JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSIS Vol. 60, No. 1, Feb. 2025, pp. 482–523

doi:10.1017/S0022109024000218

Predictability Puzzles

Bjørn Eraker
Wisconsin School of Business, Department of Finance
bjorn.eraker@wisc.edu

Abstract

Dynamic equilibrium models based on present value computation not only imply that returns are predictable but also generate particular short-term patterns of predictability in asset returns. I take advantage of this to construct a set of tests of equilibrium generated predictability (EGP). I apply the tests to document two puzzles: First, option-implied or realized measures of volatility ought to predict returns but do not; and second, the variance risk premium (VRP) predicts returns but only at long horizons. VRP fails the tests of EGP as the term structure of predictable variation is inconsistent with an equilibrium interpretation.

I. Introduction

Few topics in finance are as heavily researched and hotly contested as the predictability of asset returns. Cowles (1933), (1944) showed empirically that professional stock forecasters were no better at forecasting than a random forecast. Samuelson (1965) provided a technical proof of the random walk behavior of equilibrium stock prices, adding to existing empirical evidence (e.g., Kendall (1953), Cootner (1964)). Counter to the classic theory of random walk, Keim and Stambaugh (1986), Fama and French (1988a), and Campbell and Shiller (1988a), (1988b) document predictable variation in stock returns from price–dividend ratios. It is also well understood that some predictable variation is consistent with dynamic equilibrium under time-varying expected returns (e.g., Fama (1970), Merton (1973)).

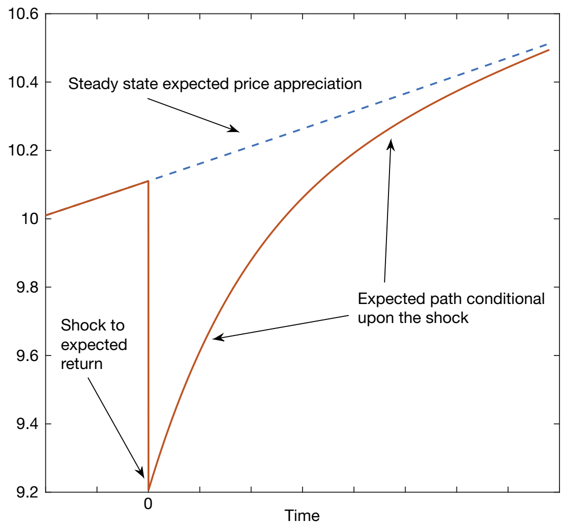
More recent work on equilibrium modeling, including Bansal and Yaron (2004), Campbell and Cochrane (1999), and Menzly, Santos, and Veronesi (2004), generates predictable variation in returns from risk-based measures such as volatility. In these models, prices are obtained as present values of future dividend payments. Expected returns, or discount factors, work so that shocks to risk have a contemporaneous negative impact on prices, creating a temporary price impact that is subsequently reversed as the shock dissipates. Figure 1 illustrates this in the form of an impulse response.

I thank Hendrik Bessembinder (the editor) and Andrea Tamoni (the referee) for valuable comments. I also thank Andrew Chen, Mikhail Chernov, Dobrislav Dobrev, Mohammad Jahan-Parvar, Dmitriy Muravyev, Neil Pearson, Cisil Sarisoy, Ivan Shaliastovich, Sang Byung Seo, Paul Whelan, Michelle Harasimowicz, Hao Zhou, and seminar participants at the 2020 Virtual Derivatives Workshop, Tsinghua University, Federal Reserve Board of Governors, University of Wisconsin–Madison, 2017 Boston University Conference on Financial Econometrics, and 2017 Midwest Finance Association Annual Meeting for helpful comments.

FIGURE 1

Equilibrium Stock Price Response of a Positive Expected Return Shock

Figure 1 shows the expected unconditional price appreciation versus price appreciation given a positive shock to expected return.



In this paper, I develop an empirical framework for testing the dynamic relationship between candidate risk variables and expected return variation, referred to as equilibrium generated predictability (EGP). The idea, loosely, is to exploit the dynamic impulse response shown in Figure 1 by imposing that shocks to risk, as measured through their contemporaneous price impact, dissipate at the same rate in the stock return data as in the risk variable data. I derive a simple algebraic relationship between the contemporaneous return–risk shocks and autocorrelation in the risk variable that imposes a strong theoretical restriction on coefficients in linear regression return forecasts. While my baseline test assumes that the risk variable follows a Gaussian AR(1) (as is common in the price ratio literature; see Stambaugh (1999), Campbell and Yogo (2006)), I derive two important generalizations: multivariate tests using multiple predictors and autoregressive moving average (ARMA) process specifications for the predictor. In the AR(1) case, I derive an analytic expression for the covariance matrix of prediction slopes at different forecast horizons. This is used to construct a joint test of the hypothesis that predictability at all horizons is consistent with equilibrium.

I apply the various tests of EGP to variables that have been shown to predict (and not predict) stock returns. In particular, I study the variance risk premium (VRP) of Bollerslev, Tauchen, and Zhou (2009). VRP predicts returns with high R^2s at the 3- to 5-month horizons. I also study option-implied variance (IV) and realized variance (RV) well as the Fear Index (FI) of Bollerslev, Todorov, and Xu (2015). The following empirical conclusions emerge:

https://doi.org/10.1017/S0022109024000218

- Tests performed on VRP reject the null of EGP. The shocks to VRP are too quickly mean-reverting to generate a hump in predictability at a 4-month horizon. Intuitively, it is not consistent with dynamic equilibrium to observe a return premium 4 months ahead generated by a shock to VRP that has dissipated by that time.
- Measures of conditional variance (IV and RV) strongly reject the null. Conditional variance measures ought to predict returns because shocks have a strong contemporaneous negative correlation with returns. But price shocks attributable to shocks in conditional variance measures appear to be permanent rather than transitory despite the fact that variance shocks clearly mean-revert.
- I cannot reject the null of EGP for Bollerslev et al.'s FI. Estimated reduced-form ordinary least-squares (OLS) predictability slopes are consistent with theory; however, the amount of predictability generated by FI is far below what is recorded in the original paper.

Bollerslev et al. (2009) and Drechsler and Yaron (2011) derive LRR models with the aim of explaining predictability R^2s that peak at the 4-month horizon. To do so, they use persistence parameters that are significantly larger (0.8 in both papers) than the first-order autocorrelation I estimate from VRP data.[1] This, in part, explains their favorable interpretation of the equilibrium story relative to my sharp statistical rejection. The tests also derive power from the null being imposed across multiple forecasting horizons simultaneously.

The tests suggested here bear similarities with tests of conditional CAPM-style models (as in, e.g., Harvey (1989), (1991)). This literature typically imposes equilibrium style restraints on linear forecasts of one1period1ahead expected returns. Here, I impose restrictions on the whole term structure of forecasts. I also do not require a specific equilibrium model. My approach is valid whether the equilibrium is generated by a long-run risk (LRR) economy, as in Bollerslev et al. (2009) and Drechsler and Yaron (2011), or a habit model, as in Bekaert, Engstrom, and Ermolov (2020).

I expand my baseline test to cover multivariate (vector autoregressive model (VAR)) state dynamics as well as ARMA dynamics. The VAR specifications still reject the null of EGP for pairs of VRP, RV, and IV.

While I am primarily looking to include ARMA dynamics to see if my baseline tests are robust, ARMA state dynamics imply that it is not optimal to forecast returns using only the state variable. Rather, under an ARMA(p,q) the econometrician should use all the ARMA components to forecast returns. I derive a test based on this principle and find that the null is still rejected consistently with what I find for the AR(1) case.

The tests I propose rely on the identification assumption that shocks to expected returns are uncorrelated with shocks to cash flows. This is assumed in virtually all of the present value-based equilibrium models.[2] I check that the main

[1]The VRP data used here, taken from Zhou (2018), have a monthly first-order autocorrelation of 0.28, with a standard error of 0.008.

[2]An incomplete list includes Bansal and Yaron (2004), Bansal, Kiku, and Yaron (2016), Bollerslev et al. (2009), Huang, Schlag, Shaliastovich, and Thimme (2019), Drechsler and Yaron (2011), and Eraker and Wu (2017).

empirical results are robust to this assumption by controlling for future earnings growth rates in the regression specifications. The test results are unaffected.

The tests are intended to work for predictors such as VRP that have been shown to predict returns at relatively short horizons. The tests do not apply to long-run prediction using price–dividend ratios, as in Fama and French (1988b), Stambaugh (1999), and many others. There are two reasons for this. First and most importantly, the tests assume that the predictor risk variables are exogenous. This precludes price–dividend ratios since the endogenous price enters the regressor. Second, option-implied or variance-related risk variables, which are the focus of this paper, are typically not persistent enough to imply that equilibrium expected returns vary much beyond a 1-year horizon.

In order to derive a test, I first consider a simple example where a single predictor follows an AR(1) process. This example has often been used in the literature on stock return predictability, including Stambaugh (1999), who shows that if returns are correlated with shocks to the predictor, the predictive slope coefficients are biased. Boudoukh, Richardson, and Whitelaw (2006) derive an analytic expression for the covariance matrix of slope coefficients under the assumption that the predictor follows an AR(1). They use the resulting estimator to construct a test of the joint null hypothesis that the predictive coefficients are zero at all horizons. In this paper, I also derive an analytic expression for the covariance matrix of the predictive slope coefficients under the AR(1) assumption, but I generalize the assumptions in Boudouk et al. to include correlated errors between the shocks to returns and the predictive variable. This allows me to construct an analytically based test of the EGP hypothesis.

The EGP restrictions can be seen as a set of nonlinear restrictions that map predictability implied by a structural model and an assumed VAR driving process for a set of state variables. A related topic is considered in the macroeconomic literature on forecasting, including Marcellino, Stock, and Watson (2006), which compare direct forecasts from period-by-period OLS regression versus multiperiod forecasts implied by AR models. A separate literature, starting with Jordà (2005), studies forecasts made from VAR or vector ARMA (VARMA) models to that of period-specific regressions, dubbed *local projections*. This paper is related to studies that impose parametric constraints on reduced-form forecasts relative to equilibrium-implied structural restrictions on VAR models for exogenous state variables. One such example is Zviadadze (2021), who compares generalized impulse response functions (IRFs) between reduced-form and equilibrium asset pricing models. Her evidence supports multiple shocks in the variance process of consumption consistent with the model in Drechsler and Yaron (2011).

The remainder of the paper is organized as follows: In Section II, I postulate a simple equilibrium relation between dividends, prices, and state variables and use this relationship to derive tests of EGP. Section III applies the test to sample data. Section IV extends the baseline test to cover the multivariate case and ARMA(p,q) dynamics for the state variable. Section V concludes.

II. A Test for EGP

In the following, I derive tests of whether -$h$period-ahead expected returns $E_t(r_{t+h})$ computed from a predictive regression on x_t is consistent with equilibrium.

To accomplish this, I derive a simple relationship between the predictor variable and the expected returns.

A. Prices and Dividends

In the following, I discuss the relationship between prices (P_t) and dividends (D_t) in dynamic equilibrium models. To maintain some analytical tractability, I will assume an economy where the price of a stock equals the present value of its future dividend payments, as

$$P_t = E_t \int_t^{\infty} D_s M_s / M_t ds = E_t^Q \int_t^{\infty} e^{-r_s s} D_s ds \tag{1}$$

where M_t is the pricing kernel and E^Q denotes the expectation under an equivalent Martingale measure, Q. I further assume that there exist a set of state variables $x_t \in \mathcal{D} \subseteq \mathbb{R}^n$. I will further assume that x_t is stationary and that the processes x_t and D_t satisfy sufficient regularity conditions such that equilibrium exists and that the economy is arbitrage-free. We then have

$$P_t = P(D_t, x_t) = D_t F(x_t). \tag{2}$$

for some function $F: \mathcal{D} \rightarrow \mathbb{R}_+$. The right-hand side of (2) states that the pricing function P is homogenous of degree 1 with respect to dividends. That is, $P(\lambda D_t, x_t) = P_t \lambda$, where λ is a scalar[3]. Taking logs gives

$$\ln P_t = \ln D_t + f(x_t) \tag{3}$$

such that $f(x_t)$ is the log-price–dividend ratio. It is clear that the (log) price–dividend ratio then will inherit the dynamic properties of x_t. In fact, equation (3) suggests that we can infer that a lot of short-term variation in prices will be almost entirely driven by variation in the (log) price–dividend ratio as prices are known to be much more volatile than dividends.

Indeed, in typical dynamic equilibrium models, the log-price–dividend ratio is approximately linear, as in the long-run-risk literature (Bansal and Yaron (2004), Bansal et al. (2016)), and others) although linearity is shown to fail over certain areas of the parameter space in Pohl, Schmedders, and Wilms (2018). Eraker and Wu (2017) derive a model with an exact linear log-price–dividend ratio.

In my empirical tests, I avoid using price–dividend ratios for two reasons. First, the tests are constructed by exploiting the contemporaneous correlation between returns and candidate risk variables, x_t, that are assumed to drive variation in expected returns. This requires x_t to be exogenous, which precludes the actual price–dividend ratio, which by construction depends on the price. Thus, the contemporaneous regressions would imply a regression of returns on returns. Second, the dividend discounting model is commonly used in endowment-based

[3]Homogeneity of degree 1 is a consequence of the law of one price: If $P_t = E_t \int_t^{\infty} D_s M_s / M_t ds$ is the price of an asset with dividend process D_s, then an asset that pays $D_s^* = \lambda D_s$ for all $s \in (t, \infty)$ has value $E_t \int_t^{\infty} \lambda D_s M_s / M_t ds = \lambda P_t$.

economies. Endowment economies are radical simplifications of actual corporate dividend policies. In actuality, corporations may or may not distribute dividends and many corporations choose not to as shareholders may prefer to plow earnings back into the corporations rather than disinvest through dividends. Dividend irrelevance (Miller and Modigliani (1961)) extends to dynamic economies if the growth in earnings equals the discount rate (expected returns) in a dividend discounting model with constant growth rates (Brennan (1971)). If firms pay a constant fraction of earnings as dividends, the price–earnings ratio is equivalent to the price–dividend ratio. Indeed, practitioners usually consider price–earnings ratios rather than price–dividend ratios as a heuristic in valuing a company.

To formulate a test of EGP, I assume that there is some exogenous state variable x_t that induces a mean-reverting component to prices. To derive the test in its simplest form, I first assume that the state variable follows an AR(1),

$$x_t = \rho x_{t-1} + w_t. \tag{4}$$

This seems restrictive but is consistent with much of the literature on long-term predictability of returns from price ratios, as, for example, Stambaugh (1999). I relax this assumption later to include multiple state variables and state variables that follow ARMA(p,q) processes.

In general, when x_t is an N dimensional process, consistent with long-run risk, habit formation, and other equilibrium models, I assume that the price–dividend ratio is log-linear,

$$\ln P_t = \ln D_t + \alpha + \beta_0 x_t, \tag{5}$$

where now x_t and β_0 are $K \times 1$ and $1 \times K$ dimensional vectors. This specification nests popular equilibrium models where multiple factors drive variation in price–dividend ratios. For the remainder of this section, I will assume that x_t is a scalar.

Since dividends contain a unit root, this equation implies that log prices and log dividends are cointegrated. The dynamics of dividends are not explicitly modeled. Equation (5) suggests that prices contain a temporary component driven by the risk variable x_t. This variable generates predictability in returns by temporarily moving the stock price away from its steady state path. To see how shocks to x_t generate time-varying expected rates of return, assume that log dividend growth rates are given by a random walk:

$$\ln D_{t+1} - \ln D_t = \mu + \varepsilon_{t+1}. \tag{6}$$

Appendix A discusses the case when expected dividend growth rates depend on x_t.

The dynamics of log capital gains follow

$$\ln P_{t+1} - \ln P_t = \mu + \beta_0 (x_{t+1} - x_t) + \varepsilon_{t+1}. \tag{7}$$

This equation suggests that an estimate of β_0 can be obtained through a regression of log capital gains onto $\Delta x_{t+1} = x_{t+1} - x_t$.

The -*h*period-ahead capital gain is given by

$$\ln P_{t+h} - \ln P_t = \mu h + \beta_0 (x_{t+h} - x_t) + \sum_{i=1}^{h} \varepsilon_{t+i} \tag{8}$$

The expected -*h*period capital gain can be found by taking expectations of (8). In the case that x follows an AR(1) with autocorrelation ρ, it is

$$E_t[\ln P_{t+h} - \ln P_t] = \mu h + \beta_0 \left(\rho^h - 1\right) x_t. \tag{9}$$

Equation (9) suggests that if one were to run the regression

$$\ln P_{t+h} - \ln P_t = \alpha_h + \beta_h x_t + u_t^h \tag{10}$$

the intercept and slope would have to satisfy

$$\alpha_h = \mu h, \tag{11}$$

$$\beta_h = \beta_0 \left(\rho^h - 1\right). \tag{12}$$

Notice that these equations impose testable restrictions. In particular, the entire term structure of predictable variation in returns is governed by the feedback coefficient β_0 and the autocorrelation, ρ, of the predictor variable.

B. A Joint Test

In this section, I derive a test for the joint hypothesis that the estimated OLS predictive coefficients are consistent with equilibrium. That is, if $\hat{b}_h$ is the OLS-estimated slope coefficient in the regression (10), I derive a test of the restriction in (12).

Hodrick (1992) observes regression coefficients from nonoverlapping data can be mapped to coefficients from cumulative overlapping returns. If we run a regression where the dependent variable is a 1-period return from $t+h-t$ to $t+h$,

$$r_{t+h} = a_h + b_h x_t + u_{t+h}^h, \tag{13}$$

it follows that

$$\beta_h = \frac{\mathrm{Cov}(r_{t:t+h}, x_t)}{\mathrm{Var}(x_t)} = \frac{\mathrm{Cov}\left(\sum_{j=1}^{h} r_{t+j}, x_t\right)}{\mathrm{Var}(x_t)} = \sum_{i=1}^{h} b_i. \tag{14}$$

Thus, a test of $\beta_h = \beta_h^*$ is equivalent to a test of $b_h = b_h^*$ for all $h = 1, .., N$. As in Hodrick (1992), I take advantage of this to reduce the residual serial correlation in the nonoverlapping return regression.

I propose a test statistic of the usual form

$$Q = \left(\hat{b} - b^*\right)' \Omega^{-1} \left(\hat{b} - b^*\right) \tag{15}$$

where $\hat{b}$ and b^* denote the N length vectors of estimated slopes, $\hat{b}_h$, and hypothesized slopes, $b^* = \beta_0\left(\rho^h - \rho^{h-1}\right)$. This expression is obtained as we are regressing r_{t+h} on x_t rather than $r_{t:t+h}$. Ω is the $N \times N$ covariance matrix of slope coefficients, $\mathrm{Cov}\left(b_j, b_h\right)$. The following result derives an asymptotic estimator for this covariance matrix.

Theorem 1. Assume that the return-generating process is given by

(16) $$r_t = \alpha_0 + \beta_0 \Delta x_t + \varepsilon_t$$

(17) $$x_t = \rho x_{t-1} + w_t$$

for known constants β_0, ρ, $\sigma_x^2 = Var(w), \sigma_x = Var(\varepsilon)$, and $Corr(w,\varepsilon) = 0$. Define

(18) $$b_h^* = \beta_0\left(\rho^h - \rho^{h-1}\right)$$

then

1. The OLS estimator $\hat{b}_h = \widehat{Cov}(r_{t+h}, x_t)/\widehat{Var}(x_t)$ is consistent for b_h^*, and the OLS estimator $\hat{\beta}_h = \widehat{Cov}(r_{t:t+h}, x_t)/\widehat{Var}(x_t)$ is consistent for $\beta_0\left(\rho^h - 1\right)$.
2. $\sqrt{T}\left(\hat{b} - b\right) \xrightarrow{d} N(0, \Omega(\beta_0, \rho, \sigma_x, \sigma_\varepsilon))$, where $\Omega = \mathrm{Cov}(b_h, b_l)$ is

(19) $$\begin{aligned}\Omega(\beta_0, \rho, \sigma_x, \sigma_\varepsilon) \approx c\beta_0^2 \sum_t \sum_s \Big(& \beta_0^2 [F(t,s,t+h,s+l) - F(t,s,t+h,s+l-1) \\ & - F(t,s,t+h-1,s+l) + F(t,s,t+h-1,s+l-1)] \\ & - \beta_0 b_l^* [F(t,s,s,t+h) - F(t,s,s,t-h-1)] \\ & - \beta_0 b_h^* [F(t,s,s,t+l) - F(t,s,s,t-l-1)] \\ & + b_h^* b_l^* F(t,t,s,s) + T\rho^{|l-h|}\sigma_x^2 \frac{\sigma_\varepsilon^2}{1-\rho^2} \Big)\end{aligned}$$

where

(20) $$c = E\left(\frac{1}{\left(\sum_t x_t^2\right)^2}\right)$$

and

(21) $$F(T_t, T_2, T_3, T_4) = G(t_1, t_2, t_3, t_4)$$

where t_i is the sorted T_j's $(i,j = 1,..4)$ and

(22) $$G(T_t, T_2, T_3, T_4) = E(x^2)\rho^{T_2 - T_1}\left(3\left[E(x^2)\rho^{2(T_2 - T_1)} + \sigma_{T_1:T_2}^2\right]\rho^{2(T_3 - T_2)} + \sigma_{T_2:T3}^2\right)\rho^{T_4 - T_3}$$

where $E(x^2) = \sigma_x^2/(1-\rho^2)$ and $T_1 \leq T_2 \leq T_3 \leq T_4$.

The covariance matrix in (19) depends on the parameters $\{\beta_0, \rho, \sigma_x, \sigma_\varepsilon\}$ and is computable from estimated values for these parameters. Thus, the test statistic

TABLE 1

Size of Q_N

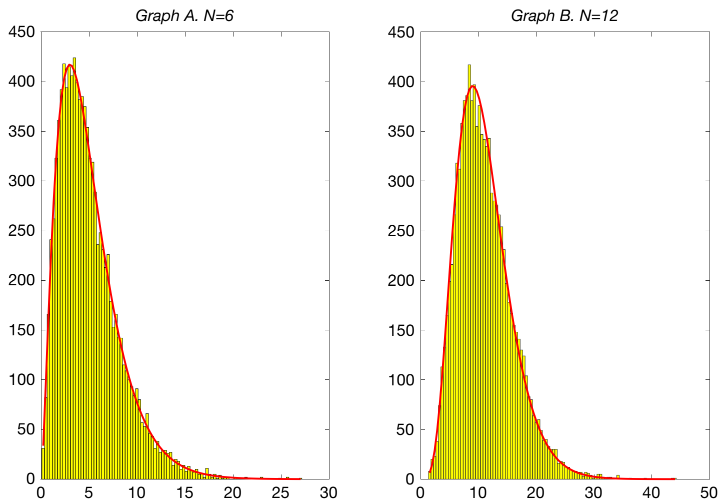
The table reports the rejection frequencies of the test of EGP for parameters $\rho = 0.25, \beta_0 = -0.762, \sigma_x = 19, \sigma_x = 3.65$, and $T = 348$ month observations. The results are based on 10,000 Monte Carlo draws.

	N	
Size	6	12
0.05	0.051	0.054
0.01	0.010	0.012

FIGURE 2

Distribution of Q_N

Figure 2 shows the distribution of the test statistic Q_N for $N = 6$ (Graph A) and $N = 12$ (Graph B) versus corresponding theoretical χ^2 densities.



$$Q_N = \left(\hat{b} - b^*\left(\hat{\beta}_0, \rho\right)\right)' \Omega^{-1}\left(\hat{b} - b^*\left(\hat{\beta}_0, \rho\right)\right) \quad (23)$$

is based on the feasible estimate of the true values $b_h^*\left(\hat{\beta}_0, \hat{\rho}\right) = \hat{\beta}_0\left(\hat{\rho}^h - \hat{\rho}^{h-1}\right)$ for $h = 1, .., N$.

I examine the performance of the test statistic Q_N, where N denotes the maximum forecasting horizon in Table 1 and Figure 2. Each simulation was based on $\rho = 0.25, \beta_0 = -0.762, \sigma_x = 19, \sigma_x = 3.65$, suggestive of monthly VRP estimates. Ω was estimated using the expression in (19) using OLS estimates of ρ, β_0, σ_x, and σ_x. The results show that for $T = 348$, the test is sized correctly since empirical rejection proportions deviate only mildly from the theoretical p-values under the theoretical χ-squared distributions.

https://doi.org/10.1017/S0022109024000218

C. *t*-Tests Based on Cumulative Returns

Most of the predictability literature is concerned with the estimation of β_hs—the slope coefficients obtained from regression predictions of cumulative returns $r_{t:t+h}$ on x_t. In order to conform with this standard, I derive the following:

Theorem 2. The covariance matrix of long-horizon regression predictability slopes β_h is related to the 1-period slopes b_h by

$$Cov(\beta_h, \beta_k) = \sum_{i=1}^{h} \sum_{j=1}^{k} Cov(b_i, b_j) \tag{24}$$

Theorem 2, whose proof is trivially a consequence of $\beta_h = \sum_{i=1}^{h} b_i$, is useful not only in tests for EGP but also in tests of predictability. The result can be applied to compute the covariance matrix of long-horizon coefficients β using any available estimator of 1-period slopes, $\text{Cov}(b_j, b_j)$. Boudoukh et al. (2006) propose a covariance matrix estimator for the predictive slope coefficients based on the assumptions that i) a scalar predictor follows a Gaussian AR(1); ii) the innovations to the AR(1) and the return shocks are uncorrelated; and iii) the null of no predictability holds. The estimator proposed here generalizes the last two assumptions.

In a well-known paper, Ang and Bekeart (2007) study the predictability of interest rates and dividend yields. Perhaps less well known, their paper (see Appendix B) also derives a generalized version of Hodrick's standard errors to include the full covariance matrix of regression slopes for different horizon overlapping regressions. Their estimator is based on the generalized method of moments (GMM) and, as such, is not based on the assumption that the predictor follows an AR(1) or uncorrelated innovations.

The estimator proposed here can also be generalized to relax the AR(1) assumption. One natural way to proceed is to obtain an estimate of $\text{Cov}(b_i, b_j)$ and then apply equation (24) to obtain the covariance matrix for the coefficients in the overlapping return regressions, $\text{Cov}(\beta_h, \beta_k)$. Obtaining an estimate of $\text{Cov}(b_i, b_j)$ is potentially easier than $\text{Cov}(\beta_h, \beta_k)$ because one can exploit the independence of returns.

III. Empirical Analysis

A. Descriptive Data

Table 2 presents descriptive statistics for the predictors/risk candidate variables. The most noteworthy part of the descriptive evidence is the difference in autocorrelation, both between variables and across lag lengths. First off, VRP exhibits relatively low first-order autocorrelation at 0.28. The most persistent series is IV, with a first-order autocorrelation of 0.81. Higher-order autocorrelations are large relative to the first order: For VRP, this is obvious, as the autocorrelation at 9 months is 0.20, the same order of magnitude as for one month.

TABLE 2

Descriptive Statistics

Table 2 shows the mean, standard deviation, and autocorrelations for VRP, IV, RV, and TI.

	VRP	IV	RV	TI
N	348	348	348	277
Mean	15.80	35.68	19.88	6.48
Std	20.19	33.07	36.28	2.57
Lag	Autocorrelations			
1	0.28	0.81	0.64	0.61
2	0.26	0.61	0.40	0.48
3	0.16	0.52	0.31	0.43
4	0.04	0.47	0.26	0.37
5	0.18	0.41	0.25	0.25
6	0.12	0.33	0.18	0.21
7	0.15	0.30	0.15	0.20
8	0.17	0.30	0.13	0.18
9	0.20	0.28	0.15	0.17
10	0.15	0.29	0.11	0.12
11	0.17	0.28	0.10	0.09

Let us consider what the long-range autocorrelations would be if the variables followed an AR(1). The 12-month autocorrelation under an AR(1) equals the 1-month autocorrelation raised to the power of 12. For VRP, we should then see $0.26^{12} = 0$ at the 12-month horizon. For IV, the number is about 0.08, and for RV and Tail Index (TI), we should see numbers less than 0.01. These are back-of-the-envelope computations that suggest that AR(1) processes may not capture long range dynamics of the time series.

B. The VRP

Since Bollerslev et al. (2009) (hereby BTZ), a substantial literature has emerged on the VRP. BTZ show that the VRP predicts stock returns with 1.07% and 6.82% R^2s at 1- and 3-month forecasting horizon and with a decreasing term structure after that. In this respect, the pattern of predictability differs from, for example, P/Dratios, which show monotonically increasing R^2s. Others have found even higher R^2s. For example, Bekaert and Hoerova (2014) report an R^2 of 13% for the quarterly forecasting horizon.

Before proceeding, note that the operational definition of the VRP matters for the empirical results to follow. I start by replicating the results in BTZ, and I, therefore, define the VRP as the difference between IV, as measured by the squared VIX, and 30-day backward-looking RV. I use publicly available data described in Zhou (2018) consisting of 348 monthly observations from Jan. 1990 to Dec. 2018. Following BTZ, the VRP is defined as

$$VRP_t = IV_t - RV_t$$

where IV_t is the squared VIX and RV_t computed from 5-minute log returns collected over the last month prior to date t.

Table 3 confirms what is well documented in the literature: VRP predicts returns with R^2s peaking at 11% for the 4-month forecasting horizon. I use three different methods for computing standard errors—Hodrick (1992) type A (see their

TABLE 3

Predictability of VRP

Table 3 reports the results of predictability regressions with cumulative returns, $r_{t:t+h} = \alpha_h + \beta_h VRP_t + u_{t,h}$. The table reports standard errors and t-statistics for the null of $\beta_h = 0$ using Hodrick (se-H and t-H), Newey–West (se-NW and t-NW), and analytical (se-A and t-A) using the estimator in (19) and (24). Parameters and standard errors are scaled by 100.

h	1	2	3	4	5	6	7	8	9	10	11	12
R^2	0.05	0.06	0.10	0.11	0.09	0.06	0.05	0.04	0.03	0.03	0.02	0.02
b_h	4.69	2.84	3.99	3.06	0.17	−0.84	−0.86	−0.07	−0.96	0.01	−0.15	0.67
β_h	4.69	7.44	11.33	14.30	14.38	13.47	12.58	12.46	11.39	11.34	11.01	11.60
se-H	(1.88)	(2.34)	(3.30)	(4.09)	(4.50)	(4.91)	(5.29)	(5.64)	(6.04)	(6.64)	(7.02)	(7.51)
se-NW	(0.76)	(1.15)	(1.41)	(1.99)	(2.24)	(2.80)	(3.37)	(3.70)	(4.31)	(4.81)	(5.37)	(5.53)
se-A	(1.06)	(1.68)	(2.17)	(2.58)	(2.93)	(3.25)	(3.54)	(3.80)	(4.05)	(4.29)	(4.51)	(4.72)
t-H	2.50	3.18	3.44	3.50	3.19	2.74	2.38	2.21	1.89	1.71	1.57	1.55
t-NW	6.14	6.49	8.01	7.19	6.42	4.81	3.74	3.36	2.64	2.36	2.05	2.10
t-A	4.43	4.42	5.21	5.54	4.90	4.14	3.56	3.27	2.81	2.65	2.44	2.46

p. 362) using 12 lags, Newey and West (1987) using 12 lags, and the analytical standard errors described in the previous section. Hodrick's method generally provides the more conservative standard errors, rendering coefficients estimated for horizons longer than 8 months insignificant. All three methods produce standard errors that increase monotonically with horizon, which is to be expected given equation (24). Note that the long-horizon predictability vanishes as the 1-period, marginal coefficients b_h turn negative for $h = 6 - 9$[4].

Table 4 presents the results from the EGP tests. First off, the estimated parameters for the data-generating process (DGP) for VRP and returns yield $\rho = 0.28$ and $\beta_0 = -0.03$. The former suggests that VRP is quickly mean-reverting, while the latter suggests that the equilibrium response to VRP shocks is negative, leading to a positive shock to expected returns. The estimated 1-period coefficients, b_h, are positive up until and including the 5-month horizon. This leads to a monotonic increase in the β_h's for $h = 1$ through 5. The theoretical coefficients b_h^* and β_h^* are smaller in magnitude than their OLS counterparts, but most importantly, b_h decays toward zero quickly. At 1 month, $b_1 - b_1^* = 2.44$ with an associated t-statistic of 2.3. At $h = 3$ months, the gap increases to 3.81, and at $h = 4$ months, the gap is 3.00. These differences are statistically significant and large. In essence, the low autocorrelation of VRP means that any shock to VRP will die out quickly enough that it will not warrant a risk premium of any magnitude beyond a 1- or 2-month horizon. This is particularly evident at the $h = 4$ month horizon: Here, predictability of cumulative returns is at its highest with 11% R^2, and the gap between β_h and β_h^* is at its highest, leading to a t-statistic of 4.33.

Table 4 also reports the multivariate tests of the null: $b_h = b_h^*, h = 1,..,N$ for $N = 6$ and 12. Both reject the null. It is clear that the $N = 6$ case presents a stronger case against the null hypothesis than the $N = 12$ case because the evidence of

[4]Notice that the OLS estimates of b_h do not sum to β_h exactly, as suggested in (14), which holds in population. In order to impose that the sum would hold for small samples, we must run regression (13) with observations $t = 1,..,T - N$ for any $h \leq N$, where N is the maximum forecast horizon. This trims the sample length for $h < H$. Since the main focus here is on the b_h coefficients, which can be estimated with $T - h$ observations, I choose to use longer samples rather than truncating all regression sample sizes to $T - N$.

TABLE 4

Test of EGP: VRP

Table 4 reports the results of tests of EGP using the VRP. The theoretical slope is $b_h^* = \beta_0(\rho^h - \rho^{h-1})$ for 1-period returns and $\beta_h = \beta_0(\rho^h - 1)$ for overlapping returns. The VRP is defined as the difference between 1-month IV and 1-month backward-looking RV. t-statistics correspond to the null that EGP holds at horizon h using the covariance matrix in (19) and (24).

					Parameter Estimates							
					ρ	β_0	σ_ε	σ_x				
					0.28	−0.03	4.04	19.37				
h	1	2	3	4	5	6	7	8	9	10	11	12
One-Period Returns												
b_h	4.69	2.84	3.99	3.06	0.17	−0.84	−0.86	−0.07	−0.96	0.01	−0.15	0.67
b_h^*	2.25	0.64	0.18	0.05	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$b_h - b_h^*$	2.44	2.21	3.81	3.00	0.16	−0.84	−0.87	−0.07	−0.96	0.01	−0.15	0.67
t–stat	2.30	2.08	3.59	2.83	0.15	−0.79	−0.82	−0.07	−0.90	0.01	−0.14	0.63
Multiperiod Returns												
β_h	4.69	7.44	11.33	14.30	14.38	13.47	12.58	12.46	11.39	11.34	11.01	11.60
β_h^*	2.25	2.89	3.07	3.12	3.13	3.14	3.14	3.14	3.14	3.14	3.14	3.14
$\beta_h - \beta_h^*$	2.44	4.55	8.26	11.18	11.25	10.33	9.44	9.32	8.25	8.20	7.87	8.46
t–stat	2.30	2.70	3.80	4.33	3.83	3.18	2.67	2.45	2.04	1.91	1.74	1.79
					Multivariate Tests							
					N	6	12					
					Q_N	21.19	22.95					
					p–val	0.00	0.02					

predictability weakens with longer horizons. In general, the "structural model" would imply no additional power beyond the 5- to 6-month horizons as the b_h^* is essentially zero.

C. Implied and RV

The VIX index is computed from S&P 500 cash index options by the Chicago Board Options Exchange (CBOE). In theory, the square of the VIX index represents a 1-month forward-looking option-implied estimate of the *risk-neutral* variance of the logarithmic return for the underlying S&P 500 index. There are strong theoretical reasons to think that the VIX index contains information about expected returns. In 1-factor models based on dynamic present value computation, including Bansal and Yaron (2004), a single economy-wide volatility factor drives expected excess returns. One-factor models also imply that the VRP is proportional to the volatility factor. This again means that risk-neutral and objective conditional variance are both scaled versions of the same underlying macro-factor and therefore work equally well in predicting returns. Multifactor models of conditional variance also imply that Q expected variance predicts returns. For example, in BTZ's model, objective measure conditional variance (P variance) is a strong predictor of return. In their model, the Q variance equals the P variance plus the VRP, which again depends on a separate volatility-of-volatility factor. Risk-neutral variance is a linear combination of these two factors.

Beyond the model-based theoretical justification, it is also clear that option traders look forward to known future events that can cause volatility. For example,

TABLE 5

Test of EGP: IV

Table 5 reports the results of predictability regressions and tests of EGP using IV. The equilibrium restriction is $b_h^* = \beta_0(\rho^h - \rho^{h-1})$ and $\beta_h = \beta_0(\rho^h - 1)$. t-statistics correspond to the null that EGP holds at horizon h using the covariance matrix in (19) and (24).

					Parameter Estimates							
					ρ	β_0	σ_ε	σ_x				
					0.81	−0.13	3.08	19.51				
h	1	2	3	4	5	6	7	8	9	10	11	12
Panel A. Predictability												
R^2	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.01	0.00	0.00
β_h	−0.13	0.47	0.46	0.78	1.78	2.88	3.26	3.19	3.35	3.21	3.10	3.14
se–A	(0.55)	(0.98)	(1.33)	(1.63)	(1.88)	(2.11)	(2.31)	(2.50)	(2.67)	(2.83)	(2.98)	(3.13)
t–A	−0.24	0.48	0.35	0.48	0.94	1.37	1.41	1.28	1.25	1.13	1.04	1.01
Panel B. One-Period Returns												
b_h	−0.13	0.63	−0.02	0.42	1.10	1.19	0.41	−0.02	0.20	−0.13	−0.10	0.15
b_h^*	2.56	2.07	1.67	1.35	1.09	0.88	0.71	0.57	0.46	0.38	0.30	0.24
$b_h - b_h^*$	−2.69	−1.44	−1.68	−0.93	0.01	0.31	−0.30	−0.60	−0.26	−0.50	−0.40	−0.10
se–A	(0.55)	(0.49)	(0.47)	(0.46)	(0.45)	(0.45)	(0.46)	(0.46)	(0.46)	(0.46)	(0.46)	(0.46)
t-stat	−4.91	−2.91	−3.60	−2.04	0.02	0.69	−0.65	−1.30	−0.57	−1.09	−0.86	−0.21
Panel C. Multiperiod Returns												
β_h^*	2.56	4.62	6.29	7.64	8.73	9.61	10.32	10.90	11.36	11.74	12.04	12.29
$\beta_h - \beta_h^*$	−2.69	−4.15	−5.83	−6.87	−6.96	−6.73	−7.07	−7.71	−8.02	−8.53	−8.94	−9.14
t-stat	−4.91	−4.26	−4.39	−4.22	−3.70	−3.19	−3.06	−3.09	−3.00	−3.01	−3.00	−2.92

Multivariate Tests		
N	6	12
Q_N	23.49	27.61
p–val	0.00	0.01

Federal Reserve (Federal Open Market Committee (FOMC)) meetings are known to move prices. Naturally, options whose maturity window contains an FOMC meeting should be more expensive than those that do not. Empirical evidence is mixed on the extent to which implied volatility predicts future volatility and what contribution it contains relative to physical volatility. Canina and Figlewski (1993) conclude that IV has no informational content over physical volatility, while Christensen and Prabhala (1998) reach the opposite conclusion.

The results for IV are reported in Table 5. The first thing to note is that there is no predictability as R^2s for cumulative returns never top 1% at any horizon. IV does, however, have a strong contemporaneous negative correlation with returns (−0.66), reflected in $\beta_0 = -0.13$. This suggests that IV *should* predict returns: b_h^* goes from 2.57 to 1.35 for $h = 1$- to 4-month forecasting horizon. This of course leads to a rejection of the null of EGP, as seen in all t-tests for $h = 1,..,4$ and also in the multivariate tests.

The fact that volatility, in this case as measured by implied volatility, does not predict returns is a puzzle. As it is, prices respond negatively to positive volatility shocks, but they do not subsequently revert back as volatility dissipates. This is not only true for IV, but even more so for RV. Table 6 presents the results for RV. Here, the results are even worse than for IV, as the b_h coefficients are negative at horizons $h = 1,..,4$. Their magnitudes are about the same as the ones predicted by EGP,

TABLE 6

Test of EGP: RV

Table 6 reports the results of predictability regressions and tests of EGP using RV (IV). The equilibrium restriction is $b_h^* = \beta_0(\rho^h - \rho^{h-1})$ and $\beta_h = \beta_0(\rho^h - 1)$. *t*-statistics correspond to the null that EGP holds at horizon *h* using the covariance matrix in (19) and (24).

					Parameter Estimates							
					ρ	β_0	σ_ε	σ_X				
					0.64	−0.04	3.93	27.83				
h	1	2	3	4	5	6	7	8	9	10	11	12
Panel A. Predictability												
R^2	0.02	0.01	0.02	0.03	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
β_h	−1.56	−1.91	−3.12	−3.77	−2.97	−1.76	−1.18	−1.20	−0.73	−0.83	−0.81	−0.96
	(0.49)	(0.87)	(1.20)	(1.49)	(1.75)	(1.99)	(2.21)	(2.41)	(2.60)	(2.77)	(2.94)	(3.10)
t–Me	−3.20	−2.20	−2.61	−2.54	−1.70	−0.89	−0.54	−0.50	−0.28	−0.30	−0.27	−0.31
Panel B. One-Period Returns												
b_h	−1.56	−0.36	−1.25	−0.59	0.86	1.25	0.61	0.00	0.46	−0.11	−0.03	−0.09
b_h^*	1.43	0.92	0.59	0.38	0.25	0.16	0.10	0.07	0.04	0.03	0.02	0.01
$b_h - b_h^*$	−2.99	−1.28	−1.84	−0.97	0.62	1.09	0.51	−0.06	0.42	−0.14	−0.05	−0.10
se–A	(0.49)	(0.48)	(0.48)	(0.48)	(0.48)	(0.48)	(0.48)	(0.48)	(0.48)	(0.48)	(0.48)	(0.48)
t–stat	−6.12	−2.67	−3.85	−2.04	1.29	2.28	1.06	−0.13	0.88	−0.28	−0.11	−0.20
Panel C. Multiperiod Returns												
β_h	−1.56	−1.91	−3.12	−3.77	−2.97	−1.76	−1.18	−1.20	−0.73	−0.83	−0.81	−0.96
β_h^*	1.43	2.34	2.93	3.31	3.56	3.72	3.82	3.88	3.93	3.95	3.97	3.98
$\beta_h - \beta_h^*$	−2.99	−4.25	−6.05	−7.09	−6.53	−5.48	−5.00	−5.09	−4.66	−4.79	−4.78	−4.94
t–stat	−6.12	−4.90	−5.06	−4.76	−3.73	−2.76	−2.27	−2.11	−1.79	−1.73	−1.62	−1.59
					Multivariate Tests							
					N	6	12					
					Q_N	39.24	41.76					
					p–val	0.00	0.00					

except again that the sign is opposite of what they ought to be. The difference, therefore, is large and statistically significant.

A similar conclusion holds for a longer sample. Specifically, I constructed monthly RV data from daily S&P 500 returns from 1927 to 2020 and reran the analysis. While I omit the details, the β_h coefficients were marginally negative and statistically insignificantly different from zero with R^2s never exceeding 1%. Overall, the Q_n statistics reject the null with p-values of 0.02 and 0.04 for $N = 6$ and $N = 12$, respectively.

The failure of conditional volatility measures in predicting returns is of course well known (Merton (1980), French, Schwert, and Stambaugh (1987), among others), which is puzzling, as investors appear to get a smaller risk–reward during volatile periods. In fact, if investors were to decrease their market exposure during high volatility periods, they would have earned a larger risk premium in the last almost 100-year-long sample from 1927 to 2020 (see Moreira and Muir (2017)).

D. TI

Bollerslev and Todorov (2011) argue that the VRP can be decomposed into two components for continuous, diffusive shocks and jumps, respectively. They argue that a left TI can be derived from options data alone by studying the rate of decay of far out-of-the-money (OTM) put options as the maturities of the options

https://doi.org/10.1017/S0022109024000218

https://doi.org/10.1017/S0022109024000218

TABLE 7

Test of EGP: TI

Table 7 reports the results of predictability regressions and tests of EGP using the TI. The tests of EGP assess whether estimated slopes b_h and β_h are significantly different from $b_h^* = \beta_0 (\rho^h - \rho^{h-1})$ and $\beta_h^* = \beta_0 (\rho^h - 1)$. t-statistics and Q tests are constructed from the covariance matrix in (19) and (24).

Parameter Estimates			
ρ	β_0	σ_ε	σ_x
0.61	−0.44	4.23	2.02

h	1	2	3	4	5	6	7	8	9	10	11	12
Panel A. Predictability												
R^2	0.01	0.01	0.01	0.02	0.03	0.03	0.02	0.03	0.03	0.04	0.04	0.04
β_h	20.32	22.00	30.71	51.82	71.51	77.76	76.37	89.44	99.07	114.98	123.27	135.27
	(0.49)	(0.87)	(1.20)	(1.49)	(1.75)	(1.99)	(2.21)	(2.41)	(2.60)	(2.77)	(2.94)	(3.10)
t–A	2.41	1.47	1.49	2.03	2.39	2.29	2.03	2.18	2.24	2.44	2.47	2.57
Panel B. One-Period Returns												
b_h	20.32	2.17	8.73	21.19	19.88	6.64	−0.80	13.20	9.81	15.29	7.79	11.33
b_h^*	17.16	10.40	6.30	3.82	2.31	1.40	0.85	0.52	0.31	0.19	0.11	0.07
$b_h - b_h^*$	3.17	−8.23	2.43	17.37	17.57	5.24	−1.65	12.68	9.49	15.10	7.68	11.26
se–A	(8.45)	(8.36)	(8.36)	(8.37)	(8.38)	(8.38)	(8.38)	(8.38)	(8.38)	(8.38)	(8.38)	(8.38)
t–stat	0.37	−0.98	0.29	2.08	2.10	0.63	−0.20	1.51	1.13	1.80	0.92	1.34
Panel C. Multiperiod Returns												
β_h^*	17.16	27.56	33.86	37.67	39.99	41.39	42.24	42.76	43.07	43.26	43.37	43.44
$\beta_h - \beta_h^*$	3.17	−5.55	−3.15	14.15	31.52	36.37	34.13	46.68	56.00	71.72	79.89	91.83
t–stat	0.37	−0.37	−0.15	0.55	1.05	1.07	0.91	1.14	1.27	1.52	1.60	1.75

Multivariate Tests		
N	6	12
Q_N	6.94	12.54
p–val	0.23	0.32

shrink. Bollerslev et al. (BTX) (2015) refine this idea and compute left and right jump premium measures. They label the difference a TI.

Table 7 presents the results for the TI. As seen, the index delivers a modest amount of predictability with R^2s ranging from 1 to 4% from the 1- to 12-month forecasting horizon. This is much lower than what is reported in BTX where R^2s peak at almost 15% (see BTX, Figure 6, p. 129). The primary driver of the difference appears to be the sampling period: BTX's sample period ends in Aug. 2013, while this paper uses the extended sample through Dec. 2019.

While predictability from TI is significantly lower in recent periods, the good news is that the small amounts that there are (in-sample) appear to be consistent with equilibrium: Table 7 does not reject the EGP null even at the 10% level. None of the t-tests based on the β coefficients reject the null and neither of the multivariate tests does either.

E. Stambaugh Bias

Stambaugh (1999) shows that OLS-estimated slopes in the forecasting regression (10) are biased. The bias is a function of the correlation between the shocks to the predictor and returns are correlated as well as the persistence in the predictor. Specifically, Stambaugh shows that the bias in the estimated forecasting slopes equals the bias in the AR1 parameter ρ multiplied by population regression slope in a regression of u^h on w,

(25)
$$E\left(\hat{b}_1 - b_1^*\right) = \frac{\sigma_{w,u^1}}{\sigma_w^2} E(\hat{\rho} - \rho),$$

where $u^h = \{u_{t+h}^h\}$ is the disturbances from the forecasting regression (13) $w = \{w_t\}$ is the innovations in the AR(1) for x in equation (4). Since $\sigma_{w,u^h}/\sigma_w^2$ can take on a wide range of values, the Stambaugh bias can be large even if the bias in $\hat{\rho}$ is modest.

The primary question of interest here is the extent to which Stambaugh bias impacts the tests of EGP. The Stambaugh bias is the difference between the average realization of the estimated slope coefficient, $E(b_1)$, and the true value $b_1 = \beta_0(\rho - 1)$. But since I do not observe the true values of β_0 and ρ, the empirical tests I conduct are based on the corresponding feasible estimate constructed from the estimates $\hat{\beta}_0$ and $\hat{\rho}$,

(26)
$$E\left(\hat{b}_h - \hat{\beta}_0\left(\hat{\rho}_1^h - \hat{\rho}_1^{h-1}\right)\right)$$

which is the quantity used in empirical tests of EGP. An interesting result obtains if we assume that β_0 is known.

Theorem 3. If β_0 is known the bias,

(27)
$$E\left(\hat{b}_h - \beta_0\left(\hat{\rho}^h - \hat{\rho}^{h-1}\right)\right) = \beta_0 E\left(\hat{\rho}_h - \hat{\rho}_{h-1} - \left(\hat{\rho}_1^h - \hat{\rho}_1^{h-1}\right)\right).$$

where $\rho_h = \hat{\text{C}}\text{orr}(x_{t+h}, x_t)$ is the h'th-order sample autocorrelation of x.

The bias is *0 for* $h = 1$.

Theorem 3 states that the bias in the estimated slope coefficient b_h is exactly canceled by the bias in the estimated AR1 coefficient ρ for $h = 1$. This implies that if β_0 were known we would have an unbiased test statistic for $h = 1$ even if the Stambaugh bias in the estimated slope $\hat{b}_1$ is substantial. While the result assumes that β_0 is known, in practice β_0 can be estimated with high precision because it uses contemporaneous returns and innovations in x, which are both near independent. It is also possible to construct an empirical test where the bias exactly cancels for all terms. This can be done by setting the estimate of the true value equal to $\beta_0(\hat{\rho}_h - \hat{\rho}_{h-1})$ (see the proof of Theorem 3 in Appendix C), thus replacing powers of the 1-period sample autocorrelation $\hat{\rho}$ with the autocorrelation estimated for each$h$. This has the disadvantage of requiring $N - 1$ additional autocorrelation coefficients. In multivariate VAR with M variables, the number increases by $(N - 1) \times M$, leading to a curse of dimensionality. In practice, therefore, as the bias is typically negligible, I use the test based on $\hat{\rho}^h$.

Since Theorem 3 shows that there is bias for $h > 1$, I investigate the size of the bias numerically. Table 8 reports results from simulation experiments where I draw data under the null of EGP for various parameters. I report the Stambaugh bias, as well as the deviations between the average slopes and their theoretical counterparts using OLS estimates of ρ and β_0, labeled B1 and B2, respectively. As Theorem 1 suggests, B1 is 0 for $h = 1$. B2 uses $\hat{\beta}$, and as such, it relaxes the assumption that β_0 is known.

TABLE 8

Stambaugh Bias and EGP Bias

Table 8 reports experimental results where 340 time-series observations are simulated under the null that EGP holds. I report the Stambaugh bias (SB): $E\left(\hat{b}_h - \beta_0\left(\rho^h - \rho^{h-1}\right)\right)$ as well B1 $E\left(\hat{b}_h - \beta_0\left(\hat{\rho}^h - \hat{\rho}^{h-1}\right)\right)$ and B2 $E\left(\hat{b}_h - \hat{\beta}_0\left(\hat{\rho}^h - \hat{\rho}^{h-1}\right)\right)$. B1 measures the estimated deviation between the theoretical and estimated slopes under the idealized condition that β_0 is known. B2 assumes β_0 unknown and estimated by OLS.

					ρ	β_0	σ_ε	σ_x				
					0.28	−0.03	4.04	19.37				
SB	0.00	0.00	0.00	0.00	−0.00	−0.00	−0.00	0.00	0.00	−0.00	0.00	−0.00
B1	−0.00	0.00	0.00	−0.00	−0.00	−0.00	−0.00	0.00	0.00	−0.00	0.00	−0.00
B2	0.00	0.00	0.00	−0.00	−0.00	−0.00	−0.00	0.00	0.00	−0.00	0.00	−0.00
					ρ	β_0	σ_ε	σ_x				
					0.81	−0.13	3.08	19.51				
SB	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	−0.00
B1	−0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
B2	−0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
					ρ	β_0	σ_ε	σ_x				
					0.64	−0.04	3.93	27.83				
SB	0.00	0.00	0.00	0.00	0.00	0.00	−0.00	−0.00	0.00	0.00	0.00	−0.00
B1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	−0.00
B2	−0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	−0.00
					ρ	β_0	σ_ε	σ_x				
					0.61	−0.44	4.23	2.02				
SB	0.00	0.00	0.00	0.00	0.00	−0.00	−0.00	−0.00	−0.00	0.00	−0.00	−0.00
B1	−0.00	0.00	0.00	0.00	0.00	0.00	−0.00	−0.00	−0.00	0.00	−0.00	−0.00
B2	−0.00	0.00	0.00	0.00	0.00	0.00	−0.00	−0.00	−0.00	0.00	−0.00	−0.00
					ρ	β_0	σ_ε	σ_x				
					0.9	−2	4	2				
SB	0.02	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.00	0.00	0.00	0.00
B1	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
B2	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
					ρ	β_0	σ_ε	σ_x				
					0.97	−2	4	20				
SB	0.05	0.05	0.04	0.04	0.04	0.04	0.03	0.03	0.03	0.03	0.03	0.03
B1	−0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
B2	−0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01

Entries 1–4 in Table 8 report the experimental results using parameter estimates reported in Tables 4 through 7 and thus serve to investigate the impact of Stambaugh bias on the empirical tests for EGP. The table shows that this impact is zero. There is no measurable Stambaugh bias in the reported results. B2 is also zero across all entries, suggesting that the reported empirical results in Tables 4 through 7 are not affected by Stambaugh bias. The last two entries in Table 8 use larger values of ρ and β_0 so as to generate more predictability and potentially therefore larger bias. As seen, increasing ρ to 0.9 and 0.97, respectively, does generate a nonzero Stambaugh bias. Interestingly, the bias in B1 and B2 is still zero for small h and marginally positive for larger h. The $h=1$ case numerically confirms Theorem 3 as the estimated bias $\hat{E}\left(\hat{b}_h - \beta_0\left(\hat{\rho}^h - \hat{\rho}^{h-1}\right)\right)$ is indistinguishable from zero.

Note here that the last two entries generate very high correlations between returns and innovations in the state variable. It is hard to construct an example in which the parameters represent reasonable DGPs and a nontrivial Stambaugh bias. In the end, the numerical results presented here suggest that EGP tests not only mitigate the bias in estimated slopes as suggested by Theorem 3 but also that the Stambaugh bias is negligible for the predictive variables studied in this paper.

https://doi.org/10.1017/S0022109024000218

https://doi.org/10.1017/S0022109024000218

F. Relaxing the Zero-Correlation Assumption

In the following, I consider the implications of correlated shocks in expected returns and cash flows. To see what the impact of this correlation is, I consider an example model where dividend shocks are correlated with shocks to a stochastic volatility factor. Specifically, I assume an example economy where consumption and consumption volatility are described by

$$\ln \frac{C_{t+1}}{C_t} = \mu + \sigma_t v_{t+1} \tag{28}$$

$$\sigma_{t+1}^2 = \sigma + \kappa\left(\sigma_t^2 - \sigma^2\right) + \sigma_t \sigma_w w_{t+1} \tag{29}$$

where $\mathrm{Corr}(v_{t+1}, w_{t+1}) = \delta$ is the correlation between shocks to consumption growth and its conditional variance, σ_t^2. One can solve this model easily with the usual long-run-risk framework. The linearized solution to the equilibrium price of an asset that pays aggregate consumption as its dividend is given by

$$\ln P_t = \ln C_t + A_o + A_\sigma \sigma_{t+1}^2 \tag{30}$$

where the A_σ is given by

$$A_\sigma = \frac{1 - \kappa k_1 - (1-\gamma) k_1 \sigma_w \delta - \sqrt{(\kappa k_1 - 1 + (1-\gamma) k_1 \sigma_w \delta)^2 - 2 k_1^2 \sigma_w^2 (1-\gamma)^2 / \left(1 - \frac{1}{\psi}\right)}}{2 k_1^2 \sigma_w^2 \theta} \tag{31}$$

To analyze the impact of a shock to conditional variance, consider equation (30). Before, I argued that we can identify the factor loading of capital gains or returns onto the risk variable through a regression of capital gains or returns onto changes in the risk variable, as in (10). The equivalent regression here would then be

$$\Delta \ln P_{t+1} = A + \beta_0 \Delta \sigma_{t+1}^2 + \varepsilon_{t+1} \tag{32}$$

where $\varepsilon_{t+1} = \sigma_t v_{t+1}$ is the error terms in the regression. These are interpretable as the demeaned shocks to consumption growth, and by assumption, they are correlated with shocks to the regressor. For this reason, an OLS estimate of β_0 in (32) is inconsistent for A_σ. At the same time, it is easy to verify that expected log capital gains are given by $E_t(\ln P_{t+i} - \ln P_{t+i-1}) = \mu + A_\sigma(\kappa^i - \kappa^{i-1})$, analogously to the Bansal–Yaron (BY) model example above. Thus, tests based on simple OLS regressions, as in (7) or (32), fail.

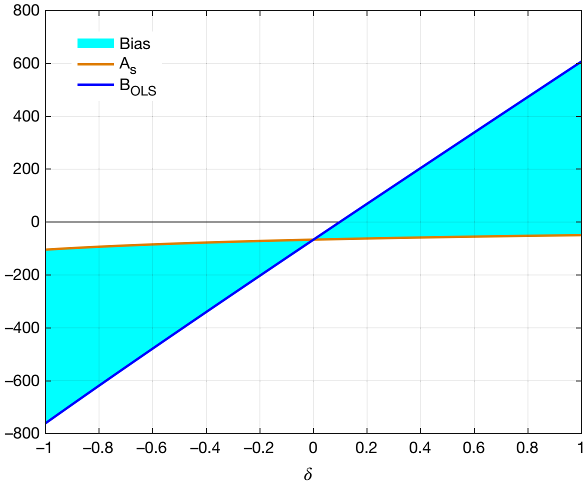
To gauge the bias in the OLS-estimated β, note that it is given by

$$\begin{aligned} \beta_{OLS} &= \frac{\mathrm{Cov}\left(\Delta \ln P_{t+1}, \Delta \sigma_{t+1}^2\right)}{\mathrm{Var}\left(\Delta \sigma_{t+1}^2\right)} = A_\sigma + \frac{\mathrm{Cov}(\sigma_t v_{t+1}, \Delta \sigma_{t+1})}{\mathrm{Var}\left(\Delta \sigma_{t+1}^2\right)} \\ &= A_\sigma + \frac{\delta}{\left[\frac{(1-\kappa)^2}{1-\kappa^2} + 1\right] \sigma_w} \end{aligned} \tag{33}$$

FIGURE 3

Bias in OLS

Figure 3 illustrates bias in the OLS estimate $\hat{\beta}_0$ in the regression $\Delta \ln P_{t+1} = a + \beta_0 \Delta \sigma^2_{t+1} + \varepsilon_{t+1}$ when the error term ε_{t+1} is correlated with the innovations in volatility, $\text{Corr}(v_{t+1}, w_{t+1}) = \delta$. The graph shows the theoretical factor loading, A_σ, and the population value of B_{OLS} as a function of δ. The plot is generated using parameter values $\gamma = 7.5, \psi = 2.5, \kappa = 0.97, k_1 = 0.999, \sigma_w = 0.0015, E(\sigma_t^2) = 0.0078^2$.



where the second term is bias in the OLS estimate B_{OLS} for A_σ. This bias can be very substantial. Its sign depends on the correlation δ between dividend and volatility shocks. If shocks to volatility are positively correlated with dividend news, the OLS estimator is upwardly biased for A_σ and vice versa.

Figure 3 shows the impact of correlation between volatility and dividend innovations on the population value of the regression in (32). The shaded turquoise area represents the bias, which is significant. Even for small positive values of δ, it is possible that volatility σ_t predicts returns even if volatility shocks have zero correlation with asset prices. It is more economically plausible, however, that shocks to dividends are negatively correlated with volatility shocks. In this case, one would observe a sharply negative correlation between volatility changes and asset returns. This is empirically relevant in the context of VIX and other measures of volatility. In particular, one routinely finds that volatility and return shocks are sharply negatively correlated, while at the same time, volatility very weakly predicts returns.

While Figure 3 indicates that the tests are invalid and biased in the case that cash flow shocks are correlated with expected return shocks, they are biased only because $\hat{\beta}_0$ will be biased. Importantly, the term structure of regression predictability regression slopes maintains the same shape but is subjected to a parallel shift up or down depending on the size of the bias. The predictability is itself unchanged, but the tests are impacted by biased estimates of $\hat{\beta}_0$. To overcome this, consider the fact that

TABLE 9

Test of EGP Without the Zero-Correlation Assumption

Table 9 reports the results of tests of EGP without using estimates of β_0. The theoretical slope coefficients are $b_h^* = b_1(\rho^h - \rho^{h-1})/(1-\rho)$ for 1-period returns using $h = 2, .., N$. The table reports multivariate test statistics, Q_N, and associated $\chi^2(N-2)$ p-values for $N = 6$ and $N = 12$ month maximum forecasting horizons.

	VRP		IV		RV		TI	
	$N=6$	$N=12$	$N=6$	$N=12$	$N=6$	$N=12$	$N=6$	$N=12$
Q_N^*	20.90	30.37	7.04	11.00	19.00	21.65	7.64	13.11
p–val	0.00	0.00	0.13	0.36	0.00	0.02	0.11	0.22

(34) $$\beta_0 = b_1/(1-\rho)$$

where, as before, b_1 is the 1-period-ahead predictability slope. We can now impose the null

(35) $$b_h^* = b_1(\rho^h - \rho^{h-1})/(1-\rho)$$

for $h = 2, .., N$. The test takes on the form $Q = \left(\hat{b} - b^*\right)' \Omega^{-1} \left(\hat{b} - b^*\right)$ as before but with $N-1$ elements corresponding to horizons $h = 2, .., N$.

Table 9 reports the results of the modified tests. Relaxing the zero-correlation assumption does not alter the conclusion of the tests for VRP, RV, and TI, but does render the EGP test insignificant for IV. To see why this is the case for IV, and not RV, note that the correlation between stock returns and shocks to IV is sharply more negative than the return-RV correlation as can be seen from the estimates of β_0, which equals -0.13 and -0.04, respectively. For IV, this implies strong equilibrium predictability. Since there is essentially no evidence of predictability from IV in the data, a highly negative estimate of β_0 drives the rejection of the EGP null in Table 5. For the other variables, β_0 is smaller in magnitude, which leads to smaller discrepancies between the tests reported previously and those in Table 9.

Empirical studies that measure the correlation between cash flow and discount rate shocks mostly find the correlation to be statistically insignificant. Vuolteenaho (2002) finds it to be positive but statistically insignificant in when using multiple lags of state variables in a VAR. Campbell and Vuolteenaho (2004) estimate the correlation to be small and statistically insignificant. Botshekan, Kraeussl, and Lucas (2012) estimate the correlation to be approximately -0.03[5], while recent work by Lockstoer and Tetlock (2020) estimates it to be -0.15 and statistically insignificant. Binsbergen and Koijen (2010) use a latent factor approach to estimate expected returns and expected dividend growth from price and dividend information. They find that shocks to the unobserved (filtered) expected returns factor are positively but statistically insignificantly correlated with dividend growth rate shocks.

[5]Table 2 reports the covariance matrix, implying the correlation to be $-0.0001/\sqrt{0.007*0.0013} = -0.03$.

TABLE 10

Tests of EGP Controlling for Future Earnings

Table 10 reports the results of predictability regressions and tests of EGP controlling for current and future Standard & Poor's 500 Index (SPX) earnings growth. Estimates of β_0 are obtained from the regression:

$$r_t = \alpha + \beta_0 \Delta x_t + \sum_{l=1}^{9} B_l e_{t+l} + \varepsilon_t$$

where $e_{t+l} = \ln(EARNINGS_{t+l}/EARNINGS_{t+l-1})$ is future S&P 500 earnings growth rates. The $R^2 - adj$ are adjusted R^2s in the regression of Δx_t on future earnings growth rates.

	VRP	IV	RV	TI
$R^2 - adj$	11.6	6.9	20.6	0.8
β_0	−0.03	−0.14	−0.04	−0.50
Q_6	23.4***	35.3***	68.3***	10.61
Q_{12}	25.5***	43.0***	72.3***	18.26

G. Controlling for Cash Flow News

To further analyze the extent to which cash flow and discount rate shock correlation is an issue with the results in the previous section, I rerun the regression (7) to estimate β_0:

$$r_t = \alpha + \beta_0 \Delta x_t + \sum_{l=1}^{L} B_l e_{t+l} + \varepsilon_t \tag{36}$$

where $e_{t+l} = \ln(Earnings_{t+l}/Earnings_{t+l-1})$ is the month-over-month future earnings growth. That is, in order to control for shocks to expected earnings, I include *future* earnings growth rates. The idea is to decompose returns into shocks due to cash flow shocks and discount rate shocks, as in Campbell and Vuolteenaho (2004). Since we do not observe expectations of future earnings, I include the actual earnings. The idea is that actual earnings can be decomposed into an expected and unexpected part, $Earnings_{t+l} = E_t(Earnings_{t+l}) + w_{t+l}$, where the shock w_{t+l} is uncorrelated with the expected return shock, Δx_t.

Table 10 reports the empirical results of the regressions where I control for future earnings growth. Firstly, I compute R^2s from regressions of changes in the candidate risk variables, Δx_t, on future earnings. The R^2s range from 0.8% to 20.6%. When these R^2s are low, there is likely no change in β_0 estimates. Estimates of β_0 are indeed similar for all the variables. For RV and VRP, the point estimates are unchanged; for IV, it differs by 0.01; and for TI, we get −0.5, versus −0.44 in Table 7. Table 10 also reports the multivariate test statistics. With the exception of TI, the tests all reject the null of EGP. In the case of TI, the test statistics are marginally smaller. In the case of RV, IV, and VRP, the test statistics are larger after controlling for future earnings.

IV. Multivariate Extensions

In this section, I study multivariate extensions of the tests considered previously. This allows for two important generalizations. First, it is common for asset pricing models to feature more than one priced state variable and therefore more than one variable that predicts returns in equilibrium. Second, generalizing the AR(1) assumption will allow us to better model the dynamics of some of the variables in question. For example, it is well documented that volatility exhibits long memory-like features (see, e.g., Bollerslev and Mikkelsen (1996), Bandi and Perron (2006)). We consider ARMA(p,q) specifications as a natural extension of the AR(1) process. Fortunately, the ARMA processes nest within the VAR framework considered next.

A. VARs

The results from the previous sections generalize straightforwardly to a multivariate setting. Consider a K dimensional state variable X_t that follows a VAR(1),

(37) $$X_t = AX_{t-1} + W_t$$

where A is a $K \times K$ dimensional VAR(1) matrix and $\mathrm{Cov}(W_t) =: \Sigma$ is an unrestricted error covariance matrix.

The equilibrium relation between returns and states is

(38) $$\Delta \ln P_t = a + \beta_0' \Delta X_t + \varepsilon_t$$

where β_0 is now an n-length vector.

The predictive regressions are now

(39) $$r_{t+h} = a_h + b_h' X_t + u_{t+h}^h$$

for -$h$periods-ahead 1-period returns and

(40) $$r_{t:t+h} = \alpha_h + \beta_h' X_t + U_{t+h}^h$$

where b_h and β_h are now n-dimensional vectors.

The structural restrictions are

(41) $$b_h = \left(A^h - A^{h-1}\right)' \beta_0$$

for the 1-period returns, and

(42) $$\beta_h = \left(A^h - I_K\right)' \beta_0$$

for cumulative returns. $I_K$is a $K \times K$ identity matrix, and A^h is the matrix A multiplied h times with itself.

In order to construct a test, let b and β denote the KN length vectors collecting the estimated predictability coefficients arranged by variable and horizon. Let $\hat{b}$ denote the OLS-estimated coefficients, while b^* denote the theoretical coefficients. I construct a quadratic test as before:

(43) $$Q = \left(\hat{b} - b^*\right)' \Omega^{-1} \left(\hat{b} - b^*\right).$$

Because Ω is too cumbersome to compute in closed form but easy to compute by simulation, I obtain a simulation-based estimate. This is done by first estimating

TABLE 11

Multivariate Tests

Table 11 reports the results of predictability regressions and tests of EGP using bivariate VAR(1) dynamics for pairs of state variables. β_0 coefficients are computed controlling for current and future SPX earnings growth.

$$r_t = \alpha + \Delta X_t \beta_0' + \sum_{l=1}^{9} B_l e_{t+l} + \varepsilon_t$$

where $e_{t+l} = \ln(EARNINGS_{t+l}/EARNINGS_{t+l-1})$ is future S&P 500 earnings growth rates. The table reports Q-tests based on 6-month (Q_6) and 12-month (Q_{12}) maximum forecasting horizons. The bottom portion of the table adds interest rates (3-month TBILL rates) as a predictive variable.

	VRP + IV	VRP + RV	IV + RV	VRP + TI
Q_6	45.9***	42.7***	25.19***	25.2**
Q_{12}	51.5***	47.2***	30.2	30.86
	VRP + IV + rf	VRP + RV + rf	IV + rf	RV + rf
Q_6	47.9***	45.5***	24.1***	43.9***
Q_{12}	55.0***	51.4***	29.1	47.1***

the parameter matrices A and $\sigma_x^2 = \text{Cov}(W_t)$, β_0, and σ_ε. I then simulate 100,000 draws from the DGP to compute $\Omega = \text{Cov}(b - b^*)$.

B. Multivariate Results

Table 11 reports the results of bivariate VAR computations. EGP is mostly rejected, especially when the maximum number of months used is 6. When using a max of 12 months, the IV + RV and VRP + TI tests fail to reject the null. The tests have roughly twice the degrees of freedom of the univariate tests (11 and 23, respectively) as they have double the number of restrictions. As a result, the tests have lower power in the 12-month case.

C. ARMA(p,q)

A potential criticism of the baseline tests considered so far is that I have assumed a simple AR(1) process for the single state variable x_t. There are three ways to imagine generalizing the AR(1) assumption. First, we could reasonably expect that an AR(p) or ARMA(p,q) dynamics for the state could provide a better approximation to the underlying DGP for x_t and therefore better forecasts x itself and, implicitly, therefore, returns. Note that both AR(p) and ARMA(p,q) are examples for processes that are non-Markovian with respect to the single state x_t.

As before, I assume that the log-price/dividend ratio is a function of a single state variable, x,

$$\ln P_t = \ln D_t + \beta_0 x_t \tag{44}$$

however with x now assumed to follow an ARMA(p,q):

$$x_t = \rho_1 x_{t-1} + \rho_2 x_{t-2} + \ldots + \rho_p x_{t-p} + w_t + \theta_1 w_{t-1} + \ldots + \theta_q w_{t-q} \tag{45}$$

where w_t is a time t shock. In order to represent this process more compactly, define $X_t = (x_t, x_{t-1}, .., x_{t-p}, w_t, w_{t-1}, .., w_{t-q})$ to be a pq dimensional state variable. While the explicit expressions are not important per se, we can represent the ARMA(p,q)

as in equation (B-1) by populating the elements of A and $\Sigma_X = \mathrm{Cov}(W_t)$ is discussed in Appendix B.

It now follows that the conditionally expected log capital gains are given by

$$E_t(\ln P_{t+h} - \ln P_{h-1}) = \beta_0 \iota' \left(A^h - A^{h-1}\right) X_t \tag{46}$$

where $\iota = (1,0,0,..,0)$.

There are three noteworthy facts about equation (46):

1. We can interpret the equation as representing a scalar β_0 multiplied by the difference in conditional expected future realization of x_t, which is also a scalar, $\iota'\left(A^h - A^{h-1}\right)X_t =: \overline{x}_t^h$. For this reason, we can forecast returns by first estimating A, or equivalently the AR and MA parameters in the ARMA model, and then essentially regress returns onto $\overline{x}_t^h$. This regression can be run for each h, say,

$$r_{t+h} = a + B_h \overline{x}_t^h + u_{t,h}. \tag{47}$$

   The structural restriction is now $B_h = \beta_0$ for all $h = 1,..,H$.
2. We can run the multivariate regression:

$$r_{t+h} = a + b_h' X_t + u_{t,h}. \tag{48}$$

   The structural restriction is now

$$b_h = \left(A^h - A^{h-1}\right)' \begin{bmatrix} \beta_0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{49}$$

   This is similar to the VAR(1) case except that the "post-multiplication vector" $\iota\beta_0 = (\beta_0, 0, .., 0)$has only one estimable parameter, β_0, instead of K. As in the AR(1) case, $\beta_0$is estimable from a simple regression of log capital gains onto changes in the state variable. This approach suffers from the curse of dimensionality as it calls for estimating $H \times p \times q$ number of predictive slope coefficients. Thus, for large values of p and q, the associated test statistic will lack power.

3. Under no scenario is the regression of future returns onto the time t value of the scalar x_t,

$$r_{t+h} = a + b_h x_t + u_{t,h}, \tag{50}$$

   justified. This is so because x_t is not a sufficient statistic in forecasting x_{t+h} and therefore does not optimally forecast returns either. The implication is that return predictability regressions constructed from a single state x_t that is reasonably more persistent than an AR(1) are essentially misspecified. This offers hope for recovering EGP from state variables that are known to be persistent beyond what can be captured by an AR(1). Put differently, we are led to reexamine EGP for state variables that have autocorrelation functions that decay at a rate slower than geometric. This is particularly true of volatility measures, as indicated in Table 2.

https://doi.org/10.1017/S0022109024000218

D. Empirical Results for ARMA Processes

In the following, I will investigate the fit of various ARMA model representations of the respective variables. In doing so, I will first present the R^2s for all the four variables in question for various combinations of max lags p and q. First, Figures 4–7 present R^2 from predictability regressions onto the components of ARMA decompositions. The ARMA parameters are estimated in-sample. For IV and TI including higher-order terms lead to modest, negligible improvements in (adjusted) R^2s. For RV, the 1-month R^2 is a bit above 5%. For VRP, the increase is relatively sizable, with a maximum R^2 of 16.6% at the 3-month cumulative return forecasting horizon for $p = 3, q = 4$. The result should be interpreted with caution as the number of regressors is large (7) and the sample size is relatively short (344) and is done in-sample.

As shown in Figures 4 to 7, predictability vanishes beyond the 3- to 4-month horizon. For this reason, I focus the joint tests of EGP under the ARMA(p,q) data-generating assumption to lags up to 6 months. Tables 12 and 13 report the EGP test results for combinations of p and q up to 4 lags for 3 and 6 months, respectively. At the 3-month horizon, with the exception of TI, EGP is mostly rejected. For VRP, there is some evidence suggesting that EGP cannot be rejected when p and q are large. For example, $p = 4$ and $q = 3$ or $q = 4$ fail to produce a large enough test statistic to reject the null at the 3-month horizon, but $p = 3, q = 3$ reject at the 6-month horizon. The evidence against EGP in IV and RV is pretty strong: At the 3-month horizon, EGP is rejected for RV for all p and q at the 1% level. There is no evidence against the null for TA at the 3-month horizon, and only three

FIGURE 4

Predictability R^2s

Figure 4 shows VRP R^2s for predictability regressions up to 6 months using ARMA components.

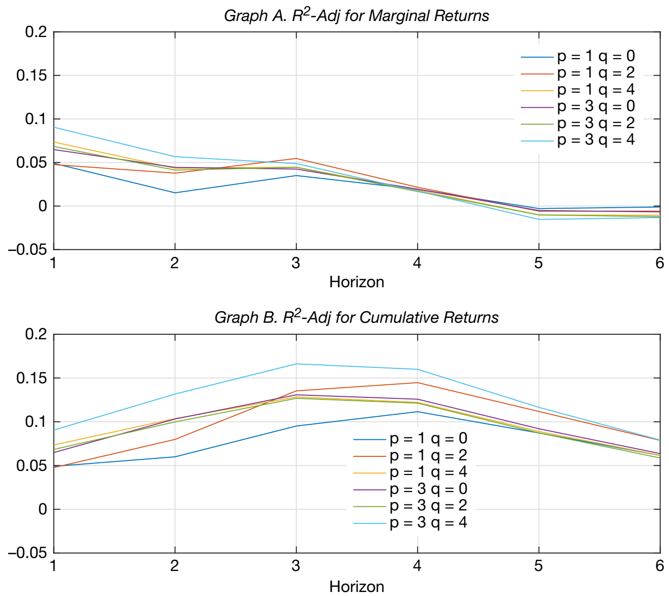


FIGURE 5

Predictability R^2s for ARMA Components for IV

Figure 5 shows marginal R^2s in predictability regressions using IV as predictor for up to 6 months for various ARMA(p,q) specifications. Graph A shows marginal returns; Graph B shows cumulative returns.

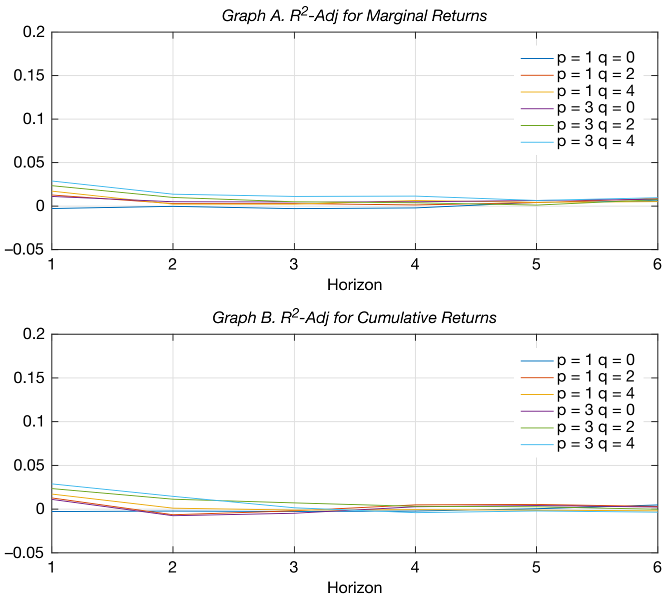


FIGURE 6

Predictability R^2 for ARMA components for RV

Figure 6 shows marginal R^2s in predictability regressions using RV as predictor for up to 6 months for various ARMA(p,q) specifications. Graph A shows marginal returns; Graph B shows cumulative returns.

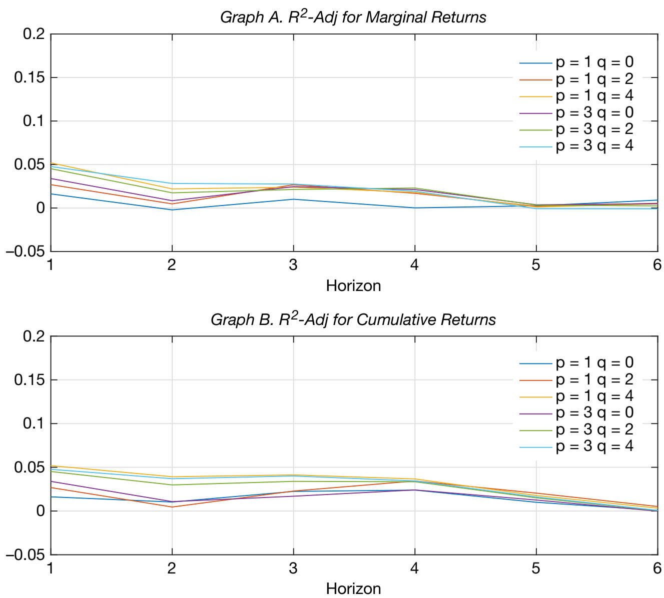
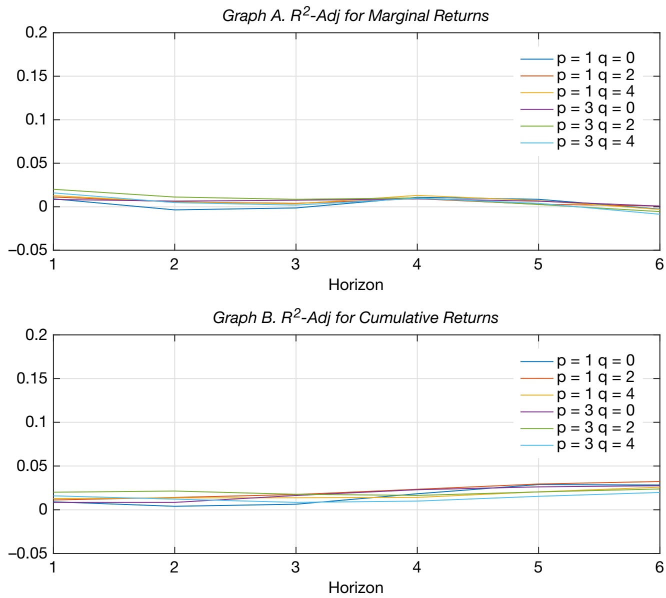


FIGURE 7

Predictability R^2s for ARMA components for TI

Figure 7 shows marginal R^2s in predictability regressions using TI as predictor for up to 6 months for various ARMA(p,q) specifications. Graph A shows marginal returns; Graph B shows cumulative returns.



combinations of large p and q pairs reject the null at the 6-month horizon. Overall, the evidence against EGP under the ARMA(p,q) assumption is pretty consistent with what was reported for the univariate, AR(1) case. EGP is pretty much rejected for most combinations of AR and MA lags for VRP, IV, and RV, but not for TI.

In order to dig a bit deeper into the question of ARMA processes as data-generating processes for the risk variables in question, I present two additional pieces of analysis. First, Table 14 presents parameter estimates for ARMA parameters estimated from VRP data. I estimate all model combinations with AR and MA lags $p = 1,.,4$, $q = 0,..,4$ for a total of 20 different combinations. What is apparent from the results is that parameters tend to vary wildly from one estimation to another. For example, the AR1 coefficient ranges -0.83 ($p = 3, q = 3$) to 1.37 ($p = 2, q = 2$). The wild behavior of the coefficient estimates across model specifications raises obvious questions of model overfitting and model misspecification. More importantly, it is pertinent to examine what these models imply for the behavior of expected returns.

To gauge which ARMA specifications may or may not be overfitting the data, consider what is reasonable behavior for expected returns. Remember that in the case of AR(1), expected returns decay monotonically in response to a shock as shown in Figure 1. This monotonicity is reasonable. By contrast, it would seem unreasonable that expected returns would oscillate from negative to positive in response to a shock. Yet, this is the implication of some of the higher-order ARMA (p,q) processes. Figure 8 plots IRFs for VRP. As seen, there are four IRFs where

TABLE 12

EGP Tests for ARMA(p,q) States for Horizon up to 3 M

Table 12 reports EGP tests under the assumption that the state variable follows an ARMA(p,q). The null hypothesis is that b_h in the forecasting regression

(51) $$r_{t+h} = a + b_h' X_t + u_{t,h}.$$

equals $\left(A^h - A^{h-1}\right)' \iota \beta_0$ for $h = 1,..,6$, where $X_t = \left(x_t, x_{t-1}, .., x_{t-p}, w_t, .., w_{t-q}\right)$ and A is the VAR matrix representation of the ARMA(p,q) process.

VRP

p \ q	0	1	2	3	4
1	16.31***	22.25***	25.66***	31.76***	21.61*
2	17.48***	26.28***	25.81***	24.72**	30.33**
3	19.59**	25.78***	25.85**	30.81**	25.74
4	20.05**	22.19*	32.50**	25.20	25.57

IV

p \ q	0	1	2	3	4
1	16.55***	25.29***	41.94***	33.32***	20.66
2	17.32***	42.20***	26.13***	23.69**	28.21**
3	19.78**	20.75**	25.22**	60.03***	30.49*
4	20.31**	21.51*	32.50**	35.35**	26.16

RV

p \ q	0	1	2	3	4
1	29.90***	32.31***	40.56***	41.08***	43.97***
2	32.03***	35.10***	39.98***	44.52***	48.74***
3	36.88***	43.80***	45.87***	43.92***	48.23***
4	41.54***	45.83***	45.59***	66.12***	52.49***

TI

p \ q	0	1	2	3	4
1	2.13	5.06	5.02	6.20	6.51
2	4.86	5.05	6.29	24.28**	9.90
3	5.48	6.47	7.42	20.83	7.55
4	6.43	9.15	7.32	14.40	11.92

the shocks dissipate monotonically. These are the ARMA(p,q) with p and q less than or equal to 2. To reduce clutter, the plot does not show all combinations of p and q larger than 2, but the IRFs all look as in the $p=4$ case: They oscillate. Accordingly, I conclude that the only ARMA(p,q) representations that generate economically plausible implications for expected returns are those that restrict $p \leq 2$ and $q \leq 2$.

I also examined the shape of the IRFs for IV, RV, and TI, and Figure 9 shows selected IRFs for RV. The results are similar across these variables and show generally monotonically decreasing IRFs with jaggedness for higher values of p and q. In general, IRFs appear smooth only for p and q equal to 2 or less, while at the same time $p=q=2$ appears to capture low-frequency dependence as well as higher-order models. While highly informal, this would suggest that the northwest quadrants of Tables 12 and 13, corresponding to, $p \leq 2, q \leq 2$ are the most robust tests.

TABLE 13

EGP Tests for ARMA(p,q) States for Horizon up to 6 M

Table 13 reports EGP tests under the assumption that the state variable follows an ARMA(p,q). The null hypothesis is that b_h in the forecasting regression

(52) $$r_{t+h} = a + b_h' X_t + u_{t,h}.$$

equals $\left(A^h - A^{h-1}\right)' \iota \beta_0$ for $h = 1,..,6$, where $X_t = \left(x_t, x_{t-1}, .., x_{t-p}, w_t, .., w_{t-q}\right)$ and A is the VAR matrix representation of the ARMA(p,q) process.

VRP

p \ Q	0	1	2	3	4
1	20.69***	51.06***	44.49***	48.11***	43.71**
2	21.78**	45.97***	44.97***	35.20	56.45**
3	24.71	27.26	35.71	190.79***	64.92**
4	25.72	27.30	74.26***	88.49***	39.40

IV

p \ q	0	1	2	3	4
1	22.07***	35.91***	38.80***	44.07***	48.30**
2	39.24***	36.61***	45.08***	47.15**	55.45**
3	39.86***	41.43**	45.38**	48.33*	71.05***
4	43.07***	45.82**	48.42*	59.90**	78.03***

RV

p \ q	0	1	2	3	4
1	38.04***	41.76***	43.59***	44.95***	50.67***
2	41.77***	45.10***	45.64***	50.95***	52.95**
3	40.97***	45.89***	51.99***	64.40***	53.17*
4	43.61***	49.91***	51.05**	83.15***	62.36*

TI

P \ q	0	1	2	3	4
1	6.72	7.30	10.74	10.28	14.07
2	7.37	10.67	10.01	74.22***	14.76
3	9.12	9.67	10.67	66.56***	18.53
4	12.15	13.71	15.71	69.13***	98.83***

E. Implications for Existing Models

It is natural to ask what the evidence presented here implies for existing models. In the following, I discuss this in the context of several models proposed in the LRR literature. I discuss the qualitative implications of the models rather than discussing whether the models, along with particular parameter calibrations, can quantitatively explain the findings.

Bansal and Yaron's (2004) model has received much attention in the literature. It is a fairly simple model in which a single AR(1) Gaussian spot variance factor, σ_t^2, drives expected excess stock market returns. Their model also features an AR(1) time-varying expected consumption growth rate, x_t. Both these factors generate time variation in price–dividend ratios and thus jointly generate the kind

TABLE 14

Estimated ARMA(*p*,*q*) Parameters for VRP

Table 14 reports parameter estimates in ARMA(*p*,*q*) models for VRP.

p				q			
AR1	*AR2*	*AR3*	*AR4*	*MA1*	*MA2*	*MA3*	*MA4*
0.28							
0.94***				−0.81***			
0.95***				−0.77***	−0.08*		
0.96***				−0.77***	0.06	−0.15***	
−0.70				0.95***	0.41***	0.28***	−0.02
0.23	0.20***						
1.06***	−0.11*			−0.87***			
1.37***	−0.39			−1.17***	0.25		
−0.62	0.13			0.87***	0.26**	0.26*	
0.14	0.79***			0.09	−0.60***	−0.05	−0.20***
0.22	0.19***	0.05					
−0.68	0.39***	0.25***		0.92***			
−0.57	−0.08	0.31*		0.81***	0.47*		
−0.83	0.90***	0.78***		1.00***	−0.70***	−0.72***	
−0.05	0.78***	0.18**		0.27	−0.55***	−0.21**	−0.23*
0.22	0.20***	0.06*	−0.06***				
−0.67	0.40***	0.23***	−0.03	0.91***			
0.17	0.97***	−0.05	−0.15***	0.03	−0.84***		
−0.09	1.02***	0.17**	−0.18	0.29***	−0.82***	−0.22**	
−0.11	0.48***	0.25	0.27*	0.33*	−0.25	−0.22	−0.47***

FIGURE 8

Impulse Response Functions for VRP

Figure 8 plots IRFs for VRP on selected ARMA specifications.

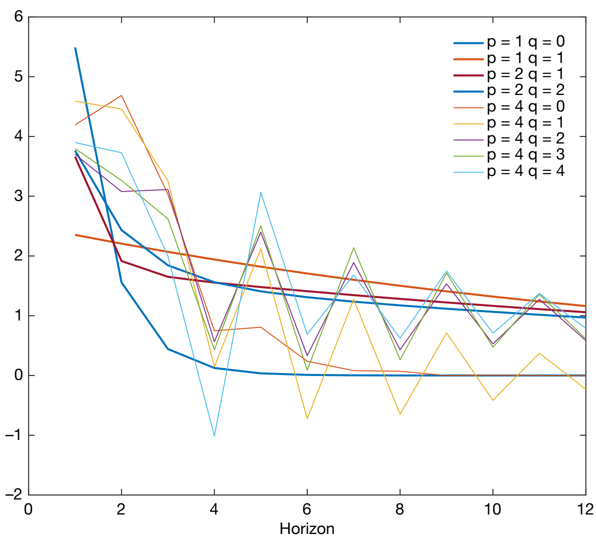
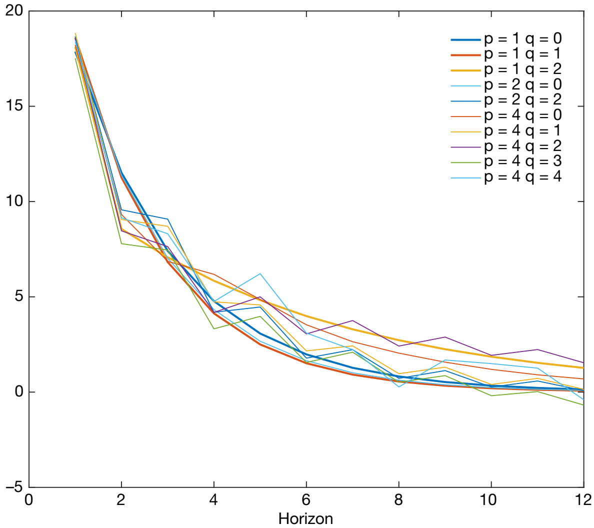


FIGURE 9

Impulse Response Functions for RV

Figure 9 shows IRFs for ARMA(p,q) models estimated for RV.



of AR(1) induced impulse response seen in Figure 1. If we were to observe x_t and σ_t^2, we could carry out a test of EGP exactly as described in the previous sections.

Since we, as econometricians, do not observe consumption spot variance σ_t^2, we need to resort to some observable quantity. Here, we then need to rely on another feature of the BY model—the fact that stock market variance is a linear function of σ_t^2[6]. It is also true that option-implied volatility is a linear function of σ_t^2[7].

I now turn to the question of whether any of the tests presented earlier are valid tests of the BY model.

Logarithmic capital gains in the BY model are given by

$$\ln P_{t+1} - \ln P_t = A_0 + A_\sigma\left(\sigma_{t+1}^2 - \sigma_t^2\right) + A_x(x_{t+1} - x_t) + e_{t+1} \tag{53}$$

where e_{t+1} is an error that is uncorrelated with changes in consumption variance, $\left(\sigma_{t+1}^2 - \sigma_t^2\right)$. We can therefore write

$$\ln P_{t+1} - \ln P_t = A_0 + A_\sigma\left(\sigma_{t+1}^2 - \sigma_t^2\right) + \hat{e}_{t+1} \tag{54}$$

That is, we can write $\hat{e}_{t+1} = A_x(x_{t+1} - x_t) + e_{t+1}$, which is uncorrelated with $\left(\sigma_{t+1}^2 - \sigma_t^2\right)$. This means that we estimate the parameter A_σ in (54) by OLS without worrying about omitted variable bias. Similarly, we can run the predictability regression.

[6]See equation (A13), p. 1505, in Bansal and Yaron (2004).

[7]See, e.g., Eraker and Yang (2022) for a detailed analysis of option-implied variance in the context of LRR models.

Since σ_t^2 and x_t are statistically independent, even univariate tests of whether σ_t^2 is EGP would be valid tests of the BY model.

V. Conclusion

While classic theories of efficient markets (e.g., Fama (1970)) postulate that returns are simple random walks with drifts, a plethora of subsequent work recognized that dynamic equilibrium generates a small predictable component of asset returns. The most theoretically appealing risk candidate is some measure of conditional variance. Disappointingly, however, measures of conditional variance do not predict returns (e.g., Merton (1973), (1980), French et al. (1987), Bollerslev, Engle, and Wooldridge (1988), Glosten, Jagannathan, and Runkle (1993)). A close cousin of conditional variance, the VRP, has been shown (e.g., Bollerslev et al. (2009)) to predict returns. Bollerslev et al. (2009) and Drechsler and Yaron (2011) derive equilibrium models of VRP based on Epstein–Zin preferences akin to the models of Bansal and Yaron (2004) and Eraker and Shaliastovich (2008). Recently, Bekaert et al. (2020) derive an equilibrium model for VRP based on a habit formulation.

Yet, there is a widespread among researchers that predictable variation in asset returns is consistent with equilibrium-based models. Bollerslev et al. (2009) and Drechsler and Yaron (2011) derive models explicitly aimed at explaining predictable variation, but it is also implied elsewhere in the predictability literature. For example, Bollerslev et al. (2015) derive an option-based TI used to predict returns and maintain that "..the new nonparametric jump risk measures proposed and analyzed here are all economically motivated, with direct analogs in popular equilibrium consumption-based asset pricing models." Kelly and Jiang (2014) state that "…tail risk plays an important role in the marginal utility of investors and in determining equilibrium asset prices" (p. 2866).

In this paper, I document that equilibrium justifications for the predictability of long- or medium-horizon returns are inconsistent with the typical equilibrium models proposed in the literature. There are two essential problems: First, equilibrium models based on persistent state variables and Epstein–Zin (EZ) preferences generate impulse responses in prices where i) the contemporaneous price impact of an increase in expected returns is negative and ii) the decay rate of the price shock is the same as the decay rate of the risk–shock. As such, any model of risk/expected return that generates a monotonically decaying impulse response will generate predictability R^2s for 1-period returns that decrease monotonically with the forecasting horizon. Both are testable implications, and in this paper, I provide a framework for testing. The results show that variables, such as VRP, option-implied, and RV measures, fail the test, while the TI studied simply does not show much predictive power.

The following are potential resolutions to these predictability puzzles: i) One can attempt to refine rational expectations equilibrium models to more carefully match the correlation and predictive correlations of priced state variables at all horizons. One such potential avenue is the decomposition of shocks to state variables into high- and low-frequency components in the hope that either will have empirical properties that match the aforementioned correlation patterns. ii) Models based on limited rationality may be useful in matching some of the features of the data. Yang (2022) proposes a model where agents are slow to update their beliefs about increases in volatility. This, along with a highly persistent state variable, provides a potential resolution of the apparent lack of aggregate return predictability from variance measures. iii) It is also conceivable that evidence of return predictability is spurious. The results in Section II can be used to derive a general-purpose test of multihorizon predictability and will be pursued elsewhere. In particular, Theorem 2 can be used to derive standard errors for overlapping regression slopes. The results can also be extended to construct a test of the null of no predictability at any horizon along the lines of Boudoukh et al. (2006) and Ang and Bekeart (2007) (see Appendix B).

This paper contains results that could be useful in comparing unconstrained OLS forecasting regressions with ones produced by VARMA or VAR models. This is a topic considered in Marcellino et al. (2006) in the context of a univariate time series. Jordà (2005) studies the performance of forecasts produced by VARMA models relative to direct, regression-based forecasts (dubbed local projections). Since the DGP in Jorda admits a linear state-space representation, there is an asymptotic equivalence between the VARMA forecast and the local projections (Plagborg-Møller and Wolf (2021)). This literature is silent on the overlapping observation problem that has been well-studied in finance. However, many economic time series, including standard measures of output, have features that resemble financial market data in that the levels or log levels contain a unit root, in which case the forecasters are typically concerned with forecasting growth rates or log growth rates. Hence, the forecaster is faced with the same choice as in the finance literature as to whether to forecast time-series cumulative or single-period growth rates. Model comparisons can then be carried out by testing for statistically significant differences between VAR (say) implied forecasting parameters and unconstrained estimates. I leave this topic for future research.

Appendix A. Returns, Capital Gains, and Predictable Dividend Growth

The paper makes two simplifying assumptions related to dividend growth and capital gains versus returns. First, I assume that expected dividend growth does not depend on x_t, and second, I treat capital gains as equivalent to returns. I am going to argue that these approximations do not alter the outcome of the analysis. To see this, consider first the addition of an expected dividend growth rate term

$$\ln D_{t+1} - \ln D_t = \mu + \omega x_t + \varepsilon_{t+1} \tag{A-1}$$

where a nonzero ω implies that the state variable drives expected dividend growth, as in Bansal—Yaron (2004).

TABLE A1

Dividend Growth Variance Ratio

Table A1 reports the fraction of variation in dividend growth relative to total return variation $\mathrm{Var}(y_t)/\mathrm{Var}(\ln R_{t+1})$, where $y_t = \ln\left(1+\frac{D_{t+1}}{P_{t+1}}\right)$. The variance ratio is shown in percentage at various sampling frequencies.

Day	1 M	1Q	1Y
0.049%	0.159%	0.157%	0.501%

The slope coefficient in the regression of cumulative log returns $r_{t:t+h}$ on the state variable x_t then satisfies

$$b_h = \left[\omega \frac{\rho^{h+1}-\rho}{\rho-1} + \beta\left(\rho^h - 1\right)\right]. \tag{A-2}$$

i.e., it has the additional term $\omega\frac{\rho^{h+1}-\rho}{\rho-1}$ relative to the baseline specification.

To assess the impact of shocks to returns, note that the linearized log return is

$$r_{t+1} = \kappa_0 + \kappa_1(\alpha + \beta x_{t+1}) - \alpha - \beta x_t + \omega x_t + \mu + \varepsilon_{t+1} \tag{A-3}$$

$$= const. + \kappa_1 \beta(x_{t+1} - x_t) - (\beta(1-\kappa_1) - \omega)x_t + \varepsilon_{t+1} \tag{A-4}$$

In the case that κ_1 is close to unity, which is the case in many applications, and especially in high-frequency data, the log returns are approximately

$$r_{t+1} \approx const. + \beta(x_{t+1} - x_t) - \omega x_t + \varepsilon_{t+1} \tag{A-5}$$

A noteworthy feature of dividend yield data is how little variation there is relative to capital gains. In Table A1, I compute the ratio of the variances of the forward-looking log dividend yield and variance of returns. It shows that at the daily frequency, dividend yield variation accounts for about five thousandths of the total variation. This is an upper bound on the R^2 that we could get from running a regression of total returns onto some predictor, which predicts only dividend yield variation. In other words, if the forward-looking log dividend yield was perfectly predictable, it could not explain more than five thousandths of the variation in returns. The numbers are larger for longer horizons, but nevertheless so small that it is clear that return predictability cannot come from the predictability of dividends at short horizons. For this reason, I assume $\omega = 0$.

Appendix B. VAR(1) Representation of ARMA(*p,q*)

The VAR(1) is

$$X_t = AX_{t-1} + W_t \tag{B-1}$$

https://doi.org/10.1017/S0022109024000218

where the elements of X_t are

(B-2)
$$X_t = \begin{bmatrix} x_t \\ x_{t-1} \\ \dots \\ x_{t-p+1} \\ w_t \\ w_{t-1} \\ \dots \\ w_{t-q+1} \end{bmatrix}.$$

The matrix A is has submatrices

(B-3)
$$A = \begin{bmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{bmatrix},$$

given by

(B-4)
$$A_{1,1} = \begin{bmatrix} \rho_1 & \rho_2 & \dots & & \rho_p \\ 1 & 0 & \dots & & 0 \\ 0 & 1 & \dots & & 0 \\ \vdots & \dots & \ddots & & \vdots \\ 0 & 0 & \vdots & 1 & 0 \end{bmatrix}_{p \times p}, \quad A_{1,2} = \begin{bmatrix} \theta_1 & \theta_2 & \dots & & \theta_q \\ 0 & 0 & \dots & & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ 0 & 0 & & \dots & 0 \end{bmatrix}_{p \times q},$$

(B-5)
$$A_{2,1} = 0_{q \times p}, A_{2,2} = \begin{bmatrix} 0 & 0 & \dots & & 0 \\ 1 & 0 & \dots & & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \dots & \dots & 1 & 0 \end{bmatrix}_{q \times q}.$$

For example, for an ARMA(1,1), $A_{1,1} = \rho_1$, $A_{1,2} = \theta_1$, and $A_{2,1} = A_{2,2} = 0$. For an ARMA(2,2),

(B-6)
$$A_{1,1} = \begin{bmatrix} \rho_1 & \rho_2 \\ 1 & 0 \end{bmatrix}, \quad A_{1,2} = \begin{bmatrix} \theta_1 & \theta_2 \\ 0 & 0 \end{bmatrix}, A_{2,1} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, A_{2,2} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

and so on.

The shock vector W_t has elements $W_t = (w_t, 0, \dots, 0, w_t, .., 0)$, where the second occurrence of the scalar shock w_t occurs in the $p+1$th position provided $q \geq 1$ provided $q > 0$. If $q = 0$, we have $W_t = (w_t, 0, \dots, 0)$. Finally, note that if the process is a pure moving average process ($p = 0$), these equations still apply as if $p = 1$ and $\rho_1 = 0$.

Appendix C. Proofs

Proof of Theorem 1. Consistency: Consider the DGP studied by Stambaugh (1999) where $r_{t+1}=\alpha_h+\beta_h x_t+u^h_{t+1}$, $x_{t+1}=\rho x_t+w_{t+1}$ with $\mathrm{Cov}(w,u)=\sigma_{u,w}$. Stambaugh gives the expression:

$$E\left(\hat{\beta}_1-\beta_1\right)=\frac{\sigma_{w,u^1}}{\sigma_w^2}(\hat{\rho}-\rho) \tag{C-1}$$

for the bias in the 1-period-ahead OLS slope. Note that $\frac{\sigma_{w,u^1}}{\sigma_w^2}=\beta_0$, which is assumed known. As such, the bias vanishes as $\lim_{T\to\infty}(\hat{\rho}-\rho)=0$.

Next, in deriving the expression for the covariance estimator for the slope coefficients I derive an expression for the function $F(T_1,T_2,T_3,T_4)$. By repeated application of the law of iterated expectations, direct computation yields

$$\begin{aligned}
&F(T_1,T_2,T_3,T_4)\\
&:=E(x_{T_1}x_{T_2}x_{T_3}x_{T_4})\\
&=E(x_{T_1}x_{T_2}x_{T_3}E_{T_3}(x_{T_4}))\\
&=E\left(x_{T_1}x_{T_2}x^2_{T_3}\right)\rho^{T_4-T_3}\\
&=E\left(x_{T_1}x_{T_2}\left(x^2_{T_2}\rho^{2(T_3-T_2)}+\sigma^2_{T_2:T3}\right)\right)\rho^{T_4-T_3}\\
&=E\left(x_{T_1}\left(x^3_{T_2}\rho^{2(T_3-T_2)}+x_{T_2}\sigma^2_{T_2:T3}\right)\right)\rho^{T_4-T_3}\\
&=E\left(x_{T_1}\left(x^3_{T_2}\rho^{2(T_3-T_2)}+x_{T_1}\rho^{T_2-T_1}\sigma^2_{T_2:T3}\right)\right)\rho^{T_4-T_3}\\
&=E\left(x_{T_1}\left(\left(x^3_{T_1}\rho^{3(T_2-T_1)}+3x_{T_1}\rho^{T_2-T_1}\sigma^2_{T_1:T_2}\right)\rho^{2(T_3-T_2)}+x_{T_1}\rho^{T_2-T_1}\sigma^2_{T_2:T3}\right)\right)\rho^{T_4-T_3}\\
&=E\left(\left(\left(x^4_{T_1}\rho^{3(T_2-T_1)}+3x^2_{T_1}\rho^{T_2-T_1}\sigma^2_{T_1:T_2}\right)\rho^{2(T_3-T_2)}+x^2_{T_1}\rho^{T_2-T_1}\sigma^2_{T_2:T3}\right)\right)\rho^{T_4-T_3}\\
&=\left(\left(E\left(x^4_{T_1}\right)\rho^{3(T_2-T_1)}+3E\left(x^2_{T_1}\right)\rho^{T_2-T_1}\sigma^2_{T_1:T_2}\right)\rho^{2(T_3-T_2)}+E\left(x^2_{T_1}\right)\rho^{T_2-T_1}\sigma^2_{T_2:T3}\right)\rho^{T_4-T_3}\\
&=\left(\left[E\left(x^4_{T_1}\right)\rho^{3(T_2-T_1)}+3E\left(x^2_{T_1}\right)\rho^{T_2-T_1}\sigma^2_{T_1:T_2}\right]\rho^{2(T_3-T_2)}+E\left(x^2_{T_1}\right)\rho^{T_2-T_1}\sigma^2_{T_2:T3}\right)\rho^{T_4-T_3}\\
&=\left(\left[3E\left(x^2_{T_1}\right)^2\rho^{3(T_2-T_1)}+3E\left(x^2_{T_1}\right)\rho^{T_2-T_1}\sigma^2_{T_1:T_2}\right]\rho^{2(T_3-T_2)}+E\left(x^2_{T_1}\right)\rho^{T_2-T_1}\sigma^2_{T_2:T3}\right)\rho^{T_4-T_3}\\
&=\left(3E(x^2)\rho^{T_2-T_1}\left[E(x^2)\rho^{2(T_2-T_1)}+\sigma^2_{T_1:T_2}\right]\rho^{2(T_3-T_2)}+E(x^2)\rho^{T_2-T_1}\sigma^2_{T_2:T3}\right)\rho^{T_4-T_3}\\
&=E(x^2)\rho^{T_2-T_1}\left(3\left[E(x^2)\rho^{2(T_2-T_1)}+\sigma^2_{T_1:T_2}\right]\rho^{2(T_3-T_2)}+\sigma^2_{T_2:T3}\right)\rho^{T_4-T_3}
\end{aligned}$$

where we now use that $E\left(x^4_{T_1}\right)=E(x^4)=3E(x^2)^2$ and $E(x^2)=\sigma_x^2/(1-\rho^2)$, and

$$\sigma^2_{T:S}=\sigma^2\frac{1-\rho^{2(T-S)}}{1-\rho^2} \tag{C-2}$$

is the conditional variance $\mathrm{Var}(x_T|\mathcal{F}_S)$.

The covariance matrix has elements:

$$
\begin{aligned}
\text{(C-3)}\quad \operatorname{Cov}(b_h, b_l) &= E\left(\frac{1}{\left(\sum_t x_t^2\right)^2}\sum_t\sum_s x_t x_s(\beta_0(x_{t+h}-x_{t+h-1})-\beta_h x_t)(\beta_0(x_{s+l}-x_{s+l-1})+\varepsilon_{s+l}-\beta_l x_s)\right.\\
&\left.+\varepsilon_{t+h}(\beta_0(x_{s+l}-x_{s+l-1})+\varepsilon_{s+l}-\beta_l x_s)\right)\\
&= E\left(\frac{1}{\left(\sum_t x_t^2\right)^2}\sum_t\sum_s x_t x_s(\beta_0(x_{t+h}-x_{t-h-1})-\beta_1 x_t)(\beta_0(x_{s+l}-x_{s+l-1})-\beta_l x_s)\right.\\
&\left.+\frac{1}{\left(\sum_t x_t^2\right)^2}E\left(\sum_t x_{t-h}x_{t-l}\sigma_\varepsilon^2\right)\right)\\
&\approx \frac{1}{\left(\sum_t x_t^2\right)^2}E\left(\sum_t\sum_s x_t x_s(\beta_0(x_{t+h}-x_{t+h-1})-\beta_1 x_t)(\beta_0(x_{s+l}-x_{s+l-1})-\beta_l x_s)+\frac{T\rho^{|l-h|}\sigma_x^2\sigma_\varepsilon^2}{\left(\sum_t x_t^2\right)^2}\right.\\
&= \frac{1}{\left(\sum_t x_t^2\right)^2}\left[E\left(\sum_t\sum_s \beta_0^2 x_t x_s(x_{t+h}-x_{t+h-1})(x_{s+l}-x_{s+l-1})\right)\right.\\
&-E\left(\sum_t\sum_s \beta_0\beta_l x_t x_s^2(x_{t+h}-x_{t+h-1})\right)\\
&-E\left(\sum_t\sum_s \beta_0\beta_h x_t^2 x_s(x_{s+l}-x_{s+l-1})\right)\\
&\left.+E\left(\sum_t\sum_s \beta_h\beta_l x_t^2 x_s^2\right)\right]+\frac{T\rho^{|l-h|}\sigma_x^2\sigma_\varepsilon^2}{\left(\sum_t x_t^2\right)^2}\\
&= c\sum_t\sum_s(\beta_0^2[F(t,s,t+h,s+l)-F(t,s,t+h,s+l-1)\\
&-F(t,s,t+h-1,s+l)+F(t,s,t+h-1,s+l-1)]\\
&-\beta_0\beta_l[F(t,s,s,t+h)-F(t,s,s,t-h-1)]\\
&-\beta_0\beta_h[F(t,s,s,t+l)-F(t,s,s,t-l-1)]\\
&+\beta_h\beta_l F(t,t,s,s)+T\rho^{|l-h|}\sigma_x^2\frac{\sigma_\varepsilon^2}{1-\rho^2})
\end{aligned}
$$

Q.E.D.

Proof of Theorem 2. Since

$$b_h=\sum_{i=1}^{h}\beta_i, \tag{C-4}$$

$$\operatorname{Cov}(b_h,b_k)=Cov\left(\sum_{i=1}^{h}\beta_i,\sum_{j=1}^{h}\beta_j\right) \tag{C-5}$$

$$=\sum_{i=1}^{h}\sum_{j=1}^{h}Cov\left(\beta_i,\beta_j\right) \tag{C-6}$$

Q.E.D.

Proof of Theorem 3. The dynamics of x can be written

$$x_{t+h}=\rho^h x_t+w_t^h \tag{C-7}$$

$$=\rho^h x_t+\sum_{i=1}^{h-1}\rho^{h-i}w_{t+i} \tag{C-8}$$

https://doi.org/10.1017/S0022109024000218

Notice that we can also write

$$x_{t+h} = \rho_h x_t + w_t^h \tag{C-9}$$

where ρ_h is defined to be the autocorrelation over h periods. It is clear that —$\rho_h = \rho^h$ the autocorrelation in x sampled at h intervals equals the 1-period autocorrelation raised to the power of h. It is also clear that the estimated h-period autocorrelation $\hat{\rho}_h$ does not have the same bias as the 1-period autocorrelation raised to the power of h: $E(\hat{\rho}_h) \neq E(\hat{\rho}^h)$.

Stambaugh (1999) derives the following expression for the 1-period ($h = 1$) bias:

$$E\left(\hat{\beta}_1 - \beta_1\right) = \frac{\mathrm{Cov}\left(u_{t+1}^1, w_{t+1}\right)}{\sigma_w^2} E(\hat{\rho} - \rho) \tag{C-10}$$

We can now prove the statement for h. The "multiplier" in Stambaugh's expression for the bias equals the contemporaneous coefficient β_0 relating the shocks in returns to changes in x,

$$\frac{\sigma_{w,u^1}}{\sigma_w^2} = \beta_0. \tag{C-11}$$

Combining with the definition of the empirical bias, we get

$$E\left(\hat{\beta}_1 - \beta_1(\hat{\rho}, \beta_0)\right) = b_1 + \beta_0(E(\hat{\rho}) - \rho) - \beta_1(\hat{\rho}, \beta)) \tag{C-12}$$

$$= \beta_1 + \beta_0(E(\hat{\rho}) - \rho) - \beta_0(E(\hat{\rho}) - 1) \tag{C-13}$$

$$= \beta_1 - \beta_0 \rho + \beta_0 = 0 \tag{C-14}$$

Since the -$h$period regression can be reinterpreted as one with h defined as a unit of time, it also follows that

$$E\left(\hat{\beta}_h - \beta_h\right) = \frac{\mathrm{Cov}\left(u_t^h, w_t^h\right)}{\mathrm{Var}\left(w^h\right)} E\left(\hat{\rho}_h - \rho^h\right). \tag{C-15}$$

The return process is

$$r_{t:t+h} = \mu h + \beta_0(x_{t+h} - x_t) + \sum_{i=1}^{h} \varepsilon_{t+i} \tag{C-16}$$

$$= \mu h + \beta_0\left(\rho^h - 1\right)x_t + \beta_0 \sum_{i=1}^{h-1} \rho^{h-i} w_{t+i} + \sum_{i=1}^{h} \varepsilon_{t+i} \tag{C-17}$$

which combined with the regression

$$r_{t+h} = \alpha_h + \beta_h x_t + u_{t+h}, \tag{C-18}$$

implies that

$$\alpha_h = \mu h \tag{C-19}$$

$$\beta_h = \beta_0 \left(\rho^h - 1 \right) \tag{C-20}$$

$$u_{t+h} = \beta_0 \sum_{i=1}^{h-1} \rho^{h-i} w_{t+i} + \sum_{i=1}^{h} \varepsilon_{t+i} \tag{C-21}$$

by matching terms. Thus, it also holds in the $h > 1$ case that

$$\frac{\mathrm{Cov}\left(u_t^h, w_t^h\right)}{\mathrm{Var}\left(w^h\right)} = \beta_0 \tag{C-22}$$

as in the $h = 1$ case. Combining the definition of the empirical bias with (C-15) gives

$$E\left(\hat{\beta}_h - \beta_h(\hat{\rho}, \beta_0)\right) = \beta_h + \beta_0 \left(E(\hat{\rho}_h) - \rho^h\right) - E(\beta_h(\hat{\rho}, \beta)) \tag{C-23}$$

$$= \beta_0 \left(\rho^h - 1\right) + \beta_0 \left(E(\hat{\rho}_h) - \rho^h\right) - E\left(\beta_0 \left(\hat{\rho}^h - 1\right)\right) \tag{C-24}$$

$$= \beta_0 E\left(\hat{\rho}_h - \hat{\rho}^h\right). \tag{C-25}$$

Since $\hat{\beta}_h = \sum_{i=1}^{h} \hat{b}_i$, we get

$$E\left(\hat{b}_h - b_h(\hat{\rho}, \beta_0)\right) = E\left(\hat{\beta}_h - \beta_h(\hat{\rho}, \beta_0)\right) - E\left(\hat{\beta}_{h-1} - \beta_{h-1}(\hat{\rho}, \beta_0)\right) \tag{C-26}$$

$$= \beta_0 E\left(\hat{\rho}_h - \hat{\rho}_{h-1} - \left(\hat{\rho}^h - \hat{\rho}^{h-1}\right)\right) \tag{C-27}$$

Q.E.D.

References

Ang, A., and G. Bekeart. "Predictability: Is It There?" *Review of Financial Studies*, 20 (2007), 651–707.

Bandi, F. M., and B. Perron. "Long Memory and the Relation Between Implied and Realized Volatility." *Journal of Financial Econometrics*, 4 (2006), 636–670.

Bansal, R.; D. Kiku; and A. Yaron. "Risks for the Long Run: Estimation and Inference." *Journal of Monetary Economics*, 82 (2016), 52–69.

Bansal, R., and A. Yaron. "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles." *Journal of Finance*, 59 (2004), 1481–1509.

Bekaert, G.; E. Engstrom; and A. Ermolov. "The Variance Risk Premium in Equilibrium Models." Working Paper, Columbia University (2020).

Bekaert, G., and M. Hoerova. "The VIX, the Variance Premium and Stock Market Volatility." *Journal of Econometrics*, 183 (2014), 181–192.

Binsbergen, J. H. V., and R. S. J. Koijen. "Predictive Regressions: A Present-Value Approach." *Journal of Finance*, 65 (2010), 1439–1471.

Bollerslev, T.; R. F. Engle; and J. M. Wooldridge. "A Capital Asset Pricing Model with Time-Varying Covariances." *Journal of Political Economy*, 96 (1988), 116–131.

Bollerslev, T., and H. O. Mikkelsen. "Modeling and Pricing Long Memory in Stock Market Volatility." *Journal of Econometrics*, 73 (1996), 151–184.

Bollerslev, T.; G. Tauchen; and H. Zhou. "Expected Stock Returns and Variance Risk Premia." *Review of Financial Studies*, 22 (2009), 4463–4492.

Bollerslev, T., and V. Todorov. "Tails, Fears, and Risk Premia." *Journal of Finance*, 66 (2011), 2165–2211.

Bollerslev, T.; V. Todorov; and L. Xu. "Tail Risk Premia and Return Predictability." *Journal of Financial Economics*, 118 (2015), 113–134.

Botshekan, M.; R. Kraeussl; and A. Lucas. "Cash Flow and Discount Rate Risk in Up and Down Markets: What Is Actually Priced?" *Journal of Financial and Quantitative Analysis*, 47 (2012), 1279–1301.

Boudoukh, J.; M. Richardson; and R. F. Whitelaw. "The Myth of Long-Horizon Predictability." *Review of Financial Studies*, 21 (2006), 1576–1605.

Brennan, M. "A Note on Dividend Irrelevance and the Gordon Valuation Model." *Journal of Finance*, 26 (1971), 1115–1121.

Campbell, J., and J. Cochrane. "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behaviour." *Journal of Political Economy*, 107 (1999), 205–251.

Campbell, J., and T. Vuolteenaho. "Bad Beta, Good Beta." *American Economic Review*, 94 (2004), 1249–1275.

Campbell, J. Y., and R. J. Shiller. "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors." *Review of Financial Studies*, 1 (1988a), 195–228.

Campbell, J. Y., and R. J. Shiller. "Stock Prices, Earnings and Expected Dividends." *Journal of Finance*, 43 (1988b), 661–676.

Campbell, J. Y., and M. Yogo. "Efficient Tests of Stock Return Predictability." *Journal of Financial Economics*, 81 (2006), 27–60.

Canina, L., and S. Figlewski. "The Informational Content of Implied Volatility." *The Review of Financial Studies*, 6 (1993), 659–681.

Christensen, B. J., and N. Prabhala. "The Relation Between Implied and Realized Volatility." *Journal of Financial Economics*, 50 (1998), 125–150.

Cootner, P. H. *The Random Character of Stock Market Prices*. MIT Press (1964).

Cowles, A. "Can Stock Market Forecasters Forecast?" *Econometrica*, 1 (1933), 309–324.

Cowles, A. "Stock Market Forecasting." *Econometrica*, 12 (1944), 206–214.

Drechsler, I., and A. Yaron. "What's Vol Got to Do with It." *Review of Financial Studies*, 24 (2011), 1–45.

Eraker, B., and I. Shaliastovich. "An Equilibrium Guide to Designing Affine Pricing Models." *Mathematical Finance*, 18 (2008), 519–543.

Eraker, B., and Y. Wu. "Explaining the Negative Returns to VIX Futures and ETNs: An Equilibrium Approach." *Journal of Financial Economics*, 125 (2017), 72–98.

Eraker, B., and A. Yang. "The Price of Higher Order Catastrophe Insurance: The Case of VIX Options." *Journal of Finance*, 77 (2022), 3289–3337.

Fama, E. F., and K. R. French. "Dividend Yields and Expected Stock Returns." *Journal of Financial Economics*, 22 (1988a), 3–25.

Fama, E. F. "Efficient Capital Markets: A Review of Theory and Empirical Work." *Journal of Finance*, 25 (1970), 383–417.

Fama, E. F., and K. R. French. "Permanent and Temporary Components of Stock Returns." *Journal of Political Economy*, 96 (1988b), 246–273.

French, K. R.; W. Schwert; and R. F. Stambaugh. "Expected Stock Returns and Volatility." *Journal of Financial Economics*, 19 (1987), 3–29.

Glosten, L. R.; R. Jagannathan; and D. E. Runkle. "On the Relation Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks." *Journal of Finance*, 5 (1993), 1779–1801.

Harvey, C. R. "Time-Varying Conditional Covariances in Tests of Asset Pricing Models." *Journal of Financial Economics*, 24 (1989), 289–317.

Harvey, C. R. "The World Price of Covariance Risk." *Journal of Finance*, 46 (1991), 111–157.

Hodrick, R. J. "Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement." *Review of Financial Studies*, 5 (1992), 357–386.

Huang, D.; C. Schlag; I. Shaliastovich; and J. Thimme. "Volatility-of-Volatility Risk." *Journal of Financial and Quantitative Analysis*, 54 (2019), 2423–2452.

Johnson, T. L. "Risk Premia and the VIX Term Structure." *Journal of Financial and Quantitative Analysis*, 52 (2017), 2461–2490.

Jordà, S. "Estimation and Inference of Impulse Responses by Local Projections." *American Economic Review*, 95 (2005), 161–182.

Keim, D. B., and R. F. Stambaugh. "Predicting Returns in the Stock and Bond Markets." *Journal of Financial Economics*, 17 (1986), 357–390.

Kelly, B., and H. Jiang. "Tail Risk and Asset Prices." *Review of Financial Studies*, 27 (2014), 2841–2871.

Kendall, M. G. "The Analysis of Economic Time-Series, Part I: Prices." *Journal of the Royal Statistical Society*, 96 (1953), 11–25.

Lockstoer, L. A., and P. C. Tetlock. "What Drives Anomaly Returns?" *Journal of Finance*, 75 (2020), 1417–1455.

Marcellino, M.; J. H. Stock; and M. W. Watson. "A Comparison of Direct and Iterated Multistep AR Methods for Forecasting Macroeconomic Time Series." *Journal of Econometrics*, 135 (2006), 499–526.

Menzly, L.; T. Santos; and P. Veronesi. "Understanding Predictability." *Journal of Political Economy*, 112 (2004), 1–47.

Merton, R. "An Intertemporal Capital Asset Pricing Model." *Econometrica*, 41 (1973), 867–887.

Merton, R. C. "On Estimating the Expected Return on the Market." *Journal of Financial Economics*, 8 (1980), 323–361.

Miller, M. H., and F. Modigliani. "Dividend Policy, Growth and the Valuation of Shares." *Journal of Business*, 34 (1961), 411–433.

Moreira, A., and T. Muir. "Volatility Managed Portfolios." *Journal of Finance*, 72 (2017), 1611–1644.

Newey, W., and K. D. West. "A Simple Positive Semi-Definite, Heteroscedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica*, 55 (1987), 703–708.

Plagborg-Møller, M., and C. K. Wolf. "Local Projections and VARs Estimate the Same Impulse Responses." *Econometrica*, 89 (2021), 955–980.

Pohl, W.; K. Schmedders; and O. Wilms. "Higher Order Effects in Asset Pricing Models with Long-Run Risks." *Journal of Finance*, 73 (2018), 1061–1111.

Samuelson, P. "Proof That Properly Anticipated Prices Fluctuate Randomly." *Industrial Management Review*, 6 (1965), 41–49.

Stambaugh, R. F. "Predictive Regressions." *Journal of Financial Economics*, 54 (1999), 375–421.

Vuolteenaho, T. "What Drives Firm-Level Stock Returns?" *Journal of Finance*, 57 (2002), 233–264.

Yang, A. "Understanding Negative Risk-Return Trade-offs." Working Paper, University of Wisconsin, Madison (2022).

Zhou, H. "Variance Risk Premia, Asset Predictability Puzzles, and Macroeconomic Uncertainty." *Annual Review of Financial Economics*, 10 (2018), 481–497.

Zviadadze, I. "Term Structure of Risk in Expected Returns." *Review of Financial Studies*, 34 (2021), 6032–6086.

https://doi.org/10.1017/S0022109024000218