



City Research Online

City, University of London Institutional Repository

Citation: Kyriakou, I., Mousavi, P., Nielsen, J. P. & Scholz, M. (2020). Longer-Term Forecasting of Excess Stock Returns—The Five-Year Case. *Mathematics*, 8(6), 927. doi: 10.3390/math8060927

This is the published version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/24325/>

Link to published version: <https://doi.org/10.3390/math8060927>

Copyright: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

Reuse: Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. 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.

City Research Online:

<http://openaccess.city.ac.uk/>

publications@city.ac.uk

Article

Longer-Term Forecasting of Excess Stock Returns—The Five-Year Case

Ioannis Kyriakou ¹, Parastoo Mousavi ¹, Jens Perch Nielsen ¹ and Michael Scholz ^{2,*}

¹ Faculty of Actuarial Science and Insurance, Cass Business School, City, University of London, 106 Bunhill Row, London EC1Y 8TZ, UK; ioannis.kyriakou@city.ac.uk (I.K.); parastoo.mousavi.2@cass.city.ac.uk (P.M.); jens.nielsen.1@city.ac.uk (J.P.N.)

² Department of Economics, University of Graz, Universitätsstraße 15/F4, 8010 Graz, Austria

* Correspondence: michael.scholz@uni-graz.at; Tel.: +43-316-380-7112

Received: 20 May 2020; Accepted: 1 June 2020; Published: 5 June 2020

Abstract: Long-term return expectations or predictions play an important role in planning purposes and guidance of long-term investors. Five-year stock returns are less volatile around their geometric mean than returns of higher frequency, such as one-year returns. One would, therefore, expect models using the latter to better reduce the noise and beat the simple historical mean than models based on the former. However, this paper shows that the general tendency is surprisingly the opposite: long-term forecasts over five years have a similar or even better predictive power when compared to the one-year case. We consider a long list of economic predictors and benchmarks relevant for the long-term investor. Our predictive approach consists of adopting and implementing a fully nonparametric smoother with the covariates and the smoothing parameters chosen by cross-validation. We consistently find that long-term forecasting performs well and recommend drawing more attention to it when designing investment strategies for long-term investors. Furthermore, our preferred predictive model did stand the test of Covid-19 providing a relatively optimistic outlook in March 2020 when uncertainty was all around us with lockdown and facing an unknown new pandemic.

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

1. Introduction

In recent years, investment planners and expert forecasters recognize the importance of the horizon when long-term predictions of (excess) stock returns are constructed. It is well-known in the financial literature that it is difficult to provide better forecasts than the simple historical long-term mean. There are also regular discussions in academic as well as practitioner circles on whether it is possible at all. This paper follows and adds to the work and insight of Lioui and Poncet [1] or Møller and Rangvid [2] which provide evidence that it is indeed possible to predict better than the trivial long-term mean when a careful validation approach is applied based on reasonable long-term economic drivers of the financial market to predict. It is also well-recognized, of course, that the longer the horizon, the less noise will disturb the prediction. Therefore, one would expect that it becomes more and more difficult to beat the long-term mean when the horizon is increased, simply because the noise involved in the long term is so low that it seems to be difficult to lower it even further through predictive modeling. In this paper, we find that long-term forecasts counter-intuitively improve as much as, or sometimes even more, than the simple one-year long-term mean forecast. We show this concretely via a comparison between the five-year and the one-year case. The extraordinary and surprisingly good results we get from the five-year prediction approach should lead to reconsiderations when deciding on the horizon to use in the design of investment strategies for long-term investors.

The linear regression model is the classical benchmark in this predictive modeling context. However, there is some evidence documented in the literature that stock return predictability is much stronger when the functional form of the relationship between stock returns and predictive variables is allowed to be nonlinear [3–5]. Along those lines, we adopt and implement the fully nonparametric local-linear smoothing technique in combination with a leave- k -out cross-validation. The linear function can be estimated through the local-linear smoother without any bias and is thus automatically embedded in our approach. We extend the single- and full-benchmarking of Kyriakou et al. [6] to the five-year horizon, since a careful imposition of structure in the statistical modeling process has shown to be promising in previous work [7–9]. Our preferred predictive model gives an optimistic outlook, even though we are at the beginning of a worldwide economic crisis caused by the Covid-19 pandemic. This surprising optimistic outlook was indeed in line with the performance of the stock market from March till June 2020.

The remainder of this paper is organized as follows. In Section 2, we present our framework for nonlinear predictive long-term regressions. We define the underlying financial model, describe the local-linear smoother with its theoretical properties, and present our validation criterion for the model and smoothing parameter selection. In Section 3, we provide the details of our dataset and a short descriptive analysis. Subsequently, we illustrate our empirical findings from two different validated scenarios: (i) a single benchmarking approach where only the dependent variable is measured in excess of the benchmark; and (ii) the case where both the independent and dependent variables are adjusted accordingly to the benchmark (full-benchmarking approach). Finally, we comment on real-income pension prediction and give one-year-ahead real-time predictions. Section 4 summarizes the key points of our analysis and concludes the paper.

2. A Method for Long-Term Prediction

The focus of our analysis lies on nonlinear predictive relationships between stock returns over the next T years in excess of a reference rate (or benchmark) and a set of explanatory variables. We aim to investigate different benchmark models and their predictability over horizons of one year and five years. We consider the four different benchmarks proposed in Kyriakou et al. [6]: the short-term interest rate, the long-term interest rate, the earnings-by-price ratio, and the inflation rate.

2.1. The One-year Case

We investigate stock returns $S_t = (P_t + D_t)/P_{t-1}$, where D_t denotes the (nominal) dividends paid during year t and P_t the (nominal) stock price at the end of year t , 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,

$$B_t^{(R)} = 1 + \frac{R_t}{100}, \quad B_t^{(L)} = 1 + \frac{L_t}{100}, \quad B_t^{(E)} = 1 + \frac{E_t}{P_t}, \quad B_t^{(C)} = \frac{CPI_t}{CPI_{t-1}}, \tag{2}$$

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

$$Y_t^{(A)} = m(X_{t-1}) + \zeta_t, \tag{3}$$

where

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

is an unknown function which we want to estimate with the local-linear smoother and ζ_t is an error term. The sequence of error terms in Equation (3) is assumed to form a martingale difference process, consisting of serially uncorrelated zero-mean random variables, given the past, of an unknown conditionally heteroscedastic form $\sigma(x)$.

Our aim is to predict the excess stock returns $Y_t^{(A)}$ using (combinations of) popular time-lagged predictive variables $X_{t-1} \in \mathbb{R}^q$ including the: (i) dividend-by-price ratio $d_{t-1} = D_{t-1}/P_{t-1}$; (ii) earnings-by-price ratio $e_{t-1} = E_{t-1}/P_{t-1}$; (iii) short-term interest rate $r_{t-1} = R_{t-1}/100$; (iv) long-term interest rate $l_{t-1} = L_{t-1}/100$; (v) inflation $\pi_{t-1} = (CPI_{t-1} - CPI_{t-2})/CPI_{t-2}$; (vi) term spread $s_{t-1} = l_{t-1} - r_{t-1}$; and (vii) excess stock return $Y_{t-1}^{(A)}$.

2.2. The T-year Case

Longer horizons are of fundamental interest to long-term investors, such as pension funds, insurance companies, other institutional investors, or market participants saving for a distant pay-off. Long-term investors are in general willing to take on more risk for higher rewards. Note that the risk is taken usually increases with the investment horizon. However, Rapach and Zhou [10] report that longer horizons also tend to produce better estimates than shorter horizons. Munk and Rangvid [11] indicates that major investors today use longer horizons (up to ten years) to stabilize and improve future predictions.

In our paper, we concentrate on the five-year view as the first compromise for a comparison between a longer horizon and the shorter one-year horizon. The choice of the five-year horizon is arbitrary and is intended to illustrate the potential of our approach. Any other combination of short- and long-term horizons could be considered. However, it seems that shorter horizons based on monthly, weekly, daily, or even intra-day data do not provide good information on the pension savers future income. Therefore, these types of short-term predictions, also known as investment robots, are not suitable when the (maximum) risk level of the pensioner is defined. More in detail, for longer horizons T we consider the sum of annual continuously compounded returns:

$$Z_t^{(A)} = \sum_{i=0}^{T-1} Y_{t+i}^{(A)} \tag{5}$$

Note that the returns $Z_t^{(A)}$ are overlapping, which requires a careful econometric modelling. Assume for illustration a linear relationship in Equation (3) between $Y_t^{(A)}$ and X_{t-1} , as well as some (linear) persistence of the forecasting variable (treating the variables as deviations from their means):

$$Y_t^{(A)} = \beta X_{t-1} + \zeta_t \quad \text{and} \quad X_t = \gamma X_{t-1} + \eta_t, \tag{6}$$

with ζ_t as in Equation (3), η_t being white noise, and slope parameters β and γ . The T -year regression problem that is implied by this pair of one-year regressions is now

$$Z_t^{(A)} = Y_t^{(A)} + \dots + Y_{t+T-1}^{(A)} = (\beta X_{t-1} + \zeta_t) + \dots + (\beta X_{t+T-2} + \zeta_{t+T-1}) \tag{7}$$

$$= \beta \sum_{i=0}^{T-1} \gamma^i X_{t-1} + \beta \sum_{i=0}^{T-1} \sum_{j=0}^{T-1-i} \gamma^j \eta_{t+i} + \sum_{i=0}^{T-1} \zeta_{t+i} = \phi X_{t-1} + \nu_t, \tag{8}$$

that is, the excess stock return for year t over the next T years can be decomposed in two parts: a predictive part depending on the variable X_{t-1} and an unpredictable error term ν_t . To avoid functional misspecification due to our simplistic assumption, we allow for nonlinearity and set up our predictive nonparametric regression model in the same fashion as in Equation (3)

$$Z_t^{(A)} = m(X_{t-1}) + \nu_t, \tag{9}$$

where

$$m(x) = \mathbb{E}(Z_t^{(A)} | X = x), \quad x \in \mathbb{R}^q, \tag{10}$$

is again an unknown smooth function. The important difference between the models (3) and (9) is now that ζ_t is a martingale difference process but v_t will be serially correlated by construction. The predictive variables under consideration collected in the q -dimensional vector X include again the dividend-by-price ratio d , earnings-by-price ratio e , short-term interest rate r , long-term interest rate l , inflation π , term spread s , and the one-year excess stock return $Y^{(A)}$.

2.3. The Local-linear Smoother for the T-year Horizon

As mentioned before, the important difference between the one-year and the T -year case is the inherited serial correlation of the error terms v_t in Equation (9). It is well documented in the statistical literature that, in the presence of correlated errors, quite fundamental problems occur: (i) while the consistency result from Theorem 3 in Kyriakou et al. [6] still holds, the left-out information of the error dependency leads to less efficient estimators [12,13]; and (ii) the commonly applied automatic smoothing parameter selection procedures, like cross-validation or plug-in, break down [14,15]. The latter problem will be discussed in detail in the next section.

More efficient estimators are proposed in the literature. For example, Xiao et al. [12] use a pre-whitening transformation of the dependent variable that has to be estimated from the data. More in detail, the residual process v_t is assumed to be stationary, zero-mean with variance σ_v^2 , and has an invertible linear process representation

$$v_t := \sum_{i=0}^{\infty} c_i \varepsilon_{t-i}, \tag{11}$$

where ε_t are i.i.d. with zero mean and variance σ_ε^2 . Define $c(L) = \sum_{i=0}^{\infty} c_i L^i$ with the usual lag operator L . By inverting $c(L)$ one gets an autoregressive representation of v_t of infinite order:

$$c(L)^{-1} = a(L) = a_0 - \sum_{i=1}^{\infty} a_i L^i \tag{12}$$

and thus $a(L)v_t = \varepsilon_t$. The transformed regression problem (9) is then

$$a(L)Z_t^{(A)} = a(L)m(X_{t-1}) + \varepsilon_t, \quad \text{or} \tag{13}$$

$$\tilde{Z}_t^{(A)} := Z_t^{(A)} - \sum_{i=1}^{\infty} a_i (Z_{t-i}^{(A)} - m(X_{t-1-i})) = m(X_{t-1}) + \varepsilon_t \tag{14}$$

with an uncorrelated error term ε_t . In practice, $\tilde{Z}_t^{(A)}$ is replaced by an approximation based on estimates of the coefficients $\{a_i\}$ and a truncation of the infinite sum. Other contributions which provide more efficient local-polynomial estimators under similar settings can be found in Su and Ullah [13], Linton and Mammen [16], or more recently in Geller and Neumann [17] (and citations therein).

We do not apply such techniques in our paper for several reasons. First, additional parameters $\{a_i\}$ have to be estimated and the infinite sum must be truncated at a meaningful value in Equation (14), or the residual process has to be adequately modeled by some parametric ARMA process or even nonparametrically, where the appropriate lag-length has to be specified. Second, most examples and simulations are given in the literature are one-dimensional. However, we adapt our local-linear smoother in a multidimensional problem and it is not clear what the efficiency gain would be in our scarce data environment. Finally, in a recent study, Scholz et al. [8] show that, in a long-term set-up with annual data, the reduction of the prediction bias is crucial as it contributes squared to the prediction

mean squared error. Our approach of imposing economic structure by using explanatory variables that are transformed according to the chosen benchmark (full benchmarking) aims in a similar direction.

Thus, we think that the more severe problem caused by autocorrelation is the misleading smoothing parameter selection for methods like cross-validation or plug-in. These will be discussed in detail in the next section.

2.4. A Principle of Validation for Model Selection and Smoothing Parameter Choice

For our nonparametric estimation technique, we require an adequate measure of the predictive power. Classical in-sample measures, such as the R^2 or adjusted R^2 , are not appropriate. Note further that in prediction, we are not interested in how well a model explains the variation inside the sample but, instead, in its out-of-sample performance. Therefore, we aim to estimate the prediction error directly.

We follow Nielsen and Sperlich [7] and use the validated R_V^2 which is based on a leave- k -out cross-validation for both models as well as optimal bandwidth (smoothing parameter) selection. This method has been shown to be suitable also in a time series context. For example, Bergmeir et al. [15] show that, in the case of uncorrelated errors, cross-validation is preferred to out-of-sample evaluation where a section from the end of the series is withheld for evaluation. For the latter, only one evaluation on a test data set is possible, whereas cross-validation performs various evaluations. This property is beneficial, especially for small data sets such as ours applied in Section 3.

The validation criterion for one-year predictions is defined as

$$R_{V,1}^2 = 1 - \frac{\sum_t (Y_t^{(A)} - \hat{m}_{-t})^2}{\sum_t (Y_t^{(A)} - \bar{Y}_{-t}^{(A)})^2}, \tag{15}$$

where leave-one-out estimators are used: \hat{m}_{-t} for the conditional mean function m from equation (3) and $\bar{Y}_{-t}^{(A)}$ for the unconditional (historical) mean of $Y_t^{(A)}$. For T -year predictions, we use a similar criterion, now based on $Z_t^{(A)}$ instead of $Y_t^{(A)}$:

$$R_{V,T}^2 = 1 - \frac{\sum_t (Z_t^{(A)} - \hat{m}_{-t})^2}{\sum_t (Z_t^{(A)} - \bar{Z}_{-t}^{(A)})^2}, \tag{16}$$

where leave- k -out estimators are used: \hat{m}_{-t} for the conditional mean function m from Equation (9) and $\bar{Z}_{-t}^{(A)}$ for the unconditional (historical) mean of $Z_t^{(A)}$. Both are computed by removing k observations around the t th time point. Here we use $k = 2T - 1$ due to the construction of the dependent variable over a horizon of T years, that is, for the five-year horizon the leave-nine-out estimator. As it is clear from the context whether we are in the one- or five-year case, we use in the following for both horizons the shorter notation R_V^2 . Note that the R_V^2 measures the predictive power of a given model compared to the cross-validated historical mean; a positive R_V^2 implies that the predictor-based regression model (3) or (9) outperforms the corresponding historical average excess stock return over 1 or T years, respectively.

Cross-validation often requires the omission of more than one single data point and our five-year scenario is one such example. It can also happen that additional corrections are necessary when the omitted fraction of data is considerable [18]. In addition, De Brabanter et al. [14] show that automatic tuning parameter selection methods, such as cross-validation or plug-in, can fail when serial correlation arises (as in our longer-horizon application) and the structure of the error terms is ignored. Here, the problem is that for increasing correlations the chosen bandwidths become smaller, and the corresponding model fits become progressively more under-smoothed [19]. This reduces the bias of the predictor which contributes in a squared fashion to the prediction mean squared error—the numerator

of the ratio in Equations (15) and (16). Consequently, the R_V^2 also increases—not for the better fit but due to the ignored correlation structure. The consequence is a misleading decision on the bandwidth or model specification (the set of preferred covariates). To avoid such problems, Chu and Marron [20] propose the use of bimodal kernel functions which are known to remove the correlation structure very effectively. Nevertheless, the estimator \hat{m} suffers from increased mean squared error [14]. To overcome the problems mentioned, De Brabanter et al. [14] propose a correlation-corrected cross-validation which consists of two steps: i) finding the amount of data k to be left out in the estimation process when a bimodal kernel function is used; and ii) applying the actual choice of the smoothing parameter using leave- k -out cross-validation with a unimodal kernel function. As we know k in our set-up by construction, we can skip the first step. For example, remember that $Z_t^{(A)} = Y_t^{(A)} + \dots + Y_{t+4}^{(A)}$ in the five-year case. Now we have to exclude all $Z_s^{(A)}$ that include any of $Y_t^{(A)}, \dots, Y_{t+4}^{(A)}$. It is easy to see from Figure 1 that this amounts to a leave-nine-out set of $Z_{t-4}^{(A)}, \dots, Z_{t+4}^{(A)}$.

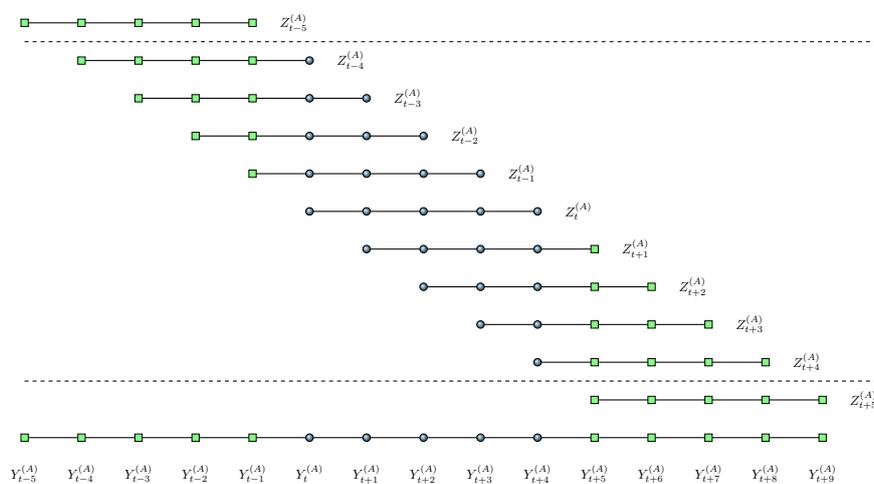


Figure 1. Illustration of the leave-nine-out set (between the dashed lines) of $Z_{t-4}^{(A)}, \dots, Z_{t+4}^{(A)}$ which include at least one element of $Y_t^{(A)}, \dots, Y_{t+4}^{(A)}$ (see bottom).

3. Empirical Results and Discussion

3.1. The Data Set

We applied our local-linear smoother to annual US stock-market data. This dataset was provided by Robert Shiller and is made available from <http://www.econ.yale.edu/~shiller/data.htm>. It includes, among other variables, long-term changes of the Standard and Poor’s (S&P) Composite Stock Price Index, consumer price index changes, and interest rate data from 1872 to 2019. This is an updated and revised version of (Shiller [21], Chapter 26), which provides a detailed description of the data.

Note that the extension of the risk-free rate series (based on the six-month commercial paper rate until 1997 and, afterward, on the six-month certificate of deposit rate, secondary market) was not possible as it was discontinued in 2013. Here, we followed the strategy of Welch and Goyal [22] and Mammen et al. [23] and replace this variable by an annual yield based on the six-month Treasury-bill rate, secondary market, from <https://fred.stlouisfed.org/series/TB6MS>. This new series was only available from 1958 to 2019. In the 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. Assuming a linear relationship and using an ordinary least squares regression, we obtained the estimated equation:

$$\text{Treasury-bill rate} = 0.0961 + 0.8648 \times \text{commercial paper rate}, \tag{17}$$

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 was 99.3%.

3.2. Descriptive Analysis

There is much research on the predictability of returns and a lot is known about the characteristics of short-horizon stock returns. For example, stylized facts about daily and monthly returns include excess kurtosis, distributions which are not normal, and volatility clustering [24]. However, less is known about distributions of long-horizon returns. However, such characteristics are of central interest to investors saving for distant pay-offs.

Figure 2 shows the time series of the one-year returns (left) and five-year returns (right), both in excess of the risk-free benchmark R , which are displayed on the same scale for the sake of comparison. The five-year series exhibits larger positive returns, which is not surprising as a longer period under risk should be paid-off with a higher risk premium. The autoregressive structure of the five-year returns can be easily seen in comparison to the assumed weak dependence of one-year returns.

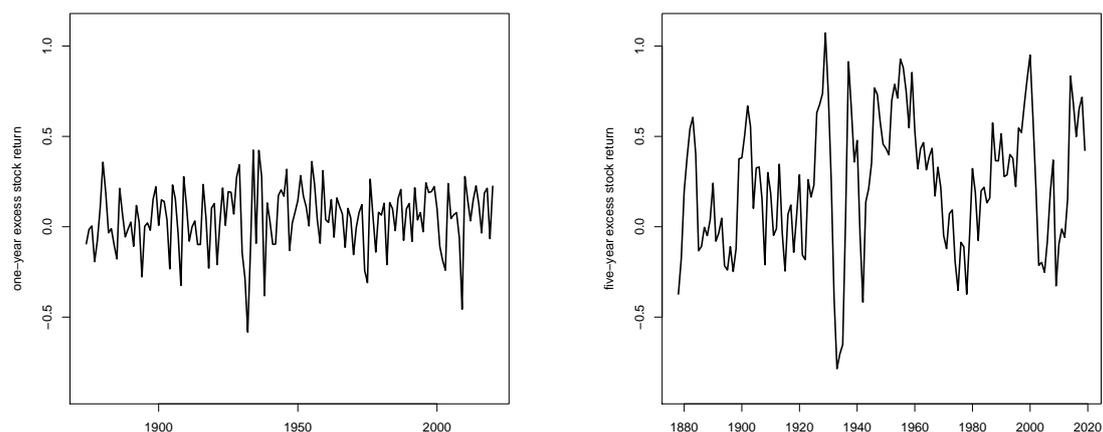


Figure 2. Left: one-year stock returns in excess of the risk-free benchmark. Right: five-year stock returns in excess of the risk-free benchmark. Period: 1872–2019. Data: annual S&P 500.

Figure 3 shows histograms of the one-year returns (left) and five-year returns (right) together with a kernel density estimate (red) and a fitted normal distribution (green). One notes again that the distribution of the five-year returns is shifted to the right but has a higher volatility. A Jarque–Bera test of the hypothesis of normality does reject for one-year returns (p -value = 0.013) but does not reject for five-year returns (p -value = 0.522). Furthermore, the density estimate for the five-year returns indicates more a mixture of normals than a single normal distribution which represents some evidence of a possible structural break in the data-generating process.

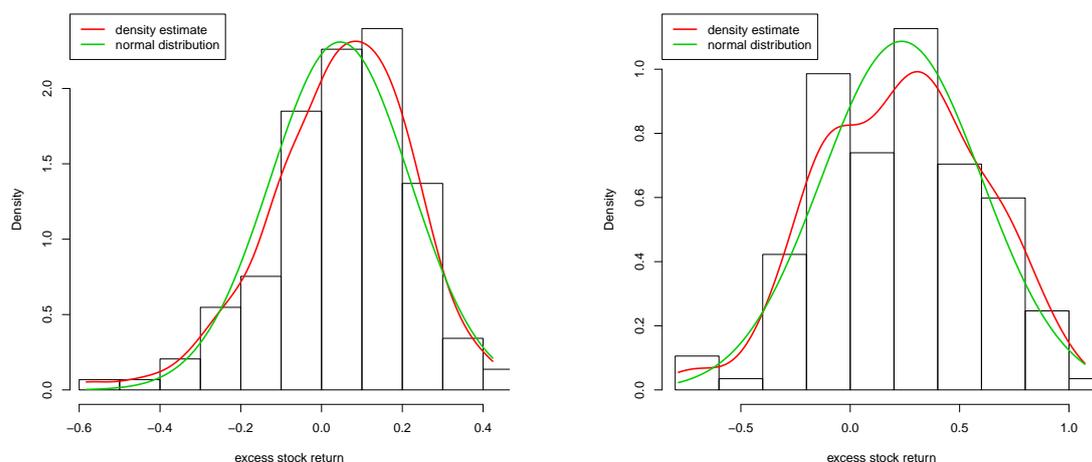


Figure 3. Histogram, kernel density estimate (red), and fitted normal distribution (green). **Left:** one-year stock returns in excess of the risk-free benchmark. **Right:** five-year stock returns in excess of the risk-free benchmark. Period: 1872–2019. Data: annual S&P 500.

Including structural changes in the modelling process could be beneficial, as shown in the literature especially for higher-frequency returns [25,26]. However, this comes with additional effort which is beyond the scope of our article and is left for future research. Several important points would have to be taken into consideration, for example: (i) it is not clear for which point in time one should incorporate a structural break in our annual data (the Great Recession, the Second World War, the Global Financial Crisis, the Bretton Woods agreement, etc.); (ii) a simple sample split would result in even smaller and not balanced data sets. From a statistical perspective, as we apply a fully nonparametric method, this would lead to mostly one-dimensional and potentially linear models. This way, we would lose the analysis of higher-dimensional models and nonlinear relationships between excess stock returns and their predictor variables.

This section is concluded with Table 1 which displays standard descriptive statistics for one-year and five-year returns as well as the available covariates. Both the one-year and five-year excess returns had a negative skewness, that is, the left tail of the distribution (large negative returns) was longer or fatter than the right tail (large positive returns). Note that this is more pronounced in the case of one-year rather than five-year returns. While one-year returns were leptokurtic (positive excess kurtosis of 0.68), five-year returns exhibited a small negative excess kurtosis of -0.37 .

Table 1. US market data (1872–2019).

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

Similar plots to those in Figures 2 and 3 and information as in Table 1 for the other benchmarks are available upon request by the authors. In the next sections, we analyzed the predictability of one-year and five-year stock returns in excess of the different benchmarks.

3.3. The Single Benchmarking Approach

In this section, we considered a single benchmarking approach as in Kyriakou et al. [6] where only the variable S_t was adjusted according to some benchmark $B_{t-1}^{(A)}$, as shown in (1), while the independent variable(s) is (are) measured on the original nominal scale. The models (3) and (9) are estimated with a local-linear kernel smoother using the quartic kernel and the optimal bandwidth is chosen by cross-validation, that is, by maximizing the R_V^2 given by (15) and (16). Moreover, it should be kept in mind that the nonparametric method can estimate linear functions without any bias, given that we apply a local-linear smoother. Thus, the linear model is automatically embedded in our approach. We study the empirical findings of R_V^2 values based on different validated scenarios shown for the one-year horizon in Table 2 and the five-year horizon in Table 3. Note that the one-year predictions may differ from those originally reported in Kyriakou et al. [6] due to the updated data set and the replacement of the commercial paper rate by the Treasury-bill rate; nevertheless, the models remain similarly ranked.

We found that in the case of the five-year returns, which was the focal point of this paper, the term spread s was, overall, the most powerful predictive variable for excess stock returns; this superior performance was also observed in the one-year case. More in detail, with the prediction constrained to using only one-dimensional covariates, the term spread is the best predictor for the one-year and five-year horizon under the short interest benchmark $B^{(R)}$ with, respectively, $R_V^2 = 9.7\%$ and 15.5% , but this also does quite well in the one-year case under the long rate and earnings-by-price benchmarks, $B^{(L)}$ and $B^{(E)}$, with $R_V^2 \in \{6.2\%, 7.5\%\}$ (for $B^{(C)}$ the best is π with 10.3%). In the five-year case under $B^{(L)}$ and $B^{(E)}$, the term spread s yields a high $R_V^2 \in \{8.0\%, 11.5\%\}$ whereas under $B^{(C)}$ the dividend-by-price ratio d gains ground with $R_V^2 = 7.6\%$.

In light of these remarks, we therefore, focused the spotlight on the relationship between the spread and the excess stock returns. We present in the top panel of Figure 4 the estimated functions \hat{m} (red solid line) for the one-year horizon (left) and the five-year horizon (right) under the single risk-free benchmark together with a corresponding linear model (dash-dotted green line), and a 45-degree line (dashed black line). Figure 4 shows thereby the three single covariates with the largest R_V^2 ($T = 1$ and $T = 5$): the term spread (9.7% and 15.5%), the short-term interest rate (3.0% and 7.8%), and the long-term interest rate (0.0% and 1.4%). Our findings for both horizons conformed to the fact that an increase in the spread corresponds to an increase in the excess stock return. While a positive spread corresponds to a positive return for the one-year case, a spread larger than -1% gave, on average, a positive five-year return. This finding is in line with, for example, Resnick and Shoesmith [27] who find that the value of the yield spread holds important information about the probability of a bear stock market. Regarding our validation procedure, Figure 4 also confirms that our approach of correcting for autocorrelation in the five-year prediction problem was successful. The estimated functions are quite smooth indicating that the chosen bandwidth is not too small and that the resulting fit and validated R^2 are reasonable.

Table 2. Predictive power measured by R_V^2 (%) for one-year excess stock returns $Y_t^{(A)}$: the single benchmarking approach.

Benchmark $B^{(A)}$	Explanatory Variable(s) X_{t-1}						
	$Y^{(A)}$	d	e	r	l	π	s
Short-term rate	-1.6	-1.1	-0.6	3.0	0.0	-1.4	9.7
Long-term rate	-1.8	-0.8	-0.4	1.9	0.0	-1.4	6.2
Earnings-by-price	-1.7	-1.2	-1.4	0.0	-0.8	-1.2	7.5
Inflation	-1.4	-0.2	-1.5	0.8	-0.8	10.3	7.2
	$(Y^{(A)}, d)$	$(Y^{(A)}, e)$	$(Y^{(A)}, r)$	$(Y^{(A)}, l)$	$(Y^{(A)}, \pi)$	$(Y^{(A)}, s)$	
Short-term rate	-2.6	-2.4	0.9	-2.4	-2.9	6.3	
Long-term rate	-2.4	-2.3	-0.2	-2.4	-3.1	2.7	
Earnings-by-price	-3.5	-3.7	-2.0	-2.8	-2.8	4.5	
Inflation	-1.6	-3.4	-0.9	-2.5	9.7	4.8	
	(d, e)	(d, r)	(d, l)	(d, π)	(d, s)		
Short-term rate	-2.9	2.1	-1.6	-2.6	9.3		
Long-term rate	-2.7	1.3	-1.3	-2.3	5.8		
Earnings-by-price	-3.7	-1.4	-2.2	-2.4	6.0		
Inflation	-1.9	0.8	-1.2	9.5	7.9		
	(e, r)	(e, l)	(e, π)	(e, s)			
Short-term rate	4.0	-1.1	-1.6	9.1			
Long-term rate	3.2	-0.5	-1.3	5.5			
Earnings-by-price	-1.4	-2.3	-2.7	5.4			
Inflation	-0.4	-2.5	10.9	5.4			
	(r, l)	(r, π)	(r, s)				
Short-term rate	8.5	1.4	10.0				
Long-term rate	4.9	0.3	6.5				
Earnings-by-price	6.0	-1.5	7.2				
Inflation	5.2	9.5	7.4				
	(l, π)	(l, s)					
Short-term rate	-2.1	10.1					
Long-term rate	-2.0	6.6					
Earnings-by-price	-2.0	7.0					
Inflation	9.9	7.4					
	(π, s)						
Short-term rate	7.7						
Long-term rate	4.1						
Earnings-by-price	5.2						
Inflation	15.4						

Back to our discussion of the results in Tables 2 and 3, in broad terms, five-year predictability improved over one-year: 67 out of 112 models achieve a larger R_V^2 , and we observed 64 (five-year) versus 52 (one-year) models with nonnegative R_V^2 , that is, our proposed predictor-based regression model for the longer forecast horizon in this application surpassed the historical average excess return in the majority of cases. In addition, combining the term spread s with the dividend-by-price d results in uplifted predictability to 26.2%; this combination is, in fact, the best-performing one for 3 out of 4 benchmarks ($B^{(R)}, B^{(L)}, B^{(C)}$). In particular, imposing an additional covariate to s results in one-year R_V^2 in the range 6–10% under $B^{(R)}$; under other benchmarks, such as $B^{(L)}$, the one-year R_V^2 is in the range 3–6.5% (approx.); changing to the $B^{(E)}$ benchmark results in R_V^2 in the range 4.5–7.5%. Interestingly, for a five-year horizon, we observed a substantially improved predictive power with our cross-validated R_V^2 ranking some two-dimensional better than one-dimensional models, in fact, more than for a one-year horizon: in particular, as possibly anticipated by the aforementioned performances

of d and s , the two-dimensional covariate (d, s) boosts R_V^2 to 26.2% under $B^{(R)}$, performs best with 21% under $B^{(L)}$, and comes second with 12% under $B^{(E)}$ being beaten by $(Y^{(E)}, s)$ with 14.1%.

Table 3. Predictive power measured by R_V^2 (%) for five-year excess stock returns $Z_t^{(A)}$: the single benchmarking approach.

Benchmark $B^{(A)}$	Explanatory Variable(s) X_{t-1}						
	$Y^{(A)}$	d	e	r	l	π	s
Short-term rate	0.9	1.1	-1.5	7.8	1.4	-1.8	15.5
Long-term rate	1.1	4.6	0.5	3.9	1.0	-1.0	8.0
Earnings-by-price	1.4	-3.8	-3.6	-4.7	-1.4	-1.4	11.5
Inflation	1.3	7.6	-3.9	-6.7	-3.5	6.8	0.8
	$(Y^{(A)}, d)$	$(Y^{(A)}, e)$	$(Y^{(A)}, r)$	$(Y^{(A)}, l)$	$(Y^{(A)}, \pi)$	$(Y^{(A)}, s)$	
Short-term rate	-1.5	-2.8	8.2	2.2	-1.7	16.4	
Long-term rate	2.4	-0.3	4.4	1.6	-0.5	9.0	
Earnings-by-price	-4.5	-3.9	-4.9	-0.6	-0.4	14.1	
Inflation	5.9	-4.8	-5.7	-3.0	7.3	2.2	
	(d, e)	(d, r)	(d, l)	(d, π)	(d, s)		
Short-term rate	-3.1	6.4	-4.1	-2.3	26.2		
Long-term rate	1.0	5.8	-0.1	1.6	21.0		
Earnings-by-price	-7.5	-12.1	-6.6	-6.1	12.0		
Inflation	3.4	0.3	1.8	10.8	13.7		
	(e, r)	(e, l)	(e, π)	(e, s)			
Short-term rate	8.1	-4.5	-1.5	19.1			
Long-term rate	7.9	-3.3	1.7	12.6			
Earnings-by-price	-11.4	-8.8	-4.5	10.4			
Inflation	-9.5	-12.5	9.7	-1.1			
	(r, l)	(r, π)	(r, s)				
Short-term rate	14.7	5.6	14.8				
Long-term rate	6.5	2.7	7.0				
Earnings-by-price	9.2	-6.5	9.0				
Inflation	-6.4	0.9	-5.9				
	(l, π)	(l, s)					
Short-term rate	-1.4	13.4					
Long-term rate	-0.5	5.8					
Earnings-by-price	-1.8	9.2					
Inflation	7.5	-5.8					
	(π, s)						
Short-term rate	15.4						
Long-term rate	8.8						
Earnings-by-price	11.0						
Inflation	8.5						

In the one-year case, quite remarkable is the predictor π , either in itself or combined with covariates $Y^{(C)}, d, e, r, l$, under the inflation benchmark $B^{(C)}$ leading to R_V^2 in the range 9.5–15.4%. In addition, when put together with the term spread, the resulting combination (π, s) under $B^{(C)}$ is the clear winner reaching up to $R_V^2 = 15.4\%$. This is probably good news in an actuarial context where the inflation benchmark can be seen as an important one in pension product applications. In the five-year case, π still does quite well in the range 6.8–10.8% (with an exception of 0.9% for (r, π)) and remains generally the best predictor under $B^{(C)}$, nevertheless, it is no longer the globally best one.

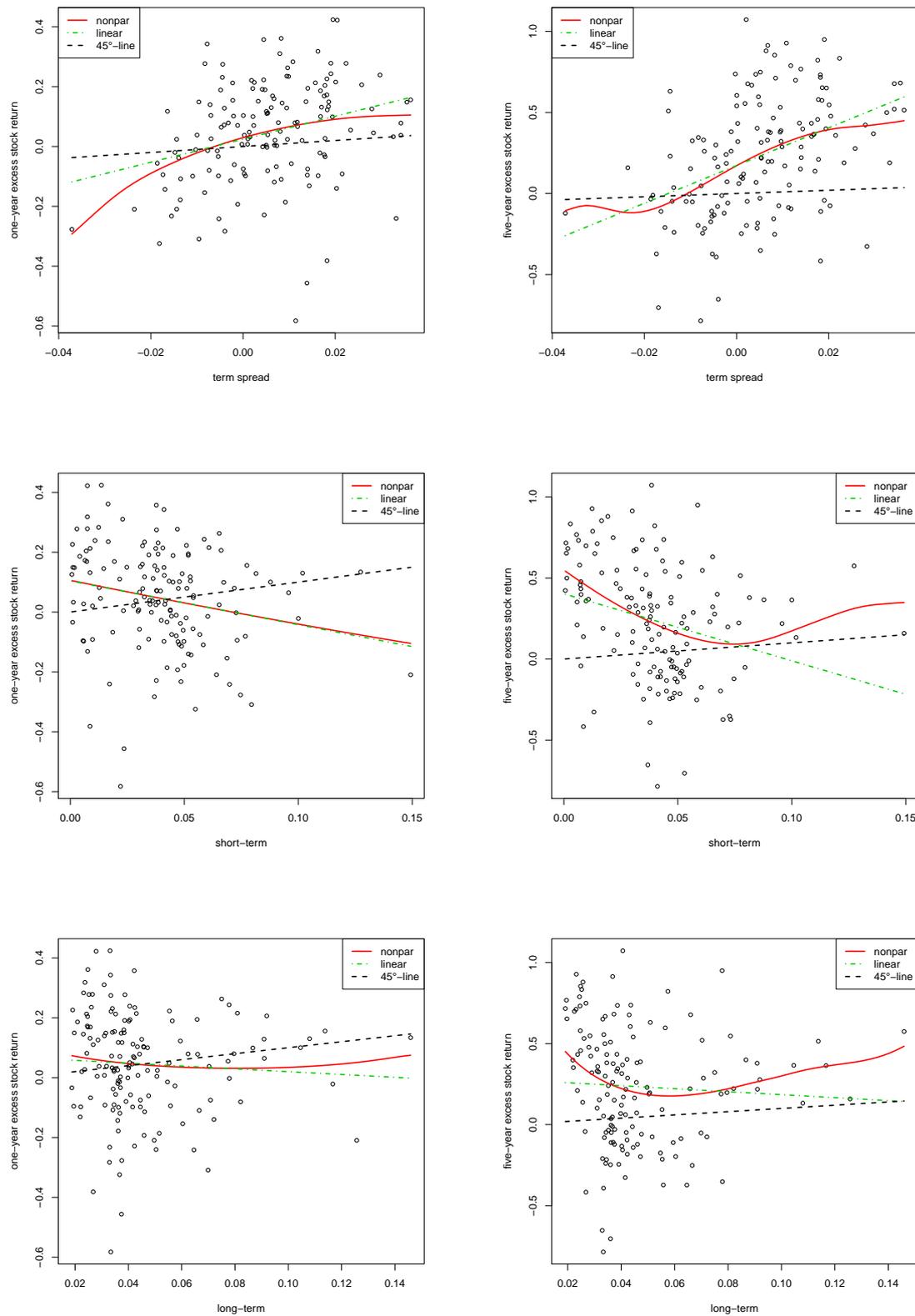


Figure 4. Single risk-free benchmark. Relation between excess stock returns and the spread (**top**), the short-term interest rate (**middle**), and the long-term interest rate (**bottom**). Estimated nonparametric function \hat{m} (red solid line), linear model (dash-dotted green line), and 45-degree line (dashed black line). Left: one-year horizon. Right: five-year horizon. Period: 1872–2019. Data: annual S&P 500.

3.4. The Full Benchmarking Approach

The next step now is to analyze whether transforming the explanatory variables can improve predictions. Recall that fully nonparametric models suffer in general by the curse of dimensionality, as in our framework where we confront sparsely distributed annual observations in higher dimensions. Importing more structure in the estimation process can help reduce or circumvent such problems.

Here, we extend the study in Section 3.3 using economic structure in the sense that we consider adjusting both the independent and dependent variables according to the same benchmark. To this end, in our full (double) benchmarking approach, the prediction problems are reformulated as

$$Y_t^{(A)} = m(X_{t-1}^{(A)}) + \xi_t, \tag{18}$$

$$Z_t^{(A)} = m(X_{t-1}^{(A)}) + \nu_t, \tag{19}$$

where we use transformed predictive variables

$$X_{t-1}^{(A)} = \begin{cases} \frac{1+X_{t-1}}{B_{t-1}^{(A)}}, & X \in \{d, e, r, l, \pi\} \\ \frac{s_{t-1}}{B_{t-1}^{(A)}} = \frac{l_{t-1}-r_{t-1}}{B_{t-1}^{(A)}} \\ Y_{t-1}^{(A)} \end{cases}, \quad A \in \{R, L, E, C\}. \tag{20}$$

This approach can be interpreted as a way of reducing dimensionality of the estimation procedure as $X_{t-1}^{(A)}$ encompasses an additional predictive variable.

Table 4. Predictive power measured by R_V^2 (%) for one-year excess stock returns $Y_t^{(A)}$: the double benchmarking approach.

Benchmark $B^{(A)}$	Explanatory Variable(s) $X_{t-1}^{(A)}$						
	$Y^{(A)}$	$d^{(A)}$	$e^{(A)}$	$r^{(A)}$	$l^{(A)}$	$\pi^{(A)}$	$s^{(A)}$
Short-term rate	-1.6	3.1	5.2	-	9.5	-1.3	9.5
Long-term rate	-1.8	-0.2	0.7	6.1	-	-1.5	6.1
Earnings-by-price	-1.7	-2.3	-	-0.2	-1.0	-0.7	7.4
Inflation	-1.4	10.4	12.2	7.2	10.5	-	6.5
	$(Y^{(A)}, d^{(A)})$	$(Y^{(A)}, e^{(A)})$	$(Y^{(A)}, r^{(A)})$	$(Y^{(A)}, l^{(A)})$	$(Y^{(A)}, \pi^{(A)})$	$(Y^{(A)}, s^{(A)})$	
Short-term rate	1.8	3.2	-	6.0	-2.9	6.0	
Long-term rate	-1.9	-1.1	2.6	-	-3.1	2.6	
Earnings-by-price	-4.1	-	-2.3	-3.2	-2.7	4.3	
Inflation	10.8	11.5	6.2	9.6	-	4.1	
	$(d^{(A)}, e^{(A)})$	$(d^{(A)}, r^{(A)})$	$(d^{(A)}, l^{(A)})$	$(d^{(A)}, \pi^{(A)})$	$(d^{(A)}, s^{(A)})$		
Short-term rate	2.4	-	9.8	1.5	9.8		
Long-term rate	-1.6	6.3	-	-1.8	6.3		
Earnings-by-price	-	-3.2	-3.6	-3.5	4.0		
Inflation	10.3	9.5	10.0	-	15.7		
	$(e^{(A)}, r^{(A)})$	$(e^{(A)}, l^{(A)})$	$(e^{(A)}, \pi^{(A)})$	$(e^{(A)}, s^{(A)})$			
Short-term rate	-	10.7	3.3	10.7			
Long-term rate	7.1	-	-0.5	7.1			
Earnings-by-price	-	-	-	-			
Inflation	11.4	11.3	-	17.8			
	$(r^{(A)}, l^{(A)})$	$(r^{(A)}, \pi^{(A)})$	$(r^{(A)}, s^{(A)})$				
Short-term rate	-	-	-				
Long-term rate	-	3.6	-				
Earnings-by-price	4.9	-2.1	5.7				
Inflation	13.9	-	14.8				

Table 4. Cont.

Benchmark $B^{(A)}$	Explanatory Variable(s) $X_{t-1}^{(A)}$	
	$(I^{(A)}, \pi^{(A)})$	$(I^{(A)}, s^{(A)})$
Short-term rate	7.2	–
Long-term rate	–	–
Earnings-by-price	–2.4	5.4
Inflation	–	14.7
	$(\pi^{(A)}, s^{(A)})$	
Short-term rate	7.2	
Long-term rate	3.6	
Earnings-by-price	5.1	
Inflation	–	

Results of this empirical study are presented for the one-year horizon in Table 4 and for the five-year horizon in Table 5. In addition, Figure 5 presents the three single covariates with the largest R_V^2 ($T = 1$ and $T = 5$) for the double inflation benchmark case: the earnings-by-price ratio (12.2% and 12.4%), the dividend-by-price ratio (10.4% and 10.9%), and the long-term interest rate (10.5% and 8.7%).

We find that, in the majority of cases, the full outruns the single benchmarking approach, even more when we consider a longer horizon, and the number of models with nonnegative R_V^2 (that is, cases of beating the historical average excess return) increases: 68 out of 82 models (full benchmarking, five-year); 55 out of 82 (full benchmarking, one-year); 64 out of 112 (single benchmarking, five-year); and 52 out of 112 (single benchmarking, one-year).

The pair $(d^{(R)}, s^{(R)})$ in the full benchmarking approach for a five-year horizon yields $R_V^2 = 21.9\%$ against 26.2% in the single benchmarking under $B^{(R)}$, whereas $(e^{(C)}, s^{(C)})$ in the full benchmarking approach for a one-year horizon yields $R_V^2 = 17.8\%$ against 15.4% using the predictor (π, s) in the single benchmarking under $B^{(C)}$. It, therefore, seems that s is an important predictor, whose power is mostly unveiled when combining with another predictor depending on the benchmark choice and the forecast horizon. In addition, although under $B^{(R)}$ and $B^{(L)}$, full benchmarking does not improve predictability, it does under $B^{(E)}$ and, especially, $B^{(C)}$ which is important if we aim to identify a likely common well-performing benchmark and predictor, that is, $(e^{(C)}, s^{(C)})$, independently of the horizon length. For $B^{(C)}$ full benchmarking, R_V^2 lies in the range 14.7–10.1% (five-year) and 17.8–11.5% (one-year), which are both an improvement from $B^{(C)}$ single benchmarking yielding R_V^2 in 13.7–0.9% (five-year) and 15.4–9.5% (one-year), that is, a maximum width reduction by a factor of almost 3 for the five-year horizon.

Overall, we conclude that the term spread is a good predictor; if we aim to homogenize our choice of predictor and benchmark with respect to the horizon length, then the earnings-by-price and the term spread under the inflation benchmark would be an ideal compromise, even if not the winning one. This is welcoming, as, for example, in pension research or other long-term saving strategies, it is sensible to look at real value and employ such a model with all returns and covariates net-of-inflation.

Table 5. Predictive power measured by R_V^2 (%) for five-year excess stock returns $Z_t^{(A)}$: the double benchmarking approach.

Benchmark $B^{(A)}$	Explanatory Variable(s) $X_{t-1}^{(A)}$						
	$\gamma^{(A)}$	$d^{(A)}$	$e^{(A)}$	$r^{(A)}$	$l^{(A)}$	$\pi^{(A)}$	$s^{(A)}$
Short-term rate	0.9	12.1	10.4	–	15.5	–2.5	15.5
Long-term rate	1.1	8.5	0.8	8.0	–	–1.6	8.0
Earnings-by-price	1.4	8.4	–	–4.9	–3.7	–0.7	11.4
Inflation	1.3	10.9	12.4	5.7	8.7	–	0.8
	$(\gamma^{(A)}, d^{(A)})$	$(\gamma^{(A)}, e^{(A)})$	$(\gamma^{(A)}, r^{(A)})$	$(\gamma^{(A)}, l^{(A)})$	$(\gamma^{(A)}, \pi^{(A)})$	$(\gamma^{(A)}, s^{(A)})$	
Short-term rate	10.8	9.9	–	16.5	–1.5	16.5	
Long-term rate	3.2	–0.1	9.2	–	–1.1	9.2	
Earnings-by-price	5.3	–	–5.3	–5.1	–0.2	14.0	
Inflation	11.1	12.4	5.8	9.0	–	2.2	
	$(d^{(A)}, e^{(A)})$	$(d^{(A)}, r^{(A)})$	$(d^{(A)}, l^{(A)})$	$(d^{(A)}, \pi^{(A)})$	$(d^{(A)}, s^{(A)})$		
Short-term rate	8.5	–	21.9	7.3	21.9		
Long-term rate	1.4	13.8	–	5.0	13.8		
Earnings-by-price	–	–1.6	4.1	5.4	15.5		
Inflation	9.5	4.1	1.7	–	13.0		
	$(e^{(A)}, r^{(A)})$	$(e^{(A)}, l^{(A)})$	$(e^{(A)}, \pi^{(A)})$	$(e^{(A)}, s^{(A)})$			
Short-term rate	–	21.4	8.2	21.4			
Long-term rate	16.4	–	2.5	16.4			
Earnings-by-price	–	–	–	–			
Inflation	8.6	4.9	–	14.7			
	$(r^{(A)}, l^{(A)})$	$(r^{(A)}, \pi^{(A)})$	$(r^{(A)}, s^{(A)})$				
Short-term rate	–	–	–				
Long-term rate	–	9.5	–				
Earnings-by-price	5.9	–6.2	6.0				
Inflation	10.8	–	10.0				
	$(l^{(A)}, \pi^{(A)})$	$(l^{(A)}, s^{(A)})$					
Short-term rate	16.0	–					
Long-term rate	–	–					
Earnings-by-price	–6.4	6.0					
Inflation	–	10.1					
	$(\pi^{(A)}, s^{(A)})$						
Short-term rate	16.0						
Long-term rate	9.5						
Earnings-by-price	11.9						
Inflation	–						

3.5. Real-Income Long-Term Pension Prediction

In long-term pension planning, real-income protection is often an important aspect [28–30]. When optimizing investment asset allocation for the long term, one therefore needs a good econometric model in real terms. Based on the research in this paper and in Kyriakou et al. [6], we are able to conclude that, in the natural double benchmark setting for real-income econometrics, it looks like earnings divided by price is the natural covariate to consider. In Table 2 of Kyriakou et al. [6] and in Table 4 of this paper, it is concluded that earnings divided by price is the best single covariate to use in the double inflation benchmark case and, in this paper’s Table 5, this is also the case in the five-year view. On balance, we therefore conclude that the intuitively appealing earnings divided by price is a good long-term predictor for real-income forecasting. In the one-year view of Kyriakou et al. [6], the nonparametric smoother estimated for the relationship between earnings divided by price and return in the inflation double benchmarking case has the exact functional form of a simple line. So, even though we consider a nonparametric estimator that can pick any functional form, the resulting

functional form is a simple line. This provides us with a strong argument for using the simple line in this case. The functional form of a line has been picked via a validation procedure against all functional forms. The linear expression is

$$\text{Real one-year stock return} = 0.004875 + 1.119 \times \text{real earnings-by-price.} \tag{21}$$

Notice that a very good long-term predictor of real income can, therefore, be expressed as a simple linear relationship, where the expected return adds first 12% to the earnings divided by price and then another 0.5%. This is a very simple relationship that is easy for long-term investors to remember. Similarly to the one-year view, our validation procedure exactly picks a line against all other functional alternatives in the five-year view. The linear form for the five-year view is

$$\text{Real five-year stock return} = 0.2068 + 2.264 \times \text{real earnings-by-price.} \tag{22}$$

The top panel of Figure 5 shows the estimated nonparametric function \hat{m} (red solid line) for the one-year horizon (left) and five-year horizon (right) under the double inflation benchmark for the earnings-by-price covariate together with the corresponding linear model (dash-dotted green line), and a 45-degree line (dashed black line). Note that the linear relationship discovered for the earnings-by-price predictor must not hold true for other covariates or their combinations. In those cases, the full benefit of our approach comes into its own. For example, the bottom panel of Figure 5 clearly shows nonlinearities for the five-year case when the long-term interest rate is considered. For a suitable statistical smoothing-based test (nonparametric versus linear model), see, for example, the test based on wild bootstrap proposed in Scholz et al. [8] or the discussion in the survey of González-Manteiga and Crujeiras [31].

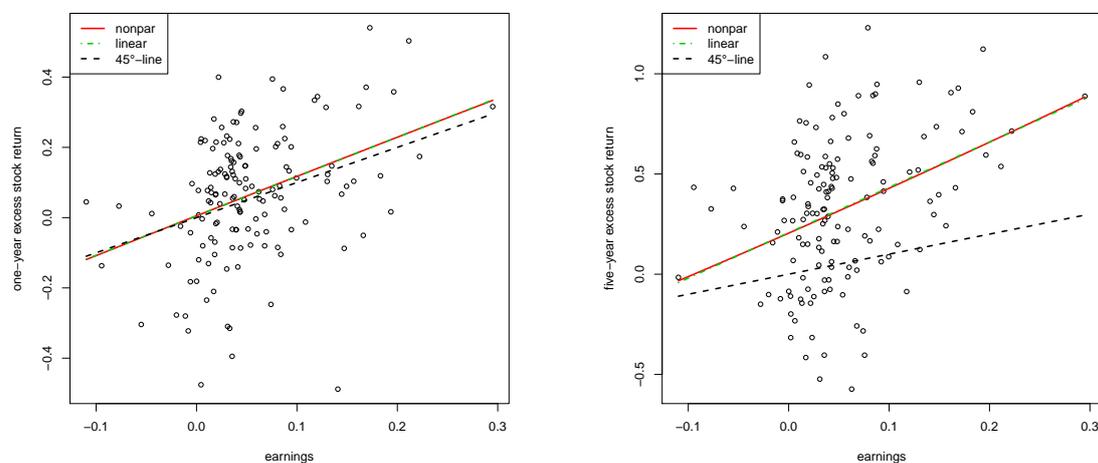


Figure 5. Cont.

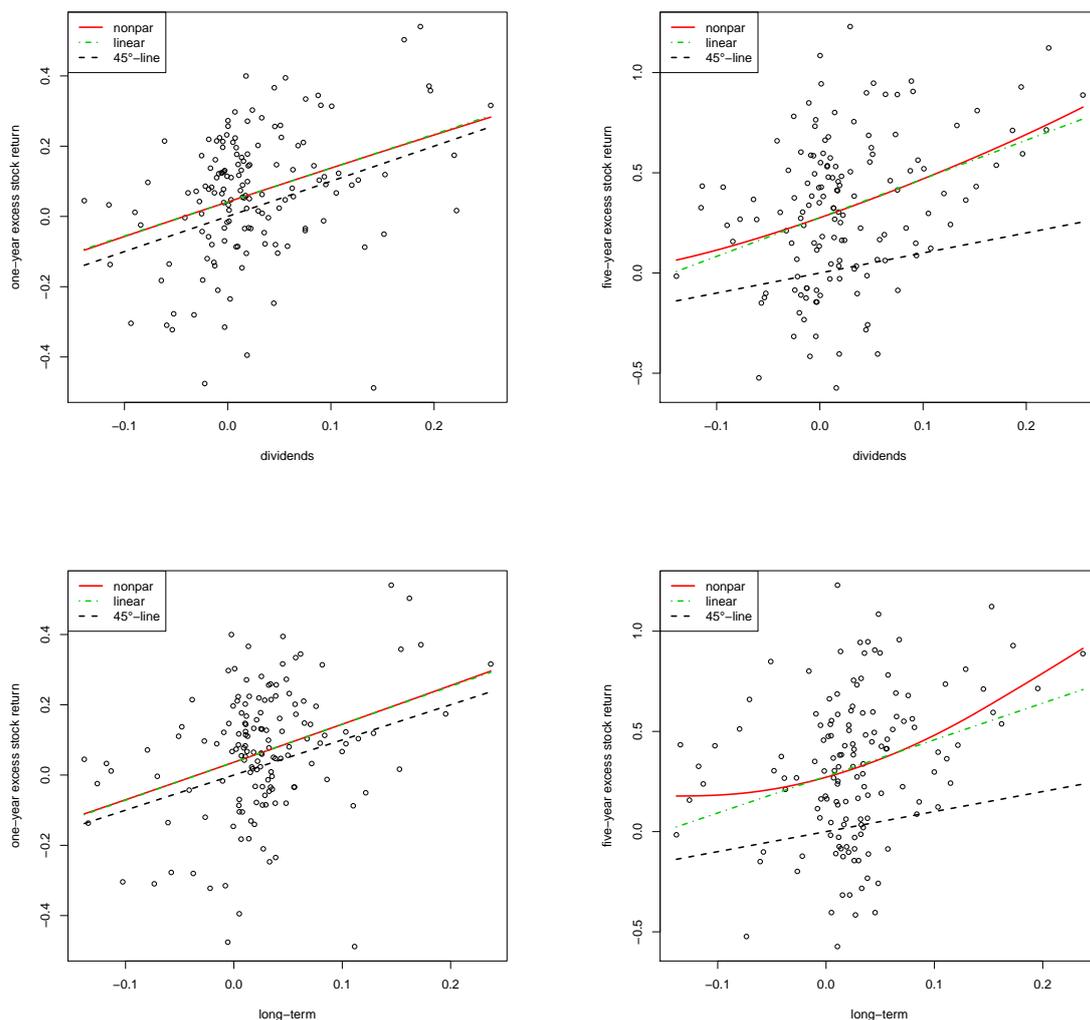


Figure 5. Double inflation benchmark. Relation between real stock returns and real earnings-by-price (**top**), real dividends-by-price (**middle**), and the real long-term interest rate (**bottom**). Estimated nonparametric function \hat{m} (red solid line), linear model (dash-dotted green line), and 45-degree line (dashed black line). Left: one-year horizon. Right: five-year horizon. Period: 1872–2019. Data: annual S&P 500.

3.6. One-Year ahead Real-Time Predictions

The four benchmarks proposed in this paper are useful in different situations. While in Section 3.5 we focused on real-income long-term predictions based on models using the inflation-double benchmark, we now want to explore the development of ‘pure’ stock returns S (without a benchmark) in the near future. Kyriakou et al. [6] found the model using the earnings benchmark with the term spread as the covariate, that is, $\hat{Y}^{(E)} = \hat{m}(s)$, to perform best in terms of R_V^2 when the predicted values are back-transformed and validated on stock returns S . We use this simple model to illustrate the usefulness of the earnings benchmark. For this purpose, we estimate this model over the full sample as before and predict the stock returns in excess of earnings using the current spread (in the period September 2018–March 2020). Finally, we back-transform those predictions to get a one-year ahead forecast for nominal stock returns $\hat{S}_{t,nom}$ from

$$\ln \hat{S}_{t,nom} = \hat{Y}_t^{(A)} + B_{t-1}^{(A)}. \tag{23}$$

As a by-product, we also calculate a prediction for the real-stock return, $\hat{S}_{t,real} = \hat{S}_{t,nom} - \pi_t$, and the risk premium, $\widehat{RP}_t = \hat{S}_{t,nom} - R_t$, of holding stocks versus a risk-free asset. Table 6 shows the results of this forecasting exercise. The considered forecasting period is of interest for two specific features: (i) the term spread is U-shaped, that is, it reduces, gets negative (an inverted yield curve in August and September 2019), and finally increases again; and (ii) the external shock to the market caused by the Covid-19 pandemic leads to large negative returns starting in March 2020.

Table 6. One-year ahead real-time forecasts: predictions from the model using the single-earnings benchmark and the term spread s as covariate ($\hat{Y}^{(E)} = \hat{m}(s)$) and estimated over the full sample period.

date	US Stock Market Data									Predictions			
	P	D	E	R	L	s	π	d	e	$\hat{Y}^{(E)}$	\hat{S}_{nom}	\hat{S}_{real}	\widehat{RP}
2018-09	2901.50	52.34	130.39	2.47	3.00	0.53	2.28	1.80	4.49	1.78	6.18	3.90	3.71
2018-10	2785.46	52.81	131.06	2.56	3.15	0.59	2.52	1.90	4.71	1.97	6.57	4.05	4.01
2018-11	2723.23	53.28	131.72	2.60	3.12	0.52	2.18	1.96	4.84	1.75	6.47	4.30	3.87
2018-12	2567.31	53.75	132.39	2.57	2.83	0.26	1.91	2.09	5.16	0.88	5.91	4.00	3.34
2019-01	2607.39	54.15	133.06	2.50	2.71	0.21	1.55	2.08	5.10	0.70	5.68	4.13	3.18
2019-02	2754.86	54.54	133.72	2.47	2.68	0.21	1.52	1.98	4.85	0.70	5.44	3.92	2.97
2019-03	2803.98	54.94	134.39	2.41	2.57	0.16	1.86	1.96	4.79	0.53	5.21	3.34	2.80
2019-04	2903.80	55.32	134.68	2.34	2.53	0.19	2.00	1.91	4.64	0.63	5.17	3.17	2.83
2019-05	2854.71	55.70	134.98	2.27	2.40	0.13	1.79	1.95	4.73	0.42	5.04	3.25	2.77
2019-06	2890.17	56.08	135.27	1.94	2.07	0.13	1.65	1.94	4.68	0.42	4.99	3.34	3.05
2019-07	2996.11	56.46	134.48	1.91	2.06	0.15	1.81	1.88	4.49	0.49	4.88	3.07	2.97
2019-08	2897.45	56.84	133.69	1.73	1.63	-0.10	1.75	1.96	4.61	-0.44	4.07	2.32	2.34
2019-09	2982.16	57.22	132.90	1.75	1.70	-0.05	1.71	1.92	4.46	-0.25	4.11	2.40	2.36
2019-10	2977.68	57.56	135.09	1.57	1.71	0.14	1.76	1.93	4.54	0.45	4.89	3.13	3.32
2019-11	3104.90	57.90	137.28	1.53	1.81	0.28	2.05	1.86	4.42	0.95	5.27	3.22	3.74
2019-12	3176.75	58.24	139.47	1.51	1.86	0.35	2.29	1.83	4.39	1.19	5.48	3.20	3.97
2020-01	3278.20	58.69	138.43	1.49	1.76	0.27	2.49	1.79	4.22	0.91	5.05	2.56	3.56
2020-02	3277.31	59.13	137.39	1.37	1.50	0.13	2.33	1.80	4.19	0.42	4.52	2.19	3.15
2020-03	2652.39	59.58	136.35	0.32	0.87	0.55	1.54	2.25	5.14	1.85	6.86	5.32	6.54

We find i) that for the (slightly) negative spread the predicted stock return in excess of earnings is also negative. Nevertheless, the predicted nominal stock return is positive (around 4.1% and mainly driven by the earnings of around 4.5%) as well as the real return (around 2.4%) and the risk-premium (around 2.3%). Note ii) that the one-year ahead predictions in March 2020 are not frightening even though we are at the beginning of a worldwide economic crisis and recession. They seem to reflect optimistic market expectations. Of course, we cannot incorporate external shocks in our model but the key variables show comforting figures as both, the term spread and the earnings-by-price, are at their second-highest values in the last 20 months. Thus, our model predicts that compared to the month before the crises started nominal one-year stock returns increase by 2.3%, real returns increase by 3.1%, and the risk-premium increases by 3.4%. Low inflation, low short-term interest rates (the latter being almost zero), and increasing prices will bring back the market to past performance such that this prediction was in line with what happened.

4. Conclusions and Outlook

In this paper, we extend the original working framework of Kyriakou et al. [6] to forecasting stock returns from a one-year to a five-year horizon in excess of different benchmarks, including the short-term rate, long-term rate, earnings-by-price ratio and inflation. We use predictors such as the dividend-by-price ratio, earnings-by-price ratio, short interest rate, long interest rate, term spread, inflation, 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.

We conclude that five-year returns can be forecasted via our economic variables. The improvement in overall variability compared to predicting a simple mean—measured via the R_V^2 —is good and

comparable to the overall forecasting improvement we have seen earlier in Kyriakou et al. [6] for the one-year view.

We find that for both one-year and five-year returns, the term spread is, overall, the most powerful predictive variable for excess stock returns. Combining this with the dividend-by-price in the five-year case boosts the predictive power to a maximum. In the one-year case, the inflation predictor is quite remarkable under the inflation benchmark either in itself, or combined with other covariates such as the term spread to achieve a best-performing pair for the given horizon. Notice that earnings seem to be the best overall predictor when working net-of-inflation. The double benchmarking approach has earnings as the best individual predictor net with the inflation benchmark, where it is almost as strong a predictor as when using earnings and spread combined. Based on the results of this paper and also of Kyriakou et al. [6], we therefore, conclude that modeling earnings-by-price is a good and relatively simple starting point when constructing forecasting models for real-value pension prognoses.

Finally, a good compromise when promoting only one set of predictors for both the one and the five-year view would be earnings-by-price and the term spread. It seems that the earnings-by-price tends to define the overall level of the return, while the term spread provides information on short-term market corrections. This is why the earnings-by-price benchmark with the term spread as a covariate was the superior combination among all considered opportunities when all forecasts were back-calculated to nominal returns (see, Kyriakou et al. [6]). The different role played by earnings-by-price and the term spread may be explaining why they work so well on aggregate. So, the overall conclusions can be (i) that one should work with these two predictors when forecasting long-term stock returns and (ii) that the good results of our approach should lead to reconsiderations, along some of the comments of Lioui and Poncet [1], when deciding on the horizon of the investment strategy for the long-term investor.

Future research might work on econometric modeling that can combine short-term and long-term predictions. One could, for example, imagine an econometric model having the same predictive mean and variance one year ahead provided by the optimal one-year forecast, while simultaneously having the same predictive mean and variance five years ahead as provided by the optimal five-year forecast. Current efforts are being undertaken in that direction by the research team behind this paper.

Author Contributions: Conceptualization, I.K., P.M., J.P.N. and M.S.; Formal analysis, I.K., P.M., J.P.N. and M.S.; Funding acquisition, J.P.N.; Investigation, P.M.; Methodology, M.S.; Software, M.S.; Supervision, J.P.N.; Writing – original draft, I.K., P.M., J.P.N. and M.S. All authors contributed equally to this work. All authors have read and agreed to the published version of the manuscript.

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

1. Lioui, A.; Poncet, P. Long horizon predictability: An asset allocation perspective. *Eur. J. Oper. Res.* **2019**, *278*, 961–975, doi:10.1016/j.ejor.2019.04.040.
2. Møller, S.; Rangvid, J. End-of-the-Year Economic Growth and Time-varying Expected Returns. *J. Financ. Econ.* **2015**, *115*, 136–154, doi:10.1016/j.jfineco.2014.08.006.
3. Lettau, M.; Van Nieuwerburgh, S. Reconciling the return predictability evidence. *Rev. Financ. Stud.* **2008**, *21*, 1601–1652.
4. Chen, Q.; Hong, Y. *Predictability of Equity Returns over Different Time Horizons: A Nonparametric Approach*; Working Paper; Department of Economics, Cornell University: Ithaca, NY, USA, 2009.
5. Cheng, T.; Gao, J.; Linton, O. *Nonparametric Predictive Regressions for Stock Return Predictions*; Cambridge Working Papers in Economics; 1932; Faculty of Economics, University of Cambridge: Cambridge, UK, 2019.
6. Kyriakou, I.; Mousavi, P.; Nielsen, J.P.; Scholz, M. Forecasting Benchmarks of Long-Term Stock Returns via Machine Learning. *Annals of Operations Research* (available as Online First at <https://doi.org/10.1007/s10479-019-03338-4>, 24 July 2019).

7. Nielsen, J.; Sperlich, S. Prediction of stock returns: A new way to look at it. *ASTIN Bull.* **2003**, *33*, 399–417.
8. Scholz, M.; Nielsen, J.; Sperlich, S. Nonparametric prediction of stock returns based on yearly data: The long-term view. *Insur. Math. Econ.* **2015**, *65*, 143–155.
9. Scholz, M.; Sperlich, S.; Nielsen, J. Nonparametric long term prediction of stock returns with generated bond yields. *Insur. Math. Econ.* **2016**, *69*, 82–96.
10. Rapach, D.; Zhou, G. Forecasting Stock Returns. In *Handbook of Economic Forecasting*, 2A ed.; Elliott, G., Timmerman, A., Eds.; Elsevier: Amsterdam, The Netherlands, 2013; pp. 328–383.
11. Munk, C.; Rangvid, J. New Assumptions of a pension forecast model: Background, level and consequences for individuals forecasted pension. *Finans/Invest* **2018**, *6*, 6–14.
12. Xiao, Z.; Linton, O.B.; Carroll, R.J.; Mammen, E. More efficient local polynomial estimation in nonparametric regression with autocorrelated errors. *J. Am. Stat. Assoc.* **2003**, *98*, 980–992.
13. Su, L.; Ullah, A. More efficient estimation in nonparametric regression with nonparametric autocorrelated errors. *Econom. Theory* **2006**, *22*, 98–126.
14. De Brabanter, K.; De Brabanter, J.; Suykens, J.; De Moor, B. Kernel regression in the presence of correlated errors. *J. Mach. Learn. Res.* **2011**, *12*, 1955–1976.
15. Bergmeir, C.; Hyndman, R.J.; Koo, B. A note on the validity of cross-validation for evaluating autoregressive time series predictions. *Comput. Stat. Data Anal.* **2018**, *120*, 70–83.
16. Linton, O.B.; Mammen, E. Nonparametric transformation to white noise. *J. Econom.* **2008**, *142*, 241–264.
17. Geller, J.; Neumann, M.H. Improved local polynomial estimation in time series regression. *J. Nonparametric Stat.* **2018**, *30*, 1–27.
18. Burman, P.; Chow, E.; Nolan, D. A cross-validatory method for dependent data. *Biometrika* **1994**, *81*, 351–358.
19. Opsomer, J.; Wang, Y.; Yang, Y. Nonparametric regression with correlated errors. *Stat. Sci.* **2001**, *16*, 134–153.
20. Chu, C.K.; Marron, J.S. Comparison of two bandwidth selectors with dependent errors. *Ann. Stat.* **1991**, *19*, 1906–1918.
21. Shiller, R. *Market Volatility*; MIT Press: Cambridge, MA, USA, 1989.
22. Welch, I.; Goyal, A. A comprehensive look at the empirical performance of equity premium prediction. *Rev. Financ. Stud.* **2008**, *21*, 1455–1508.
23. Mammen, E.; Nielsen, J.P.; Scholz, M.; Sperlich, S. Conditional Variance Forecasts for Long-Term Stock Returns. *Risks* **2019**, *7*, 1–22, doi:10.3390/risks7040113.
24. Cont, R. Empirical properties of asset returns: Stylized facts and statistical issues. *Quant. Financ.* **2001**, *1*, 223–236, doi:10.1080/713665670.
25. Pesaran, H.; Timmermann, A. Market Timing and Return Predictability under Model Instability. *J. Empir. Financ.* **2002**, *8*, 495–510.
26. Rapach, D.; Wohar, M. *Structural Change and the Predictability of Stock Returns*; Working Paper; University of Nebraska-Omaha: Omaha, NE, USA, 2004.
27. Resnick, B.G.; Shoosmith, G.L. Using the yield curve to time the stock market. *Financ. Anal. J.* **2002**, *58*, 82–90.
28. Merton, R. The crisis in retirement planning. *Harv. Bus. Rev.* **2014**, *92*, 43–50.
29. Gerrard, R.; Hiabu, M.; Kyriakou, I.; Nielsen, J.P. Communication and personal selection of pension saver's financial risk. *Eur. J. Oper. Res.* **2019**, *274*, 1102–1111.
30. Gerrard, R.; Hiabu, M.; Kyriakou, I.; Nielsen, J.P. Self-selection and risk sharing in a modern world of life-long annuities. *Br. Actuar. J.* **2019**, *23*, doi:10.1017/s135732171800020x.
31. González-Manteiga, W.; Crujeiras, R. An updated review of Goodness-of-Fit tests for regression models. *TEST Off. J. Span. Soc. Stat. Oper. Res.* **2013**, *22*, 361–411.

