sim N(0,1)$. To decide if we reject or not, we use as critical values the corresponding quantiles of the empirical distribution[13],

$$F^*(\tau) = \frac{1}{B} \sum_b \mathbb{1}_{\{\tau^b \leq \tau\}}. \tag{13}$$

The consistency of the procedure for stationary sequences is given in Kreiss et al. (2008).

An alternative version for a wild bootstrap test is the SNS-test proposed in Scholz et al. (2015). There the B bootstrap samples are constructed using the residuals under the null, $\iota_t^0 := \hat{\varepsilon}_t^2 - \overline{\hat{\varepsilon}^2}$, and u_t^b's as above, such that

$$\hat{\varepsilon}_t^{2,b} = \overline{\hat{\varepsilon}^2} + \iota_t^0 \cdot u_t^b.$$

Then, in each bootstrap repetition b, the cross-validated mean is calculated of the $\hat{\varepsilon}_t^{2,b}$, $t = 1, \ldots, T$, as well the estimates of the predictor-based model $\hat{v}_{-t}^b$ in order to get $R_{V,\nu}^{2,b}$ like in (8). Critical values are chosen from corresponding quantiles of the empirical distribution function similar to (13).

Both tests have their own merits. We expect the KNY-test to be more conservative and potentially with less power in comparison to the SNS-test but with clear and well-established asymptotic theory. For more discussion on standard smoothing based tests and other examples for tests of the variance function, see, for example, the survey of Gonzales-Manteiga and Crujeiras (2013).

3. Empirical Application: Conditional Variance Prediction for Stock Returns in Excess of Different Benchmarks

3.1. The Data

In this paper, we extend the analysis of Kyriakou et al. (2019b), who considered the forecasting of long-term stock returns, to conditional variance predictions. Thus, we base our predictions on the same annual US data set which is provided by Robert Shiller and can be downloaded from http://www.econ.yale.edu/~shiller/data.htm. It includes, among other variables, the Standard and Poor's (S&P) Composite Stock Price Index, the consumer price index, and interest rate data from 1872 to 2019. We use here an updated and revised version of Shiller (1989, chp. 26), which provides a detailed description of the data. Note that the risk-free rate in this data set (based on the six-month commercial paper rate until 1997 and afterwards on the six-month certificate of deposit rate, secondary market) was discontinued in 2013. We follow the strategy of Welch and Goyal (2008) and replace it by an annual yield that is based on the six-month Treasury-bill rate, secondary market, from https://fred.stlouisfed.org/series/TB6MS. This new series is only available from 1958 to 2019. In the

[13] The symbol $\mathbb{1}_A$ denotes the indicator function of an appropriate condition A, i.e., it is one when A is true and zero otherwise.

absence of information prior to 1958, we had to estimate it. To this end, we regressed the Treasury-bill rate on the risk-free rate from Shiller's data for the overlapping period 1958 to 2013, which yielded

$$\text{Treasury-bill rate} = 0.0961 + 0.8648 \times \text{commercial paper rate}$$

with an R^2 of 98.6%. Therefore, we instrumented the risk-free rate from 1872 to 1957 with the predicted regression equation. The correlation between the actual Treasury-bill rate and the predictions for the estimation period is 99.3%. Table 1 displays standard descriptive statistics for one-year and five-year returns as well as the available covariates.

Table 1. US market data (1872–2019).

	Max	Min	Mean	Sd	Skew	Exc. kurt
S&P stock price index P	2789.80	3.25	277.58	558.13	2.43	5.50
Dividend accruing to index D	53.75	0.18	6.04	10.56	2.45	6.00
Earnings accruing to index E	132.39	0.16	13.96	26.31	2.43	5.35
Dividend-by-price d	9.88	1.17	4.31	1.71	0.46	0.25
Earnings-by-price e	17.75	1.72	7.28	2.75	1.05	1.39
Short-term interest rate r	14.93	0.07	3.97	2.50	0.96	2.34
Long-term interest rate l	14.59	1.88	4.53	2.27	1.81	3.63
Inflation π	20.69	−15.65	2.23	5.96	0.26	1.60
Spread s	3.64	−3.71	0.56	1.32	−0.05	0.02
One-year excess stock returns $Y^{(R)}$	42.39	−58.26	4.58	17.28	−0.57	0.68
One-year excess stock returns $Y^{(C)}$	54.04	−48.81	6.41	18.05	−0.40	0.64
Five-year excess stock returns $Z^{(R)}$	107.27	−78.54	23.49	36.69	−0.14	−0.37
Five-year excess stock returns $Z^{(C)}$	122.96	−57.34	32.34	36.42	−0.05	−0.40

3.2. Single Benchmarking Approach

In this section, we consider a single benchmarking approach as in Kyriakou et al. (2019a, 2019b), i.e., only the dependent variable S_t is benchmark adjusted, as shown in (1), while the independent variable(s) is (are) measured on the original (nominal) scale. The models (2) and (5) are estimated in both steps with a local-linear kernel smoother using the quartic kernel. The optimal bandwidths are chosen by cross-validation, i.e., by maximizing the corresponding validation measure given by (8). Given that we apply a local-linear smoother, it should be kept in mind that the nonparametric method can estimate linear functions without any bias. Thus, the linear model is automatically embedded in our approach. This is an important observation as the linear model is the usual benchmark in financial applications. In addition, in case that the true (but in advance) unknown function is really linear, our approach would exactly pick the line against all other functional alternatives. We study the $R^2_{V,\nu}$ values based on different validated scenarios shown for the one-year horizon in Table 2 and the five-year horizon in Table 3. Here, the same predictive variables X_{t-1} are used in both steps of our approach. Note that we have only about 150 observations in our records. The small sample size clearly limits the complexity of our analysis in the sense of using higher dimensional vectors of explanatory variables. In what follows, we consider only one- and two-dimensional models. For a discussion on sparsely distributed annual observations in higher dimensions and ways to circumvent the curse-of-dimensionality, see, for example, Kyriakou et al. (2019a).

Overall, we find for the one-year horizon that only a few variables have small positive validated $R^2_{V,\nu}$'s and thus possibly some low explanatory power. For example, for the benchmarks $B^{(R)}$, $B^{(L)}$, and $B^{(E)}$, the excess stock return has the largest validated $R^2_{V,\nu}$ values for one-dimensional models (2.2%, 2.4%, and 1.5%). This finding would support an ARCH-type variance structure. For the inflation benchmark $B^{(C)}$, the model with the long-term interest rate produces the largest validated $R^2_{V,\nu}$ of 0.5%. When we apply the bootstrap tests introduced in Section 2.5, the KNY-test does not reject the null of no predictability for all cases at the 5%-level. The SNS-test rejects the null only for the $Y^{(A)}_{t-1}$

covariate under the benchmarks $B^{(R)}$, $B^{(L)}$ and $B^{(E)}$ at the 5%-level.[14] Note that the two-dimensional models do not add predictive power as the validated $R^2_{V,\nu}$ values remain in the same low range.

Table 2. Predictive power for the variance of one-year excess stock returns $Y_t^{(A)}$: the single benchmarking approach. The prediction problem is defined in (2). The same predictive variables X_{t-1} are used in the predictions for the conditional mean and variance function. The predictive power (%) is measured by $R^2_{V,\nu}$ as defined in (8). The benchmarks $B^{(A)}$ considered are based on the short-term interest rate ($A \equiv R$), long-term interest rate ($A \equiv L$), earnings-by-price ratio ($A \equiv E$), and consumer price index ($A \equiv C$). The predictive variables used are X_{t-1}, given by the dividend-by-price ratio d_{t-1}, earnings-by-price ratio e_{t-1}, short-term interest rate r_{t-1}, long-term interest rate l_{t-1}, inflation π_{t-1}, term spread s_{t-1}, excess stock return $Y_{t-1}^{(A)}$, or the possible different pairwise combinations as indicated.

Benchmark $B^{(A)}$	**Explanatory Variable(s) X_{t-1}**						
	$Y^{(A)}$	d	e	r	l	π	s
Short-term rate	2.2	−1.1	−0.6	−0.3	0.3	−1.2	−0.1
Long-term rate	2.4	−1.2	−0.6	0.3	0.6	−1.4	−0.1
Earnings-by-price	1.5	−1.3	−0.7	−0.1	0.5	−1.4	0.1
Inflation	0.2	0.1	−1.3	−0.4	0.5	−1.2	−0.6
	$(Y^{(A)}, d)$	$(Y^{(A)}, e)$	$(Y^{(A)}, r)$	$(Y^{(A)}, l)$	$(Y^{(A)}, \pi)$	$(Y^{(A)}, s)$	
Short-term rate	2.4	1.9	1.1	2.2	0.1	0.3	
Long-term rate	1.5	1.4	1.1	2.1	−0.2	0.1	
Earnings-by-price	1.6	1.4	0.9	2.0	−0.2	0.1	
Inflation	−1.0	−1.1	−0.6	0.6	−2.1	−1.0	
	(d, e)	(d, r)	(d, l)	(d, π)	(d, s)		
Short-term rate	−2.1	−1.5	−0.8	−2.4	−1.5		
Long-term rate	−2.0	−1.1	−0.6	−2.2	−1.5		
Earnings-by-price	−1.9	−1.4	−0.7	−2.3	−1.5		
Inflation	−0.4	−1.0	−0.2	−2.3	−1.3		
	(e, r)	(e, l)	(e, π)	(e, s)			
Short-term rate	−1.0	−0.4	−2.3	−0.8			
Long-term rate	−0.6	−0.2	−2.2	−0.8			
Earnings-by-price	−1.0	−0.2	−2.2	−0.8			
Inflation	−1.7	−0.9	−2.2	−1.6			
	(r, l)	(r, π)	(r, s)				
Short-term rate	1.3	−1.5	1.4				
Long-term rate	1.3	−1.0	1.4				
Earnings-by-price	1.4	−1.5	1.6				
Inflation	1.3	−1.5	1.2				
	(l, π)	(l, s)					
Short-term rate	−1.2	1.4					
Long-term rate	−0.9	1.4					
Earnings-by-price	−1.0	1.6					
Inflation	−0.9	1.3					
	(π, s)						
Short-term rate	0.2						
Long-term rate	0.2						
Earnings-by-price	−0.6						
Inflation	−0.1						

Contrary to the mean prediction, where Kyriakou et al. (2019b) find that five-year predictability improves over the one-year case, we observe that the majority of predictor based volatility models do

[14] The tests were conducted with 1000 repetitions at the 5% significance level for a selected number of cases. We do not present the p-values of the tests to save space. The results are available upon request by the authors.

not surpass the constant volatility alternative for the five-year horizon. Even though some models produce small positive $R^2_{V,\nu}$ values, this time both the SNS- and the KNY-test do not reject the null of no predictability. Note that our results are in line with Christoffersen and Diebold (2000) who conclude that volatility forecastability may be much less important at longer horizons.

Table 3. Predictive power for the variance of five-year excess stock returns $Z_t^{(A)}$: the single benchmarking approach. The prediction problem is defined in (5). The same predictive variables X_{t-1} are used in the predictions for the conditional mean and variance function. Additional notes: see Table 2.

Benchmark $B^{(A)}$	**Explanatory Variable(s) X_{t-1}**						
	$Y^{(A)}$	d	e	r	l	π	s
Short-term rate	0.6	−1.7	−1.7	−1.2	−1.0	−2.0	−3.0
Long-term rate	0.0	−1.5	−1.3	−1.2	−1.1	−1.2	−2.7
Earnings-by-price	0.8	−1.8	−1.1	−1.8	−2.7	−0.3	−3.8
Inflation	−1.0	−3.8	−4.7	−0.7	−1.5	1.4	0.5
	$(Y^{(A)},d)$	$(Y^{(A)},e)$	$(Y^{(A)},r)$	$(Y^{(A)},l)$	$(Y^{(A)},\pi)$	$(Y^{(A)},s)$	
Short-term rate	−2.8	−2.5	−1.7	−1.7	−1.7	−3.9	
Long-term rate	−2.5	−2.1	−1.6	−1.8	−1.2	−3.4	
Earnings-by-price	−2.3	−2.1	−1.2	−4.1	0.4	−3.4	
Inflation	−5.1	−4.7	−1.5	−2.6	0.4	−0.9	
	(d,e)	(d,r)	(d,l)	(d,π)	(d,s)		
Short-term rate	−3.6	−3.1	−2.2	−2.8	−4.1		
Long-term rate	−3.1	−3.2	−2.7	−2.3	−4.3		
Earnings-by-price	−4.1	−4.0	−5.3	−2.3	−4.9		
Inflation	−5.2	−5.0	−8.9	−2.5	−3.2		
	(e,r)	(e,l)	(e,π)	(e,s)			
Short-term rate	−3.3	−3.3	−3.5	−4.9			
Long-term rate	−2.8	−3.3	−2.9	−4.9			
Earnings-by-price	−4.5	−5.5	−2.7	−6.5			
Inflation	−8.5	−7.8	−4.9	−6.4			
	(r,l)	(r,π)	(r,s)				
Short-term rate	−3.8	−1.7	−3.9				
Long-term rate	−4.1	−1.3	−4.2				
Earnings-by-price	−5.3	−1.9	−5.4				
Inflation	−3.9	0.3	−1.9				
	(l,π)	(l,s)					
Short-term rate	−1.7	−3.9					
Long-term rate	−1.3	−4.2					
Earnings-by-price	−2.6	−5.4					
Inflation	−1.2	−1.8					
	(π,s)						
Short-term rate	−4.4						
Long-term rate	−3.5						
Earnings-by-price	−4.8						
Inflation	−0.1						

3.3. Full Benchmarking Approach

In the next step, we consider the double benchmarking approach of Kyriakou et al. (2019a, 2019b) to analyze now whether transforming the explanatory variables can improve the predictions for the volatility function. Recall that fully nonparametric models suffer in general by the curse of dimensionality. Problems with sparsely distributed annual observations in higher dimensions, as in our framework, could be reduced or circumvented by importing more structure in the estimation process.

Here, we extend the study presented in Section 3.2 transforming both the dependent and independent variables according to the same benchmark. To this end, in our full (double) benchmarking approach, the prediction problems are reformulated as

$$Y_t^{(A)} = m(X_{t-1}^{(A)}) + \nu(X_{t-1}^{(A)})^{1/2}\zeta_t, \tag{14}$$

$$Z_t^{(A)} = m(X_{t-1}^{(A)}) + \nu(X_{t-1}^{(A)})^{1/2}\omega_t, \tag{15}$$

where we use transformed predictive variables

$$X_{t-1}^{(A)} = \begin{cases} \frac{1+X_{t-1}}{B_{t-1}^{(A)}}, & X \in \{d, e, r, l, \pi\} \\ \frac{s_{t-1}}{B_{t-1}^{(A)}} = \frac{l_{t-1}-r_{t-1}}{B_{t-1}^{(A)}} & , \quad A \in \{R, L, E, C\}. \\ Y_{t-1}^{(A)} \end{cases} \tag{16}$$

This approach can be interpreted as a simple way of reducing the dimensionality of the estimation procedure. The adjusted variable $X_{t-1}^{(A)}$ includes now an additional predictive variable, the benchmark itself. Results of this empirical study are presented for the one-year horizon in Table 4 and for the five-year horizon in Table 5.

We find that, in comparison to the single-benchmarking approach in the one-year case, the double benchmarking improves in 15 out of 82 models (in the sense of producing a positive and higher $R_{V,\nu}^2$ as before). However, predictability is still questionable. The best model under the long-term interest rate benchmark $B^{(L)}$ uses the pair $(Y_{t-1}^{(L)}, e_{t-1}^{(L)})$ and yields $R_{V,\nu}^2 = 3.0$, while the best model under $B^{(E)}$ uses the pair $(Y_{t-1}^{(E)}, l_{t-1}^{(E)})$ and yields $R_{V,\nu}^2 = 2.5$. The SNS-test rejects for both the null of no predictability, while the KNY-test does not. For the rest of the new combinations of predictive variables in all benchmarks, both tests again do not reject.

For the five-year case, we find that in comparison to the single-benchmarking the double benchmarking improves in 11 out of 82 models. The best model under $B^{(E)}$ uses $d_{t-1}^{(E)}$ and yields $R_{V,\nu}^2 = 1.8$, while under $B^{(C)}$ the covariates $d_{t-1}^{(C)}$ and $l_{t-1}^{(C)}$ both yield $R_{V,\nu}^2 = 1.6$. Nevertheless, we do not find any combination of covariates with statistically significant predictive power.

Table 4. Predictive power for the variance of one-year excess stock returns $Y_t^{(A)}$: the double benchmarking approach. The prediction problem is defined in (14). The same predictive variables $X_{t-1}^{(A)}$ are used in the predictions for the conditional mean and variance. The predictive power (%) is measured by $R^2_{V,\nu}$ as defined in (8). The benchmarks $B^{(A)}$ considered are based on the short-term interest rate ($A \equiv R$), long-term interest rate ($A \equiv L$), earnings-by-price ratio ($A \equiv E$), and consumer price index ($A \equiv C$). The predictive variables used are $X_{t-1}^{(A)}$ using the indicated benchmark $B_{t-1}^{(A)}$ as shown in (16). X_{t-1} are given by the dividend-by-price ratio d_{t-1}, earnings-by-price ratio e_{t-1}, short-term interest rate r_{t-1}, long-term interest rate l_{t-1}, inflation π_{t-1}, term spread s_{t-1}, excess stock return $Y_{t-1}^{(A)}$, or the possible different pairwise combinations as indicated. "–" are not applicable cases of matched covariate with benchmark. Note: $s^{(R)}$ and $l^{(R)}$ (and their combinations with Y, d, e, π) have the same R^2_V by construction of the transformed spread according to (16). For example, $s_{t-1}^{(R)} = (l_{t-1} - r_{t-1})/B_{t-1}^{(R)} = (1 + l_{t-1})/(1 + r_{t-1}) - 1$ and $l_{t-1}^{(R)} = (1 + l_{t-1})/(1 + r_{t-1})$. The case of $s^{(L)}$ and $r^{(L)}$ is similar.

Benchmark $B^{(A)}$	**Explanatory Variable(s) X_{t-1}**						
	$Y^{(A)}$	$d^{(A)}$	$e^{(A)}$	$r^{(A)}$	$l^{(A)}$	$\pi^{(A)}$	$s^{(A)}$
Short-term rate	2.2	−0.3	0.7	–	−0.2	0.1	−0.2
Long-term rate	2.4	0.2	−0.5	−0.1	–	−0.2	−0.1
Earnings-by-price	1.5	−0.2	–	0.6	−0.2	−0.7	0.0
Inflation	0.2	−0.9	−1.2	−0.3	−0.2	–	−0.7
	$(Y^{(A)}, d^{(A)})$	$(Y^{(A)}, e^{(A)})$	$(Y^{(A)}, r^{(A)})$	$(Y^{(A)}, l^{(A)})$	$(Y^{(A)}, \pi^{(A)})$	$(Y^{(A)}, s^{(A)})$	
Short-term rate	0.8	0.7	–	0.2	0.1	0.2	
Long-term rate	1.3	3.0	0.1	–	−0.3	0.1	
Earnings-by-price	0.2	–	0.7	2.5	0.0	0.1	
Inflation	−3.1	−1.4	−1.5	−1.9	–	−1.0	
	$(d^{(A)}, e^{(A)})$	$(d^{(A)}, r^{(A)})$	$(d^{(A)}, l^{(A)})$	$(d^{(A)}, \pi^{(A)})$	$(d^{(A)}, s^{(A)})$		
Short-term rate	−1.3	–	0.9	0.0	0.9		
Long-term rate	−1.0	0.9	–	−0.7	0.9		
Earnings-by-price	–	−0.3	-0.8	−1.8	0.4		
Inflation	−1.9	0.7	1.6	–	−0.7		
	$(e^{(A)}, r^{(A)})$	$(e^{(A)}, l^{(A)})$	$(e^{(A)}, \pi^{(A)})$	$(e^{(A)}, s^{(A)})$			
Short-term rate	–	−0.4	−2.6	−0.4			
Long-term rate	−0.6	–	−2.5	−0.6			
Earnings-by-price	–	–	–	–			
Inflation	−1.6	−1.5	–	−1.6			
	$(r^{(A)}, l^{(A)})$	$(r^{(A)}, \pi^{(A)})$	$(r^{(A)}, s^{(A)})$				
Short-term rate	–	–	–				
Long-term rate	–	−1.2	–				
Earnings-by-price	−0.5	−2.1	−0.3				
Inflation	−1.9	–	−1.6				
	$(l^{(A)}, \pi^{(A)})$	$(l^{(A)}, s^{(A)})$					
Short-term rate	−1.4	–					
Long-term rate	–	–					
Earnings-by-price	−2.5	−0.5					
Inflation	–	−1.7					
	$(\pi^{(A)}, s^{(A)})$						
Short-term rate	−1.4						
Long-term rate	−1.2						
Earnings-by-price	−1.6						
Inflation	–						

Table 5. Predictive power for the variance of five-year excess stock returns $Z_t^{(A)}$: the double benchmarking approach. The prediction problem is defined in (15). The same predictive variables $X_{t-1}^{(A)}$ are used in the predictions for the conditional mean and variance. Additional notes: see Table 4.

Benchmark $B^{(A)}$	**Explanatory Variable(s) X_{t-1}**						
	$Y^{(A)}$	$d^{(A)}$	$e^{(A)}$	$r^{(A)}$	$l^{(A)}$	$\pi^{(A)}$	$s^{(A)}$
Short-term rate	0.6	−2.2	−3.2	–	−3.1	−3.2	−3.1
Long-term rate	0.0	−3.4	−2.8	−2.8	–	−1.3	−2.8
Earnings-by-price	0.8	1.8	–	−2.3	−3.2	0.6	−3.8
Inflation	−1.0	1.6	0.3	0.6	1.6	–	0.3
	$(Y^{(A)}, d^{(A)})$	$(Y^{(A)}, e^{(A)})$	$(Y^{(A)}, r^{(A)})$	$(Y^{(A)}, l^{(A)})$	$(Y^{(A)}, \pi^{(A)})$	$(Y^{(A)}, s^{(A)})$	
Short-term rate	−2.1	−4.3	–	−4.0	−1.2	−4.0	
Long-term rate	−3.8	−3.2	−3.6	–	−1.1	−3.6	
Earnings-by-price	1.1	–	−2.8	−3.8	−0.5	−3.4	
Inflation	0.3	−0.8	−0.3	0.4	–	−1.0	
	$(d^{(A)}, e^{(A)})$	$(d^{(A)}, r^{(A)})$	$(d^{(A)}, l^{(A)})$	$(d^{(A)}, \pi^{(A)})$	$(d^{(A)}, s^{(A)})$		
Short-term rate	−3.7	–	−5.4	−2.1	−5.4		
Long-term rate	−4.2	−5.8	–	−3.3	−5.8		
Earnings-by-price	–	−0.4	−2.6	0.3	−3.3		
Inflation	−4.3	−0.2	−0.8	–	−0.8		
	$(e^{(A)}, r^{(A)})$	$(e^{(A)}, l^{(A)})$	$(e^{(A)}, \pi^{(A)})$	$(e^{(A)}, s^{(A)})$			
Short-term rate	–	−5.9	−4.9	−5.9			
Long-term rate	−6.1	–	−4.1	−6.1			
Earnings-by-price	–	–	–	–			
Inflation	−4.8	−4.1	–	−2.1			
	$(r^{(A)}, l^{(A)})$	$(r^{(A)}, \pi^{(A)})$	$(r^{(A)}, s^{(A)})$				
Short-term rate	–	–	–				
Long-term rate	–	−2.3	–				
Earnings-by-price	−6.3	−3.2	−6.1				
Inflation	−1.0	–	0.5				
	$(l^{(A)}, \pi^{(A)})$	$(l^{(A)}, s^{(A)})$					
Short-term rate	−3.4	–					
Long-term rate	–	–					
Earnings-by-price	−3.6	−6.2					
Inflation	–	0.5					
	$(\pi^{(A)}, s^{(A)})$						
Short-term rate	−3.4						
Long-term rate	−2.3						
Earnings-by-price	−4.6						
Inflation	–						

3.4. Real-Income Long-Term Pension Prediction

In long-term pension planning or other asset allocation problems optimized with regard to real-income protection (Gerrard et al. (2019a, 2019b); (Merton 2014)), the econometric models should reflect those needs and use covariates net-of-inflation. Therefore, we take the inflation benchmark $B^{(C)}$ and analyse in more detail the best model found by Kyriakou et al. (2019b), which uses the earnings-by-price variable for the mean prediction and produced a $R^2_{V,m} = 12.2$ for the one-year horizon and $R^2_{V,m} = 12.4$ for the five-year horizon (see Kyriakou et al. 2019b, Tables 4 and 5) in the double benchmarking case. For this specific model, we are now interested in finding the set of

covariates that best predicts the conditional variance.[15,16] The empirical findings in terms of $R^2_{V,\nu}$ are shown for the one-year horizon in Table 6 and the five-year horizon in Table 7. For the one-year horizon, we find in the double benchmarking approach when inflation is the benchmark, $B^{(C)}$ that the dividend-by-price $d^{(C)}$ together with the short-term interest-rate $r^{(C)}$ or the long-term interest-rate $l^{(C)}$ are chosen as best predictive variables in terms of $R^2_{V,\nu}$ (2.9% and 2.0%). Note that these values are rather low and that the SNS-test does reject the null of no predictability for both models, while the KNY-test does not reject. For all other combinations and also the five-year case, we do not find evidence for statistical significant predictability of the conditional variance. Therefore, we conclude that the constant volatility model is appropriate for practical purposes.

Note further that the ratio in our validation criterion for the mean prediction, $R^2_{V,m}$, in (8) compares the sample variance of the estimated residuals from our model based on earnings-by-price (the numerator) with the sample variance of the benchmarked stock returns (the denominator). For the one-year case, we find from Table 1 the latter to be equal to $0.1805^2 = 0.03258$. A simple calculation using the corresponding $R_{V,m} = 12.2\%$ leads then to $0.03258(1 - 0.122) = 0.02861$ or a standard deviation of 16.91% for returns based on the earnings-model. This means that the linear expression of real stock returns in terms of real earnings-by-price presented in Kyriakou et al. (2019b) as

$$\text{Real one-year stock return} = 0.004875 + 1.119 \times \text{real earnings-by-price} \tag{17}$$

gives on average 2.4% higher returns at the same risk as the historical mean $\bar{Y}^{(C)}$.[17] Similarly, for the five-year case, we get from Table 1 that $0.3642^2 = 0.1326$. From the $R_{V,m} = 12.4\%$, we obtain then $0.1326(1 - 0.122) = 0.1162$ or a standard deviation of 34.08% for returns based on the earnings-model. Thus, the linear expression of real stock returns in terms of real earnings-by-price presented in Kyriakou et al. (2019b) as

$$\text{Real five-year stock return} = 0.2068 + 2.264 \times \text{real earnings-by-price} \tag{18}$$

gives on average 6.1% higher returns at the same risk as the historical mean $\bar{Y}^{(C)}$.[18] Figure 1 shows the estimated nonparametric function $\hat{m}$ (red solid line) for the one-year horizon (left) and the five-year horizon (right) under the double inflation benchmark for the earnings-by-price covariate together with the corresponding historical mean (dashed green line). Figure 2 depicts histograms and a kernel density estimate (red solid line) of the standardized predicted returns for the one-year horizon (left) and the five-year horizon (right). The similarity for both horizons is striking and driven by the fact that the ratio of the slope of the regression lines in (17) and (18) with the corresponding standard deviation given above yields almost the same value of 6.63.

---

[15] Note that until now we have used the same set of covariates in both steps of our analysis to reduce the overwhelming number of models. It is also clear that not all combinations of variables are practically relevant. Now, we relax this restriction for the model with the highest predictive power for the returns.

[16] Tables 6 and 7 also present the results for the short- and long-term interest benchmarks $B^{(R)}$ and $B^{(L)}$. However, it is again hard to find predictability at all in these cases. Note that the benchmark using the earnings-by-price variable $B^{(E)}$ is not applicable since it matches the covariate and the benchmark in the first step.

[17] Here, we use the Sharpe-ratio for the comparison. From Table 1, we get $\bar{Y}^{(C)} = 6.41\%$ and divide it either by 18.05% or by 16.91%. We obtain 0.355 and 0.379, which corresponds to a difference of 2.4% points.

[18] Here, we use again the Sharpe-ratio for the comparison. From Table 1, we get $\bar{Y}^{(C)} = 32.34\%$ and divide it either by 36.42% or by 34.08%. We obtain 0.888 and 0.949, which corresponds to a difference of 6.1% points.

Finally, we consider a simple mean-reverting autoregressive model of order one for the real earnings-by-price—the main drivers of real returns in Equations (17) and (18)—and estimate it with ordinary least squares (OLS) [19]:

$$\begin{aligned}&\text{Change in real earnings-by-price}\\&\quad = -0.715 \times (\text{real earnings-by-price} - \text{mean of real earnings-by-price}).\end{aligned} \tag{19}$$

Note that, for the whole sample period (1872–2019), the mean and standard deviation of real earnings-by-price are 0.0524 and 0.0595, resp. Moreover, using the current (30/09/2019) value of real earnings-by-price of 0.0278, model (19) predicts a change in real earnings-by-price of 0.0176, i.e., an expected value of real earnings-by-price of 0.0454 for 2020Q3, which is still below the long-term average.[20]

We subsequently calculate the correlation between the estimated residuals of models (17) and (19) to be -0.014. A standard stationary block-bootstrap (Politis and Romano 1994) based on 10,000 repetitions and a block-length of 12 suggests that this correlation is not statistically significantly different from zero. The correlation structure between returns and their drivers is important while searching for optimal investment strategies in a dynamic market, see Kim and Omberg (1996). Gerrard et al. (2019c) follow the approach of Kim and Omberg (1996) in a long-term return setting and show that the above correlation is very hard to estimate with precision. Sometimes, it is negative and, with a slight change of data, it is positive, and a test would almost always provide that zero correlation cannot be rejected. When this added insight is provided that zero correlation significantly simplifies that technical calculation of the optimal dynamic strategy while significantly reducing parameter uncertainty, the conclusion seems clear: we should work with zero correlation unless there is a strong argument not to do that. In our case—which is a discrete analogue to the continuous models considered in Gerrard et al. (2019c) and Kim and Omberg (1996)—it is, therefore, comforting that we can provide a simple zero-correlation econometric model to guide the market dynamics. In further work, we expect the simple econometric model of this paper to be used while generalizing the non-dynamic new approach to pension products of Gerrard et al. (2019a, 2019b).

---

[19] The estimated coefficient is significant at the 0.1%-level (with a corresponding standard error of 0.08), the residual standard error of the regression is 0.0572, and its R^2 has a value of 0.357.

[20] The following values are used for the calculation of the current real earnings-by-price: $P = 2976.74$, $E = 135.53$, $B^{(C)} = 1.0173$.

Table 6. Predictive power for the variance of one-year excess stock returns $Y_t^{(A)}$: the double benchmarking approach for the conditional mean model with earnings-by price as single covariate. The prediction problem is defined in (14). The predictive power (%) is measured by $R^2_{V,\nu}$ as defined in (8). The benchmarks $B^{(A)}$ considered are based on the short-term interest rate ($A \equiv R$), long-term interest rate ($A \equiv L$), and consumer price index ($A \equiv C$). The predictive variables used are $X_{t-1}^{(A)}$ using the indicated benchmark $B_{t-1}^{(A)}$ as shown in (16). X_{t-1} are given by the dividend-by-price ratio d_{t-1}, earnings-by-price ratio e_{t-1}, short-term interest rate r_{t-1}, long-term interest rate l_{t-1}, inflation π_{t-1}, term spread s_{t-1}, excess stock return $Y_{t-1}^{(A)}$, or the possible different pairwise combinations as indicated. "–" are not applicable cases of matched covariate with benchmark. Note: $s^{(R)}$ and $l^{(R)}$ (and their combinations with Y, d, e, π) have the same $R^2_{V,\nu}$ by construction of the transformed spread according to (16). For example, $s_{t-1}^{(R)} = (l_{t-1} - r_{t-1})/B_{t-1}^{(R)} = (1 + l_{t-1})/(1 + r_{t-1}) - 1$ and $l_{t-1}^{(R)} = (1 + l_{t-1})/(1 + r_{t-1})$. Similar is the case of $s^{(L)}$ and $r^{(L)}$.

Benchmark $B^{(A)}$	**Explanatory Variable(s) X_{t-1}**						
	$Y^{(A)}$	$d^{(A)}$	$e^{(A)}$	$r^{(A)}$	$l^{(A)}$	$\pi^{(A)}$	$s^{(A)}$
Short-term rate	1.0	0.3	0.7	–	0.1	−0.4	0.1
Long-term rate	1.4	0.1	−0.5	0.9	–	−0.1	0.9
Inflation	0.4	−0.6	−1.2	−0.4	−0.1	–	0.8
	$(Y^{(A)}, d^{(A)})$	$(Y^{(A)}, e^{(A)})$	$(Y^{(A)}, r^{(A)})$	$(Y^{(A)}, l^{(A)})$	$(Y^{(A)}, \pi^{(A)})$	$(Y^{(A)}, s^{(A)})$	
Short-term rate	0.6	0.7	–	0.2	−0.5	0.2	
Long-term rate	0.7	2.0	0.7	–	−0.6	0.7	
Inflation	−1.7	−1.6	−1.5	−1.7	–	−0.4	
	$(d^{(A)}, e^{(A)})$	$(d^{(A)}, r^{(A)})$	$(d^{(A)}, l^{(A)})$	$(d^{(A)}, \pi^{(A)})$	$(d^{(A)}, s^{(A)})$		
Short-term rate	0.0	–	−0.5	−0.4	−0.5		
Long-term rate	−1.0	0.3	–	−1.4	0.3		
Inflation	−1.9	2.9	2.0	–	1.5		
	$(e^{(A)}, r^{(A)})$	$(e^{(A)}, l^{(A)})$	$(e^{(A)}, \pi^{(A)})$	$(e^{(A)}, s^{(A)})$			
Short-term rate	–	0.5	−2.2	0.5			
Long-term rate	−0.7	–	−2.5	−0.7			
Inflation	−0.9	−1.7	–	−0.3			
	$(r^{(A)}, l^{(A)})$	$(r^{(A)}, \pi^{(A)})$	$(r^{(A)}, s^{(A)})$				
Short-term rate	–	–	–				
Long-term rate	–	−0.4	–				
Inflation	−0.5	–	0.7				
	$(l^{(A)}, \pi^{(A)})$	$(l^{(A)}, s^{(A)})$					
Short-term rate	0.1	–					
Long-term rate	–	–					
Inflation	–	−0.2					
	$(\pi^{(A)}, s^{(A)})$						
Short-term rate	0.1						
Long-term rate	−0.4						
Inflation	–						

Table 7. Predictive power for the variance of five-year excess stock returns $Z_t^{(A)}$: the double benchmarking approach for the conditional mean model with earnings-by price as single covariate. The prediction problem is defined in (15). Additional notes: see Table 6.

Benchmark $B^{(A)}$	**Explanatory Variable(s) X_{t-1}**						
	$Y^{(A)}$	$d^{(A)}$	$e^{(A)}$	$r^{(A)}$	$l^{(A)}$	$\pi^{(A)}$	$s^{(A)}$
Short-term rate	0.1	−1.8	−3.2	–	−4.5	−2.5	−4.5
Long-term rate	0.6	−3.9	−2.8	−4.2	–	−1.1	−4.2
Inflation	0.0	−0.1	0.3	−0.4	−0.1	–	−2.6
	$(Y^{(A)}, d^{(A)})$	$(Y^{(A)}, e^{(A)})$	$(Y^{(A)}, r^{(A)})$	$(Y^{(A)}, l^{(A)})$	$(Y^{(A)}, \pi^{(A)})$	$(Y^{(A)}, s^{(A)})$	
Short-term rate	−1.7	−4.6	–	−5.7	−3.7	−5.7	
Long-term rate	−4.5	−4.5	−4.2	–	−2.5	−4.2	
Inflation	−1.9	−1.8	−1.9	−1.7	–	−3.9	
	$(d^{(A)}, e^{(A)})$	$(d^{(A)}, r^{(A)})$	$(d^{(A)}, l^{(A)})$	$(d^{(A)}, \pi^{(A)})$	$(d^{(A)}, s^{(A)})$		
Short-term rate	−6.2	–	−7.1	−4.3	−7.1		
Long-term rate	−4.5	−7.9	–	−5.2	−7.9		
Inflation	−3.9	−2.1	−3.2	–	−2.8		
	$(e^{(A)}, r^{(A)})$	$(e^{(A)}, l^{(A)})$	$(e^{(A)}, \pi^{(A)})$	$(e^{(A)}, s^{(A)})$			
Short-term rate	–	−8.1	−5.8	−8.1			
Long-term rate	−6.6	–	−4.9	−6.6			
Inflation	−2.8	−3.4	–	−2.6			
	$(r^{(A)}, l^{(A)})$	$(r^{(A)}, \pi^{(A)})$	$(r^{(A)}, s^{(A)})$				
Short-term rate	–	–	–				
Long-term rate	–	−5.7	–				
Inflation	−3.0	–	−3.1				
	$(l^{(A)}, \pi^{(A)})$	$(l^{(A)}, s^{(A)})$					
Short-term rate	−6.5	–					
Long-term rate	–	–					
Inflation	–	−3.0					
	$(\pi^{(A)}, s^{(A)})$						
Short-term rate	−6.5						
Long-term rate	−5.7						
Inflation	–						

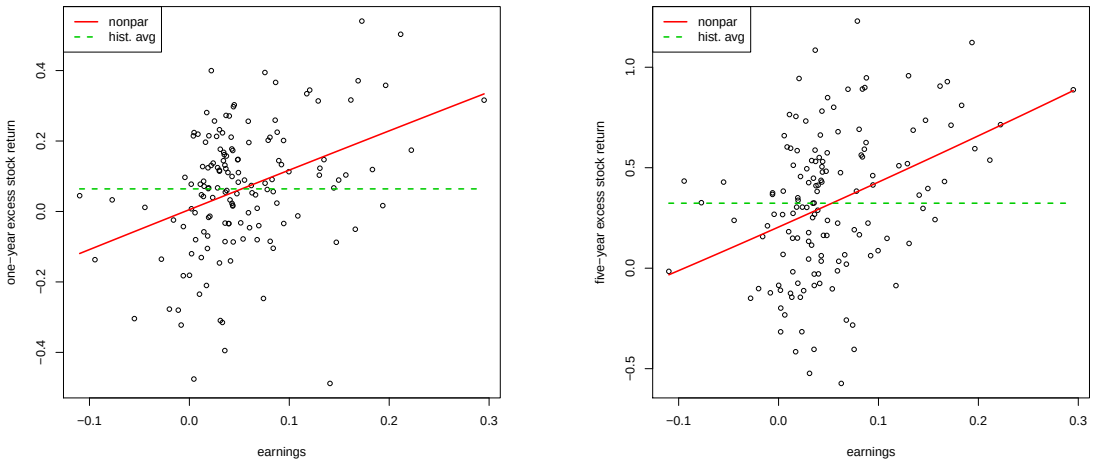

Figure 1. Double inflation benchmark. Relation between real stock returns and real earnings-by-price. Estimated nonparametric function $\hat{m}$ (red solid line) and historical average (dashed green line). **Left**: one-year horizon. **Right**: five-year horizon. Period: 1872–2019. Data: annual S&P 500.

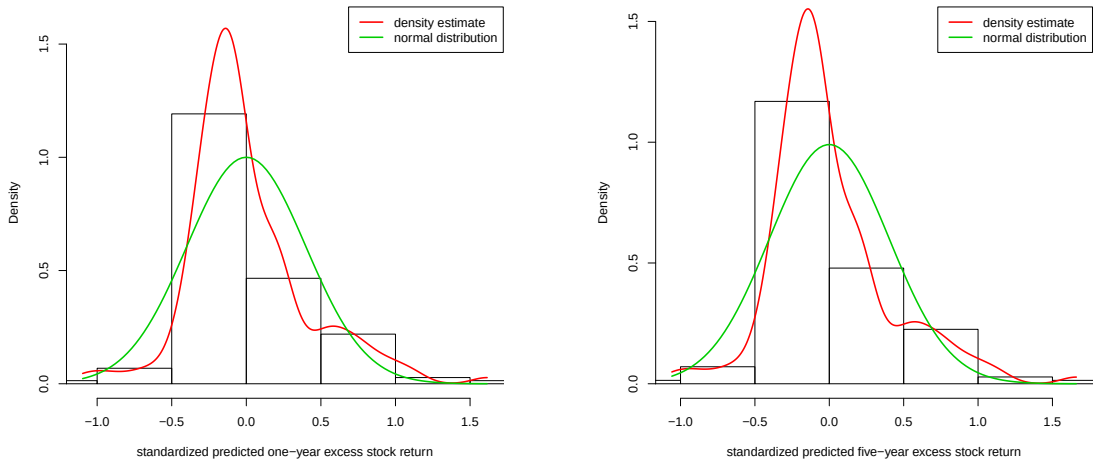

Figure 2. Standardized predicted stock returns in excess of the inflation benchmark (based on the model using earnings-by-price as covariate for mean-prediction; double benchmarking). Histogram, kernel density estimate (red), and fitted normal distribution (green). **Left**: one-year horizon. **Right**: five-year horizon. Period: 1872–2019. Data: annual S&P 500.

4. Conclusions

In this paper, we extend the original working framework of Kyriakou et al. (2019a, 2019b) of forecasting stock returns to modelling their conditional variance and test for predictability in this context. We consider returns of one-year and five-year horizons in excess of different benchmarks, considering the short- and long-term rate, the earnings-by-price ratio, and the inflation rate. We use popular explanatory variables with predictive power such as the dividend-by-price ratio, the earnings-by-price ratio, the short- and long-term interest rates, the term spread, the inflation rate, as well as the lagged excess stock return, in one- and two-dimensional settings, with the returns benchmarked or also the covariates used to predict them.

In our analysis, we find only little to no evidence of model-based volatility predictability for the one-year and five-year horizon. Only for a few of the models considered under different benchmarks, we get validation measures that are positive and significantly different from zero but of a rather small magnitude. We thus conclude that volatility forecastability is much less important at longer horizons regardless of the chosen combination of explanatory variables. The homoscedastic historical average of the squared return prediction errors gives an adequate approximation of the unobserved realised conditional variance for both the one-year and five-year horizon.

In the practically important double inflation benchmarking case, we find that the model with the largest predictive power is not only of linear functional form based on real earnings-by-price but also has a constant variance for both horizons. A simple mean-reverting linear AR1-model for the real-earnings-by-price allows then to analyse the correlation structure between returns and their main drivers. We find zero correlation which significantly simplifies the econometric modelling to guide market dynamics. This is an important observation and a relatively simple starting point when constructing forecasting models for real-value pension prognoses for long-term saving strategies.

Author Contributions: All authors contributed equally to this work.

Funding: The authors thank: (i) the Institute and Faculty of Actuaries in the UK for funding this research through the grant "Minimizing Longevity and Investment Risk while Optimizing Future Pension Plans", and (ii) the University of Graz for the Open Access Funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and Heiko Ebens. 2001. The distribution of realized stock return volatility. *Journal of Finacial Economics* 61: 43–76. [CrossRef]

Burman, Prabir, Edmond Chow, and Deborah Nolan. 1994. A cross-validatory method for dependent data. *Biometrika* 81: 351–58. [CrossRef]

Chen, Qingqing, and Yongmiao Hong. 2010. *Predictability of Equity Returns Over Different Time Horizons: A Nonparametric Approach*. Working Paper. Ithaca: Cornell University/Department of Economics.

Cheng, Tingting, Jiti Gao, and Oliver Linton. 2019. *Nonparametric Predictive Regressions for Stock Return Predictions*. Cambridge Working Papers in Economics: 1932. Cambridge: Faculty of Economics, University of Cambridge.

Christoffersen, Peter F., and Francis X. Diebold. 2000. How relevant is volatility forecasting for financial risk management? *Review of Economics and Statistics* 82: 12–22. [CrossRef]

Chu, C. K., and J. S. Marron. 1991. Comparison of two bandwidth selectors with dependent errors. *The Annals of Statistics* 19: 1906–18. [CrossRef]

De Brabanter, Kris, Jos De Brabanter, Johan A.K. Suykens, and Bart De Moor. 2011. Kernel regression in the presence of correlated errors. *Journal of Machine Learning Research* 12: 1955–76.

Fan, Jianqing, and Qiwei Yao. 1998. Efficient estimation of conditional variance functions in stochastic regression. *Biometrika* 85: 645–60. [CrossRef]

Galbraith, John W. 2003. Content horizons for univariate time-series forecasts. *International Journal of Forecasting* 19: 43–55. [CrossRef]

Galbraith, John W., and Turgut Kisinbay. 2005. Content horizons for conditional variance forecasts. *International Journal of Forecasting* 21: 249–60. [CrossRef]

Geller, Juliane, and Michael H. Neumann. 2018. Improved local polynomial estimation in time series regression. *Journal of Nonparametric Statistics* 30: 1–27. [CrossRef]

Gerrard, Russell, Munir Hiabu, Ioannis Kyriakou, and Jens Perch Nielsen. 2019a. Communication and personal selection of pension saver's financial risk. *European Journal of Operational Research* 274: 1102–11. [CrossRef]

Gerrard, Russell, Munir Hiabu, Ioannis Kyriakou, and Jens Perch Nielsen. 2019b. Self-selection and risk sharing in a modern world of life-long annuities. *British Actuarial Journal* 23: e30. [CrossRef]

Gerrard, Russell, Munir Hiabu, Jens Perch Nielsen, and Peter Vodicka. 2019c. *Long-Term Real Dynamic Investment Planning*. Working Paper. London: Cass Business School.

Geweke, John F. 1996. Monte carlo simulation and numerical integration. In *Handbook of Computational Economics*. Edited by Hans M. Amman, David A. Kendrick and John Rust. Amsterdam: Elsevier, vol. I, pp. 731–800.

Glad, Ingrid K. 1998. Parametrically guided non-parametric regression. *Scandinavian Journal of Statistics* 25: 649–68. [CrossRef]

Gonzales-Manteiga, Wenceslao, and Rosa M. Crujeiras. 2013. An updated review of goodness-of-fit tests for regression models. *Test* 22: 361–411. [CrossRef]

Härdle, Wolfgang K., and Enno Mammen. 1993. Comparing nonparametric versus parametric regression fits. *Annals of Statistics* 21: 1926–47. [CrossRef]

Härdle, Wolfgang K., and Alexandre B. Tsybakov. 1997. Local polynomial estimators of the volatility function in nonparametric autoregression. *Journal of Econometrics* 81: 223–42. [CrossRef]

Kim, Tong Suk, and Edward Omberg. 1996. Dynamic nonmyopic portfolio behavior. *The Review of Financial Studies* 9: 141–61. [CrossRef]

Kothari, S. P., Jonathan Lewellen, and Jerold B. Warner. 2006. Stock returns, aggregate earnings surprises, and behavioral finance. *Journal of Financial Economics* 79: 537–68. [CrossRef]

Kreiss, Jens-Peter, Michael H. Neumann, and Qiwei Yao. 2008. Bootstrap tests for simple structures in nonparametric time series regression. *Statistics and Its Interfaces* 1: 367–80. [CrossRef]

Kyriakou, Ioannis, Parastoo Mousavi, Jens Perch Nielsen, and Michael Scholz. 2019a. Forecasting benchmarks of long-term stock returns via machine learning. *Annals of Operations Research*. doi:10.1007/s10479-019-03338-4. [CrossRef]

Kyriakou, Ioannis, Parastoo Mousavi, Jens Perch Nielsen, and Michael Scholz. 2019b. *Machine Learning for Forecasting Excess Stock Returns—The Five-year View*. Graz Economics Papers 2019-06. Graz: University of Graz, Departmemt of Economics.

Lettau, Martin, and Stijn Van Nieuwerburgh. 2008. Reconciling the return predictability evidence. *Review of Financial Studies* 21: 1607–52. [CrossRef]

Linton, Oliver, and Benoit Perron. 2003. The shape of the risk premium: Evidence from a semiparametric generalized autoregressive conditional heteroscedasticity model. *Journal of Business & Economic Statistics* 21: 354–67.

Linton, Oliver B., and Yang Yan. 2011. Semi- and nonparametric arch processes. *Journal of Probability and Statistics* 2011: 906212. [CrossRef]

Linton, Oliver B., and Enno Mammen. 2008. Nonparametric transformation to white noise. *Journal of Econometrics* 142: 241–64. [CrossRef]

Merton, Robert C. 1973. An intertemporal capital asset pricing model. *Econometrica* 41: 867–87. [CrossRef]

Merton, Robert C. 2014. The crisis in retirement planning. *Harvard Business Review* 92: 43–50.

Mishra, Santosh, Liangjun Su, and Aman Ullah. 2010. Semiparametric estimator of time series conditional variance. *Journal of Business & Economic Statistics* 28: 256–74.

Munk, Claus, and Jesper Rangvid. 2018. New assumptions of a pension forecast model: Background, level and consequences for individuals forecasted pension. *Finans/Invest* 6: 6–14.

Nielsen, Jens Perch, and Stefan Sperlich. 2003. Prediction of stock returns: A new way to look at it. *ASTIN Bulletin* 33: 399–417. [CrossRef]

Opsomer, Jean, Yuedong Wang, and Yuhong Yang. 2001. Nonparametric regression with correlated errors. *Statistical Science* 16: 134–53.

Pagan, Adrian R., and Yong-Sik Hong. 1991. Nonparametric estimation and the risk premium. In *Nonparametric and Semiparametric Methods in Econometrics and Statistics*. Edited by William A. Barnett, James Powell and George E. Tauchen. Cambridge: Cambridge University Press, pp. 51–76.

Pagan, Adrian R., and Aman Ullah. 1988. The econometric analysis of models with risk terms. *Journal of Applied Econometrics* 3: 87–105. [CrossRef]

Politis, Dimitris N., and Joseph P. Romano. 1994. The stationary bootstrap. *Journal of the American Statistical Association* 89: 1303–13. [CrossRef]

Rapach, David, and Guofu Zhou. 2013. Forecasting stock returns. In *Handbook of Economic Forecasting*. Edited by Graham Elliott and Allan Timmerman. Amsterdam: Elsevier, vol. 2A, pp. 328–83.

Scholz, Michael, Jens Perch Nielsen, and Stefan Sperlich. 2015. Nonparametric prediction of stock returns based on yearly data: The long-term view. *Insurance: Mathematics and Economics* 65: 143–55. [CrossRef]

Scholz, Michael, Stefan Sperlich, and Jens Perch Nielsen. 2016. Nonparametric long term prediction of stock returns with generated bond yields. *Insurance: Mathematics and Economics* 69: 82–96. [CrossRef]

Shiller, Robert J. 1989. *Market Volatility*. Cambridge: MIT Press.

Su, Liangjun, and Aman Ullah. 2006. More efficient estimation in nonparametric regression with nonparametric autocorrelated errors. *Econometric Theory* 22: 98–126. [CrossRef]

Wang, Lie, Lawrence D. Brown, T. Tony Cai, and Michael Levine. 2008. Effect of mean on variance function estimation in nonparametric regression. *The Annals of Statistics* 36: 646–64. [CrossRef]

Wang, WenWu, and Ping Yu. 2017. Asymptotically optimal differenced estimators of error variance in nonparametric regression. *Computational Statistics and Data Analysis* 105: 125–43. [CrossRef]

Welch, Ivo, and Amit Goyal. 2008. A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies* 21: 1455–508. [CrossRef]

Xiao, Zhijie, Oliver B. Linton, Raymond J. Carroll, and Enno Mammen. 2003. More efficient local polynomial estimation in nonparametric regression with autocorrelated errors. *Journal of the American Statistical Association* 98: 980–92. [CrossRef]

Xu, Ke Li, and Peter C. B. Phillips. 2011. Tilted nonparametric estimation of volatility functions with empirical applications. *Journal of Business & Economic Statistics* 29: 518–28.

Yu, Keming, and M.C. Jones. 2004. Likelihood-based local linear estimation of the conditional variance function. *Journal of the American Statistical Association* 99: 139–44. [CrossRef]

Ziegelmann, Flavio A. 2002. Nonparametric estimation of volatility functions: The local exponential estimator. *Econometric Theory* 18: 985–91. [CrossRef]

© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).