Graz Economics Papers – GEP

GEP 2020–20

Short-Term Exuberance and long-term stability: A simultaneous optimization of stock return predictions for short and long horizons

Ioannis Kyriakou, Parastoo Mousavi, Jens Perch Nielsen and Michael Scholz

December 2020

Department of Economics
Department of Public Economics
University of Graz

Short-term exuberance and long-term stability: A simultaneous optimization of stock return predictions for short and long horizons

A Preprint

Ioannis Kyriakou
Parastoo Mousavi
Jens Perch Nielsen
Cass Business School
City, University of London
106 Bunhill Row
London EC1Y 8TZ, UK
ioannis.kyriakou@city.ac.uk
parastoo.mousavi.2@cass.city.ac.uk
jens.nielsen.1@city.ac.uk

Michael Scholz *
Department of Economics
University of Graz
Universitätsstraße 15/F4
8010 Graz, Austria
michael.scholz@uni-graz.at

December 2, 2020

Abstract

The fundamental interest of investors in econometric modelling for excess stock returns usually focuses either on short- or long-term predictions to reduce individually the investment risk. In this paper, we present a new simple model that accounts contemporaneously for short- and long-term predictions. By combining the different horizons, we can exploit the lower long-term variance to further reduce the short-term variance which is susceptible to speculative exuberance. Different combinations of short and long horizons as well as definitions of excess returns, for example, concerning the traditional short-term interest rate but also the inflation, are easily accommodated in our model. We show that the estimated relationship between excess stock returns and the predictive variables under the inflation benchmark is stable across different horizons, which is especially important for long-term real savings for pension products. We conclude the paper with a study of stock market predictions during the recent Covid-19 pandemic.

Keywords finance · investment analysis · stock returns · cross-validation · variation reduction
JEL C14 · C53 · C58 · G17 · G22

*Corresponding author.

1 Introduction

There is considerable practical and theoretical effort being channelled into understanding the movements of the stock market. This is natural as this is, perhaps, the most significant driver of returns providing long-term savers with sufficient wealth at retirement. Long-term predictability does impact strongly investors' welfare, as pointed out, for example, by Lioui and Poncet (2019), and recent years have witnessed the emergence of research on pension products. This marks the need to provide an econometric model when planning long-term savings (e.g., see Merton, 2014, Gerrard *et al.*, 2019, 2020). Such a model should be able to forecast the future to facilitate an individual long-term saver's planning for retirement and help them act dynamically on market information. A proper econometric approach to long-term savings needs to be able to account for both the long and the short term concurrently aiming to line up the long-term projections while circumventing inaccurate trading due to short-term bubbles in the market. It is worth noting that long-horizon predictability has also been studied in other contexts. For example, this has been favoured by Carmona *et al.* (2012), who analyze return predictability effects on the fair value of long-term executive stock options, but also Bodnar *et al.* (2015) who study the multi-period (long-run) portfolio choice problem under return predictability.

Our paper provides a general strategy in support of such a novel econometric model. We concretely consider the standard case of returns in excess of the short-term interest rate and the, perhaps, more relevant case of returns in excess of inflation (i.e., real returns) as led by Merton (2014). In our empirical application we consider a short-term period of one year and a long-term period of five years. Nevertheless, by its universality, our approach lends itself to any benchmark, so not just short-term interest rate or inflation, and can fit in with any assumption of a short and a long term. The baseline is a model for the earnings-by-price, which is an intuitive, attractive quantity that can be compared with interest rates and other returns and one of the most important drivers of return predictability (see also Kyriakou *et al.*, 2020). A final correction is made to ensure that the model is capable of capturing the returns trend.

An accurate model does not only provide a better understanding of the expected return, but also a reduced variation. Our contribution is twofold. First, the application of predictive regressions for two different horizons reduces individually the noise for short- and long-term investments. Second, by combining predictions of different horizons we reduce further the noise for the short-term investment. This confirms the view, as put forward by Lioui and Poncet (2019), of long-horizon predictability defenders that the use of long-term returns reduces the noise in asset returns. The reason is that even in, for example, one-year returns, a large amount of speculative variation is still included. This is clearly reduced in longer-horizon investments. However, our simple model is able to optimize the one-year investments according to the bubble-free long-term variance and achieve variation reduction for the short-term predictions after year one. We put the spotlight on two findings which are particularly interesting and intuitively appealing. With the aid of our optimal predictive model, we are able to perceptibly reduce the standard deviation of the one-year returns by 10% from about 18% (see also Mammen *et al.*, 2019) to 16%. The advantage of prediction with incorporated long-term modelling does not stop there: it turns out that the standard deviation beyond the short term is only around 14%. Therefore, a long-term investor that optimizes pensions or other long-term savings should rely on this, based on the available information, rather than the typically used standard deviation of 16% or even 18%. In general, a larger, out-of-line standard deviation would lead to an over-conservative portfolio with implied potential upside reductions for the long-term saver given their optimal risk appetite.

The last part of the paper is dedicated to a study of the stock market dynamics during the Covid-19 pandemic. In particular, we show that, whereas earnings have dropped dramatically from November 2019 to June 2020, the future outlook for the long-term real returns seems to be unaffected. This is attributed to the quantitative easing resulting in a short-term interest rate slump in the same period. Thus, the good performance of the stock market seems to be rather advanced by political decisions for as long as possible in a protracted Covid-19 crisis. A practical implication is, therefore, that the stock market dynamics might be more fragile and volatile going forward.

The remainder of the paper is structured as follows. In Section 2, we present a gradual built-up of our proposed framework from short-term and long-term nonlinear predictive modelling to their merge to a single model that aims to achieve reduction of the investment risk. Section 3 focuses on our empirical application to one- and five-year excess stock returns based on historical US market data. Section 4 touches on stock market predictions during the Covid-19 pandemic based on our optimal model. Section 5 concludes the paper.

2 A framework for combined nonlinear predictive regressions

Linear regression models are popular in predictive modelling as this classical benchmark is easy to estimate and to interpret. However, the fixed functional form of the relationship between stock returns and predictive variables leads to inferior predictive power compared with nonlinear approaches (Lettau and Van Nieuwerburgh, 2008; Chen and Hong, 2010; Scholz *et al.*, 2015, 2016; Cheng *et al.*, 2019). Therefore, we focus on potentially nonlinear predictive relationships between returns over the next T years in excess of a reference rate (or benchmark) and a set of economic predictors relevant for the long-term investor using a fully nonparametric smoother. We analyze the two most important benchmark models of Kyriakou et al. (2019, 2020): the short-term interest rate and the inflation rate. Note that the former directly corresponds to the prediction of the *risk premium* (over a risk-free investment), whereas the latter refers to the forecast of *real returns*. We aim, first, to investigate their predictability over horizons of one year and five years separately and, then, provide an intuitive *single* econometric model which combines both predictive horizons.

2.1 One-year predictions

We start with annual nominal stock returns defined by $S_t := (P_t + D_t)/P_{t-1}$, where P_t is the stock price at the end of year t and D_t the dividends paid during year t. We focus on returns in excess (log-scale) of a given benchmark $B_{t-1}^{(A)}$ with $A \in \{R, C\}$:

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

where $B_t^{(R)} := 1 + R_t/100$ and $B_t^{(C)} := 1 + \pi_t$ with R_t denoting the short-term interest rate and $\pi_t := (CPI_t - CPI_{t-1})/CPI_{t-1}$ the inflation rate for CPI_t the consumer price index for year t.

Our predictive nonparametric regression model for the one-year ($1y$) excess returns defined in equation (1) is now given by

$$Y_t^{(A)} = m_{1y}\left(X_{t-1}^{(A)}\right) + \xi_t. \tag{2}$$

Note that the conditional mean in equation (2),

$$m_{1y}(x^{(A)}) := \mathbb{E}\left(Y^{(A)} | X^{(A)} = x^{(A)}\right),\ x^{(A)} \in \mathbb{R}^q, \tag{3}$$

is unknown and its functional form is not predetermined, for example, to be linear, but can take any shape. Our preferred nonparametric method to estimate this function m_{1y} is the local-linear smoother because of its flexibility and well-known statistical properties. For example, the linear function can be estimated without any bias and is, thus, automatically embedded in our analysis, that is, in case that the data-generating process is linear, we will expose this simple functional form. Note further that the error terms ξ_t in equation (2) form a martingale difference process, i.e., ξ_t are serially uncorrelated random variables with zero mean, given the past, and unknown conditionally heteroscedastic variance of the form $\sigma_{1y}^2\left(x^{(A)}\right)$. The elements of the q-dimensional vector $X_{t-1}^{(A)}$ in equation (2), which collects the explanatory variables, are also transformed under the chosen benchmark A according to

$$X_{t-1}^{(A)} := \begin{cases} \frac{1+X_{t-1}}{B_{t-1}^{(A)}}, & X_{t-1} \in \{d_{t-1}, e_{t-1}, r_{t-1}, l_{t-1}, \pi_{t-1}\} \\ \frac{s_{t-1}}{B_{t-1}^{(A)}} = \frac{l_{t-1} - r_{t-1}}{B_{t-1}^{(A)}} & \end{cases}. \tag{4}$$

Therefore, $X_{t-1}^{(A)}$ contains (combinations of transformed) popular time-lagged predictive variables based on the: (i) dividend-by-price ratio $d_{t-1} = D_{t-1}/P_{t-1}$; (ii) earnings-by-price ratio $e_{t-1} = E_{t-1}/P_{t-1}$ where E_t denotes the earnings accruing to the index in year t; (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 rate π_{t-1}; and (vi) term spread $s_{t-1} = l_{t-1} - r_{t-1}$. The use of such a transformation is one example of careful imposition of additional structure in the statistical modelling process, which has shown promising results in previous works of Nielsen and Sperlich (2003), Scholz *et al.* (2015) and Scholz *et al.* (2016). This adjustment of both the independent and dependent variables according to the same benchmark we call double (or full) benchmarking.

2.2 Longer-horizon predictions

A main contribution of our article is the combination of short-term and long-term predictions in one single model. Hence, we introduce, in addition to the short one-year predictions, our version of long-horizon predictions. We want to highlight three important points which distinguish both cases fundamentally from each other: first, the autoregressive behaviour of the underlying predictive variable in equation (6), which will be used as the building block of our econometric model in Section 2.4 as well; second, the more complicated error structure (serial correlation by construction) in the predictive relationship (8); and, third, closely related to the latter point, a more complicated smoothing parameter selection for the correct estimation of m_{Ty} in equation (9).

For longer horizons T, with $T > 1$, we consider the sum of annual continuously compounded returns defined in equation (1), that is,

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

Here, a careful econometric modelling is necessary because of the overlapping nature of the returns $Z_t^{(A)}$ (refer also to the appendix). For ease of illustration, assume a linear model for $Y_t^{(A)}$ in equation (2) as well as some linear and autoregressive behaviour of order one for the forecasting variable $X_{t-1}^{(A)}$:

$$Y_t^{(A)} = \beta_0 + \beta_1 X_{t-1}^{(A)} + \xi_t \quad \text{and} \quad X_t^{(A)} = \gamma_0 + \gamma_1 X_{t-1}^{(A)} + \eta_t, \tag{6}$$

with ξ_t as in equation (2), η_t a white noise, and regression parameters $\beta_0, \beta_1, \gamma_0$ and γ_1. A simple linear model for the T-year (Ty) regression problem which directly follows from equations (5) and (6) is then

$$Z_t^{(A)} = \phi_0 + \phi_1 X_{t-1}^{(A)} + \nu_t, \tag{7}$$

with parameters ϕ_0 and ϕ_1, and error terms ν_t (more details are deferred to the appendix). Equation (7) shows that the excess stock return for year t over the next T years can be decomposed in two parts: a predictive linear part dependent only on the variable $X_{t-1}^{(A)}$, the same predictive variable as in the one-year case, and unpredictable error terms ν_t which are now serially correlated by construction.

As the linear setup of equation (6) could be misspecified and, thus, not account for important nonlinearities, we model the functional relationship between the predictive variable $X_{t-1}^{(A)}$ and T-year excess stock returns $Z_t^{(A)}$ in a more flexible nonparametric way in analogy to equation (2)

$$Z_t^{(A)} = m_{Ty}\left(X_{t-1}^{(A)}\right) + \nu_t, \tag{8}$$

where

$$m_{Ty}(x^{(A)}) := \mathbb{E}\left(Z_t^{(A)} | X^{(A)} = x^{(A)}\right),\ x^{(A)} \in \mathbb{R}^q, \tag{9}$$

is an unknown smooth function. Note again the important difference between the error terms of model (2) and model (8): while ξ_t is a martingale difference process, ν_t is serially correlated by construction. This property has to be taken into account when estimating the unknown conditional mean function m_{Ty}, otherwise quite fundamental problems occur: the estimators are still consistent but less efficient than such ones correcting for autocorrelation (Xiao *et al.*, 2003; Su and Ullah, 2006; Linton and Mammen, 2008; Geller and Neumann, 2018); and, more importantly, the commonly applied automatic smoothing parameter selection procedures (such as cross-validation and plug-in) break down (De Brabanter *et al.*, 2011; Bergmeir *et al.*, 2018). In the empirical part of our paper, we overcome the aforementioned problems by using a special leave-l-out cross-validation strategy which is closely related to our way in measuring predictive power. Our approach to this issue will be discussed in detail in the next section.

Before we proceed, let's summarize what we have discussed so far: the nonparametric models (2) and (8) for one-year and T-year returns; the autoregressive behaviour of order one for the predictive variable in (6); and, the necessity of a leave-l-out cross-validation in the estimation procedure.

2.3 Predictive power, variable selection, and smoothing parameter choice

For our nonparametric one- and T-year models defined earlier, we need an adequate measure that (a) quantifies and validates the predictive power; (b) allows for comparisons and ranking of models when different sets of explanatory variables are used (variable selection); and (c) best selects the bandwidth(s) and, thus, determines the functional form of the conditional mean for the given predictive variables (smoothing parameter choice). In our work, we apply the *validated R-squared* (R_V^2) of Nielsen and Sperlich (2003) which conforms to these requirements. It directly aims to estimate the k-year (ky) ahead prediction error based on a leave-l-out cross-validation (with $l := 2k - 1$) and can thus be used for both variable as well as smoothing parameter selection. In our notation, the validated R-squared is defined as

$$R_{V,ky}^2 = 1 - \frac{\sum_t (W_t - \hat{m}_{-t,ky})^2}{\sum_t (W_t - \bar{W}_{-t})^2}, \tag{10}$$

where such estimators are used that leave out l observations around the tth point in time: $\hat{m}_{-t,ky}$ for the conditional mean function m_{ky} from equations (2) or (8) with $k \in \{1, T\}$ and $\bar{W}_{-t}$ for the unconditional (historical) mean of W_t, that is, the k-year return to predict (equal to $Y_t^{(A)}$ for $k = 1$ and $Z_t^{(A)}$ for $k = T$). In order to keep the notation simple, we drop an extra subscript for the bandwidth h used in the calculation of $\hat{m}_{-t,ky}$, as we always choose h in the numerator in equation (10) so that the prediction error is minimized and, thus, the largest possible R_V^2 is achieved for the given predictive variables. Note that R_V^2 measures the predictive power of a given model against a benchmark (here, the cross-validated historical mean). This means for our setup that when R_V^2 is positive, the predictor-based regression model (2) or (8) outperforms the corresponding historical mean forecast.

In a time-series context, often out-of-sample evaluations are proposed where a fraction of the data from the end of the time-series is not used for estimation but is withheld for evaluation. In the case of uncorrelated errors, Bergmeir *et al.* (2018) show that cross-validation, as proposed in this section, is preferred to out-of-sample evaluation. Another advantage is that cross-validation involves various evaluations, whereas out-of-sample analysis can test the data only once. This property is especially beneficial when the number of recorded observations is small as in our case with annual stock market data. When errors are correlated, as discussed in Section 2.2 for our T-year predictions, it can be necessary to omit more than a single point and apply leave-l-out cross-validation (with $l > 1$). This strategy avoids model fits that are progressively under-smoothed caused by too small bandwidths (Opsomer *et al.*, 2001). Alternative approaches, for example, use bimodal kernels (Chu and Marron, 1991) or the correlation-corrected cross-validation (De Brabanter *et al.*, 2011). Note that in the case of a large fraction of skipped data, additional corrections might be required (Burman *et al.*, 1994).

2.4 An econometric model for combined short- and long-term predictions

In this section, we present a simple way of combining short- and long-term predictions. Our model builds on the autoregressive development of the earnings variable $e^{(A)}$ or, more precisely, on the change in earnings growth which has been identified as one of the key drivers of stock prices P. Other important factors, such as the dividend yield $d^{(A)}$, can be easily incorporated in our model as well, for example, as covariates in the one- or five-year conditional mean regressions (2) or (8) which will be used to calibrate our model. The important contribution of our approach is twofold. First, the application of predictive regressions for two different horizons reduces individually the noise or risk for short- and long-term investments. Second, the combination of predictions of different horizons further reduces the noise or risk for the short-term investment. The reason is that even in, for example, one-year returns, a large amount of

speculative variation is still included. This is clearly reduced in longer-horizon investments. Using now such T-year predictions in combination with the one-year ones, the latter benefit from the former as they are forced to sum up to the long-term
which are equal to the conditional mean forecasts based on regressions (2) and (8) (and thus with an interpolation argument also for the horizons in-between), and reduces the variation for the short-term predictions after year one.

We start with the linear formulations of the autoregressive behaviour of order one of the predictive variable and the linear model version of one-year return predictions in equation (6). Here, we consider the earnings variable $e^{(A)}$ to be this special predictor and estimate the linear models by ordinary least squares (OLS). We get in a first step:

$$e_t^{(A)} - e_{t-1}^{(A)} = \rho\,(e_{t-1}^{(A)} - \bar{e}^{(A)}) + \eta_t \quad \Leftrightarrow \quad e_t^{(A)} = \gamma_0 + \gamma_1 e_{t-1}^{(A)} + \eta_t \tag{11}$$

with unknown parameters $\gamma_0 := -\rho\bar{e}^{(A)}$ and $\gamma_1 := \rho + 1$, sample average of earnings $\bar{e}^{(A)}$, and i.i.d. error terms η_t. The OLS estimates of γ_0 and γ_1 shall be denoted by c_0 and c_1, respectively. In a second step, we apply the linear version of equation (2) for the earnings variable $e^{(A)}$:

$$Y_{t+1}^{(A)} = \beta_0 + \beta_1 e_t^{(A)} + \xi_{t+1} \tag{12}$$

with unknown parameters β_0 and β_1 which will be estimated again by OLS; their estimates are denoted by b_0 and b_1, respectively. Remember that we have n observations in our records. Thus, with equations (11) and (12) and the corresponding OLS estimates, which we will keep fixed in the following steps, we can now forecast out-of-sample $\hat{Y}_{n+1}^{(A)}, \hat{Y}_{n+2}^{(A)}, \ldots, \hat{Y}_{n+T}^{(A)}$.

Our aim is to construct an econometric model that reflects one-year and T-year predictions (from the preferred models (2) and (8) at hand) at the same time. For this reason, we will correct $\hat{Y}_{n+1}^{(A)}, \hat{Y}_{n+2}^{(A)}, \ldots, \hat{Y}_{n+T}^{(A)}$ in the following linear way:

$$\hat{Y}_{n+1}^{(A),c} = \alpha_0 + \alpha_1 \hat{Y}_{n+1}^{(A)} + \varepsilon_{n+1} \tag{13}$$

$$\hat{Y}_{n+2}^{(A),c} = \alpha_0 + \alpha_1 \hat{Y}_{n+2}^{(A)} + \varepsilon_{n+2} \tag{14}$$

$$\vdots$$

$$\hat{Y}_{n+T}^{(A),c} = \alpha_0 + \alpha_1 \hat{Y}_{n+T}^{(A)} + \varepsilon_{n+T}, \tag{15}$$

where α_0 and α_1 are unknown parameters, $\varepsilon_{n+1} \sim \mathcal{N}(0, \sigma_1^2)$ and $\varepsilon_{n+2}, \ldots \varepsilon_{n+T} \sim \mathcal{N}(0, \sigma_2^2)$ are independent error terms with unknown variances σ_1^2 and σ_2^2. Note that we allow for a different variation in the first corrected one-year ahead prediction $\hat{Y}_{n+1}^{(A),c}$ in equation (13) compared with the second to Tth corrected one-year ahead predictions $\hat{Y}_{n+2}^{(A),c}, \ldots, \hat{Y}_{n+T}^{(A),c}$ in equations (14) to (15). This way, our model can account for the lower variation in longer-horizon returns relative to one-year returns. In other words, after calibrating the model, we expect σ_2^2 to be smaller than σ_1^2. Note further that, from equations (13)–(15), we directly get an expression for the corrected T-year return $Z_{n+T}^{(A),c}$:

$$Z_{n+T}^{(A),c} = \sum_{k=1}^{T} \hat{Y}_{n+k}^{(A),c} = \alpha_0 T + \alpha_1 \sum_{k=1}^{T} \hat{Y}_{n+k}^{(A)} + \sum_{k=1}^{T} \varepsilon_{n+k}. \tag{16}$$

Next, we adequately calibrate equations (13)–(16), i.e., choose the model parameters α_0, α_1, σ_1^2 and σ_2^2, and based on these obtain the corrected one-year and T-year returns. Here, we use the recursive representation of the earnings $e^{(A)}$ from equation (11) with the starting value $e_n^{(A)}$ (the last earnings observation in our records) together with the linear predictive model (12) and the corresponding OLS estimates c_0, c_1, b_0, b_1. Plugging-in gives for the corrected

$\hat{Y}_{n+1}^{(A)}, \ldots, \hat{Y}_{n+T}^{(A)}$ and $Z_{n+T}^{(A),c}$:

$$\hat{Y}_{n+1}^{(A),c} = \alpha_0 + \alpha_1(b_0 + b_1 e_n^{(A)}) + \varepsilon_{n+1} \tag{17}$$

$$\hat{Y}_{n+2}^{(A),c} = \alpha_0 + \alpha_1(b_0 + b_1(c_0 + c_1 e_n^{(A)})) + \varepsilon_{n+2} \tag{18}$$

$$\vdots$$

$$\hat{Y}_{n+T}^{(A),c} = \alpha_0 + \alpha_1 \left(b_0 + c_0 b_1 \sum_{i=0}^{T-2} c_1^i + c_1^{T-1} b_1 e_n^{(A)} \right) + \varepsilon_{n+T} \tag{19}$$

and

$$Z_{n+T}^{(A),c} = \alpha_0 T + \alpha_1 b_0 T + \alpha_1 c_0 b_1 \sum_{k=2}^{T} \sum_{i=0}^{k-2} c_1^i + \alpha_1 b_1 e_n^{(A)} \sum_{i=0}^{T-1} c_1^i + \sum_{k=1}^{T} \varepsilon_{n+k}. \tag{20}$$

Now, we can fix the first and second moments of $\hat{Y}_{n+1}^{(A),c}$ and $Z_{n+T}^{(A),c}$ with the estimated values from our preferred (best) one- and T-year predictive models (2) and (8). By doing so, we obtain a linear equation system with four equations, which can be easily solved for the four unknown parameters α_0, α_1, σ_1^2 and σ_2^2. For this purpose, let $\hat{\mu}_{1y}$ and $\hat{\sigma}_{1y}^2$ be the conditional mean forecast and its estimated variance from equation (2), and $\hat{\mu}_{Ty}$ and $\hat{\sigma}_{Ty}^2$ the conditional mean forecast and its estimated variance from equation (8). Note that $\hat{\sigma}_{1y}^2$ and $\hat{\sigma}_{Ty}^2$ can be readily calculated from the R_V^2 of the predictive regressions (2) and (8). A closer inspection of equation (10) shows that the ratio in our validation criterion compares the sample variance of the estimated residuals from the preferred predictive model (the numerator) with the sample variance of the benchmarked returns (the denominator). Algebraically, we therefore have that $R_{V,1y}^2 = 1 - \hat{\sigma}_{1y}^2/\sigma_{Y^{(A)}}^2$ and $R_{V,Ty}^2 = 1 - \hat{\sigma}_{Ty}^2/\sigma_{Z^{(A)}}^2$, and

$$\hat{\sigma}_{1y}^2 = (1 - R_{V,1y}^2)\sigma_{Y^{(A)}}^2, \tag{21}$$

$$\hat{\sigma}_{Ty}^2 = (1 - R_{V,Ty}^2)\sigma_{Z^{(A)}}^2. \tag{22}$$

Given

$$\begin{aligned} \mathbb{E}(\hat{Y}_{n+1}^{(A),c}) = \hat{\mu}_{1y}, &\quad \mathbb{E}(\hat{Z}_{n+T}^{(A),c}) = \hat{\mu}_{Ty}, \\ \mathrm{Var}(\hat{Y}_{n+1}^{(A),c}) = \hat{\sigma}_{1y}^2, &\quad \mathrm{Var}(\hat{Z}_{n+T}^{(A),c}) = \hat{\sigma}_{Ty}^2, \end{aligned} \tag{23}$$

the solution of the equation system (23) is

$$\alpha_0 = \hat{\mu}_{1y} - \alpha_1 \left(b_0 + b_1 e_n^{(A)} \right), \tag{24}$$

$$\alpha_1 = \frac{\hat{\mu}_{Ty} - T\hat{\mu}_{1y}}{S - b_0 T - b_1 T e_n^{(A)}}, \tag{25}$$

where

$$S := b_0 T + c_0 b_1 \sum_{k=2}^{T} \sum_{i=0}^{k-2} c_1^i + b_1 e_n^{(A)} \sum_{i=0}^{T-1} c_1^i,$$

and

$$\sigma_1^2 = \hat{\sigma}_{1y}^2, \tag{26}$$

$$\sigma_2^2 = \frac{1}{T-1} \left(\hat{\sigma}_{Ty}^2 - \hat{\sigma}_{1y}^2 \right). \tag{27}$$

The *a priori* expectations about our model are the following. First, when the autoregressive behaviour of the earnings in model (11) and the linear model for stock returns (12) give reasonable predictions $\hat{Y}_{n+1}^{(A)}, \hat{Y}_{n+2}^{(A)}, \ldots, \hat{Y}_{n+T}^{(A)}$, only a marginal correction is necessary, i.e., α_0 is close to zero and α_1 close to one. Second, when $T\hat{\mu}_{1y} > \hat{\mu}_{Ty}$, one-year returns should diminish over time (as the sum of $\hat{Y}_{n+1}^{(A)}, \ldots, \hat{Y}_{n+T}^{(A)}$ still has to be equal to $\hat{\mu}_{Ty}$) and α_1 gets negative. Now α_0 takes the role of an upper limit (larger than $\hat{\mu}_{1y}$), from which increasing values (over time) are subtracted to

Table 1: US market data (1872–2020).

	Max	Min	Mean	Sd	Skew	Exc. kurt
S&P stock price index	3278.20	3.25	297.62	607.94	2.57	6.62
Dividend accruing to index	58.24	0.18	6.39	11.36	2.52	6.35
Earnings accruing to index	139.47	0.16	14.80	28.17	2.47	5.66
Short-term interest rate	14.93	0.07	3.99	2.49	0.94	2.27
Long-term interest rate	14.59	1.76	4.51	2.25	1.80	3.68
Consumer price index	257.97	6.47	60.32	74.02	1.34	0.35

match the T-year prediction $\hat{\mu}_{Ty}$. Finally, note that, by construction, $\sigma_1^2 > \sigma_2^2$ if and only if $T\hat{\sigma}_{1y}^2 > \hat{\sigma}_{Ty}^2$, that is, the cumulated risk over T periods of short-term investments exceeds the risk of a T-year investment (as discussed earlier).

3 Empirical illustration: combined prediction of excess stock returns over one-year and five-year horizons

3.1 Data sources and descriptive statistics

Our empirical application is based on historical US stock market data on the annual frequency. The dataset includes, among other variables, the Standard and Poor's (S&P) Composite Stock Price Index, dividends and earnings accruing to the index, as well as macroeconomic measures like the short-term interest rate, the long-term interest rate, and the consumer price index covering the period from 1872 to 2020. Table 1 exhibits their basic descriptive statistics.

We use, here, an updated and revised version of Shiller's (1989, Chapter 26) data which are available from `http://www.econ.yale.edu/~shiller/data.htm`. Note that a simple extension of the risk-free rate series was not possible because the underlying 6-month certificate of deposit rate (secondary market) was discontinued in 2013. We follow thus the strategy of Welch and Goyal (2008) and replace this variable by an annual risk-free rate based on the 6-month Treasury-bill rate (secondary market) from `https://fred.stlouisfed.org/series/TB6MS`. As this series is available only from 1958, we had to estimate the information prior to 1958 using results from an OLS regression of the Treasury-bill rate on the risk-free rate from Shiller's data for the overlapping period 1958 to 2013. With the estimated linear model (R^2 of 98.6%) of

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

we finally instrumented the risk-free rate from 1872 to 1957. The high correlation of 99.3% between the actual Treasury-bill rate and the predictions for the estimation period ensures the usefulness of such an approach.

This section is concluded with Table 2 which displays standard descriptive statistics for the transformed variables according to equations (1), (4), and (5). It appears that the predictive variables under the inflation benchmark are spread out more with a wider range and a higher standard deviation than the variables under the risk-free rate benchmark. This property of the inflation benchmark could be beneficial for the estimation process because a larger variability in the regressors usually leads to a more efficient predictor.

However, the returns transformed with the two benchmarks differ only slightly. A small upward shift under the inflation benchmark is noticeable in Figure 1 which shows density plots of the benchmarked returns for both the one-year and five-year horizons.

In the next section, we summarize our analysis of the predictability of stock returns in excess of the different benchmarks separately for the one-year and the five-year case.

Table 2: Summary statistics of transformed variables (in percentage). Panel (a) shows the available variables transformed according to the short-term interest rate, e.g., excess returns corresponding to the risk premium. Panel (b) shows the available variables net of inflation, i.e., in real terms. $l^{(R)}$ equals $s^{(R)}$ by construction as explained in the footnote in Section 3.2.

	Max	Min	Mean	Sd	Skew	Exc. kurt
(a) Benchmark: short-term interest rate ($A \equiv R$)						
One-year excess stock returns $Y^{(R)}$	42.39	-58.26	4.71	17.28	-0.58	0.68
Five-year excess stock returns $Z^{(R)}$	107.27	-78.54	23.69	36.65	-0.16	-0.37
Dividend-by-price $d^{(R)}$	7.26	-8.96	0.37	2.78	-0.15	0.15
Earnings-by-price $e^{(R)}$	13.25	-3.29	3.22	3.07	0.96	1.18
Long-term interest rate $l^{(R)}$	3.55	-3.46	0.55	1.27	0.01	-0.04
Inflation $\pi^{(R)}$	17.00	-19.16	-1.64	5.62	0.24	1.79
Spread $s^{(R)}$	3.55	-3.46	0.55	1.27	0.01	-0.04
(b) Benchmark: inflation rate ($A \equiv C$)						
One-year excess stock returns $Y^{(C)}$	54.04	-48.81	6.52	18.04	-0.41	0.64
Five-year excess stock returns $Z^{(C)}$	122.96	-57.34	32.47	36.33	-0.06	-0.39
Dividend-by-price $d^{(C)}$	25.49	-13.90	2.38	6.51	0.93	1.77
Earnings-by-price $e^{(C)}$	29.50	-10.98	5.23	5.93	0.94	2.13
Short-term interest rate $r^{(C)}$	23.70	-14.53	2.00	5.85	0.40	1.84
Long-term interest rate $l^{(C)}$	23.70	-13.81	2.55	5.78	0.23	2.17
Spread $s^{(C)}$	3.51	-3.45	0.55	1.28	-0.02	-0.05

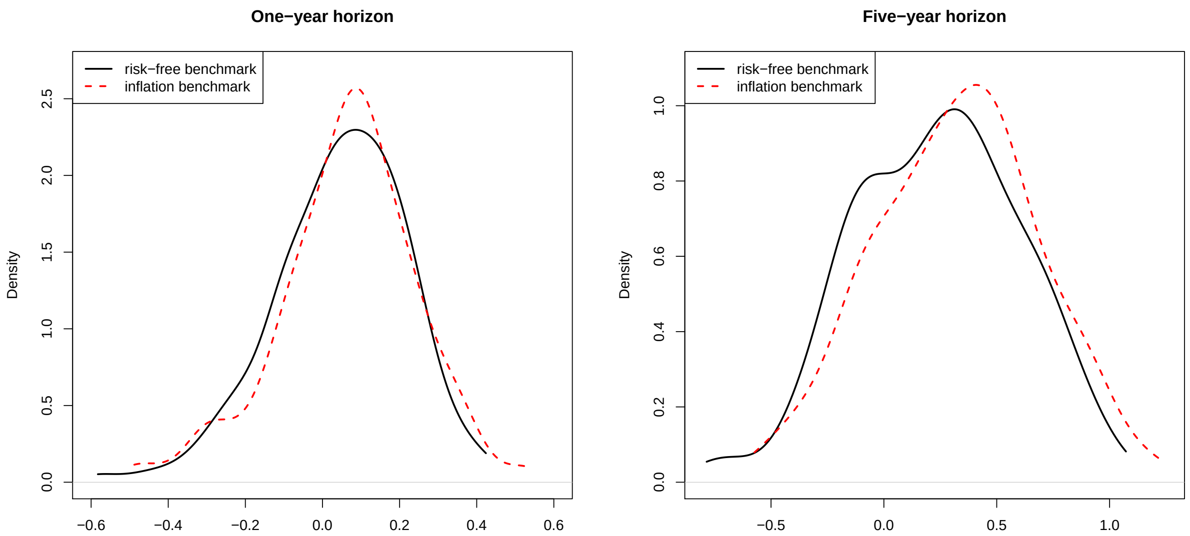

Figure 1: Kernel density estimates of the probability density function of returns transformed with the risk-free rate benchmark (solid) and the inflation benchmark (dotted). Left: one-year horizon. Right: five-year horizon. Period: 1872–2020. Data: annual S&P 500.

Table 3: Predictive power (in %) for the one-year excess stock returns $Y_t^{(A)}$ corresponding to the prediction problem defined in (2). The predictive power is measured by R_V^2 as defined in (10). The benchmarks $B^{(A)}$ considered are based on the short-term interest rate ($A \equiv R$) and the inflation rate ($A \equiv C$). The predictive variables used are $X_{t-1}^{(A)}$ using the indicated benchmark $B_{t-1}^{(A)}$ as shown in (4). "–" are not applicable cases of matched covariate with benchmark. $l^{(R)}$ equals $s^{(R)}$ by construction as explained in the footnote in Section 3.2.

Benchmark $B^{(A)}$	Explanatory variable(s) $X_{t-1}^{(A)}$					
	$d^{(A)}$	$e^{(A)}$	$r^{(A)}$	$l^{(A)}$	$\pi^{(A)}$	$s^{(A)}$
Short-term rate	3.0	5.0	–	9.6	-1.3	9.6
Inflation	10.2	12.0	7.1	10.4	–	6.6
	$(d^{(A)}, e^{(A)})$	$(d^{(A)}, r^{(A)})$	$(d^{(A)}, l^{(A)})$	$(d^{(A)}, \pi^{(A)})$	$(d^{(A)}, s^{(A)})$	
Short-term rate	2.3	–	9.8	1.4	9.8	
Inflation	10.1	9.3	9.8	–	15.4	
	$(e^{(A)}, r^{(A)})$	$(e^{(A)}, l^{(A)})$	$(e^{(A)}, \pi^{(A)})$	$(e^{(A)}, s^{(A)})$		
Short-term rate	–	10.7	3.3	10.7		
Inflation	11.2	11.1	–	17.5		
	$(r^{(A)}, l^{(A)})$	$(r^{(A)}, \pi^{(A)})$	$(r^{(A)}, s^{(A)})$			
Short-term rate	–	–	–			
Inflation	13.6	–	14.7			
	$(l^{(A)}, \pi^{(A)})$	$(l^{(A)}, s^{(A)})$				
Short-term rate	7.2	–				
Inflation	–	14.6				
	$(\pi^{(A)}, s^{(A)})$					
Short-term rate	7.2					
Inflation	–					

3.2 One- and five-year excess stock return predictability

In what follows, we apply the double benchmarking approach introduced in Section 2 to the annual US stock market data. The models (2) and (8) are estimated with a local-linear kernel smoother using the quartic kernel. The optimal bandwidths are chosen by cross-validation, that is, by maximizing the R_V^2 introduced in equation (10). Note that the linear model is automatically embedded in our approach because of the ability of the local-linear smoother to estimate this simple functional form without any bias. Remember also that the R_V^2 value compares the predictive power of a specific model (as a combination of predictive variables) with the predictive power of the historical mean. Thus, the largest positive R_V^2 under each benchmark indicates our favoured model with the highest predictive power. We study the empirical findings of R_V^2 values based on different validated scenarios shown for each of the one-year and five-year horizon in Table 3 and 4, respectively.

We find for almost all the variable combinations in the one- and five-year case as well as under both benchmarks a positive R_V^2, that is, a better predictive power compared with the historical mean. Only the inflation rate as a single covariate under the short-term benchmark has for both horizons a negative R_V^2 and thus no predictive power.

When comparing one- with five-year predictions for the risk-free rate benchmark, we confirm the findings of Rapach and Zhou (2013) that longer horizon predictions give better estimates than shorter horizons. All considered combinations of predictive variables have higher R_V^2 values for the five-year case. However, for the inflation benchmark, we observe the contrary, that is, almost all models for the one-year horizon have a higher predictive power. The only exception is the earnings-by-price variable with a slightly increased R_V^2 value in the five-year case.

Under the short interest benchmark $B^{(R)}$, the term spread $s^{(R)}$ is the most powerful predictive variable for excess stock returns. More in details, with the prediction constrained to using only single covariates, the term spread is the best

Table 4: Predictive power (in %) for the five-year excess stock returns $Z_t^{(A)}$ corresponding to the prediction problem defined in (8). For additional notes, refer to 3.

Benchmark $B^{(A)}$	Explanatory variable(s) $X_{t-1}^{(A)}$					
	$d^{(A)}$	$e^{(A)}$	$r^{(A)}$	$l^{(A)}$	$\pi^{(A)}$	$s^{(A)}$
Short-term rate	11.7	10.6	–	15.9	-2.4	15.9
Inflation	10.8	12.4	5.5	8.6	–	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	8.6	–	22.2	7.4	22.2	
Inflation	9.5	4.2	1.7	–	13.1	
	$(e^{(A)}, r^{(A)})$	$(e^{(A)}, l^{(A)})$	$(e^{(A)}, \pi^{(A)})$	$(e^{(A)}, s^{(A)})$		
Short-term rate	–	21.8	8.4	21.8		
Inflation	8.6	4.9	–	14.9		
	$(r^{(A)}, l^{(A)})$	$(r^{(A)}, \pi^{(A)})$	$(r^{(A)}, s^{(A)})$			
Short-term rate	–	–	–			
Inflation	10.8	–	10.1			
	$(l^{(A)}, \pi^{(A)})$	$(l^{(A)}, s^{(A)})$				
Short-term rate	16.3	–				
Inflation	–	10.2				
	$(\pi^{(A)}, s^{(A)})$					
Short-term rate	16.3					
Inflation	–					

Table 5: Summarized optimal combinations of predictive variables and their predictive power R_V^2 (in %).

Horizon	Risk-free rate benchmark		Inflation benchmark	
	Variables	R_V^2	Variables	R_V^2
One-year	$(e^{(R)}, s^{(R)})$	10.7	$(e^{(C)}, s^{(C)})$	17.5
Five-year	$(d^{(R)}, s^{(R)})$	22.2	$(e^{(C)}, s^{(C)})$	14.9

predictor for the one-year and five-year horizon with, respectively, $R_V^2 = 9.6\%$ and 15.9%.[2] Considering now also the models with combined predictive variables, we find in the one-year case that $(e^{(R)}, s^{(R)})$ yields $R_V^2 = 10.7\%$, whereas in the five-year case $(d^{(R)}, s^{(R)})$ and $(e^{(R)}, s^{(R)})$ perform closely with, respectively, $R_V^2 = 22.2\%$ and 21.8%; for both cases there is, thus, an increased predictive power compared to the best model with the single term spread covariate.

Under the inflation benchmark $B^{(C)}$, the earnings variable $e^{(C)}$ is the most powerful single predictor for the one-year and five-year horizon with, respectively, $R_V^2 = 12.0\%$ and 12.4%. In the one-year case the pair $(e^{(C)}, s^{(C)})$ boosts further the predictive power to $R_V^2 = 17.5\%$, while for the five-year horizon we find the same variable combination to be the most predictive model with $R_V^2 = 14.9\%$.

In our model which combines both one-year and five-year predictions, we will use for each benchmark and horizon the optimal combination of predictive variables. For convenience, Table 5 summarizes again the best models. For consistency but also to examine the robustness of our results, we will additionally consider the second best set of predictors under the risk-free rate benchmark for the five-year horizon, that is, $(e^{(R)}, s^{(R)})$.

To get deeper insights into the relationship between excess stock returns and the predictive variables for the different benchmarks and horizons discussed above, Figures 2 and 3 show the estimated nonparametric function $\hat{m}$ (light blue

[2]Note that $s^{(R)}$ and $l^{(R)}$ (and their combinations with $d^{(R)}, e^{(R)}, \pi^{(R)}$) have the same R_V^2 by construction of the transformed spread according to (4). 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 $l_{t-1}^{(R)} = (1+l_{t-1})/(1+r_{t-1})$. Both differ by a constant shift of one which has no impact in the estimation process with the local-linear smoother.

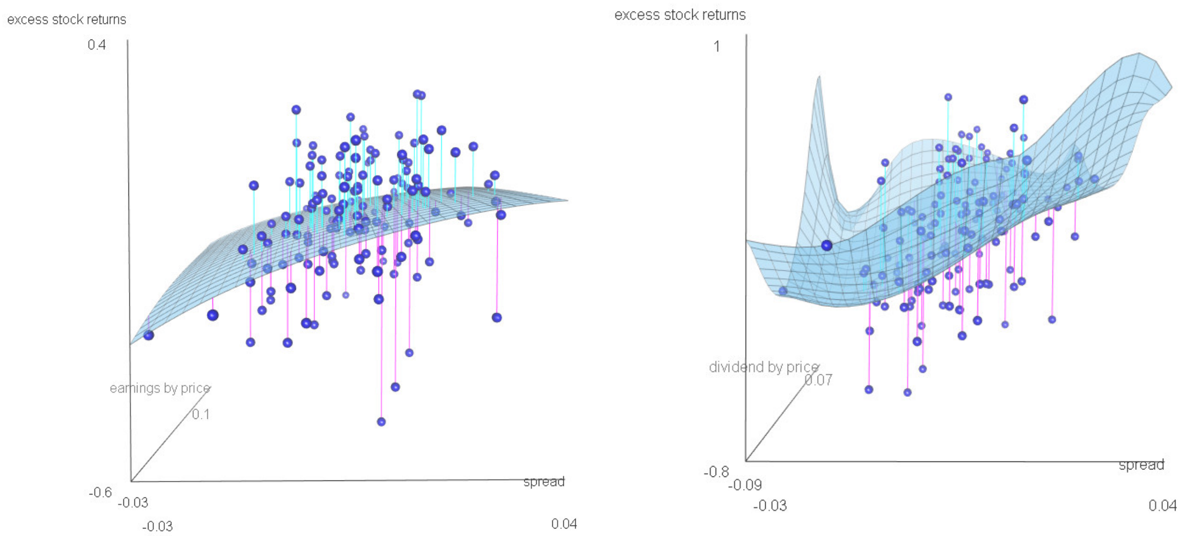

Figure 2: Risk-free rate benchmark. Relation between excess stock returns and predictive variables: the earnings-by-price ratio and the spread, one-year horizon (left); the dividend-by-price ratio and the spread, five-year horizon (right). Estimated nonparametric function $\hat{m}$ (light blue surface), observations (dark blue balls). Period: 1872–2020. Data: annual S&P 500.

surface) together with the underlying observations (dark blue balls). Especially for the risk-free rate benchmark, a nonlinear relationship is notable. However, the estimated function seems to be more stable under the inflation benchmark, that is, it is very similar for the one- and five-year horizon. All four plots indicate that with an increase in the spread also the predicted return increases, holding other factors fixed. Also an increase in the earnings predicts an increase in the return. Note that this effect is stronger for the inflation than for the risk-free rate benchmark. The dividend-by-price versus excess stock return relation for a fixed spread under the risk-free rate benchmark and the five-year horizon is U-shaped.

3.3 Short-term exuberance and long-term stability: combining predictions of short and long horizons

In this section, we illustrate the main empirical contribution of our paper. Recall that the simple econometric model introduced in Section 2.4, which combines short- and long-term predictions, builds on (a) the predictive power of earnings for excess stock returns with its linear model formulation of equation (12) and (b) the autoregressive development of order one of the earnings in equation (11). Table 6 shows the estimated OLS coefficients and regression summaries of the linear models (11) and (12). We find that the earnings variable has more predictive power for excess stock returns under the inflation benchmark than the short-term benchmark: the R^2 of the former is more than twice as large as the R^2 of the latter (13.4% versus 5.8%). The autoregressive behaviour of the earnings is stronger in terms of R^2 for the short-term benchmark than the inflation benchmark (63.8% versus 8.2%). However, the much smaller estimated coefficient under the inflation benchmark (0.286 versus 0.798) indicates a more stable variation of the earnings around the scaled historical mean. The intercept is also estimated significantly different from zero for both benchmarks. Figures 4 and 5 show the linear models (11) and (12) for both benchmarks (solid red line) together with estimates of the local-linear smoother (dashed green line) and the $45°$-line (dotted black line). These illustrations assure the usefulness of using linear functions in our econometric model of Section 2.4.

The next step in running our model is its calibration to the conditional mean and variance estimates for the one-year and five-year horizon (the right-hand side values in equation (23)); Table 7 shows those estimates for both benchmarks. Note that we use out-of-sample predictions from the optimal models discussed in Section 3.2 for both horizons (see also

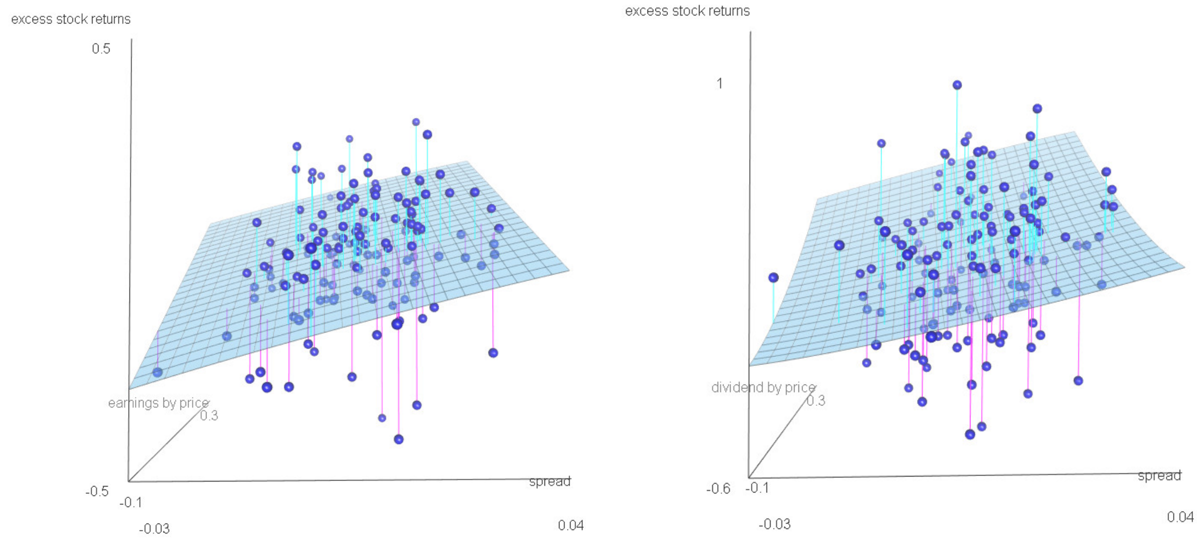

Figure 3: Inflation benchmark. Relation between excess stock returns and predictive variables: the earnings-by-price ratio and the spread, one-year horizon (left); the earnings-by-price ratio and the spread, five-year horizon (right). Estimated nonparametric function $\hat{m}$ (light blue surface), observations (dark blue balls). Period: 1872–2020. Data: annual S&P 500.

Table 6: Estimated parameters (and standard errors in parentheses) of the linear models (11) and (12) used for the econometric model of Section 2.4 under the short-term interest rate benchmark or the inflation benchmark. R^2 denotes the standard coefficient of determination of a linear model, Adj. R^2 the adjusted R^2, "Num. obs." the number of observations used in the regression, and RMSE the root mean square error.

Benchmark	Short-term interest rate		Inflation rate	
Dependent variable	$e_{t+1}^{(R)}$	$Y_{t+1}^{(R)}$	$e_{t+1}^{(C)}$	$Y_{t+1}^{(C)}$
Intercept	0.0066^{**}	0.0035	0.0373^{***}	0.0069
	(0.0022)	(0.0201)	(0.0063)	(0.0186)
$e_t^{(A)}$	0.7976^{***}	1.3522^{**}	0.2859^{***}	1.1144^{***}
	(0.0500)	(0.4531)	(0.0799)	(0.2350)
R^2	0.6384	0.0579	0.0817	0.1343
Adj. R^2	0.6359	0.0514	0.0753	0.1283
Num. obs.	146	147	146	147
RMSE	0.0186	0.1683	0.0572	0.1684

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

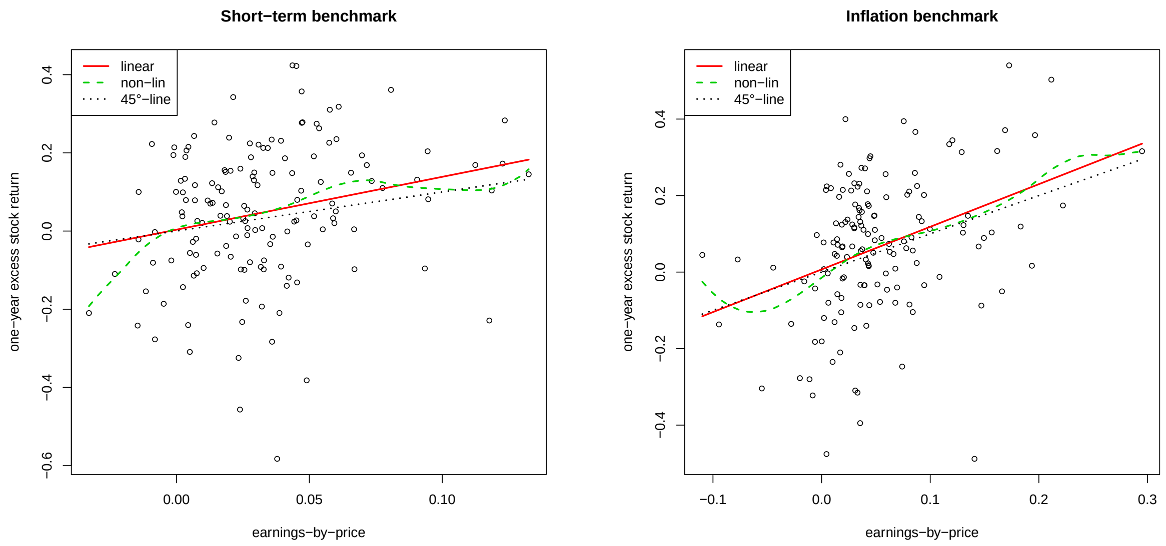

Figure 4: Relation between excess stock returns and the earnings-by-price ratio. Left: Short-term interest rate benchmark. Right: Inflation benchmark. Shown are estimates of a linear function (solid red), the local-linear smoother (dashed green), and the 45°-line (dotted black). Period: 1872–2020. Data: annual S&P 500.

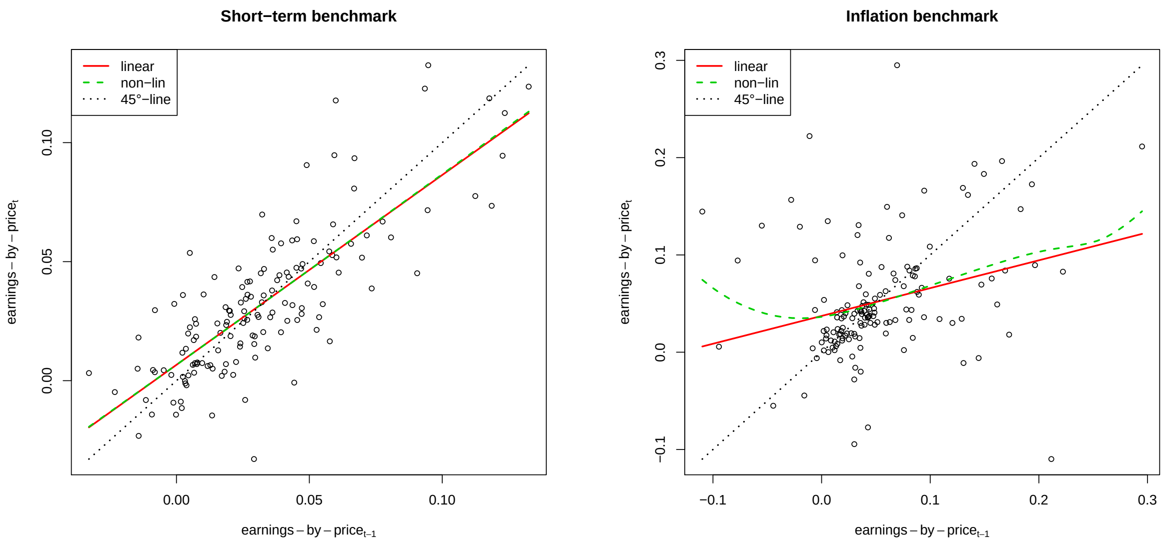

Figure 5: Autoregressive behaviour of the earnings-by-price ratio. Left: Short-term interest rate benchmark. Right: Inflation benchmark. Shown are estimates of a linear function (solid red), the local-linear smoother (dashed green), and the 45°-line (dotted black). Period: 1872–2020. Data: annual S&P 500.

Table 7: Estimated parameters of the econometric model under the short-term interest rate benchmark or the inflation benchmark (in %). For (conditional) predictions of the mean $\hat{\mu}_{ky}$, the variable combination with the largest $R^2_{V,ky}$ is used (see Table 5). $\hat{\sigma}_{ky}$ denotes the estimated standard deviation of the predictions $\hat{Y}^{(A),c}$ or $\hat{Z}^{(A),c}$, σ is the sample standard deviation of $Y^{(A)}$ or $Z^{(A)}$ (see also Table 2). $\hat{\alpha}_0$, $\hat{\alpha}_1$, $\hat{\sigma}_1$, $\hat{\sigma}_2$ are the parameter estimates of the econometric model of Section 2.4.

Benchmark	Short-term interest rate				Inflation rate			
	$\hat{\mu}_{ky}$	$\hat{\sigma}_{ky}$	$R^2_{V,ky}$	σ	$\hat{\mu}_{ky}$	$\hat{\sigma}_{ky}$	$R^2_{V,ky}$	σ
One-year ($k = 1$)	4.30	16.34	10.67	17.28	4.15	16.38	17.53	18.04
Five-year ($k = 5$)	18.81	32.33	22.18	36.65	27.41	33.52	14.85	36.33
	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\sigma}_1$	$\hat{\sigma}_2$	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\sigma}_1$	$\hat{\sigma}_2$
Parameter estimate	13.70	-2.30	16.34	13.95	-0.17	0.95	16.38	14.62

Table 5), that is, $\hat{\mu}_{1y}$ and $\hat{\mu}_{5y}$ are based on the newest predictive variables in our records (corresponding to December 2019 values). For the short-term benchmark, the optimal models predict returns of 4.30 ($1y$) and 18.81 ($5y$). Note that the average annual return for the five-year horizon of $18.81/5 = 3.76$ is smaller than the predicted return for the one-year horizon of 4.30. The econometric model should be able to capture adequately such a decline in annual returns and we exactly achieve this, as we show later, via the simple linear correction proposed in Section 2.4. For the inflation benchmark, the corresponding predictions are 4.15 ($1y$) and 27.41 ($5y$). While the picture is quite similar for the one-year horizon, the behaviour of the five-year predictions is different. We forecast an increase in one-year real returns as the average annual return for the five-year horizon of $27.41/5 = 5.48$, which is now larger than the predicted return for the one-year horizon of 4.15.

From the upper panel of Table 7, the standard deviations of predicted one-year and five-year returns, respectively, $\hat{\sigma}_{1y}$ and $\hat{\sigma}_{5y}$ appear reduced, compared to the standard deviation σ of observed returns, through the statistical modelling process for both benchmarks (see equations (21) and (22)). Under the short-term benchmark, we get a reduction from 17.28 to 16.34 ($1y$) and 36.65 to 32.33 ($5y$), while under the inflation benchmark from 18.04 to 16.38 ($1y$) and 36.33 to 33.52 ($5y$). Note that our model, combining the one- and the five-year horizon, further reduces the uncertainty and thus the risk for one-year returns under both benchmarks, as we explain in the following lines.

Using the estimated coefficients of the linear models (11) and (12) (c_0, c_1, b_0, b_1) as well as the predicted one-year and five-year returns and estimated variation ($\hat{\mu}_{1y}$, $\hat{\mu}_{5y}$, $\hat{\sigma}^2_{1y}$, $\hat{\sigma}^2_{5y}$), we can solve the equation system (23) and obtain the estimates $\hat{\alpha}_0$, $\hat{\alpha}_1$, $\hat{\sigma}_1$, $\hat{\sigma}_2$ (see equations (24)–(27)) reported in the lower panel of Table 7 for both benchmarks. For the inflation benchmark, we get $\hat{\alpha}_0 = -0.17$ and $\hat{\alpha}_1 = 0.95$, i.e., an intercept in our simple linear correction of predicted one-year returns (13)–(15) which is close to zero and a slope nearby one. This implies that only a slight correction suffices in combining optimal one-year and five-year stock return predictions. However, under the short-term benchmark, a much stronger correction is necessary to model the decline of the one-year returns over time: $\hat{\alpha}_0 = 13.70$ and $\hat{\alpha}_1 = -2.30$. Table 8 shows the development of the one-year returns for the periods of interest, i.e., from $n + 1$ to $n + 5$ for both benchmarks. Note that the corrected risk premium $\hat{Y}^{(R),c}$ equals 4.30 in period $n + 1$ by construction, reduces over time to 3.33 in period $n + 5$, and sums up over the five-year horizon to 18.81 again by construction. Similarly, the corrected real return $\hat{Y}^{(C),c}$ equals 4.15 in period $n + 1$ by construction, increases from year to year to 5.98 in period $n + 5$, and sums up to the five-year prediction of 27.41. The underlying development of the earnings and the simple return predictions from models (11) and (12) are also shown in Table 8.

Another important outcome of our econometric model is the additionally reduced variation of the corrected one-year returns of the periods $n + 2$ to $n + 5$. Table 7 reports for the short-term benchmark a reduction from $\hat{\sigma}_1 = 16.34$ to $\hat{\sigma}_2 = 13.95$, that is, a drop in variation of 19.2% compared with the sample standard deviation of one-year returns of $\sigma = 17.28$ – the starting point of our analysis. Similarly, for the inflation benchmark, $\hat{\sigma}_1 = 16.38$ reduces to $\hat{\sigma}_2 = 14.62$, that is, a decrease in variation of 19.0% from the sample standard deviation of one-year returns of $\sigma = 18.04$. The fact that $\hat{\sigma}_1 > \hat{\sigma}_2$ tells us that in predicted 'pure' one-year returns (i.e., ignoring the long-term view)

Table 8: Predicted excess stock returns from the econometric model of Section 2.4 under the short-term interest rate benchmark or the inflation rate benchmark. $e_n^{(A)}$ is the last earnings-by-price observation in our records (transformed according to the benchmark A) and corresponds to December 2019. $\hat{Y}^{(A)}$ denotes the one-year predictions of excess stock returns from the linear model (12) (parameter estimates in Table 6) and $\hat{Y}^{(A),c}$ their corrected counterparts based on (13)–(15) (parameter estimates in Table 7).

Benchmark	Short-term interest rate			Inflation rate		
Period	$e^{(R)}$	$\hat{Y}^{(R)}$	$\hat{Y}^{(R),c}$	$e^{(C)}$	$\hat{Y}^{(C)}$	$\hat{Y}^{(C),c}$
n	2.76	–	–	3.47	–	–
$n+1$	2.87	4.08	4.30	4.72	4.56	4.15
$n+2$	2.95	4.23	3.97	5.08	5.95	5.47
$n+3$	3.02	4.34	3.71	5.18	6.35	5.85
$n+4$	3.07	4.43	3.50	5.21	6.47	5.95
$n+5$	–	4.50	3.33	–	6.50	5.98

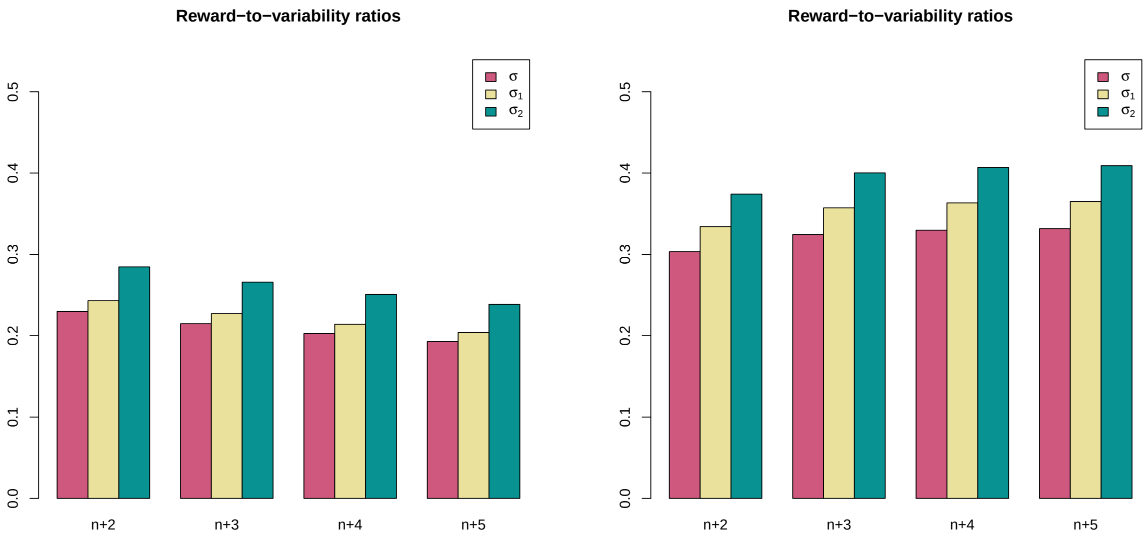

Figure 6: Comparison of reward-to-variability ratios based on corrected one-year predictions for periods $n+2, \ldots, n+5$ and σ (red), σ_1 (yellow) and σ_2 (green). Left: Short-term interest rate benchmark. Right: Inflation benchmark.

still a sort of bubble is present. In other words, even short-term predictions of the one-year horizon are prone to speculative exuberance. However, our simple model optimizes the one-year investments according to the bubble-free long-term variance reducing the variation/risk. This finding is relevant for the long-term investors (above one year), i.e., for the majority of us via our pension. Figure 6 illustrates this discussion and shows reward-to-variability ratios, for each benchmark A, based on the corrected one-year predictions $\hat{Y}^{(A),c}$ for the periods $n+2$ to $n+5$ and the three standard deviations in the model σ, σ_1, σ_2.

Finally, we repeat the model calibration for an alternative set of predictors under the risk-free rate benchmark for the five-year case, that is, the combination of earnings and term spread, aiming for congruity in the choice of the baseline set of predictors across benchmarks and horizons. In analogy to the reports in Tables 7–8, we present in Table 9 our estimates, which are minimally affected by this choice, whereas in Table 10 we exhibit the development of the earnings and return predictions which remains qualitatively similar.

Table 9: Estimated parameters of the econometric model under the short-term interest rate benchmark or the inflation benchmark based on common predictive variables (earnings and spread) for one- and five-year horizons. Changes are given in boldface. See also notes in Table 7.

Benchmark	Short-term interest rate				Inflation rate			
	$\hat{\mu}_{ky}$	$\hat{\sigma}_{ky}$	$R^2_{V,ky}$	σ	$\hat{\mu}_{ky}$	$\hat{\sigma}_{ky}$	$R^2_{V,ky}$	σ
One-year ($k = 1$)	4.30	16.34	10.67	17.28	4.15	16.38	17.53	18.04
Five-year ($k = 5$)	**20.93**	**32.41**	**21.78**	36.65	27.41	33.52	14.85	36.33
	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\sigma}_1$	$\hat{\sigma}_2$	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\sigma}_1$	$\hat{\sigma}_2$
Parameter estimate	**6.27**	**-0.48**	16.34	**14.00**	-0.17	0.95	16.38	14.62

Table 10: Table 8 continued: the case of common predictive variables (earnings and spread) under both benchmarks for one- and five-year predictions. Changes are given in boldface.

Benchmark	Short-term interest rate			Inflation rate		
Period	$e^{(R)}$	$\hat{Y}^{(R)}$	$\hat{Y}^{(R),c}$	$e^{(C)}$	$\hat{Y}^{(C)}$	$\hat{Y}^{(C),c}$
n	2.76	–	–	3.47	–	–
$n+1$	2.87	4.08	4.30	4.72	4.56	4.15
$n+2$	2.95	4.23	**4.23**	5.08	5.95	5.47
$n+3$	3.02	4.34	**4.17**	5.18	6.35	5.85
$n+4$	3.07	4.43	**4.13**	5.21	6.47	5.95
$n+5$	–	4.50	**4.10**	–	6.50	5.98

3.4 A final comment on the performance and the choice of the benchmark

Notice that our underlying estimates when considering the inflation benchmark are more stable than in the equivalent short-term interest case. The autoregressive earnings model is also more stable in the inflation case compared to the short-term interest case with a much higher mean-reversion. The modelling itself of excess returns shows a linear shape in the inflation case, see Figure 3, but comes with a lot of variability in the short-term interest case, see Figure 2. Also, adjusting under the inflation benchmark from a one-year model to a five-year model is non-dramatic, quite contrary to the complete change involved in the short-term interest case. While both models do similarly after validating with the short-term interest rate being ahead in the long-term case, we might have a tendency to prefer to work with the stable and intuitive inflation benchmark when providing our long-term and short-term model of stock returns. Of course, the choice of the benchmark depends on the ultimate application at hand. If one follows, for example, one of the key messages of Merton (2014), then forecast, especially for pensions, should be net of inflation.

4 Stock market predictions during the Covid-19 pandemic

The last part of the paper is dedicated to recent stock market predictions during the Covid-19 pandemic based on our optimal models from Section 3.2. Table 11 shows monthly US stock market data (in the period September 2019–June 2020), the transformed inputs in our models for one-year and five-year predictions (here, the dividend-by-price $d^{(A)}$, earnings-by-price $e^{(A)}$, and the term spread $s^{(A)}$), and the predicted model outcomes (here, the one-year risk premium $\widehat{RP}$, the one-year real return $\hat{Y}_{real}$, the one-year nominal return $\hat{Y}_{nom}$, and the five-year benchmarked return $\hat{Z}^{(A)}$) for both benchmarks. Note that the 'pure' nominal stock return is simply calculated by back-transforming the equation (1):

$$\hat{Y}_{nom,t} = \widehat{\ln S_t} = \hat{Y}_t^{(A)} + \ln B_{t-1}^{(A)}. \tag{28}$$

Several interesting insights into the dynamics of the stock market can be provided. First, the predicted one-year and five-year returns for the two months of (slightly) negative spread in August and September 2019 (an inverted yield curve) do not seem to be dramatically lower compared with other 'normal' months. However, considering the average

Table 11: One-year and five-year ahead real-time forecasts based on recent monthly US stock market variables (as defined in Section 2.1) under the short-term interest rate benchmark or the inflation rate benchmark. For the predictions $\hat{Y}^{(A)}$ and $\hat{Z}^{(A)}$, the variable combination with the largest R^2_V under each benchmark A is used (see Table 5). $\hat{Y}_{nom}$ denotes the predictions of the one-year nominal stock returns calculated from the back-transformation (28). $\hat{Y}_{real}$ is the one-year real stock return, whereas $\widehat{RP}$ is the estimated risk premium. (By construction, $\hat{Y}^{(R)}$ is a direct estimate of the risk premium $\widehat{RP}$, as well as $\hat{Y}^{(C)}$ of the real stock returns $\hat{Y}_{real}$). † indicates estimates of 12 month earnings per share instead of actuals due to reporting lags in quarterly earnings data.

Monthly US stock market data							Short-term interest rate							Inflation rate						
Date	P	D	E	R	L	π	$d^{(R)}$	$e^{(R)}$	$s^{(R)}$	$\widehat{RP}$	$\hat{Y}_{nom}$	$\hat{Y}_{real}$	$\hat{Z}^{(R)}$	$d^{(C)}$	$e^{(C)}$	$s^{(C)}$	$\widehat{RP}$	$\hat{Y}_{nom}$	$\hat{Y}_{real}$	$\hat{Z}^{(C)}$
2018-09	2901.50	52.34	130.39	2.47	3.00	2.28	-0.65	1.98	0.52	3.62	6.06	3.78	19.83	-0.46	2.17	0.52	2.64	5.11	2.86	24.79
2018-10	2785.46	52.81	131.06	2.56	3.15	2.52	-0.65	2.09	0.58	3.95	6.47	3.95	18.73	-0.61	2.13	0.58	2.95	5.51	3.02	25.07
2018-11	2723.23	53.28	131.72	2.60	3.12	2.18	-0.63	2.18	0.51	3.82	6.39	4.21	22.38	-0.22	2.60	0.51	2.89	5.49	3.34	25.75
2018-12	2567.31	53.75	132.39	2.57	2.83	1.91	-0.46	2.52	0.25	3.36	5.90	3.99	22.75	0.18	3.19	0.26	2.43	5.00	3.11	25.48
2019-01	2607.39	54.15	133.06	2.50	2.71	1.55	-0.41	2.54	0.20	3.21	5.68	4.13	25.46	0.52	3.50	0.21	2.34	4.84	3.30	25.90
2019-02	2754.86	54.54	133.72	2.47	2.68	1.52	-0.48	2.33	0.20	2.96	5.40	3.88	25.24	0.45	3.28	0.21	2.09	4.56	3.05	25.40
2019-03	2803.98	54.94	134.39	2.41	2.57	1.86	-0.44	2.33	0.16	2.78	5.17	3.30	20.52	0.10	2.88	0.16	1.82	4.23	2.39	24.13
2019-04	2903.80	55.32	134.68	2.34	2.53	2.00	-0.42	2.25	0.19	2.79	5.10	3.11	18.60	-0.09	2.59	0.19	1.79	4.13	2.15	23.65
2019-05	2854.71	55.70	134.98	2.27	2.40	1.79	-0.31	2.40	0.13	2.77	5.02	3.23	19.98	0.16	2.89	0.13	1.79	4.06	2.29	23.97
2019-06	2890.17	56.08	135.27	1.94	2.07	1.65	0.00	2.69	0.13	3.11	5.03	3.38	19.31	0.29	2.98	0.13	2.10	4.04	2.40	24.19
2019-07	2996.11	56.46	134.48	1.91	2.06	1.81	-0.03	2.53	0.15	2.99	4.89	3.07	17.17	0.07	2.63	0.15	1.94	3.85	2.06	23.49
2019-08	2897.45	56.84	133.69	1.73	1.63	1.75	0.23	2.84	-0.10	2.46	4.18	2.43	14.08	0.21	2.82	-0.10	1.38	3.11	1.38	22.36
2019-09	2982.16	57.22	132.90	1.75	1.70	1.71	0.17	2.66	-0.05	2.44	4.17	2.46	15.06	0.20	2.70	-0.05	1.37	3.12	1.42	22.40
2019-10	2977.68	57.56	135.09	1.57	1.71	1.76	0.36	2.92	0.14	3.42	4.98	3.21	15.63	0.17	2.72	0.14	2.31	3.88	2.14	23.65
2019-11	3104.90	57.90	137.28	1.53	1.81	2.05	0.33	2.85	0.28	3.82	5.33	3.28	13.17	-0.18	2.32	0.27	2.66	4.19	2.16	23.59
2019-12	3176.75	58.24	139.47	1.51	1.86	2.29	0.32	2.84	0.34	4.04	5.54	3.25	10.79	-0.44	2.06	0.34	2.85	4.36	2.10	23.41
2020-01	3278.20	58.69	131.76	1.49	1.76	2.49	0.30	2.49	0.27	3.37	4.85	2.36	6.82	-0.68	1.50	0.26	2.11	3.60	1.15	21.61
2020-02	3277.31	59.13	124.04	1.37	1.50	2.33	0.43	2.38	0.13	2.75	4.11	1.78	6.33	-0.52	1.42	0.13	1.49	2.86	0.56	20.56
2020-03	2652.39	59.58	116.33	0.32	0.87	1.54	1.92	4.05	0.55	6.04	6.36	4.82	19.26	0.70	2.80	0.54	4.90	5.22	3.69	26.43
2020-04	2761.98	59.72	110.87†	0.18	0.66	0.33	1.98	3.83	0.48	5.58	5.76	5.43	29.99	1.83	3.67	0.48	4.64	4.82	4.49	28.06
2020-05	2919.61	59.87	105.41†	0.16	0.67	0.12	1.89	3.44	0.51	5.26	5.42	5.31	31.36	1.93	3.49	0.51	4.34	4.50	4.38	27.82
2020-06	3104.66	60.01	99.95†	0.18	0.73	0.65	1.75	3.03	0.55	4.94	5.12	4.47	26.02	1.28	2.56	0.55	3.89	4.07	3.42	25.88

returns for the five-year period under the short-term interest rate, one finds similar to the inverted yield curve that short-term investments are more profitable compared with longer-term investments during the period October 2019 to March 2020 ($15.63/5 < 3.42, \ldots, 19.26/5 < 6.04$). This inversion ends in April 2020 with the dramatic fall in stock prices caused by the Covid-19 pandemic.

Second, whereas earnings have dropped dramatically from November 2019 until June 2020, the future outlook for the long-term stock returns net of inflation seems to be unaffected. The reason for this is the quantitative easing policy of the Federal Reserve resulting in a dramatic drop of the short-term interest rate in the same period. Thus, the good performance of the stock market during the crisis seems to be facilitated by political decisions (including those decisions of the central bank) rather than market dynamics. Are we to look forward into 2021 when it is no longer election year in the US and when interest rates cannot drop much more, then it seems clear that it can be hard to find political tools to keep the hand in a prolonged Covid-19 crisis or the emergence of a completely new crisis. A practical conclusion of our study is therefore that the stock market dynamics might be more fragile and volatile going forward than what we have seen in the recent past.

Finally, as in the previous section, we repeat here the same exercise based on the alternative model estimates in Table 9, which result in a similar qualitative behaviour as observed in Table 12.

5 Concluding remarks

In this paper, we combine optimal long-term and short-term information to provide a state-of-the-art econometric model that serves short-term market timing as well as long-term asset-allocation strategy for the long-term saver. We apply to US stock market excess returns and common predictors, based on the short-term interest rate or the inflation benchmark. As part of our study, we interestingly find a high short-term (one-year) standard deviation insinuating presence of bubbles in the returns. This suggests that long-term investors should disregard short-term econometric models when deciding on their long-term asset allocation. We conclude that, for a given risk appetite level, the ability of adding equity exposure to result in increased long-term savers' portfolio return is significant as it provides better pensions for everyone (see also Merton, 2014, Gerrard *et al.*, 2019). We favour the inflation benchmark, which expresses everything in real terms, in the current climate as it perfectly links with Merton's (2014) pension vision and provides good predictive power based on our empirical results. In fact, we find the short-term interest rate benchmark, which is often used, to be more unstable. We also find earnings to be the most important predictor of stock returns.

Appendix

In Section 2.2, we introduce our setup for the T-year predictions. Here, we describe in some more detail important single steps. Equation (5) defines $Z_t^{(A)}$ as the sum of annual continuously compounded returns which are of an overlapping nature:

$$\begin{aligned}
Z_1^{(A)} &= Y_1^{(A)} + Y_2^{(A)} + Y_3^{(A)} + \ldots + Y_T^{(A)} \\
Z_2^{(A)} &= Y_2^{(A)} + Y_3^{(A)} + Y_4^{(A)} + \ldots + Y_{T+1}^{(A)} \\
Z_3^{(A)} &= Y_3^{(A)} + Y_4^{(A)} + Y_5^{(A)} + \ldots + Y_{T+2}^{(A)} \\
&\vdots \\
Z_{n-T+1}^{(A)} &= Y_{n-T+1}^{(A)} + Y_{n-T+2}^{(A)} + Y_{n-T+3}^{(A)} + \ldots + Y_n^{(A)},
\end{aligned}$$

Table 12: Table 11 continued: the case of common predictive variables (earnings and spread) under both benchmarks for one- and five-year predictions. Changes are given in boldface.

Monthly US stock market data							Short-term interest rate							Inflation rate						
Date	P	D	E	R	L	π	$d^{(R)}$	$e^{(R)}$	$s^{(R)}$	$\widehat{RP}$	$\hat{Y}_{nom}$	$\hat{Y}_{real}$	$\hat{Z}^{(R)}$	$d^{(C)}$	$e^{(C)}$	$s^{(C)}$	$\hat{Y}_{real}$	$\hat{Y}_{nom}$	$\widehat{RP}$	$\hat{Z}^{(C)}$
2018-09	2901.50	52.34	130.39	2.47	3.00	2.28	-0.65	1.98	0.52	3.62	6.06	3.78	**20.57**	-0.46	2.17	0.52	2.86	5.11	2.64	24.79
2018-10	2785.46	52.81	131.06	2.56	3.15	2.52	-0.65	2.09	0.58	3.95	6.47	3.95	**20.03**	-0.61	2.13	0.58	3.02	5.51	2.95	25.07
2018-11	2723.23	53.28	131.72	2.60	3.12	2.18	-0.63	2.18	0.51	3.82	6.39	4.21	**23.63**	-0.22	2.60	0.51	3.34	5.49	2.89	25.75
2018-12	2567.31	53.75	132.39	2.57	2.83	1.91	-0.46	2.52	0.25	3.36	5.90	3.99	**24.72**	0.18	3.19	0.26	3.11	5.00	2.43	25.48
2019-01	2607.39	54.15	133.06	2.50	2.71	1.55	-0.41	2.54	0.20	3.21	5.68	4.13	**27.26**	0.52	3.50	0.21	3.30	4.84	2.34	25.90
2019-02	2754.86	54.54	133.72	2.47	2.68	1.52	-0.48	2.33	0.20	2.96	5.40	3.88	**26.44**	0.45	3.28	0.21	3.05	4.56	2.09	25.40
2019-03	2803.98	54.94	134.39	2.41	2.57	1.86	-0.44	2.33	0.16	2.78	5.17	3.30	**21.79**	0.10	2.88	0.16	2.39	4.23	1.82	24.13
2019-04	2903.80	55.32	134.68	2.34	2.53	2.00	-0.42	2.25	0.19	2.79	5.10	3.11	**19.68**	-0.09	2.59	0.19	2.15	4.13	1.79	23.65
2019-05	2854.71	55.70	134.98	2.27	2.40	1.79	-0.31	2.40	0.13	2.77	5.02	3.23	**21.09**	0.16	2.89	0.13	2.29	4.06	1.79	23.97
2019-06	2890.17	56.08	135.27	1.94	2.07	1.65	0.00	2.69	0.13	3.11	5.03	3.38	**20.28**	0.29	2.98	0.13	2.40	4.04	2.10	24.19
2019-07	2996.11	56.46	134.48	1.91	2.06	1.81	-0.03	2.53	0.15	2.99	4.89	3.07	**17.87**	0.07	2.63	0.15	2.06	3.85	1.94	23.49
2019-08	2897.45	56.84	133.69	1.73	1.63	1.75	0.23	2.84	-0.10	2.46	4.18	2.43	**15.25**	0.21	2.82	-0.10	1.38	3.11	1.38	22.36
2019-09	2982.16	57.22	132.90	1.75	1.70	1.71	0.17	2.66	-0.05	2.44	4.17	2.46	**15.68**	0.20	2.70	-0.05	1.42	3.12	1.37	22.40
2019-10	2977.68	57.56	135.09	1.57	1.71	1.76	0.36	2.92	0.14	3.42	4.98	3.21	**16.27**	0.17	2.72	0.14	2.14	3.88	2.31	23.65
2019-11	3104.90	57.90	137.28	1.53	1.81	2.05	0.33	2.85	0.28	3.82	5.33	3.28	**14.03**	-0.18	2.32	0.27	2.16	4.19	2.66	23.59
2019-12	3176.75	58.24	139.47	1.51	1.86	2.29	0.32	2.84	0.34	4.04	5.54	3.25	**12.07**	-0.44	2.06	0.34	2.10	4.36	2.85	23.41
2020-01	3278.20	58.69	131.76	1.49	1.76	2.49	0.30	2.49	0.27	3.37	4.85	2.36	**7.56**	-0.68	1.50	0.26	1.15	3.60	2.11	21.61
2020-02	3277.31	59.13	124.04	1.37	1.50	2.33	0.43	2.38	0.13	2.75	4.11	1.78	**5.99**	-0.52	1.42	0.13	0.56	2.86	1.49	20.56
2020-03	2652.39	59.58	116.33	0.32	0.87	1.54	1.92	4.05	0.55	6.04	6.36	4.82	**14.26**	0.70	2.80	0.54	3.69	5.22	4.90	26.43
2020-04	2761.98	59.72	110.87†	0.18	0.66	0.33	1.98	3.83	0.48	5.58	5.76	5.43	**23.88**	1.83	3.67	0.48	4.49	4.82	4.64	28.06
2020-05	2919.61	59.87	105.41†	0.16	0.67	0.12	1.89	3.44	0.51	5.26	5.42	5.31	**24.70**	1.93	3.49	0.51	4.38	4.50	4.34	27.82
2020-06	3104.66	60.01	99.95†	0.18	0.73	0.65	1.75	3.03	0.55	4.94	5.12	4.47	**18.16**	1.28	2.56	0.55	3.42	4.07	3.89	25.88

where n is the number of observations of one-year returns. Using the relations stated in equation (6), one gets easily

$$\begin{aligned} Z_t^{(A)} &= (\beta_0 + \beta_1 X_{t-1}^{(A)} + \xi_t) + \ldots + (\beta_0 + \beta_1 X_{t+T-2}^{(A)} + \xi_{t+T-1}) \\ &= \beta_0 T + \beta_1 \gamma_0 \sum_{i=0}^{T-1} \sum_{j=0}^{T-2-i} \gamma_1^j + \beta_1 X_{t-1}^{(A)} \sum_{i=0}^{T-1} \gamma_1^i + \beta_1 \sum_{i=0}^{T-1} \sum_{j=0}^{T-2-i} \gamma_1^j \eta_{t+i} + \sum_{i=0}^{T-1} \xi_{t+i} \\ &= \phi_0 + \phi_1 X_{t-1}^{(A)} + \nu_t, \end{aligned}$$

where

$$\begin{aligned} \phi_0 &:= \beta_0 T + \beta_1 \gamma_0 \sum_{i=0}^{T-1} \sum_{j=0}^{T-2-i} \gamma_1^j, \\ \phi_1 &:= \beta_1 \sum_{i=0}^{T-1} \gamma_1^i, \\ \nu_t &:= \beta_1 \sum_{i=0}^{T-1} \sum_{j=0}^{T-2-i} \gamma_1^j \eta_{t+i} + \sum_{i=0}^{T-1} \xi_{t+i}. \end{aligned}$$

References

Bergmeir, C., Hyndman, R. J. and Koo, B. (2018) A note on the validity of cross-validation for evaluating autoregressive time series predictions. *Computational Statistics and Data Analysis*, **120**, 70–83.

Bodnar, T., Parolya, N. and Schmid, W. (2015) On the exact solution of the multi-period portfolio choice problem for an exponential utility under return predictability. *European Journal of Operational Research*, **246**, 528–542.

Burman, P., Chow, E. and Nolan, D. (1994) A cross-validatory method for dependent data. *Biometrika*, **81**, 351–358.

Carmona, J., León, A. and Vaello-Sebastià, A. (2012) Does stock return predictability affect ESO fair value? *European Journal of Operational Research*, **223**, 188–202.

Chen, Q. and Hong, Y. (2010) Predictability of equity returns over diferent time horizons: a nonparametric approach. Cornell University/Department of Economics. Working Paper.

Cheng, T., Gao, J. and Linton, O. (2019) Nonparametric predictive regressions for stock return predictions. Cambridge Working Papers in Economics: 1932.

Chu, C. K. and Marron, J. S. (1991) Comparison of two bandwidth selectors with dependent errors. *The Annals of Statistics*, **19**, 1906–1918.

De Brabanter, K., De Brabanter, J., Suykens, J. and De Moor, B. (2011) Kernel regression in the presence of correlated errors. *Journal of Machine Learning Research*, **12**, 1955–1976.

Geller, J. and Neumann, M. H. (2018) Improved local polynomial estimation in time series regression. *Journal of Nonparametric Statistics*, **30**, 1–27.

Gerrard, R., Hiabu, M., Kyriakou, I. and Nielsen, J. P. (2019) Communication and personal selection of pension saver's financial risk. *European Journal of Operational Research*, **274**, 1102–1111.

Gerrard, R., Hiabu, M., Nielsen, J. and Vodička, P. (2020) Long-term real dynamic investment planning. *Insurance: Mathematics and Economics*, **92**, 90–103.

Kyriakou, I., Mousavi, P., Nielsen, J. P. and Scholz, M. (2019) Forecasting benchmarks of long-term stock returns via machine learning. URL `http://link.springer.com/article/10.1007/s10479-019-03338-4`. Annals of Operations Research.

Kyriakou, I., Mousavi, P., Nielsen, J. P. and Scholz, M. (2020) Longer-term forecasting of excess stock returns–the five-year case. *Mathematics*, **8**, 1–20.

Lettau, M. and Van Nieuwerburgh, S. (2008) Reconciling the return predictability evidence. *Review of Financial Studies*, **21**, 1601–1652.

Linton, O. B. and Mammen, E. (2008) Nonparametric transformation to white noise. *Journal of Econometrics*, **142**, 241–264.

Lioui, A. and Poncet, P. (2019) Long horizon predictability: An asset allocation perspective. *European Journal of Operational Research*, **278**, 961–975.

Mammen, E., Nielsen, J. P., Scholz, M. and Sperlich, S. (2019) Conditional variance forecasts for long-term stock returns. *Risks*, **7**, 1–22.

Merton, R. C. (2014) The crisis in retirement planning. *Harvard Business Review*, **92**, 43–50.

Nielsen, J. P. and Sperlich, S. (2003) Prediction of stock returns: A new way to look at it. *ASTIN Bulletin*, **33**, 399–417.

Opsomer, J., Wang, Y. and Yang, Y. (2001) Nonparametric regression with correlated errors. *Statistical Science*, **16**, 134–153.

Rapach, D. and Zhou, G. (2013) Forecasting stock returns. In *Handbook of Economic Forecasting* (eds. G. Elliott and A. Timmerman), 328–383. Amsterdam: Elsevier, 2a edn.

Scholz, M., Nielsen, J. P. and Sperlich, S. (2015) Nonparametric prediction of stock returns based on yearly data: The long-term view. *Insurance: Mathematics and Economics*, **65**, 143–155.

Scholz, M., Sperlich, S. and Nielsen, J. P. (2016) Nonparametric long term prediction of stock returns with generated bond yields. *Insurance: Mathematics and Economics*, **69**, 82–96.

Shiller, R. J. (1989) *Market Volatility*. Cambridge: MIT Press.

Su, L. and Ullah, A. (2006) More efficient estimation in nonparametric regression with nonparametric autocorrelated errors. *Econometric Theory*, **22**, 98–126.

Welch, I. and Goyal, A. (2008) A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies*, **21**, 1455–1508.

Xiao, Z., Linton, O. B., Carroll, R. J. and Mammen, E. (2003) More efficient local polynomial estimation in nonparametric regression with autocorrelated errors. *Journal of the American Statistical Association*, **98**, 980–992.

Graz Economics Papers

For full list see:
http://ideas.repec.org/s/grz/wpaper.html
Address: Department of Economics, University of Graz,
Universitätsstraße 15/F4, A-8010 Graz

20–2020	**Ioannis Kyriakou, Parastoo Mousavi, Jens Perch Nielsen and Michael Scholz**: Short-Term Exuberance and long-term stability: A simultaneous optimization of stock return predictions for short and long horizons
19–2020	**Jie Chen, Yu Chen, Robert J. Hill, and Pei Hu**: The User Cost of Housing and the Price-Rent Ratio in Shanghai
18–2020	**Robert J. Hill, Miriam Steurer and Sofie R. Waltl**: Owner Occupied Housing, Inflation and Monetary Policy
17–2020	**Norbert Pfeifer and Miriam Steurer**: Early Real Estate Indicators during the Covid-19 Crisis - A Tale of Two Cities
16–2020	**Noha Elboghdadly and Michael Finus**: Non-Cooperative Climate Policies among Asymmetric Countries: Production- versus Consumption-based Carbon Taxes
15–2020	**Daniel Reiter**: Socioeconomic Integration through Language: Evidence from the European Union
14–2020	**Robert J. Hill, Michael Scholz, Chihiro Shimizu and Miriam Steurer**: Rolling-Time-Dummy House Price Indexes: Window Length, Linking and Options for Dealing with the Covid-19 Shutdown
13–2020	**Stefan Borsky and Martin Jury**: The role of global supply chains in the transmission of weather induced production shocks
12–2020	**Stefan Borsky, Hannah Hennighausen, Andrea Leiter and Keith Williges**: CITES and the Zoonotic Disease Content in International Wildlife Trade
11–2020	**Noha Elboghdadly and Michael Finus**: Enforcing Climate Agreements: The Role of Escalating Border Carbon Adjustments
10–2020	**Alejandro Caparros and Michael Finus**: The Corona-Pandemic: A Game-theoretic Perspective on Regional and Global Governance

09–2020 **Stefan Borsky and Hannah B. Hennighausen**: Public flood risk mitigation and the homeowner's insurance demand response

08–2020 **Robert Hill and Radoslaw Trojanek**: House Price Indexes for Warsaw: An Evaluation of Competing Methods

07–2020 **Noha Elboghdadly and Michael Finus**: Strategic Climate Policies with Endogenous Plant Location: The Role of Border Carbon Adjustments

06–2020 **Robert J. Hill, Norbert Pfeifer, and Miriam Steurer**: The Airbnb Rent-Premium and the Crowding-Out of Long-Term Rentals

05–2020 **Thomas Schinko, Birgit Bednar-Friedl, Barbara Truger, Rafael Bramreiter, Nadejda Komendantova, Michael Hartner**: Economy-wide benefits and costs of local-level energy transition in Austrian Climate and Energy Model Regions

04–2020 **Alaa Al Khourdajie and Michael Finus**: Measures to Enhance the Effectiveness of International Climate Agreements: The Case of Border Carbon Adjustments

03–2020 **Katja Kalkschmied**: Rebundling Institutions

02–2020 **Miriam Steurer and Caroline Bayr**: Measuring Urban Sprawl using Land Use Data

01–2020 **Thomas Aronsson, Sugata Ghosh and Ronald Wendner**: Positional Preferences and Efficiency in a Dynamic Economy

14–2019 **Nicole Palan, Nadia Simoes, and Nuno Crespo**: Measuring Fifty Years of Trade Globalization

13–2019 **Alejandro Caparrós and Michael Finus**: Public Good Agreements under the Weakest-link Technology

12–2019 **Michael Finus, Raoul Schneider and Pedro Pintassilgo**: The Role of Social and Technical Excludability for the Success of Impure Public Good and Common Pool Agreements: The Case of International Fisheries

11–2019 **Thomas Aronsson, Olof Johansson-Stenman and Ronald Wendner**: Charity as Income Redistribution: A Model with Optimal Taxation, Status, and Social Stigma

10–2019 **Yuval Heller and Christoph Kuzmics**: Renegotiation and Coordination with Private Values

09–2019 **Philipp Külpmann and Christoph Kuzmics**: On the Predictive Power of Theories of One-Shot Play

08–2019 **Enno Mammen, Jens Perch Nielsen, Michael Scholz and Stefan Sperlich**: Conditional variance forecasts for long-term stock returns

07–2019 **Christoph Kuzmics, Brian W. Rogers and Xiannong Zhang**: Is Ellsberg behavior evidence of ambiguity aversion?

06–2019 **Ioannis Kyriakou, Parastoo Mousavi, Jens Perch Nielsen and Michael Scholz**: Machine Learning for Forecasting Excess Stock Returns The Five-Year-View

05–2019 **Robert J. Hill, Miriam Steurer and Sofie R. Waltl**: Owner-Occupied Housing, Inflation, and Monetary Policy

04–2019 **Thomas Aronsson, Olof Johansson-Stenman and Ronald Wendner**: Charity, Status, and Optimal Taxation: Welfarist and Paternalist Approaches

03–2019 **Michael Greinecker and Christoph Kuzmics**: Limit Orders under Knightian Uncertainty

02–2019 **Miriam Steurer and Robert J. Hill**: Metrics for Measuring the Performance of Machine Learning Prediction Models: An Application to the Housing Market

01–2019 **Katja Kalkschmied and Jörn Kleinert**: (Mis)Matches of Institutions: The EU and Varieties of Capitalism

21–2018 **Nathalie Mathieu-Bolh and Ronald Wendner**: We Are What We Eat: Obesity, Income, and Social Comparisons

20–2018 **Nina Knittel, Martin W. Jury, Birgit Bednar-Friedl, Gabriel Bachner and Andrea Steiner**: The implications of climate change on Germanys foreign trade: A global analysis of heat-related labour productivity losses

19–2018 **Yadira Mori-Clement, Stefan Nabernegg and Birgit Bednar-Friedl**: Can preferential trade agreements enhance renewable electricity generation in emerging economies? A model-based policy analysis for Brazil and the European Union

18–2018 **Stefan Borsky and Katja Kalkschmied**: Corruption in Space: A closer look at the world's subnations

17–2018 **Gabriel Bachner, Birgit Bednar-Friedl and Nina Knittel**: How public adaptation to climate change affects the government budget: A model-based analysis for Austria in 2050

16–2018 **Michael Günther, Christoph Kuzmics and Antoine Salomon**: A Note on Renegotiation in Repeated Games [Games Econ. Behav. 1 (1989) 327360]

15–2018 **Meng-Wei Chen, Yu Chen, Zhen-Hua Wu and Ningru Zhao**: Government Intervention, Innovation, and Entrepreneurship

14–2018 **Yu Chen, Shaobin Shi and Yugang Tang**: Valuing the Urban Hukou in China: Evidence from a Regression Discontinuity Design in Housing Price

13–2018 **Stefan Borsky and Christian Unterberger**: Bad Weather and Flight Delays: The Impact of Sudden and Slow Onset Weather Events

12–2018 **David Rietzke and Yu Chen**: Push or Pull? Performance-Pay, Incentives, and Information

11–2018 **Xi Chen, Yu Chen and Xuhu Wan**: Delegated Project Search

10–2018 **Stefan Nabernegg, Birgit Bednar-Friedl, Pablo Muñoz, Michaela Titz and Johanna Vogel**: National policies for global emission reductions: Effectiveness of carbon emission reductions in international supply chains

09–2018 **Jonas Dovern and Hans Manner**: Order Invariant Tests for Proper Calibration of Multivariate Density Forecasts

08–2018 **Ioannis Kyriakou, Parastoo Mousavi, Jens Perch Nielsen and Michael Scholz**: Choice of Benchmark When Forecasting Long-term Stock Returns

07–2018 **Joern Kleinert**: Globalization Effects on the Distribution of Income

06–2018 **Nian Yang, Jun Yang and Yu Chen**: Contracting in a Continuous-Time Model with Three-Sided Moral Hazard and Cost Synergies

05–2018 **Christoph Kuzmics and Daniel Rodenburger**: A case of evolutionary stable attainable equilibrium in the lab

04–2018 **Robert J. Hill, Alicia N. Rambaldi, and Michael Scholz**: Higher Frequency Hedonic Property Price Indices: A State Space Approach

03–2018 **Reza Hajargasht, Robert J. Hill, D. S. Prasada Rao, and Sriram Shankar**: Spatial Chaining in International Comparisons of Prices and Real Incomes

02–2018 **Christoph Zwick**: On the origin of current account deficits in the Euro area periphery: A DSGE perspective

01–2018 **Michael Greinecker and Christopher Kah**: Pairwise stable matching in large economies