Internet Appendix for "Pockets of Predictability"

LELAND E. FARMER, LAWRENCE SCHMIDT, and ALLAN TIMMERMANN*

ABSTRACT

Appendix I establishes analytically a general set of conditions under which the constant-coefficient return prediction model conventionally used in empirical work, holds to a close approximation. Appendix II gives further details on our nonparametric estimation approach such as choice of kernel function. Appendix III contains a set of cumulative sum of squared error differential plots comparing forecasts from our local kernel regressions versus forecasts from the benchmark prevailing mean model. Appendix IV discusses the Stambaugh bias in return prediction regressions. Appendix V contains further details of simulations from the macro-finance models that allow for time-varying risk premia, including how we calibrate the parameters of a range of canonical asset pricing models studied in the paper. Appendix VI further explains some of the challenges faced by classical asset pricing models fitted to predictive return regressions at short horizons. Appendix VII reports impulse response models for a baseline model with sticky expectations and a rational expectations model with the same choice of parameters but without any stickiness. Related to this, Appendix VIII explains how we calibrate the parameters for the sticky expectations model.

*Citation format: Farmer Leland E., Lawrence Schmidt, and Allan Timmermann, Internet Appendix for "Pockets of Predictability," *Journal of Finance,* DOI: 10.1111/jofi.13229. Please note: Wiley is not responsible for the content or functionality of any additional information provided by the authors. Any queries (other than missing material) should be directed to the authors of the article.

I. Pockets and Time-Varying Risk Premia

This appendix establishes a set of conditions under which the conventional constant-coefficient return prediction model (1) holds almost exactly within a fairly general class of endowment economies nesting many canonical asset pricing specifications considered in the literature. We parameterize cash flow risks and investor preferences in the economy, allowing for time-variation in either the quantity or the price of risk. To this end, let z_t be an $L \times 1$ vector of state variables capturing the aggregate state of the economy. We assume that this evolves according to the following law of motion.

ASSUMPTION IA.1: *The aggregate state of the economy follows a stationary VAR process,*

$$z_{t+1} = \mu + F z_t + \epsilon_{t+1}, \tag{IA.1}$$

with z_0 given, where all of the eigenvalues of the $L \times L$ matrix F are inside the unit circle and $E[\epsilon_{t+1}] = 0$. Moreover, the log of aggregate dividend growth, Δd_{t+1}, equals $S_d' z_{t+1}$ for some $L \times 1$ vector S_d.

Assumption IA.1, which is quite standard, states that aggregate dividend growth can be captured by a linear combination of the elements of a finite-dimensional, stationary vector autoregressive process, z_t. We place further restrictions on the vector of innovations below.

In addition to the restrictions on the cash flow process in Assumption IA.1, we restrict investor preferences. In particular, Assumption IA.2 imposes that the log risk-free rate and pricing kernel are "essentially" affine functions of the z_t vector that summarizes the aggregate state of the economy, possibly with time-varying prices of risk.

ASSUMPTION IA.2: *The continuously compounded risk-free rate, $r_{f,t+1}$, satisfies*

$$r_{f,t+1} = A_{0,f} + A_f' z_t, \tag{IA.2}$$

and the continuously compounded return on any financial asset, $r_{a,t+1}$, satisfies the Euler equation

$$1 = E_t[\exp(-\Lambda_t' \epsilon_{t+1} - \log E_t \exp[-\Lambda_t' \epsilon_{t+1}] + r_{a,t+1} - r_{f,t+1})], \tag{IA.3}$$

where Λ_t is an $L \times 1$ vector of risk prices.

A large class of models have risk-free rates and pricing kernels that fit into this class. For example, Assumption IA.2 holds approximately in a representative agent model where agents have Epstein and Zin (1989) preferences when aggregate consumption growth is also an affine function of the state vector.[2] Our results will therefore apply to many of the specifications considered in the literature on consumption-based asset pricing models with long-run risks and rare disasters. This property also holds in an incomplete markets setting with state-dependent higher moments of uninsurable idiosyncratic shocks.[3] We also allow, with some restrictions discussed below, for time-variation in the price of risk, Λ_t, which enables our results to nest many models that have been used to characterize the term structure of interest rates as well as the log-linearized stochastic discount factor of the Campbell-Cochrane habit formation model.

Finally, we provide two alternative sets of restrictions on risk prices and quantities that ensure that, up to a log-linear approximation, price-dividend ratios and market returns are exponential affine functions of z_t.[4] We also define a partition of the set of state variables z_t in a way that will be useful later.

ASSUMPTION IA.3: *Partition the state vector $z_t = [z'_{1t}, z'_{2t}]'$, where $dim(z_{1t}) = L_1 \leq L$. One of the following sets of conditions is satisfied:*

1. *Risk prices are constant: $\Lambda_t = \Lambda$. In addition, for any $\gamma \in \mathbb{R}^L$, $E_t[\epsilon_{t+1}] = 0$ and the conditional Laplace transform of ϵ_{t+1} satisfies*

$$\log E_t[\exp(\gamma' \epsilon_{t+1}) | z_t] = f(\gamma) + g(\gamma)' z_{1t}, \tag{IA.4}$$

   where $f(\gamma) \colon \mathbb{R}^L \to \mathbb{R}$ and $g(\gamma) \colon \mathbb{R}^L \to \mathbb{R}^{L_1}$.

2. *Risk prices satisfy $\Lambda_t = \Lambda_0 + \Lambda_1 z_{1t}$, where Λ_1 is an $L \times L_1$ matrix, and $\epsilon_{t+1} \overset{iid}{\sim} MVN(0, \Sigma)$, where Σ is a positive semi-definite matrix.*

Assumption IA.3 characterizes two sets of assumptions that are commonly made to get affine valuation ratios. First, we assume that risk prices are constant but risk quantities are time-varying. In this case, z_{1t} is the subset of variables (e.g., stochastic volatility and/or Poisson jump intensities) that are useful for predicting the quantity of risk, while z_{2t} contain

[2] See, for example, Bansal and Yaron (2004), Hansen et al. (2008), Eraker and Shaliastovich (2008), and Drechsler and Yaron (2011).

[3] See, for example, Constantinides and Duffie (1996), Constantinides and Ghosh (2017), Schmidt (2020), and Herskovic et al. (2015).

[4] Note that we can get exact exponential affine expressions for the price-dividend ratio and returns of dividend strips, that is, the value as of time t of a dividend paid at time $t+k$ for any k. The linearization is only necessary because the market return is a weighted average of these individual dividend strip returns that is not exactly affine in the state vector. Some authors, such as Lettau and Van Nieuwerburgh (2007), have elected to work with the exact dividend strip formulas.

additional variables useful for predicting cash flows or the risk-free rate. We summarize our main restriction on the distribution of ϵ_{t+1} in terms of its cumulant generating function, which is the logarithm of its moment-generating function. The affine structure greatly facilitates analytical tractability and is satisfied for a wide class of distributions used in the theoretical asset pricing literature.[5] For instance, suppose that $\epsilon_{t+1} \sim MVN(0, \sigma_t^2 \Sigma)$ for some positive semi-definite matrix Σ. Then $f(\gamma) = 0$ and $g(\gamma)' z_{1t} = \frac{1}{2}\gamma'\Sigma\gamma$ with $z_{1t} = \sigma_t^2$.

Second, we allow risk prices to be affine in a subset of the state variables, z_{1t}, but we restrict the innovations ϵ_{t+1} to be homoskedastic and multivariate normally distributed.[6] In this case, z_{1t} indicates the subset of variables that characterize time-variation in the price of risk Λ_t. These assumptions are quite common in the bond pricing literature as well as for models featuring time-varying risk aversion and are identical to those in Lustig et al. (2013), among others.

To solve for asset prices in this economy, we apply the Campbell and Shiller (1988) log-linearization of the stock market return, $r_{s,t+1}$, in excess of the risk-free rate, $r_{f,t+1}$, as a function of the log-dividend growth rate, Δd_{t+1}, and the log price-dividend ratios at time $t+1$ and t, pd_{t+1} and pd_t:

$$r_{s,t+1} \approx c + \Delta d_{t+1} + \rho \cdot pd_{t+1} - pd_t. \tag{IA.5}$$

Here, c and $\rho < 1$ are linearization constants. Using this linearization and Assumptions IA.1 to IA. 3, we can show the following result.

PROPOSITION IA.1: *Suppose Assumptions IA.1, IA.2, and IA.3 are valid and that a solution exists to the log-linearized asset pricing model. Then the following properties are satisfied:*

(i) The market price-dividend ratio is

$$pd_t = A_{0,m} + A_m' z_t;$$

[5]For example, the property holds for affine jump-diffusion models, for example, Eraker and Shaliastovich (2008) and Drechsler and Yaron (2011). In these models, ϵ_{t+1} is the sum of Gaussian and jump components and the variance-covariance matrix for the Gaussian shocks and the arrival intensities for the jump shocks are affine functions of z_t. See also Bekaert and Engstrom (2017) and Creal and Wu (2016) for alternative stochastic processes with affine cumulant generating functions.

[6]Creal and Wu (2016) provide some restrictions that permit both risk prices and quantities to vary while keeping valuation ratios in the affine class. We do not detail these assumptions here, but note that the constant-coefficient result should obtain for this more general case as well.

(ii) The expected excess return is

$$E_t[r_{s,t+1}] - r_{f,t+1} = \beta_0 + \beta' z_{1t},$$

where $A_{0,m}, A_{0,f}, and \beta_0$ *are scalars,* $A_m \in \mathbb{R}^L$*, and* $\beta \in \mathbb{R}^d$*.*

Part (i) of Proposition IA.1 shows that the log price-dividend ratio is an affine function of the aggregate state vector, which immediately implies that the log-linearized market return is also an affine function of z_t and ϵ_{t+1}. Part (ii) of the proposition characterizes the extent of return predictability. It shows that risk premia–expected log excess returns–are an affine function of z_{1t}, variables used to forecast cash flows and the risk-free rate. For a set of predictors x_t chosen to be elements of the underlying state variables (z_{1t}), Proposition IA.1 justifies using constant-coefficient linear return prediction models of the form in (1).

Part (ii) of Proposition IA.1 also indicates the extent to which the theory allows for some degree of dimension reduction. In principle, one could allow for a very large number of state variables to predict cash flow growth, each of which could have innovations which may even be priced. Nonetheless, if these variables do not predict time-variation in the quantity of risk (under the conditions of Assumption IA. 3, part 1) or the price of risk (under the conditions of Assumption IA. 3, part 2), they may safely be omitted from the predictive regression. In contrast, if the true state variables z_{1t} are not spanned by the choice of predictors, x_t, included in the return regression, as could be the case if there are additional drivers of risk prices or quantities omitted from the regression, it need not necessarily be the case that the projection of $r_{s,t+1} - r_{f,t+1}$ on the empirical proxies would have constant coefficients.

As is the case for many asset pricing tests, it is worth emphasizing that we can only test the joint hypothesis that the model is correctly specified (i.e., we have the correct predictors) and the theoretical restrictions (constant coefficients) hold. Thus, an important caveat on interpretations of our results is that any evidence that is inconsistent with the null of constant coefficients could potentially be explained by omitted factors, as opposed to our incomplete cash flow learning story.

Proof: To show part (i) of Proposition IA.1, we conjecture and verify that the price-dividend ratio is $pd_t = A_{0,m} + A'_m z_t$.

By Assumption IA.1, $\Delta d_t = S'_d z_t$. Suppose that Assumption IA.3, part 1 holds. Using $r_{s,t+1} \approx k + \rho(p_{t+1} - d_{t+1}) + \Delta d_{t+1} + d_t - p_t$ and plugging the log-linearized return into the

Euler equation, we have

$$\begin{aligned}
1 &= \exp[-A_{0,f} - A'_f z_t - \log E_t \exp[-\Lambda'_t \epsilon_{t+1}] + \kappa + (\rho - 1)A_{0,m} - A'_m z_t] \\
&\times \; E_t \left[\exp\left\{-\Lambda' \epsilon_{t+1} + [S'_d + \rho A'_m] z_{t+1}\right\}\right] \\
0 &= -A_{0,f} - A'_f z_t + \kappa + (\rho - 1)A_{0,m} - A'_m z_t + [S'_d + \rho A'_m](\mu + F z_t) \\
&+ \; [f(-\Lambda' + S'_d + \rho A'_m) - f(-\Lambda)] + [\tilde{g}(-\Lambda' + S'_d + \rho A'_m)' - \tilde{g}(-\Lambda')'] z_t,
\end{aligned}$$

where $\tilde{g}(u) \equiv [g(u)', \mathbf{0}']'$ and the second line takes logs and applies Assumption IA.1(ii). Rearranging yields the $(L+1)$-dimensional system of equations in $A_{0,m}$ and A_m:

$$\begin{aligned}
f(-\Lambda + S_d + \rho A_m) - f(-\Lambda) - A_{0,f} + \kappa + (\rho - 1)A_{0,m} + (S'_d + \rho A'_m)\mu &= 0, \\
\tilde{g}(-\Lambda + S_d + \rho A_m) - \tilde{g}(-\Lambda) - A_f - (I - \rho F')A_m + F' S_d &= 0.
\end{aligned}$$

This system does not have an analytical solution in the general case, but, it is relatively straightforward to solve the system numerically. We note that Assumption IA.3.part 2 for the data-generating process is identical to those in Lustig et al. (2013). Therefore, we refer the interested reader to the proof of their Proposition 1 for full derivations of the $A_{0,m}$ and A_m coefficients in that case.

To show part ii, we follow a very similar argument to Drechsler and Yaron (2011). We can write expected returns as follows (using the normalization $E_t[\epsilon_{t+1}] = 0$):[7]

$$\begin{aligned}
E_t[\exp(r_{s,t+1})] &= \exp[E_t r_{s,t+1}] E_t[\exp([S'_d + \rho A'_m]\epsilon_{t+1})] \equiv \exp[E_t r_{s,t+1}] E_t[\exp(B'_m \epsilon_{t+1})] \\
\exp(-r_{f,t+1}) &\equiv \exp[E_t m_{t+1}] E_t[\exp(-\Lambda'_t \epsilon_{t+1})].
\end{aligned}$$

Next, using the Euler equation in (IA.3) and the law of iterated expectations, we have

$$\begin{aligned}
1 &= \exp[E_t r_{s,t+1}] \exp[E_t m_{t+1}] E_t \exp[(-\Lambda'_t + B'_m)\epsilon_{t+1}] \\
\frac{E_t \exp[B'_m \epsilon_{t+1}] E_t \exp[-\Lambda_t \epsilon_{t+1}]}{E_t \exp[(-\Lambda'_t + B'_m)\epsilon_{t+1}]} &= \exp[E_t r_{s,t+1}] \exp[E_t m_{t+1}] E_t \exp[B'_m \epsilon_{t+1}] E_t \exp[-\Lambda'_t \epsilon_{t+1}] \\
&= E_t[\exp(r_{s,t+1})] \exp(-r_{f,t+1}).
\end{aligned}$$

[7] Note that this normalization is for convenience. Given our assumptions on the relationship between the distribution of ϵ_{t+1} and the state vector, the mean of ϵ_{t+1} would be affine in z_t in the absence of this normalizing assumption. Therefore, we could always include this additional term in μ and F in equation (IA.1).

Taking logs and noting that $E_t r_{s,t+1} = \log E_t \exp(r_{s,t+1}) - \log E_t \exp[B'_m \epsilon_{t+1}]$, we get

$$E_t[r_{s,t+1}] - r_{f,t+1} = \log E_t \exp[-\Lambda_t \epsilon_{t+1}] - \log E_t \exp[(-\Lambda'_t + B'_m)\epsilon_{t+1}]. \tag{IA.6}$$

Suppose that Assumption IA.3, part 1 holds. Then (IA.6) simplifies to

$$E_t[r_{s,t+1}] - r_{f,t+1} = f(-\Lambda) - f(-\Lambda + B'_m) + [g(-\Lambda) - g(-\Lambda + B'_m)]' z_{1t}, \tag{IA.7}$$

which establishes the claim. If Assumption IA.3 part 2 holds, we can evaluate each of the expressions in (IA.6) using the cumulant generating function of the normal distribution:

$$E_t[r_{s,t+1}] - r_{f,t+1} = -\tfrac{1}{2} B'_m \Sigma B_m + B'_m \Sigma \Lambda_t = -\tfrac{1}{2} B'_m \Sigma B_m + B'_m \Sigma [\Lambda_0 + \Lambda_1 z_{1t}], \tag{IA.8}$$

which also establishes the claim. The first term is due to Jensen's inequality, while the second captures the covariance between the market return and the priced risk factors. Collecting terms in front of z_{1t} in the two equations above yields the expressions for β under the two sets of assumptions. $\square$

II. Details on the Nonparametric Estimation

This appendix describes our nonparametric estimation approach. Robinson (1989) and Cai (2007) consider local constant and local linear approximations of β, respectively, but this approach can easily be generalized to accommodate polynomials of arbitrary order. In particular, we can approximate the function β_t as a p^{th}-order Taylor expansion about the point $\frac{t}{T}$ (where $p \geq 0$). To this end, define the quantities

$$\begin{aligned}
\mathbf{W}_{st} &= \left(1, \frac{s-t}{T}, \ldots, \left(\frac{s-t}{T}\right)^p\right)', \\
K_{st} &= K\left(\frac{s-t}{hT}\right), \\
\mathbf{Q}_{st} &= \mathbf{W}_{st} \otimes x_s,
\end{aligned}$$

for $s, t = 1, \ldots, T$, where K is a kernel function and $h \equiv h(T)$ is the bandwidth. More formally, $K : [-1, 1] \to \mathbb{R}^+$ is a function that is symmetric about zero and integrates to one, and $h \in [0, 1]$ satisfies $h \to 0$ and $hT \to \infty$ as $T \to \infty$.

The local polynomial estimator $\beta = \left(\beta_0', \beta_1', \ldots, \beta_p'\right)'$ is obtained by solving

$$\min_{\beta \in \mathbb{R}^{pd}} \sum_{s=t-\lfloor hT \rfloor}^{t+\lfloor hT \rfloor} K_{st} \left[r_{s+1} - \beta_0' x_s - \beta_1' \left(\frac{s-t}{T} \right) x_s - \ldots - \beta_p' \left(\frac{s-t}{T} \right)^p x_s \right]^2$$
$$= \sum_{s=t-\lfloor hT \rfloor}^{t+\lfloor hT \rfloor} K_{st} \left(r_{s+1} - \beta' \mathbf{Q}_{st} \right)^2 .$$

Solving this optimization problem for β gives the solution

$$\hat{\beta}_t = \left(\sum_{s=t-\lfloor Th \rfloor}^{t+\lfloor Th \rfloor} K_{st} \mathbf{Q}_{st} \mathbf{Q}_{st}' \right)^{-1} \sum_{s=t-\lfloor Th \rfloor}^{t+\lfloor Th \rfloor} K_{st} \mathbf{Q}_{st} r_{s+1}. \tag{IA.9}$$

Our object of interest, $\hat{\beta}_{1t}$, is the first element of $\hat{\beta}_t$, which is given by

$$\hat{\boldsymbol{\beta}}_{1t} = \left(\mathbf{e}_1' \otimes \mathbf{I}_d \right) \hat{\beta}_t,$$

where $\mathbf{e}_1$ is the first standard basis vector of $\mathbb{R}^{p+1}$, $\mathbf{I}_d$ is a $(d \times d)$ identity matrix, and d is the dimension of x_t. This can also be thought of as the OLS estimator of β_0 in the transformed model

$$K_{st}^{1/2} y_{s+1} = K_{st}^{1/2} x_s' \sum_{q=0}^{p} \beta_q + \varepsilon_{s+1}.$$

The asymptotic properties of these estimators are studied in Robinson (1989) and Cai (2007). Under various regularity conditions, it can be shown that the estimator $\hat{\boldsymbol{\beta}}_t$ in (IA.9) is consistent and asymptotically normal.

Our main empirical results adopt a local constant (Nadarya-Watson) estimation procedure and so set $p = 0$. The motivation behind this choice is that the nonparametric procedures require very large amounts of data to perform well in finite samples and every additional degree of approximation requires that we estimate dT additional parameters.

III. Cumsum Plots

To get a sense of how predictive accuracy evolves over time, Figure IA.1 follows Welch and Goyal (2008) and plots the cumulative sum of squared forecast error differentials using real-time forecasts from our local kernel regressions versus forecasts from the prevailing mean model, that is,

$$CSSED_t = \sum_{\tau=t_0}^{t} \left(\bar{e}^2_{\tau|\tau-1} - \hat{e}^2_{\tau|\tau-1}\right).$$

Here, $\bar{e}^2_{\tau|\tau-1}$ and $\hat{e}^2_{\tau|\tau-1}$ are the squared forecast errors from the prevailing mean and local kernel regression models, respectively, and t_0 is the initial data point in the (out-of-sample) test period. Positive and increasing values of $CSSED_t$ indicate periods in which the local kernel regression produces smaller squared forecast errors than the prevailing mean and thus is more accurate; periods with declining (negative) values show the reverse. Pockets are marked in grey vertical bars.

Panels in the top row show how in-pocket predictive accuracy evolves by letting the $CSSED_t$ line be flat outside pockets while panels in the bottom row do the reverse, flatlining the $CSSED_t$ curve in the pockets and tracking how it evolves outside the pockets. For the univariate T-bill rate model, the $CSSED_t$ curve rises inside most pockets (top panel) while it systematically declines and is negative outside the pockets (bottom panel).

IV. Stambaugh Bias

In cases in which the predictor variable follows a highly persistent process and the correlation between innovations to the predictor variable and shocks to the return equation is large, Stambaugh (1999) shows that the estimated slope coefficient in equation (2) can be subject to a potentially large finite-sample bias. Both conditions are satisfied in our return regression that uses the dividend-price ratio; in particular, the estimated persistence of the daily dividend-price ratio series is 0.9995.

The Stambaugh bias affects inference based on the estimated slope coefficient $\hat{\beta}_t$. However, it does not lead our approach to spuriously identify out-of-sample pockets. Biases in the local regression estimates of β_t will tend to reduce the accuracy of our time-varying return forecasts, leading to fewer periods in which $SED_t > 0$ and *fewer* pockets. Rather than making the pockets that we identify spurious, this reduces their number.

Still, biases in estimated slope coefficients could affect which pockets get identified through its effect on our $\widehat{SED}_t$ measure, so we next explore this point through Monte Carlo simulations.

First, we generate joint standard normal random variables with a correlation of $\rho_{r,x}$,

$$\begin{bmatrix} v_{r,t+1} \\ v_{x,t+1} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho_{r,x} \\ \rho_{r,x} & 1 \end{bmatrix}\right), \tag{IA.10}$$

where $\rho_{r,x}$ takes values of $[-0.5, -0.8, -0.9, -0.95, -0.99]$. We next convert these normal draws to uniform random variables by evaluating the standard normal cdf of each series, $\{\Phi(v_{r,t+1}), \Phi(v_{x,t+1})\}$. Let $\hat{Q}_r$ and $\hat{Q}_x$ denote the empirical quantile functions of the normalized residuals from the estimated EGARCH(1,1) model (11), $\{\hat{u}_{r,t+1}\}$ and $\{\hat{u}_{x,t+1}\}$, respectively. We convert the uniform random variables to bootstrap samples of the normalized residuals by evaluating them at their respective empirical quantile functions

$$\left\{\hat{Q}_r(\Phi(v_{r,t+1})), \hat{Q}_x(\Phi(v_{x,t+1}))\right\}.$$

Simulation results (reported in Table IA.IV) show that the simulated statistical models become slightly worse at matching the number of pockets and the fraction of significant observations inside pockets as the correlation parameter is reduced from -0.5 to -0.99. For example, for the EGARCH model the p-value of the alpha t-statistic decreases from 0.10 when $\rho_{r,x} = -0.50$ to 0.01 when $\rho_{r,x} = -0.99$.

Overall, however, changes in the correlation $\rho_{r,x}$ have only a modest effect on the simulation results. Thus, while the assumed correlation is important to *inference* about the significance of predictive return regressions, it matters far less to out-of-sample forecasting performance. This is because, in practice, the bias in $\widehat{\beta}_t$ is small relative to the variation in local return predictability picked up by our local return regressions.

V. Details on Simulations from Macro-Finance Models with Time-Varying Risk Premia

In this appendix, we discuss details related to the simulation exercises described in Section IV of the main text, in which we generate sample paths of daily asset returns and state variables from four workhorse asset pricing models with time-varying risk premia. In each case, we focus on versions of these models that are solved in continuous time, making it straightforward to discretize to a daily frequency. Below we provide details on the individual models, but we begin by discussing some features of our analysis that are common across all the models we consider.

A. *Overview of Simulation Procedure*

In each of the models we consider, standard no-arbitrage conditions hold, and thus there exists a unique stochastic discount factor Λ_t that prices all shocks in each economy. Three of the four models we consider satisfy assumptions for a representative agent to exist, while the fourth model has dynamically complete markets. In this latter case, optimal risk-sharing conditions pin down the functional form of the stochastic discount factor.

Characterizing the solution to these models usually proceeds in two steps. First, we need to characterize the properties of the stochastic discount factor (SDF) given the primitives of the problem. For three of the four models we consider, this requires solving a set of partial differential equations (PDEs). The functional form for the SDF also allows us to characterize the risk-free rate (except for the rare disaster model, in which there is a wedge between the short-term interest rate and the mean of the SDF). Second, we price the set of cash flows associated with the stock market, that is, the aggregate dividend, D_t. Defining the price-dividend ratio by ξ, the usual no-arbitrage argument implies that it satisfies the equation

$$\Lambda_t D_t dt + E_t[d(\Lambda_t \xi_t D_t)] = 0, \tag{IA.11}$$

which gives us a PDE that characterizes the behavior of the price-dividend ratio, a second state variable that we consider for our predictive regressions. Then, given the price-dividend ratio, it is straightforward to compute excess stock returns. Given a time series of realized returns, we can then compute realized volatility, which gives us a third state variable ($rvar$) to use in our simulation exercises. For the first three of the four models under consideration, we use the EconPDE Julia package developed by Matthieu Gomez, as well as his codes that compute the solutions to each.[8] In the final case (Wachter (2013)), we run replication codes from the original paper kindly shared with us by Jessica Wachter.

For each of the four models we consider, we generate 1 million years of daily observations (i.e., 252 million "trading day" observations). Since the model is stationary, we then randomly select a starting point in this history from which to extract a daily time series that has the same length as the sample period we use for our out-of-sample empirical analyses in the data. We then run our same codes on these simulated data points to assess the extent to which these models can replicate our evidence.

Next, we discuss the basic setup of the four models we consider. As the analysis is somewhat more transparent for the Campbell and Cochrane (1999) model, we illustrate each of these steps fairly explicitly. We proceed analogously for the other models, but refer the readers to the relevant papers for more explicit characterizations of the associated PDEs.

[8] We are extremely grateful to Matthieu Gomez for making these codes available, as they greatly facilitated our work on this project.

B. Campbell and Cochrane (1999)

We begin by discussing the model of Campbell and Cochrane (1999), where investors have preferences that feature "habit formation," which features a single state variable capturing investors' "habit level" of consumption that generates time-variation in the effective risk aversion.[9] The model is described by one state variable, namely, the surplus consumption ratio $S_t = \frac{C_t - X_t}{C_t}$, $s_t = \log S_t$, where X_t is a reference point for consumption. The surplus ratio enters the utility of the agent and the SDF has the form $\Lambda_t = e^{-\rho t} \left(S_t C_t\right)^{-\gamma}$. Following a sequence of bad shocks, risk aversion and risk premia rise, lowering asset prices.

Specifically, aggregate consumption growth follows a geometric Brownian motion with constant drift. The log surplus ratio is a mean-reverting process with state-dependent volatility σ_s:

$$\begin{aligned} d\log C &= \mu dt + \sigma dW_t \\ ds &= \underbrace{-\kappa_S(s-\bar{s})}_{\mu_s} dt + \underbrace{\lambda(s-\bar{s})\cdot\sigma}_{\sigma_s} dW_t, \end{aligned}$$

where $\bar{s} = \log \bar{S}$, $\bar{S} = \sigma \cdot \sqrt{\frac{\gamma}{\kappa_S - \frac{b}{\gamma}}}$. A set of restrictions on state s gives $\lambda(s-\bar{s}) = \frac{1}{\bar{S}}\sqrt{1-2\cdot(s-\bar{s})} - 1$, $s \in (-\infty, \bar{S}]$, where $s^{max} = \bar{s} + \frac{1}{2}(1-\bar{S}^2)$. For simplicity, we simply assume that the stock market represents a claim on the aggregate stock market. Thus, general equilibrium is established by $D_t = C_t$. (For simplicity, we follow this approach rather than characterize the price of a levered claim on aggregate consumption, which is also standard.)

We then conjecture the following form of the SDF,

$$d\log\Lambda_t = (-r - \frac{\kappa^2}{2})dt - \kappa dW_t,$$

and an Ito process for the price-dividend ratio ξ,

$$\frac{d\xi}{\xi} = \mu_\xi dt + \sigma_\xi dW_t.$$

Plugging processes for s_t and $c_t := \log C_t$ into the SDF and applying Ito's Lemma yields

[9] We use the continuous-time version of the calibration from Wachter (2005), which also allows habit to affect the risk-free interest rate. We refer readers to that paper for further details.

equations for the market price of risk and the interest rate:

$$\kappa = \gamma(\sigma_s + \sigma),$$
$$r = \rho + \gamma(\mu_s + \mu) - \frac{\kappa^2}{2},$$

or

$$\kappa = \frac{\gamma\sigma}{\bar{S}}\sqrt{1 - 2\cdot(s - \bar{s})} = \sqrt{(\gamma\kappa_S - b)(1 - 2\cdot(s - \bar{s}))}$$
$$r = \rho + \gamma(\mu - \frac{\kappa_s}{2}) + \frac{b}{2} + b(\bar{s} - s).$$

Plugging the processes for C_t, ξ_t, and Λ_t into (IA.11), we get

$$0 = \frac{1}{\xi} - r + \mu_\xi + (\mu + \frac{\sigma^2}{2}) - (\sigma_\xi + \sigma)\kappa + \sigma_\xi\sigma.$$

Hence, we are looking for a solution of the PDE $\dot{\xi} = 0$ with

$$\frac{\dot{\xi}}{\xi} = \frac{1}{\xi} - r + \mu_\xi + (\mu + \frac{\sigma^2}{2}) - (\sigma_\xi + \sigma)\kappa + \sigma_\xi\sigma,$$

where μ_ξ and σ_ξ are identified by Ito's Lemma,

$$\mu_\xi = \frac{\partial\xi}{\partial s}\cdot\frac{\mu_s}{\xi} + \frac{1}{2}\cdot\frac{\partial^2\xi}{\partial s^2}\cdot\frac{\sigma_s^2}{\xi} \qquad \sigma_\xi = \frac{\partial\xi}{\partial s}\cdot\frac{\sigma_s}{\xi}.$$

Our calibration of the associated parameters comes from the solution of the discrete-time Campbell-Cochrane model by Wachter (2005), which are translated to continuous time as follows:

- Consumption growth drift $\mu = 0.022$
- Consumption growth volatility $\sigma = 0.0086$
- Relative risk aversion $\gamma = 2$
- Rate of time preference parameter $\rho = 0.072$
- Surplus consumption parameters $\kappa_s = 0.11$ and $b = 0.011$

The parameters identify $\bar{S} = 0.376$ and $\bar{s} = -3.28$. We find the solution numerically on a grid with 798 points.

C. *Bansal and Yaron (2004)*

Next, we consider a continuous-time version of the long-run risk model of Bansal and Yaron (2004), which is well described in Chen et al. (2009). In the model, investors have recursive preferences, $U_t = \mathbb{E}_t\left[\int_t^\infty f(c_u, U_u)du\right]$,with aggregator

$$f(c,U) = \frac{1}{1-\psi^{-1}}\left[\frac{\rho c^{1-\psi^{-1}}}{[(1-\gamma)U]^{(\gamma-\psi^{-1})/(1-\gamma)}} - \rho(1-\gamma)U\right],$$

the properties of which are analyzed in Duffie and Epstein (1992). The model is characterized by processes for the consumption growth drift and stochastic volatility:

$$\begin{aligned}
d\mu &= k_\mu(\bar{\mu}-\mu)dt + \nu_\mu\sqrt{v}(dW_t^1),\\
dv &= k_v(1-v)dt + \nu_v\sqrt{v}dW_t^2,\\
dC/C &= \mu dt + \sqrt{v}(\nu_{c,1}dW_t^1 + \nu_{c,3}dW_t^3).
\end{aligned}$$

Hence, the model is described by two state variables, μ and v. Following Hansen et al. (2018), who solve a version of the same model in continuous time as part of their mfrSuite package, we allow shocks to consumption growth and μ to be contemporaneously correlated. For the stock return, we price a levered claim on aggregate consumption that pays $D_t = C_t^\phi$.

Define $x := (\mu, v)$. We conjecture Ito processes for the SDF,

$$\frac{d\Lambda_t}{\Lambda_t} = -rdt - \kappa_\mu dW_t^1 - \kappa_v dW_t^2 - \kappa_c dW_t^3,$$

and the price-dividend ratio,

$$\frac{d\xi(x)}{\xi(x)} = \mu_\xi dt + \sigma_{\xi,\mu}dW_t^1 + \sigma_{\xi,v}dW_t^2.$$

Then, we can use the properties of the utility function to characterize a PDE in x that yields the prices of risk. We omit these details for brevity. We then price the aggregate dividend ratio by applying the no-arbitrage condition above.

The state space has two dimensions, so we solve it on a two-dimensional grid. The grid G_x is a product of grids for μ and v, $G_x = G_\mu \otimes G_v$. We choose a G_μ consisting of 90 points for μ and a G_v consisting of 90 points for v, giving us $8,100$ grid points in total. We adapt the Julia codes from the EconPDE package to use the calibrated parameters from Hansen et al. (2018).

Specifically, we assume the following annualized values for each of the parameters above:

- Consumption growth $\bar{\mu} = 0.0015 \times 12$
- Average shock variance $\bar{v} = 1$
- Annual persistence coefficients $k_\mu = 0.021 \times 12$ and $k_v = 0.013 \times 12$
- Average shock volatilities $\nu_\mu = 0.000344384 \times \sqrt{12} \times 12$, $\nu_v = 0.038 \times \sqrt{12}$, $\nu_{c,1} = 0.000011615 \times \sqrt{12}$, and $\nu_{c,3} = -0.00778202 \times \sqrt{12}$
- Rate of time preference parameter $\rho = 0.024$
- Relative risk aversion $\gamma = 7.5$ and elasticity of intertemporal substitution $\psi = 1.5$.

D. *Gârleanu and Panageas (2015)*

Gârleanu and Panageas (2015) consider an overlapping generations model with two types of agents, with types $i = A, B$ characterizing heterogeneity in their preferences. A agents constitute a smaller fraction of more risk-loving agents.[10] Aggregate consumption follows a geometric Brownian motion with drift. We omit most details of the model because we follow the original paper, which is already in continuous time, as closely as possible. Agents can write contracts to share risks associated with fluctuations in the aggregate endowment and have access to a set of annuity contracts that insure against longevity risk.

For the purpose of solving the model, the key is that the model has only one state variable X_t, which is the consumption share of type-A agents, where total consumption is aggregated over all generations of type-A agents. Intuitively, the effective level of risk aversion in the economy is lower when X_t is high, which generates time-variation in the price of risk on shocks to the aggregate endowment. The state variable follows an Ito process that is driven by one shock only (the shock to the aggregate endowment),

$$dX_t = \mu_X(X_t)dt + \sigma_X(X_t)dW_t.$$

As in Bansal and Yaron (2004), each type of agent has recursive Duffie and Epstein (1992) preferences. Lifetime utility of every agent of type i with wealth W can be represented as

$$U(W, x) = \frac{W^{1-\gamma^i}(X_t)}{1-\gamma^i} \cdot \left(g^i(X_t)\right)^{-\frac{\psi^{-1}(1-\gamma^i)}{(1-(\psi^i)^{-1})}},$$

[10] For instance, we might interpret such agents as entrepreneurs in other models.

where $g^i(X_t)$ is the consumption-to-wealth ratio.[11] For more robust convergence of the finite-difference method, we solve the model in terms of the inverse of the consumption-to-wealth ratio $\varsigma^i(x) = (g^i(x))^{-1}$ for each type of agent. To define all state-dependent parameters we need to solve four functions, namely, the wealth-to-consumption ratios for both agents $\{\varsigma(x)^i\}_{i=A,B}$ and functions $\{\phi(x)^j\}_{j=1,2}$, which can be interpreted as capturing the price of a claim on a pre-specified cash flow.[12] Following the solution approach in the paper, we can get a set of PDEs that capture each one of these functions, prices of risk, and the value of a claim on aggregate capital income. For simplicity, we assume that the stock market in the model corresponds to the value of an unlevered claim on capital income.

E. Wachter (2013)

Wachter (2013) considers a representative agent economy in which aggregate consumption and the dividend on the aggregate stock market are exposed to the risk of rare disasters, that is, large downward jumps in the aggregate endowment. As in the Bansal-Yaron model above, investors have Epstein-Zin preferences. (As earlier, we omit expressions for the SDF and PDEs for valuation ratios in this subsection for brevity.)

Aggregate consumption evolves according to

$$dC_t = \mu C_{t^-} dt + \sigma C_{t^-} dB_t + (e^{Z_t} - 1) C_{t^-} dN_t, \tag{IA.12}$$

where B_t is a standard Brownian motion and N_t is a Poisson process with time-varying intensity λ_t, which evolves according to

$$d\lambda_t = \kappa(\bar{\lambda} - \lambda_t)dt + \sigma_\lambda \sqrt{\lambda_t} dB_{\lambda,t}, \tag{IA.13}$$

where $B_{\lambda,t}$ is also a Brownian motion, and B_t, $B_{\lambda,t}$, and N_t are mutually independent. Dividends are modeled as a levered claim on consumption, that is, $D_t = C_t^\phi$, where $\phi = 2.6$. In addition, the model allows for partial default on government debt if a disaster occurs.

Our parameter values and solution approach are identical to those in Wachter (2013); accordingly, we refer the reader to that paper for further technical details.[13] Consistent with conventions in the rare disaster literature, we consider two sets of simulation exercises: one

[11] For example, $g^A = \frac{C_t^A}{W_t^A} = (\varsigma^A)^{-1}$.

[12] Specifically, each captures the price at time t of a claim on $B_j \times e^{-(\pi+\delta_j)(s-t)} \frac{Y_s}{Y_t}$ for every $s \geq t$, where Y_s is aggregate consumption/production, and other parameters come from the paper.

[13] We are extremely grateful to Jessica Wachter for kindly providing the replication codes, from which we simulated data from the model.

in which we include sample paths with disasters, and another in which we focus exclusively on sample paths in which no disaster occurs.

VI. Moments of Classical Asset Pricing Models

In this appendix we discusses the challenges classical asset pricing models face in estimating predictive regressions at short horizons, especially with fairly short data samples as is the case of our kernel regressions. Table IA.XIV helps illustrate these challenges by reporting a number of moments from the asset pricing models covered in our analysis.

First, there is potential for model misspecification in simple univariate return forecasts, which occurs because the observed state variable(s) may not map one-to-one into the equity risk premium and because this mapping may not be linear. For instance, the price-dividend ratio encodes information about future expected cash flow growth, risk premia, and real interest rates in addition to the current equity risk premium, introducing an errors-in-variables problem in univariate predictive regressions of excess returns on the dividend-price ratio. As an example of this, the primary driver of the risk premium in the Bansal-Yaron model is a variable capturing stochastic volatility (ν), which has a correlation of only 14% and -3% with the dp ratio and risk-free rate, respectively, and a higher though still imperfect 75% correlation with $rvar$. (In contrast, correlations are quite high with the expected growth rate μ in this calibration.) The Campbell-Cochrane, Garleanu-Panageas, and Wachter models all feature a single state variable, so only nonlinearities can bring correlations below unity in absolute value. In any case, the dp ratio tracks the relevant state variable with correlations often exceeding 90% (Table IA.XIV, Panel A).

Second, the signal-to-noise ratio is extremely low, especially at a daily frequency. To illustrate this, Panel B in Table IA.XIV reports the daily R^2 associated with univariate regressions of each of the simulated predictor variables, as well as the true risk premium, for each of the models in question at various horizons. While there is a modest amount of predictability over the span of multiple years, signal-to-noise ratios tend to be extremely low at short horizons. Therefore, given that regressors are quite persistent at a daily frequency, one would generally expect to see very poor finite-sample performance of regressions estimated with only a few years of daily data.

Finally, because many of the models considered here replicate the third property mentioned above (discount rates vary mostly due to changes in risk premia rather than risk-free rates), increases in risk premia captured by the state variables in each model are associated with sharply negative realized returns. As a result, to compound the challenges of a low signal-to-noise ratio, the Stambaugh (1999) bias is a very serious concern for regressions of returns on the price-dividend ratio, especially in very short samples. Since many of these models involve only a single state variable, other variables such as the risk-free rate are often subject to nontrivial biases coming from a correlation between shocks to the predictor variables and realized returns (even if data analogs to the correlations relevant for assessing the magnitude of the Stambaugh bias suggest that the problem should be less pronounced for these variables). Predictors are usually quite persistent and there are often strong negative correlations between realized returns and our state variables of interest (Table IA.XIV, Panels C and D). These factors combine to suggest that in-sample predictive regression coefficients would likely be associated with substantial attenuation bias.

VII. Impulse Responses

Figure IA.5 simulates impulse responses resulting from very large shocks to $z_{dr,t}$ and $z_{cf,t}$. Specifically, we consider the response with a size equal to $\sqrt{252/4} = \sqrt{63}$ times the standard deviation of a daily shock to each variable, which corresponds roughly to the amount of variation the model would generate in a single quarter. We then consider two configurations of the model, namely, baseline calibrated model with sticky expectations (we discuss our calibration approach below) and a rational expectations model with the same parameters except the stickiness parameter λ is set to zero. Responses to shocks to subjective risk premia (left panel) and orthogonal shocks to the risk-free rate $\epsilon_{tp,t}$ (not pictured) are identical across the two models. In both cases, large upward revisions in subjective discount rates trigger large negative return realizations that are gradually offset by modest increases in expected returns over the medium term. Such a pattern creates substantial scope for Stambaugh (1999) bias.

In contrast, the two models differ substantially in terms of how expected returns and state variables respond to a large shock to expected cash flows $\epsilon_{cf,t}$ (right panel). In the rational expectations model, such a shock generates a one-time, large realized return and a very modest change in the risk-free rate. However, in the sticky expectations model, responses of both the dividend-price ratio and risk-free rate are hump-shaped, where the gap between the rational and sticky model impulse response functions closes essentially to zero within about half a year. These sluggish adjustments yield a modest amount of predictability in expected returns that decays towards zero fairly quickly, contrasting sharply with the spike obtained in the rational expectations model.

It is also easy to see why performance of predictive regressions can be unstable in this environment. Both dp_t and $r_{f,t+1}$ load linearly, albeit with different weights, on $z_{dr,t}$, agents' subjective beliefs of cash flow growth $F_t[\Delta d_{t+1}] = z_{cf,t} - \vartheta_t$, as well as other factors. On average, $F_t[\Delta d_{t+1}]$ is positively correlated with ϑ_t, since both are moving averages of $\{\epsilon_{cf,t-j}\}_{j=0}^{\infty}$ with strictly positive weights. Depending on the sequence of shocks experienced, recent level changes in each state variable may reflect different combinations of these factors at different times.

VIII. Calibration of Parameters for Sticky Expectations Model

To assess the potential quantitative importance of a sticky beliefs mechanism in explaining our empirical results, we calibrate the parameters of the simple model outlined in equations (17) to (23). We first fix parameters related to the annualized means of dividend growth, the risk free rate, and expected returns at 5.3%, 1.5%, and 7%, respectively, to match the sample average of pd_t, $E[r_{f,t+1}]$, and $E[r_{t+1}]$. We set the linearization point at $E[pd_t]$ when selecting values of κ and ρ.

Given the central importance of the sticky expectations channel, we discipline parameters governing the degree of stickiness via external estimates from the literature. In our numerical experiments below, we fix a value of λ ex-ante using empirical results from Coibion and Gorodnichenko (2015). These authors argue that in two classes of models of information rigidity, the degree of information rigidity can be consistently estimated using microdata from professional forecasters. We take the estimated degree of information rigidity computed using quarterly forecasts of real output (from Table VI, column (3) in the paper), choosing a daily information rigidity parameter that implies a similar degree of mean reversion at a quarterly frequency. Specifically, we set $\lambda = 0.3^{4/252} \approx 0.981$, which is also similar to the implied degree of rigidity found by Bouchaud et al. (2019). The associated degree of information rigidity is fairly mild; in particular, if objective expectations were to increase today by 10%, subjective expectations would have already increased by about 3.6% within a month and 7% within a quarter.

Likewise, the time-series properties of ϑ_t depend crucially on the extent of objective cash flow predictability in the model, which is governed by the parameters ρ_{cf} and $Std[\epsilon_{cf,t}]$. Two related papers, Schorfheide et al. (2018) and Pettenuzzo et al. (2020), both leverage fairly high-frequency data to estimate the parameters of a latent, persistent component in expected dividend growth. Specifically, Schorfheide et al. (2018) use cash flows measured at annual, quarterly, and monthly frequencies, whereas Pettenuzzo et al. (2020) exploit daily data on cash flows for all companies in the U.S. stock market. Both papers uncover moderately persistent estimates for the AR(1) coefficient of cash flow dynamics (ρ_{cf}); Pettenuzzo et al. (2020) obtain annualized estimates of ρ_{cf} ranging between 0.6 and 0.77, while the posterior median estimate of Schorfheide et al. (2018) (Table VI) is 0.67. That said, their estimates of shock volatilities imply quite different unconditional volatilities of the persistent component of expected cash flow growth.[14] Since our objective function is fairly flat in the parameter ρ_{cf}, we fix the persistence at the posterior median estimate of Schorfheide et al. (2018) but allow other cash flow volatility parameters to be internally calibrated to match additional volatility and covariance targets.

All remaining parameters are set to match a sequence of asset pricing moments calculated over the sample for which we have computed our out-of-sample results. First, we target unconditional volatilities of daily log excess returns, the annualized one-period risk-free rate, as well as the log dividend-price ratio. Next, we seek to match the monthly autocorrelations of the latter two of these variables, both of which are quite persistent, as well as the correlation between them.[15] Finally, we seek to match the full sample OLS coefficients from regressions of log excess returns on pd_t and rf_{t+1}, respectively, as well as the correlations between AR(1) innovations in each of these variables and forecast errors from these predictive regressions. These additional moments are intended to ensure that the model generates potential Stambaugh biases that are consistent with the data, so we refer to them as "Stambaugh correlations." Table IA.XVIII summarizes the calibrated parameters as well as a comparison of data versus model-implied moments obtained from these exercises. In general, our calibrated model matches these targets fairly well.

Before discussing our simulations, we pause to discuss what is *not* targeted in these calibrations. We deliberately fix the degree of information rigidity based on estimates from the literature. The asset pricing moments selected are fairly standard and hence not explicitly tied to any evidence related to pockets of predictability. Therefore, we view our examination of the model's ability (or lack thereof) to match evidence related to pockets as a nontargeted validation test of the model.

[14] Whereas the Schorfheide et al. (2018) calibration implies that the cash flow growth component has a volatility of around 11% (in annualized units), the Pettenuzzo et al. (2020) data-generating process has a smaller unconditional volatility. This difference largely reflects that Pettenuzzo et al. (2020) explicitly filter out jump components in cash flows that naturally reduces the volatility estimates.

[15] Rather than directly target the AR(1) coefficients, we compute the absolute value of the difference between the data and the model-implied half-life of a shock in our objective function that compares data and model-implied moments.

As additional points of comparison, our simulations below also consider two alternative models. The first is a rational expectations version of our model that has the same true cash flow dynamics but no information rigidities $\lambda = 0$. The second is an additional rational expectations model whose parameters are recalibrated with $\lambda = 0$. Since the effects of sticky expectations on unconditional asset pricing moments are fairly modest, these recalibrated parameters are similar to those from our baseline model.

Table IA.I
Pocket Return Statistics

This table reports the mean, standard deviation, autocorrelation, skewness, and kurtosis of excess returns (measured in percentage points) in- versus out-of-pocket. Coefficients are estimated using a one-sided kernel with a 2.5-year effective sample size and pockets are determined as periods in which a fitted squared forecast error differential (relative to a prevailing mean forecast and estimated using a one-sided kernel with a one-year effective sample size) is above zero in the preceding period.

Panel A: Out-of-pocket

	Daily				Monthly			
Statistics	**dp**	**tbl**	**tsp**	**rvar**	**dp**	**tbl**	**tsp**	**rvar**
Mean	0.03	0.03	0.03	0.03	0.79	0.76	0.78	0.72
Standard deviation	1.08	1.09	1.06	1.12	5.35	5.24	5.14	5.38
First-order autocorrelation	0.07	0.05	0.05	0.06	0.12	0.11	0.11	0.13
Skewness	0.03	-0.03	-0.02	0.02	0.58	0.62	0.58	0.64
Kurtosis	19.39	19.24	19.58	18.42	11.78	12.21	12.16	11.83

Panel B: In-pocket

	Daily				Monthly			
Statistics	**dp**	**tbl**	**tsp**	**rvar**	**dp**	**tbl**	**tsp**	**rvar**
Mean	0.03	0.03	0.04	0.04	0.31	0.41	-0.24	0.69
Standard deviation	0.89	0.83	0.94	0.78	4.27	4.86	5.66	4.30
First-order autocorrelation	0.04	0.22	0.22	0.12	-0.02	0.15	0.14	0.07
Skewness	-0.49	0.09	0.06	-0.39	-0.45	-0.37	-0.17	-0.68
Kurtosis	9.25	5.42	4.66	9.23	3.45	4.44	4.08	3.92

Table IA.II

OOS Statistical Model Simulations (Zero Predictability Null)

This table reports Monte Carlo simulation results for the empirical one-sided kernel empirical findings. We bootstrap the fitted residuals from a zero coefficient predictive regression model for excess returns and an AR(1) model for the predictor using three approaches: (i) an i.i.d. heteroskedastic bootstrap, (ii) a stationary block bootstrap where the optimal block length is chosen according to Politis and White (2004), (iii) an EGARCH(1,1) with t-distributed shocks. All residuals are resampled jointly to preserve the cross-sectional correlation between the innovations to the predictor and excess returns. We generate 1,000 bootstrap samples of the same sample size as is available for each predictor in the data. A pocket is classified as a period in which a fitted squared forecast error differential (estimated using a one-sided kernel with a one-year effective sample size) is above zero in the preceding period. We report six statistics. The first three are Clark and West (2007) t-statistics relative to a prevailing mean benchmark in the full sample, in-pocket, and out-of-pocket. The second three are economic statistics associated with returns on a portfolio that uses the time-varying coefficient model forecast in-pocket and the prevailing mean forecast out-of-pocket to allocate between the risk-free asset and the market (portfolio weights are limited to be between zero and two): the annualized estimated alpha in percentage points, the HAC t-statistic associated with that estimated alpha, and the annualized Sharpe ratio of the portfolio. The second column presents the corresponding statistics from the data for reference.

	Unrestricted							**+ excess return forecasts**							**All sign restrictions**						
		i.i.d		**Block**		**EGARCH**			**i.i.d.**		**Block**		**EGARCH**			**i.i.d.**		**Block**		**EGARCH**	
Statistic	**Actual**	**Avg.**	**p-val**	**Avg.**	**p-val**	**Avg.**	**p-val**	**Actual**	**Avg.**	**p-val**	**Avg.**	**p-val**	**Avg.**	**p-val**	**Actual**	**Avg.**	**p-val**	**Avg.**	**p-val**	**Avg.**	**p-val**
Panel A: dp																					
CW_{FS}	−0.74	−0.11	0.73	−0.19	0.72	−0.27	0.67	0.40	−0.08	0.33	0.07	0.37	0.02	0.36	0.68	−0.08	0.23	−0.09	0.22	0.04	0.26
CW_{IP}	3.00	−0.11	0.00	0.12	0.00	−0.33	0.00	3.79	−0.06	0.00	0.40	0.00	−0.01	0.00	4.03	−0.05	0.00	0.33	0.00	−0.01	0.00
CS_{OOP}	−1.62	−0.07	0.94	−0.38	0.90	−0.11	0.93	−1.94	−0.08	0.97	−0.30	0.95	0.02	0.98	−1.84	−0.08	0.96	−0.37	0.93	0.05	0.98
$\hat{\alpha}$	1.69	−0.20	0.06	0.41	0.10	0.17	0.03	2.50	−0.21	0.01	0.52	0.05	0.57	0.08	2.95	−0.29	0.00	0.44	0.03	0.46	0.04
$t_{\hat{\alpha}}$	2.09	−0.19	0.01	0.37	0.04	0.28	0.03	2.89	−0.18	0.00	0.40	0.00	0.42	0.01	3.26	−0.24	0.00	0.32	0.00	0.33	0.00
SR	0.47	0.46	0.47	0.46	0.48	0.54	0.67	0.54	0.45	0.28	0.46	0.30	0.56	0.57	0.57	0.45	0.19	0.45	0.22	0.54	0.41
Panel B: tbl																					
CW_{FS}	0.68	−0.08	0.23	0.06	0.28	−0.26	0.18	1.98	−0.03	0.02	0.21	0.03	0.01	0.03	2.03	0.03	0.02	0.29	0.03	0.21	0.03
CW_{IP}	3.28	−0.06	0.00	0.22	0.00	−0.25	0.00	4.75	−0.06	0.00	0.32	0.00	−0.06	0.00	4.69	−0.06	0.00	0.34	0.00	0.08	0.00
CS_{OOP}	−1.58	−0.08	0.93	−0.18	0.91	−0.18	0.89	−1.33	0.00	0.90	−0.07	0.90	0.05	0.92	−1.21	0.07	0.90	0.04	0.90	0.21	0.92
$\hat{\alpha}$	3.57	−0.20	0.00	0.26	0.00	0.67	0.00	6.47	−0.24	0.00	0.31	0.00	1.21	0.00	6.06	−0.29	0.00	0.24	0.00	1.31	0.00
$t_{\hat{\alpha}}$	4.35	−0.22	0.00	0.22	0.00	0.97	0.00	5.56	−0.23	0.00	0.24	0.00	1.00	0.00	5.36	−0.28	0.00	0.18	0.00	1.03	0.00
SR	0.79	0.50	0.02	0.50	0.03	0.57	0.08	0.94	0.50	0.00	0.50	0.00	0.57	0.01	0.92	0.50	0.00	0.50	0.00	0.56	0.01
Panel C: tsp																					
CW_{FS}	0.15	−0.10	0.41	0.08	0.47	−0.27	0.35	0.95	−0.04	0.17	0.19	0.21	0.15	0.24	0.79	0.04	0.22	0.15	0.27	0.24	0.31
CW_{IP}	3.04	−0.11	0.00	0.19	0.00	−0.28	0.00	4.52	−0.08	0.00	0.27	0.00	0.10	0.00	4.18	−0.03	0.00	0.31	0.00	0.05	0.00
CS_{OOP}	−1.52	−0.06	0.93	−0.12	0.91	−0.14	0.91	−1.54	−0.00	0.94	−0.04	0.93	0.10	0.93	−0.21	0.05	0.60	−0.03	0.58	0.24	0.65
$\hat{\alpha}$	3.14	−0.26	0.01	0.20	0.01	0.48	0.00	5.69	−0.35	0.00	0.17	0.00	0.96	0.00	5.03	−0.24	0.00	0.27	0.00	0.76	0.00
$t_{\hat{\alpha}}$	4.26	−0.26	0.00	0.15	0.00	0.68	0.00	4.95	−0.31	0.00	0.11	0.00	0.79	0.00	4.45	−0.22	0.00	0.20	0.00	0.64	0.00
SR	0.77	0.41	0.00	0.41	0.01	0.46	0.02	0.85	0.41	0.00	0.41	0.00	0.45	0.01	0.84	0.40	0.00	0.42	0.00	0.45	0.01
Panel D: rvar																					
CW_{FS}	−1.49	−0.15	0.90	−0.28	0.91	−0.39	0.88	−0.79	−0.12	0.74	0.05	0.79	−0.16	0.71	−0.44	−0.09	0.64	0.10	0.70	0.12	0.70
CW_{IP}	2.88	−0.14	0.00	−0.06	0.00	−0.41	0.00	3.93	−0.14	0.00	0.24	0.00	−0.17	0.00	3.10	−0.08	0.00	0.28	0.00	−0.11	0.00
CS_{OOP}	−1.77	−0.10	0.95	−0.34	0.94	−0.20	0.94	−1.07	−0.05	0.84	−0.21	0.81	−0.08	0.84	−0.97	−0.08	0.81	−0.12	0.81	0.20	0.89
$\hat{\alpha}$	2.31	−0.23	0.02	0.65	0.05	0.36	0.01	2.88	−0.26	0.01	0.81	0.05	0.78	0.05	2.69	−0.13	0.02	0.87	0.06	0.57	0.04
$t_{\hat{\alpha}}$	3.62	−0.23	0.00	0.64	0.00	0.53	0.00	3.46	−0.24	0.00	0.64	0.00	0.70	0.00	3.03	−0.13	0.00	0.69	0.01	0.51	0.01
SR	0.68	0.45	0.05	0.46	0.07	0.56	0.23	0.71	0.45	0.04	0.46	0.05	0.55	0.16	0.54	0.45	0.25	0.47	0.31	0.55	0.52

Table IA.III
Individual Pocket p-values

This table reports p-values for individual pockets estimated using the unrestricted forecasts from the time-varying coefficient model. p-values are computed as the fraction of pockets from the EGARCH(1,1) model with t-distributed shocks that have integral R^2 values greater than the integral R^2 from the individual pocket in the data.

Pocket num	**dp**	**tbl**	**tsp**	**rvar**
1	0.05	3.93***	0.28	0.61*
	(0.42)	(0.00)	(0.13)	(0.07)
2	0.49	0.86**	0.69**	1.90***
	(0.12)	(0.05)	(0.05)	(0.01)
3	0.37	9.29***	7.54***	0.49*
	(0.15)	(0.00)	(0.00)	(0.09)
4	4.76***	0.54*	5.87***	2.54***
	(0.00)	(0.08)	(0.00)	(0.01)
5	3.69***	8.73***	1.77***	1.22**
	(0.00)	(0.00)	(0.01)	(0.03)
6	0.06	1.93***	1.68***	16.42***
	(0.38)	(0.01)	(0.01)	(0.00)
7	0.29	11.69***	2.62***	1.88***
	(0.17)	(0.00)	(0.00)	(0.01)
8	-0.24	-0.24		4.70***
	(1.00)	(1.00)		(0.00)
9	1.92***	3.47***		1.48**
	(0.01)	(0.00)		(0.02)
10	0.33	2.42***		-0.87
	(0.16)	(0.01)		(1.00)
11	0.13	1.05**		5.22***
	(0.27)	(0.03)		(0.00)
12	3.90***	0.72*		3.45***
	(0.00)	(0.06)		(0.00)
13	0.90*			0.94**
	(0.06)			(0.04)
14	2.92***			0.49*
	(0.00)			(0.09)
15	1.68**			1.93***
	(0.02)			(0.01)
16	0.22			1.92***
	(0.21)			(0.01)
17	4.56***			
	(0.00)			
18	1.11**			
	(0.05)			

Table IA.IV
Correlation Robustness of Dividend-Price Ratio Results

This table reports Monte Carlo simulation results for the one-sided kernel empirical findings for the dp model. We bootstrap the fitted residuals from a constant coefficient predictive regression model and an AR(1) model for the predictor using three approaches: (i) an i.i.d. heteroskedastic bootstrap, (ii) a stationary block bootstrap where the optimal block length is chosen according to Politis and White (2004), (iii) an EGARCH(1,1) with t-distributed shocks. All fitted residuals are made uniform using their empirical CDFs and then resampled using a Gaussian copula to achieve a particular correlation before transforming back. A pocket is classified as a period in which a fitted squared forecast error differential (estimated using a one-sided kernel with a one-year effective sample size) is above zero in the preceding period. We report six statistics. The first three are Clark-West t-statistics relative to a prevailing mean benchmark in the full sample, in pockets, and out of pockets. The second three are economic statistics associated with returns on a portfolio that uses the time-varying coefficient model forecast in-pocket and the prevailing mean forecast out-of-pocket to allocate between the risk-free asset and the market (portfolio weights are limited to be between zero and two): the annualized estimated alpha in percentage points, the HAC t-statistic associated with that alpha, and the annualized Sharpe ratio of the portfolio. The second column presents the corresponding statistics from the data for reference.

		i.i.d.			Block			EGARCH		
Statistic	**Actual**	**Avg.**	**Std. err.**	**p-val**	**Avg.**	**Std. err.**	**p-val**	**Avg.**	**Std. err.**	**p-val**
Panel A: Correlation -0.5										
CW_{FS}	−0.74	−0.03	1.04	0.77	0.13	0.99	0.82	−0.32	0.96	0.67
CW_{IP}	3.00	−0.06	1.02	0.00	0.36	0.96	0.00	−0.35	0.95	0.00
CS_{OOP}	−1.62	−0.01	1.04	0.94	−0.20	0.99	0.92	−0.15	1.00	0.93
$\hat{\alpha}$	1.69	−0.15	1.19	0.06	0.84	1.10	0.22	0.47	0.90	0.08
$t_{\hat{\alpha}}$	2.09	−0.15	1.03	0.01	0.72	0.96	0.07	0.65	1.10	0.10
SR	0.47	0.40	0.12	0.31	0.45	0.14	0.44	0.47	0.15	0.50
Panel B: Correlation -0.8										
CW_{FS}	−0.74	−0.10	1.05	0.72	−0.25	0.95	0.69	−0.32	0.96	0.65
CW_{IP}	3.00	−0.12	1.01	0.00	0.07	0.96	0.00	−0.37	0.90	0.00
CS_{OOP}	−1.62	−0.06	1.06	0.93	−0.41	0.97	0.90	−0.13	1.03	0.92
$\hat{\alpha}$	1.69	−0.19	1.18	0.06	0.35	1.05	0.12	0.11	0.81	0.03
$t_{\hat{\alpha}}$	2.09	−0.18	1.03	0.02	0.31	0.96	0.03	0.20	0.99	0.03
SR	0.47	0.40	0.10	0.28	0.44	0.11	0.41	0.45	0.12	0.46
Panel C: Correlation -0.9										
CW_{FS}	−0.74	−0.12	1.03	0.74	−0.45	0.96	0.60	−0.35	0.98	0.66
CW_{IP}	3.00	−0.15	0.97	0.00	−0.17	0.90	0.00	−0.34	0.92	0.00
CS_{OOP}	−1.62	−0.06	1.04	0.93	−0.45	0.99	0.89	−0.20	1.06	0.91
$\hat{\alpha}$	1.69	−0.21	1.13	0.04	−0.01	1.01	0.04	−0.06	0.84	0.02
$t_{\hat{\alpha}}$	2.09	−0.19	0.97	0.01	−0.01	0.93	0.01	0.01	1.01	0.02
SR	0.47	0.40	0.09	0.24	0.43	0.10	0.35	0.45	0.11	0.47
Panel D: Correlation -0.95										
CW_{FS}	−0.74	−0.14	1.00	0.71	−0.55	0.94	0.57	−0.35	0.98	0.65
CW_{IP}	3.00	−0.15	0.99	0.00	−0.29	0.95	0.00	−0.35	0.91	0.00
CS_{OOP}	−1.62	−0.10	1.02	0.93	−0.48	0.98	0.88	−0.19	1.03	0.92
$\hat{\alpha}$	1.69	−0.21	1.16	0.04	−0.23	1.03	0.04	−0.16	0.86	0.02
$t_{\hat{\alpha}}$	2.09	−0.19	1.01	0.01	−0.22	0.95	0.01	−0.14	1.01	0.01
SR	0.47	0.40	0.09	0.23	0.43	0.09	0.35	0.45	0.11	0.46
Panel E: Correlation -0.99										
CW_{FS}	−0.74	−0.18	1.00	0.71	−0.63	0.96	0.54	−0.35	0.99	0.64
CW_{IP}	3.00	−0.18	1.00	0.00	−0.37	0.94	0.00	−0.33	0.90	0.00
CS_{OOP}	−1.62	−0.11	1.01	0.93	−0.53	1.00	0.85	−0.20	1.00	0.92
$\hat{\alpha}$	1.69	−0.23	1.16	0.05	−0.37	1.00	0.02	−0.25	0.85	0.01
$t_{\hat{\alpha}}$	2.09	−0.20	1.00	0.01	−0.36	0.93	0.01	−0.24	0.98	0.01
SR	0.47	0.40	0.08	0.22	0.44	0.09	0.35	0.45	0.10	0.45

Table IA.V

Out-of-Sample Measures of Forecasting Performance
(Daily, Pockets Identified Relative to a Local Prevailing Mean (lpm) Benchmark)

Panel A reports the Clark and West (2007) test statistics for out-of-sample return predictability measured relative to a local prevailing mean forecast. Panel B reports three measures of economic significance associated with returns on a portfolio that uses the time-varying coefficient model forecast in-pocket and the local prevailing mean forecast out-of-pocket to allocate between the risk-free asset and the market (portfolio weights are limited to be between zero and two): the annualized estimated alpha in percentage points, the HAC t-statistic for the estimated alpha, and the annualized Sharpe ratio of the portfolio. We use a purely backward-looking kernel with an effective sample size of 2.5 years to compute forecasts. "pc" is a recursively computed first principal component of the four predictor variables. "mv" is a four-variable multivariate forecast estimated using a product kernel. "comb1," "comb2," and "comb3" refer to using a simple average of the univariate forecasts. "comb1" sets an individual predictor's forecast to the time-varying coefficient model forecast during a pocket and to the prevailing mean otherwise. "comb2" is the same as "comb1" except it ignores individual predictor forecasts when that variable is not in a pocket but at least one other variable is in a pocket. "comb3" makes no distinction between in-pocket and out-of-pocket periods and always uses the simple equal-weighted average of all four univariate models. The CW test statistics approximately follow a normal distribution with positive values indicating more accurate out-of-sample return forecasts than the local prevailing mean benchmark and negative values indicating the opposite. A pocket is classified as a period in which a fitted squared forecast error differential (estimated using a one-sided kernel with a one-year effective sample size) is above zero in the preceding period. $*$ $**$, and $***$ represent statistical significance at the 10%, 5%, and 1% level from a hypothesis test of $\beta > 0$. $\dagger$ $\dagger\dagger$, and $\dagger\dagger\dagger$ represent statistical significance at the 10%, 5%, and 1% level from a hypothesis test of $\beta < 0$.

Panel A: Clark-West statistics

	Unrestricted			+ excess return forecasts			All sign restrictions		
Variable	**Full sample**	**In-pocket**	**Out-of-pocket**	**Full sample**	**In-pocket**	**Out-of-pocket**	**Full sample**	**In-pocket**	**Out-of-pocket**
dp	0.01	3.79^{***}	-1.24	1.52^{*}	4.50^{***}	-1.12	1.96^{**}	4.84^{***}	-0.84
tbl	1.32^{*}	4.52^{***}	$-2.46^{\dagger\dagger\dagger}$	2.64^{***}	4.04^{***}	-0.56	2.98^{***}	4.62^{***}	-0.68
tsp	0.85	3.58^{***}	$-1.67^{\dagger\dagger}$	1.64^{*}	3.52^{***}	-0.99	0.30	3.37^{***}	-0.75
rvar	$-1.51^{\dagger}$	2.88^{***}	$-1.87^{\dagger\dagger}$	-0.69	2.57^{***}	$-1.35^{\dagger}$	-0.33	2.82^{***}	-1.19
mv	-0.69	4.21^{***}	$-1.36^{\dagger}$	0.34	4.42^{***}	-1.18	0.34	4.42^{***}	-1.18
pc	1.77^{**}	3.97^{***}	-1.17	2.50^{***}	3.93^{***}	-0.15	2.50^{***}	3.93^{***}	-0.15
comb1	5.58^{***}	5.73^{***}	–	5.33^{***}	5.44^{***}	–	5.80^{***}	5.86^{***}	–
comb2	6.17^{***}	6.34^{***}	–	6.05^{***}	6.18^{***}	–	6.54^{***}	6.60^{***}	–
comb3	-0.75	0.80	$-1.85^{\dagger\dagger}$	1.13	1.76^{**}	$-1.30^{\dagger}$	1.25	1.56^{*}	-0.61

Panel B: Economic significance

	Unrestricted			+ excess return forecasts			All sign restrictions		
Variable	$\hat{\alpha}$	$t_{\hat{\alpha}}$	**Sharpe ratio**	$\hat{\alpha}$	$t_{\hat{\alpha}}$	**Sharpe ratio**	$\hat{\alpha}$	$t_{\hat{\alpha}}$	**Sharpe ratio**
dp	2.02^{**}	1.71	0.44	2.91^{**}	2.10	0.47	3.10^{**}	2.28	0.48
tbl	3.67^{***}	3.04	0.61	4.60^{***}	3.07	0.60	5.17^{***}	3.37	0.63
tsp	2.50^{**}	2.05	0.50	4.18^{***}	2.66	0.55	3.15^{**}	1.94	0.49
rvar	1.58^{*}	1.49	0.44	2.21^{*}	1.60	0.45	2.07^{**}	1.73	0.45
mv	3.23^{***}	2.84	0.58	5.33^{***}	3.89	0.71	5.33^{***}	3.89	0.71
pc	3.28^{***}	2.63	0.56	4.54^{***}	2.91	0.57	4.54^{***}	2.91	0.57
comb1	3.40^{***}	2.99	0.58	5.30^{***}	3.46	0.63	5.35^{***}	3.51	0.65
comb2	5.37^{***}	4.45	0.71	6.99^{***}	5.01	0.79	7.03^{***}	5.15	0.78
comb3	0.76^{*}	1.32	0.43	2.35^{**}	1.88	0.46	2.73^{**}	2.08	0.48
pm	0.04	0.03	0.38	0.33	0.18	0.38	0.33	0.18	0.38

Table IA.VI

Out-of-Sample Measures of Forecasting Performance
(Monthly, Pockets Identified Relative to a Local Prevailing Mean (lpm) Benchmark)

Panel A reports the Clark and West (2007) test statistics for out-of-sample return predictability measured relative to a local prevailing mean forecast. Panel B reports three measures of economic significance associated with returns on a portfolio that uses the time-varying coefficient model forecast in-pocket and the local prevailing mean forecast out-of-pocket to allocate between the risk-free asset and the market (portfolio weights are limited to be between zero and two): the annualized estimated alpha in percentage points, the HAC t-statistic for the estimated alpha, and the annualized Sharpe ratio of the portfolio. We use a purely backward-looking kernel with an effective sample size of 2.5 years to compute forecasts. "pc" is a recursively computed first principal component of the four predictor variables. "mv" is a four-variable multivariate forecast estimated using a product kernel. "comb1," "comb2," and "comb3" refer to simple average of the univariate forecasts. "comb1" sets an individual predictor's forecast to the time-varying coefficient model forecast during a pocket and to the prevailing mean otherwise. "comb2" is the same as "comb1" except it ignores individual predictor forecasts when that variable is not in a pocket but at least one other variable is in a pocket. "comb3" makes no distinction between in-pocket and out-of-pocket periods and always uses the simple equal-weighted average of all four univariate models. The CW test statistics approximately follow a normal distribution with positive values indicating more accurate out-of-sample return forecasts than the local prevailing mean benchmark and negative values indicating the opposite. A pocket is classified as a period in which a fitted squared forecast error differential (estimated using a one-sided kernel with a one-year effective sample size) is above zero in the preceding period. $*$ $**$, and $***$ represent statistical significance at the 10%, 5%, and 1% level from a hypothesis test of $\beta > 0$. $\dagger$ $\dagger\dagger$, and $\dagger\dagger\dagger$ represent statistical significance at the 10%, 5%, and 1% level from a hypothesis test of $\beta < 0$.

Panel A: Clark-West statistics

	Unrestricted			+ excess return forecasts			All sign restrictions		
Variable	**Full sample**	**In-pocket**	**Out-of-pocket**	**Full sample**	**In-pocket**	**Out-of-pocket**	**Full sample**	**In-pocket**	**Out-of-pocket**
dp	2.21^{**}	4.98^{***}	-0.60	2.62^{***}	4.53^{***}	$-1.32^{\dagger}$	2.98^{***}	4.83^{***}	-1.24
tbl	1.57^{*}	2.81^{***}	-0.40	1.79^{**}	2.85^{***}	-1.21	3.02^{***}	3.43^{***}	-0.06
tsp	1.02	2.01^{**}	-0.80	1.00	2.73^{***}	0.09	0.14	2.43^{***}	-0.21
rvar	0.70	3.01^{***}	-0.33	0.90	2.64^{***}	$-2.24^{\dagger\dagger}$	1.39^{*}	3.21^{***}	$-2.01^{\dagger\dagger}$
mv	2.71^{***}	3.43^{***}	1.46^{*}	2.23^{**}	4.29^{***}	-0.68	2.23^{**}	4.29^{***}	-0.68
pc	1.60^{*}	2.23^{**}	0.08	1.64^{*}	2.42^{***}	0.53	1.64^{*}	2.42^{***}	0.53
comb1	3.81^{***}	4.08^{***}	–	4.33^{***}	4.51^{***}	–	4.62^{***}	4.79^{***}	–
comb2	4.64^{***}	5.06^{***}	–	5.45^{***}	5.68^{***}	–	5.89^{***}	6.14^{***}	–
comb3	1.99^{**}	2.49^{***}	0.28	2.46^{***}	2.98^{***}	$-1.87^{\dagger\dagger}$	2.44^{***}	2.67^{***}	-1.26

Panel B: Economic significance

	Unrestricted			+ excess return forecasts			All sign restrictions		
Variable	$\hat{\alpha}$	$t_{\hat{\alpha}}$	**Sharpe ratio**	$\hat{\alpha}$	$t_{\hat{\alpha}}$	**Sharpe ratio**	$\hat{\alpha}$	$t_{\hat{\alpha}}$	**Sharpe ratio**
dp	3.84^{***}	3.11	0.59	3.75^{***}	2.58	0.53	4.01^{***}	2.95	0.56
tbl	2.09^{*}	1.47	0.46	3.95^{**}	2.27	0.56	5.89^{***}	3.82	0.71
tsp	1.58	1.17	0.43	2.08	1.09	0.42	1.68	0.89	0.41
rvar	2.11^{**}	1.66	0.47	2.20^{*}	1.43	0.45	3.19^{**}	2.01	0.51
mv	2.80^{**}	2.16	0.51	4.41^{***}	3.15	0.60	4.41^{***}	3.15	0.60
pc	1.87^{*}	1.29	0.45	2.37^{*}	1.29	0.44	2.37^{*}	1.29	0.44
comb1	3.72^{***}	2.72	0.60	4.66^{***}	2.65	0.59	5.66^{***}	3.38	0.67
comb2	6.22^{***}	4.68	0.80	8.16^{***}	5.34	0.90	8.95^{***}	5.73	0.93
comb3	1.38^{*}	1.49	0.46	2.19^{*}	1.32	0.44	4.48^{***}	3.21	0.62
pm	-0.22	-0.17	0.38	-0.17	-0.09	0.38	-0.17	-0.09	0.38

Table IA.VII

Out-of-Sample Measures of Forecasting Performance
(Daily, Pockets Identified Relative to a Global Prevailing Mean Benchmark)

Panel A reports the Clark and West (2007) test statistics for out-of-sample return predictability measured relative to a local prevailing mean forecast. Panel B reports three measures of economic significance associated with returns on a portfolio that uses the time-varying coefficient model forecast in-pocket and the local prevailing mean forecast out-of-pocket to allocate between the risk-free asset and the market (portfolio weights are limited to be between zero and two): the annualized estimated alpha in percentage points, the HAC t-statistic for the estimated alpha, and the annualized Sharpe ratio of the portfolio. We use a purely backward-looking kernel with an effective sample size of 2.5 years to compute forecasts. "pc" is a recursively computed first principal component of the four predictor variables. "mv" is a four-variable multivariate forecast estimated using a product kernel. "comb1," "comb2," and "comb3" refer to a simple average of the univariate forecasts. "comb1" sets an individual predictor's forecast to the time-varying coefficient model forecast during a pocket and to the prevailing mean otherwise. "comb2" is the same as "comb1" except it ignores individual predictor forecasts when that variable is not in a pocket but at least one other variable is in a pocket. "comb3" makes no distinction between in-pocket and out-of-pocket periods and always uses the simple equal-weighted average of all four univariate models. The CW test statistics approximately follow a normal distribution with positive values indicating more accurate out-of-sample return forecasts than the local prevailing mean benchmark and negative values indicating the opposite. A pocket is classified as a period in which a fitted squared forecast error differential (estimated using a one-sided kernel with a one-year effective sample size) is above zero in the preceding period. $*$ $**$, and $***$ represent statistical significance at the 10%, 5%, and 1% level from a hypothesis test of $\beta > 0$. $\dagger$ $\dagger\dagger$, and $\dagger\dagger\dagger$ represent statistical significance at the 10%, 5%, and 1% level from a hypothesis test of $\beta < 0$.

Panel A: Clark-West statistics

	Unrestricted			+ excess return forecasts			All sign restrictions		
Variable	**Full sample**	**In-pocket**	**Out-of-pocket**	**Full sample**	**In-pocket**	**Out-of-pocket**	**Full sample**	**In-pocket**	**Out-of-pocket**
dp	0.01	3.11^{***}	-0.91	1.52^{*}	4.22^{***}	-0.37	1.96^{**}	4.30^{***}	0.02
tbl	1.32^{*}	3.44^{***}	-1.04	2.64^{***}	3.84^{***}	0.11	2.98^{***}	4.10^{***}	0.45
tsp	0.85	3.38^{***}	$-1.31^{\dagger}$	1.64^{*}	3.43^{***}	-0.34	0.30	3.34^{***}	-0.55
rvar	$-1.51^{\dagger}$	3.21^{***}	$-1.72^{\dagger\dagger}$	-0.69	2.66^{***}	-0.85	-0.33	2.29^{**}	-0.71
mv	-0.69	3.35^{***}	-1.18	0.34	4.10^{***}	-1.05	0.34	4.10^{***}	-1.05
pc	1.77^{**}	2.73^{***}	0.25	2.50^{***}	3.43^{***}	0.84	2.50^{***}	3.43^{***}	0.84
comb1	4.84^{***}	4.97^{***}	–	5.75^{***}	5.94^{***}	–	5.90^{***}	6.04^{***}	–
comb2	5.52^{***}	5.69^{***}	–	6.12^{***}	6.36^{***}	–	6.47^{***}	6.71^{***}	–
comb3	-0.75	3.02^{***}	$-1.80^{\dagger\dagger}$	1.13	1.20	0.13	1.25	2.73^{***}	-0.39

Panel B: Economic significance

	Unrestricted			+ excess return forecasts			All sign restrictions		
Variable	$\hat{\alpha}$	$t_{\hat{\alpha}}$	**Sharpe ratio**	$\hat{\alpha}$	$t_{\hat{\alpha}}$	**Sharpe ratio**	$\hat{\alpha}$	$t_{\hat{\alpha}}$	**Sharpe ratio**
dp	2.06^{**}	1.75	0.45	2.31^{*}	1.56	0.44	2.16^{*}	1.48	0.44
tbl	2.57^{**}	2.29	0.52	3.99^{***}	2.56	0.55	4.24^{***}	2.68	0.57
tsp	2.53^{**}	2.20	0.51	3.59^{**}	2.27	0.50	2.83^{**}	1.71	0.47
rvar	0.98	0.87	0.40	1.16	0.63	0.39	0.93	0.68	0.39
mv	2.52^{**}	2.23	0.51	4.80^{***}	3.54	0.67	4.80^{***}	3.54	0.67
pc	2.21^{**}	1.86	0.47	2.84^{**}	1.70	0.46	2.84^{**}	1.70	0.46
comb1	2.94^{***}	2.55	0.54	4.69^{***}	2.82	0.56	4.50^{***}	2.70	0.56
comb2	5.24^{***}	4.41	0.73	7.80^{***}	5.39	0.84	7.38^{***}	4.98	0.81
comb3	0.76^{*}	1.32	0.43	2.35^{**}	1.88	0.46	2.73^{**}	2.08	0.48
pm	0.04	0.03	0.38	0.33	0.18	0.38	0.33	0.18	0.38

Table IA.VIII

Out-of-Sample Measures of Forecasting Performance

(Monthly, Pockets Identified Relative to a Global Prevailing Mean Benchmark)

Panel A reports the Clark and West (2007) test statistics for out-of-sample return predictability measured relative to a local prevailing mean forecast. Panel B reports three measures of economic significance associated with returns on a portfolio that uses the time-varying coefficient model forecast in-pocket and the local prevailing mean forecast out-of-pocket to allocate between the risk-free asset and the market (portfolio weights are limited to be between zero and two): the annualized estimated alpha in percentage points, the HAC t-statistic for the estimated alpha, and the annualized Sharpe ratio of the portfolio. We use a purely backward-looking kernel with an effective sample size of 2.5 years to compute forecasts. "pc" is a recursively computed first principal component of the four predictor variables. "mv" is a four-variable multivariate forecast estimated using a product kernel. "comb1," "comb2," and "comb3" refer to a simple average of the univariate forecasts. "comb1" sets an individual predictor's forecast to the time-varying coefficient model forecast during a pocket and to the prevailing mean otherwise. "comb2" is the same as "comb1" except it ignores individual predictor forecasts when that variable is not in a pocket but at least one other variable is in a pocket. "comb3" makes no distinction between in-pocket and out-of-pocket periods and always uses the simple equal-weighted average of all four univariate models. The CW test statistics approximately follow a normal distribution with positive values indicating more accurate out-of-sample return forecasts than the local prevailing mean benchmark and negative values indicating the opposite. A pocket is classified as a period in which a fitted squared forecast error differential (estimated using a one-sided kernel with a one-year effective sample size) is above zero in the preceding period. $*$ $**$, and $***$ represent statistical significance at the 10%, 5%, and 1% level from a hypothesis test of $\beta > 0$. $\dagger$ $\dagger\dagger$, and $\dagger\dagger\dagger$ represent statistical significance at the 10%, 5%, and 1% level from a hypothesis test of $\beta < 0$.

Panel A: Clark-West statistics

	Unrestricted			+ excess return forecasts			All sign restrictions		
Variable	Full sample	In-pocket	Out-of-pocket	Full sample	In-pocket	Out-of-pocket	Full sample	In-pocket	Out-of-pocket
dp	2.21^{**}	4.10^{***}	0.53	2.62^{***}	4.22^{***}	0.32	2.98^{***}	4.22^{***}	0.78
tbl	1.57^{*}	2.54^{***}	0.08	1.79^{**}	2.44^{***}	0.26	3.02^{***}	3.37^{***}	1.07
tsp	1.02	2.37^{***}	−0.14	1.00	2.59^{***}	0.33	0.14	1.46^{*}	−0.05
rvar	0.70	2.98^{***}	−0.10	0.90	2.66^{***}	$-2.30^{\dagger\dagger}$	1.39^{*}	3.22^{***}	$-1.87^{\dagger\dagger}$
mv	2.71^{***}	2.98^{***}	2.20^{**}	2.23^{**}	3.95^{***}	−0.16	2.23^{**}	3.95^{***}	−0.16
pc	1.60^{*}	2.94^{***}	0.07	1.64^{*}	2.40^{***}	0.58	1.64^{*}	2.40^{***}	0.58
comb1	3.87^{***}	4.44^{***}	–	3.61^{***}	3.87^{***}	–	3.78^{***}	4.06^{***}	–
comb2	4.20^{***}	4.87^{***}	–	4.33^{***}	4.69^{***}	–	4.92^{***}	5.44^{***}	–
comb3	1.99^{**}	2.47^{***}	−0.99	2.46^{***}	2.30^{**}	1.11	2.44^{***}	1.73^{**}	2.11^{**}

Panel B: Economic significance

	Unrestricted			+ excess return forecasts			All sign restrictions		
Variable	$\hat{\alpha}$	$t_{\hat{\alpha}}$	Sharpe ratio	$\hat{\alpha}$	$t_{\hat{\alpha}}$	Sharpe ratio	$\hat{\alpha}$	$t_{\hat{\alpha}}$	Sharpe ratio
dp	2.07^{*}	1.50	0.45	2.80^{*}	1.59	0.46	2.80^{*}	1.59	0.46
tbl	1.81	1.27	0.45	2.46^{*}	1.38	0.45	4.10^{**}	2.29	0.55
tsp	0.72	0.61	0.39	1.65	0.84	0.41	0.61	0.33	0.38
rvar	1.44	1.13	0.42	2.24^{*}	1.46	0.45	3.11^{**}	1.98	0.50
mv	0.78	0.62	0.39	3.60^{***}	2.54	0.54	3.60^{***}	2.54	0.54
pc	1.70	1.25	0.44	1.92	0.99	0.42	1.92	0.99	0.42
comb1	2.55^{**}	1.94	0.50	3.53^{**}	1.92	0.50	4.09^{**}	2.25	0.54
comb2	5.06^{***}	3.44	0.72	6.17^{***}	3.86	0.74	7.18^{***}	4.08	0.79
comb3	1.38^{*}	1.49	0.46	2.19^{*}	1.32	0.44	4.48^{***}	3.21	0.62
pm	−0.22	−0.17	0.38	−0.17	−0.09	0.38	−0.17	−0.09	0.38

Table IA.IX

Out-of-Sample Measures of Forecasting Performance

(Daily, Pockets Identifed Relative to a Global Prevailing Mean that Are Not Identified By the Local Prevailing Mean)

Panel A reports the Clark and West (2007) test statistics for out-of-sample return predictability measured relative to a global prevailing mean forecast. Panel B reports three measures of economic significance associated with returns on a portfolio that uses the time-varying coefficient model forecast in-pocket and the global prevailing mean forecast out-of-pocket to allocate between the risk-free asset and the market (portfolio weights are limited to be between zero and two): the annualized estimated alpha in percentage points, the HAC t-statistic for the estimated alpha, and the annualized Sharpe ratio of the portfolio. We use a purely backward-looking kernel with an effective sample size of 2.5 years to compute forecasts. "pc" is a recursively computed first principal component of the four predictor variables. "mv" is a four-variable multivariate forecast estimated using a product kernel. "comb1," "comb2," and "comb3" refer to a simple average of the univariate forecasts. "comb1" sets an individual predictor's forecast to the time-varying coefficient model forecast during a pocket and to the prevailing mean otherwise. "comb2" is the same as "comb1" except it ignores individual predictor forecasts when that variable is not in a pocket but at least one other variable is in a pocket. "comb3" makes no distinction between in-pocket and out-of-pocket periods and always uses the simple equal-weighted average of all four univariate models. The CW test statistics approximately follow a normal distribution with positive values indicating more accurate out-of-sample return forecasts than the global prevailing mean benchmark and negative values indicating the opposite. A pocket is classified as a period in which a fitted squared forecast error differential (estimated using a one-sided kernel with a one-year effective sample size) is above zero in the preceding period. $*$ $**$, and $***$ represent statistical significance at the 10%, 5%, and 1% level from a hypothesis test of $\beta > 0$. $\dagger$ $\dagger\dagger$, and $\dagger\dagger\dagger$ represent statistical significance at the 10%, 5%, and 1% level from a hypothesis test of $\beta < 0$.

Panel A: Clark-West statistics

	Unrestricted			+ excess return forecasts			All sign restrictions		
Variable	**Full sample**	**In-pocket**	**Out-of-pocket**	**Full sample**	**In-pocket**	**Out-of-pocket**	**Full sample**	**In-pocket**	**Out-of-pocket**
dp	-0.74	2.11^{**}	-1.26	0.40	1.99^{**}	-0.37	0.68	2.20^{**}	-0.18
tbl	0.68	1.29^{*}	0.11	1.98^{**}	3.39^{***}	0.07	2.03^{**}	3.57^{***}	-0.02
tsp	0.15	1.83^{**}	-0.52	0.95	3.00^{***}	-0.31	0.79	2.73^{***}	0.30
rvar	$-1.49^{\dagger}$	2.61^{***}	$-1.64^{\dagger}$	-0.79	3.04^{***}	-0.91	-0.44	2.39^{***}	-0.59
mv	-0.99	2.93^{***}	$-1.36^{\dagger}$	-0.01	3.45^{***}	-1.02	-0.01	3.45^{***}	-1.02
pc	0.99	1.01	0.62	1.85^{**}	3.54^{***}	0.54	1.85^{**}	3.54^{***}	0.54
comb1	3.05^{***}	3.02^{***}	–	4.58^{***}	4.63^{***}	–	4.53^{***}	4.60^{***}	–
comb2	3.31^{***}	3.30^{***}	–	4.67^{***}	4.79^{***}	–	4.62^{***}	4.74^{***}	–
comb3	-1.03	-0.18	-1.00	0.23	-0.12	0.63	0.66	1.29^{*}	0.05

Panel B: Economic significance

	Unrestricted			+ excess return forecasts			All sign restrictions		
Variable	$\hat{\alpha}$	$t_{\hat{\alpha}}$	**Sharpe ratio**	$\hat{\alpha}$	$t_{\hat{\alpha}}$	**Sharpe ratio**	$\hat{\alpha}$	$t_{\hat{\alpha}}$	**Sharpe ratio**
dp	0.67	0.91	0.40	0.68	1.06	0.40	0.97^{*}	1.44	0.42
tbl	1.83^{***}	2.63	0.55	2.70^{***}	3.66	0.69	2.83^{***}	3.62	0.70
tsp	1.86^{***}	2.95	0.58	2.28^{***}	3.13	0.61	1.68^{***}	2.90	0.56
rvar	0.91^{**}	1.95	0.48	0.87^{**}	1.80	0.47	0.94^{**}	1.69	0.45
mv	2.36^{***}	3.17	0.63	3.81^{***}	4.32	0.78	3.81^{***}	4.32	0.78
pc	1.65^{**}	2.32	0.52	2.18^{***}	3.36	0.65	2.18^{***}	3.36	0.65
comb1	3.22^{***}	3.86	0.67	3.01^{***}	4.72	0.76	2.96^{***}	4.56	0.73
comb2	3.86^{***}	3.91	0.65	4.80^{***}	4.70	0.74	4.54^{***}	4.62	0.72
comb3	0.76^{*}	1.32	0.43	2.35^{**}	1.88	0.46	2.73^{**}	2.08	0.48

Table IA.X
Economic Measures of Forecasting Performance Controlling for Time-Varying Variance (Daily Benchmark Specification)

This table reports three measures of economic significance associated with returns on a portfolio that uses the time-varying coefficient model forecast in-pocket and the prevailing mean forecast out-of-pocket to allocate between the risk-free asset and the market controlling for local volatility using realized variance (portfolio weights are limited to be between zero and two): the annualized estimated alpha in percentage points, the HAC t-statistic for the estimated alpha, and the annualized Sharpe ratio of the portfolio. We use a purely backward-looking kernel with an effective sample size of 2.5 years to compute forecasts. “pc” is a recursively computed first principal component of the four predictor variables. “mv” is a four-variable multivariate forecast estimated using a product kernel. “comb1,” “comb2,” and “comb3” refer to a simple average of the univariate forecasts. “comb1” sets an individual predictor’s forecast to the time-varying coefficient model forecast during a pocket and to the prevailing mean otherwise. “comb2” is the same as “comb1” except it ignores individual predictor forecasts when that variable is not in a pocket but at least one other variable is in a pocket. “comb3” makes no distinction between pocket and nonpocket periods and always uses the simple equal-weighted average of all four univariate models. A pocket is classified as a period in which a fitted squared forecast error differential (estimated using a one-sided kernel with a one-year effective sample size) is above zero in the preceding period. $*$ $**$, and $***$ represent statistical significance at the 10%, 5%, and 1% level from a hypothesis test of $\beta > 0$. † ††, and † † † represent statistical significance at the 10%, 5%, and 1% level from a hypothesis test of $\beta < 0$.

	Unrestricted			+ excess return forecasts			All sign restrictions		
Variable	$\hat{\alpha}$	$t_{\hat{\alpha}}$	**Sharpe ratio**	$\hat{\alpha}$	$t_{\hat{\alpha}}$	**Sharpe ratio**	$\hat{\alpha}$	$t_{\hat{\alpha}}$	**Sharpe ratio**
dp	2.24**	1.94	0.47	3.70**	2.15	0.51	4.01**	2.32	0.53
tbl	4.69***	2.74	0.62	9.10***	4.07	0.72	9.81***	4.30	0.75
tsp	2.72**	1.91	0.50	8.40***	3.88	0.70	7.34***	3.66	0.73
rvar	3.98***	2.89	0.63	4.67***	3.12	0.70	6.54***	3.48	0.68
mv	2.04**	2.08	0.51	4.46***	2.59	0.57	4.46***	2.59	0.57
pc	2.99**	1.86	0.50	7.90***	3.59	0.67	7.90***	3.59	0.67
comb1	5.95***	3.48	0.69	1.66*	1.58	0.47	1.73**	1.66	0.49
comb2	8.45***	4.12	0.77	10.25***	4.46	0.76	8.26***	3.84	0.71
comb3	3.71*	1.46	0.44	2.65*	1.50	0.43	2.96**	1.77	0.49
lm	4.85***	3.01	0.64	7.86***	3.44	0.65	7.86***	3.44	0.65

Table IA.XI

Economic Forecasting Performance Robustness Under Transaction Costs (Daily Benchmark Specification)

This table reports annualized estimated alphas in percentage points associated with returns on a portfolio that uses the time-varying coefficient model forecast in-pocket and the prevailing mean forecast out-of-pocket to allocate between the risk-free asset and the market (portfolio weights are limited to be between zero and two). We consider four amounts of proportional transaction costs: none, 1 bp, 2 bps, and 10 bps. Significance of the estimated alpha is assessed using a t-statistic estimated using HAC standard errors that are reported below each alpha estimate. We use a purely backward-looking kernel with an effective sample size of 2.5 years to compute forecasts. "pc" is a recursively computed first principal component of the four predictor variables. "mv" is a four-variable multivariate forecast estimated using a product kernel. "comb1," "comb2," and "comb3" refer to a simple average of the univariate forecasts. "comb1" sets an individual predictor's forecast to the time-varying coefficient model forecast during a pocket and to the prevailing mean otherwise. "comb2" is the same as "comb1" except it ignores individual predictor forecasts when that variable is not in a pocket but at least one other variable is in a pocket. "comb3" makes no distinction between pocket and nonpocket periods and always uses the simple equal-weighted average of all four univariate models. A pocket is classified as a period in which a fitted squared forecast error differential (estimated using a one-sided kernel with a one-year effective sample size) is above zero in the preceding period. $*$ $**$, and $***$ represent statistical significance at the 10%, 5%, and 1% level from a hypothesis test of $\beta > 0$. $\dagger$ $\dagger\dagger$, and $\dagger\dagger\dagger$ represent statistical significance at the 10%, 5%, and 1% level from a hypothesis test of $\beta < 0$.

	Unrestricted				**+ excess return forecasts**				**All sign restrictions**			
Variable	**None**	**1 bp**	**2 bps**	**10 bps**	**None**	**1 bp**	**2 bps**	**10 bps**	**None**	**1 bp**	**2 bps**	**10 bps**
dp	1.69**	1.66**	1.63**	1.40**	2.51***	2.47***	2.44***	2.18***	2.95***	2.90***	2.85***	2.46***
	(2.10)	(2.06)	(2.03)	(1.75)	(2.89)	(2.85)	(2.82)	(2.53)	(3.26)	(3.21)	(3.16)	(2.74)
tbl	3.57***	3.54***	3.52***	3.30***	6.48***	6.44***	6.41***	6.13***	6.07***	6.03***	6.01***	5.80***
	(4.35)	(4.32)	(4.28)	(4.03)	(5.56)	(5.53)	(5.50)	(5.27)	(5.37)	(5.34)	(5.31)	(5.13)
tsp	3.15***	3.12***	3.10***	2.94***	5.70***	5.66***	5.63***	5.41***	5.04***	5.01***	4.98***	4.79***
	(4.26)	(4.23)	(4.21)	(4.00)	(4.95)	(4.92)	(4.90)	(4.71)	(4.45)	(4.43)	(4.41)	(4.25)
rvar	2.31***	2.28***	2.26***	2.09***	2.89***	2.86***	2.84***	2.66***	2.69***	2.65***	2.61***	2.29***
	(3.63)	(3.59)	(3.56)	(3.29)	(3.47)	(3.44)	(3.41)	(3.20)	(3.03)	(2.99)	(2.94)	(2.57)
mv	2.59***	2.56***	2.52***	2.25***	4.79***	4.71***	4.64***	4.03***	4.79***	4.71***	4.64***	4.03***
	(3.37)	(3.33)	(3.29)	(2.96)	(4.97)	(4.89)	(4.82)	(4.21)	(4.97)	(4.89)	(4.82)	(4.21)
pc	3.43***	3.39***	3.36***	3.11***	5.87***	5.83***	5.79***	5.47***	5.87***	5.83***	5.79***	5.47***
	(3.97)	(3.93)	(3.89)	(3.61)	(5.01)	(4.97)	(4.94)	(4.67)	(5.01)	(4.97)	(4.94)	(4.67)
comb1	6.38***	6.34***	6.29***	5.95***	6.72***	6.67***	6.63***	6.30***	6.69***	6.64***	6.59***	6.20***
	(6.11)	(6.07)	(6.03)	(5.71)	(6.71)	(6.66)	(6.62)	(6.30)	(6.46)	(6.41)	(6.37)	(6.02)
comb2	6.10***	6.05***	6.00***	5.64***	8.53***	8.47***	8.41***	7.94***	8.36***	8.29***	8.22***	7.66***
	(5.66)	(5.62)	(5.58)	(5.25)	(6.69)	(6.64)	(6.60)	(6.23)	(6.51)	(6.46)	(6.40)	(5.99)
comb3	0.76*	0.72	0.67	0.32	2.34**	2.25**	2.14**	1.31	2.72**	2.64**	2.55**	1.86*
	(1.32)	(1.25)	(1.17)	(0.55)	(1.87)	(1.79)	(1.71)	(1.04)	(2.07)	(2.01)	(1.95)	(1.41)
pm	$-0.25^{\dagger}$	$-0.26^{\dagger}$	$-0.26^{\dagger\dagger}$	$-0.29^{\dagger\dagger}$	$-0.25^{\dagger}$	$-0.26^{\dagger}$	$-0.26^{\dagger\dagger}$	$-0.29^{\dagger\dagger}$	$-0.25^{\dagger}$	$-0.26^{\dagger}$	$-0.26^{\dagger\dagger}$	$-0.29^{\dagger\dagger}$
	(−1.58)	(−1.64)	(−1.67)	(−1.85)	(−1.58)	(−1.64)	(−1.67)	(−1.85)	(−1.58)	(−1.64)	(−1.67)	(−1.85)

Table IA.XII

OOS Statistical Model Simulations (Monthly)

This table reports Monte Carlo simulation results for the empirical one-sided kernel empirical findings. We consider three ways of bootstrapping the fitted residuals from a constant coefficient predictive regression model for excess returns and an AR(1) model for the predictor: (i) an i.i.d. heteroskedastic bootstrap, (ii) a stationary block bootstrap where the optimal block length is chosen according to Politis and White (2004), (iii) an EGARCH(1,1) with t-distributed shocks. All residuals are resampled jointly to preserve the cross-sectional correlation between the innovations to the predictor and excess returns. We generate 1,000 bootstrap samples of the same sample size as is available for each predictor in the data. A pocket is classified as a period in which a fitted squared forecast error differential (estimated using a one-sided kernel with a one-year effective sample size) is above zero in the preceding period. We report six statistics. The first three are Clark and West (2007) t-statistics relative to a prevailing mean benchmark in the full sample, in-pocket, and out-of-pocket. The second three are economic statistics associated with returns on a portfolio that uses the time-varying coefficient model forecast in-pocket and the prevailing mean forecast out-of-pocket to allocate between the risk-free asset and the market (portfolio weights are limited to be between zero and two): the annualized estimated alpha in percentage points, the HAC t-statistic associated with that estimated alpha, and the annualized Sharpe ratio of the portfolio. The second column presents the corresponding statistics from the data for reference.

	Unrestricted							**+ excess return forecasts**							**All sign restrictions**						
		i.i.d.		**Block**		**EGARCH**			**i.i.d.**		**Block**		**EGARCH**			**i.i.d.**		**Block**		**EGARCH**	
Statistics	**Actual**	**Avg.**	***p*-val**	**Avg.**	***p*-val**	**Avg.**	***p*-val**	**Actual**	**Avg.**	***p*-val**	**Avg.**	***p*-val**	**Avg.**	***p*-val**	**Actual**	**Avg.**	***p*-val**	**Avg.**	***p*-val**	**Avg.**	***p*-val**
Panel A: dp																					
CW_{FS}	0.96	0.00	0.17	−0.17	0.15	−0.14	0.14	1.03	−0.03	0.15	0.02	0.14	0.12	0.20	1.13	0.06	0.14	0.00	0.15	0.19	0.18
CW_{IP}	4.05	−0.23	0.00	−0.45	0.00	−0.24	0.00	4.14	−0.22	0.00	−0.17	0.00	−0.15	0.00	4.14	−0.14	0.00	−0.11	0.00	−0.09	0.00
CS_{OOP}	−0.09	0.02	0.53	−0.09	0.49	−0.08	0.51	−3.12	0.02	1.00	0.06	1.00	0.17	1.00	−3.01	0.09	1.00	0.04	1.00	0.21	1.00
$\hat{\alpha}$	2.37	−0.29	0.00	−0.26	0.00	−0.15	0.00	4.13	−0.32	0.00	−0.27	0.00	−0.19	0.00	4.13	−0.26	0.00	−0.30	0.00	−0.18	0.00
$t_{\hat{\alpha}}$	2.53	−0.75	0.00	−0.57	0.00	−0.32	0.00	3.26	−0.50	0.00	−0.42	0.00	−0.29	0.00	3.26	−0.39	0.00	−0.42	0.00	−0.25	0.00
SR	0.55	0.42	0.07	0.42	0.07	0.46	0.21	0.69	0.41	0.00	0.41	0.00	0.47	0.03	0.69	0.41	0.00	0.42	0.00	0.47	0.03
Panel B: tbl																					
CW_{FS}	1.25	0.33	0.20	0.82	0.35	0.14	0.14	1.23	0.43	0.23	1.00	0.44	0.47	0.27	2.40	0.82	0.06	1.53	0.18	0.73	0.06
CW_{IP}	3.55	−0.66	0.00	−0.13	0.00	−0.73	0.00	4.38	−0.03	0.00	0.26	0.00	−0.10	0.00	4.47	0.25	0.00	0.71	0.00	0.14	0.00
CS_{OOP}	−0.83	0.32	0.85	0.81	0.94	0.19	0.84	−1.99	0.41	0.99	0.90	0.99	0.44	0.98	−1.26	0.75	0.97	1.33	0.99	0.67	0.97
$\hat{\alpha}$	3.77	−0.14	0.00	0.05	0.00	0.09	0.01	6.08	−0.06	0.00	0.28	0.00	0.37	0.00	6.22	0.13	0.00	0.62	0.00	0.55	0.00
$t_{\hat{\alpha}}$	3.36	−0.54	0.00	−0.17	0.00	−0.13	0.01	4.45	−0.33	0.00	0.08	0.00	0.12	0.00	4.43	−0.05	0.00	0.47	0.00	0.28	0.00
SR	0.76	0.52	0.12	0.53	0.13	0.60	0.21	0.86	0.51	0.05	0.52	0.06	0.62	0.13	0.87	0.52	0.06	0.53	0.07	0.61	0.11
Panel C: tsp																					
CW_{FS}	0.78	0.36	0.38	0.58	0.45	0.18	0.29	0.28	0.50	0.61	1.00	0.77	0.60	0.62	0.46	0.64	0.57	0.81	0.63	0.73	0.58
CW_{IP}	2.44	−0.28	0.01	−0.56	0.01	−0.49	0.01	4.75	−0.08	0.00	0.16	0.00	−0.71	0.00	4.93	0.02	0.00	0.01	0.00	0.23	0.00
CS_{OOP}	−1.15	0.36	0.92	0.59	0.94	0.21	0.89	−1.45	0.48	0.97	0.95	0.99	0.56	0.96	−0.20	0.59	0.77	0.74	0.81	0.67	0.76
$\hat{\alpha}$	2.22	−0.38	0.01	−0.22	0.01	−0.17	0.01	5.14	−0.24	0.00	−0.02	0.00	0.03	0.00	4.11	−0.12	0.00	−0.06	0.00	0.17	0.00
$t_{\hat{\alpha}}$	2.66	−0.77	0.00	−0.52	0.00	–	0.00	3.54	−0.51	0.00	−0.22	0.00	−0.20	0.00	3.24	−0.32	0.00	−0.20	0.00	0.06	0.00
SR	0.65	0.44	0.08	0.44	0.09	–	0.05	0.76	0.43	0.01	0.43	0.02	0.38	0.01	0.71	0.43	0.03	0.43	0.04	0.37	0.02
Panel D: rvar																					
CW_{FS}	0.64	−0.09	0.26	0.27	0.45	−0.09	0.30	0.40	−0.07	0.33	0.05	0.44	0.13	0.44	1.04	−0.10	0.14	−0.06	0.14	0.12	0.19
CW_{IP}	3.28	−0.44	0.00	−0.12	0.00	−0.23	0.00	3.18	−0.22	0.00	0.03	0.00	−0.06	0.00	3.58	−0.22	0.00	−0.07	0.00	−0.12	0.00
CS_{OOP}	0.00	−0.05	0.48	0.24	0.63	0.07	0.49	−2.73	−0.03	1.00	−0.03	1.00	0.16	1.00	−2.10	−0.09	0.98	−0.08	0.99	0.12	0.99
$\hat{\alpha}$	2.21	−0.27	0.00	−0.07	0.01	0.09	0.01	3.53	−0.24	0.00	0.03	0.00	0.17	0.01	3.73	−0.20	0.00	−0.02	0.00	0.15	0.01
$t_{\hat{\alpha}}$	2.73	−0.68	0.00	−0.18	0.00	0.00	0.01	3.27	−0.45	0.00	−0.04	0.00	0.04	0.00	3.68	−0.37	0.00	−0.11	0.00	−0.02	0.00
SR	0.55	0.46	0.25	0.45	0.24	0.49	0.35	0.65	0.46	0.10	0.45	0.10	0.49	0.18	0.66	0.45	0.05	0.44	0.08	0.49	0.17

Table IA.XIII
Pocket Statistics (Daily Fama French Factor Returns)

This table reports statistics on the duration of pockets (in days) and the integral R^2 of pockets. Coefficients are estimated using a one-sided kernel with a 2.5-year effective sample size, and pockets are determined as periods in which a fitted squared forecast error differential (relative to a prevailing mean forecast and estimated using a one-sided kernel with a one-year effective sample size) is above zero in the preceding period.

	SMB				**HML**			
Statistics	**dp**	**tbl**	**tsp**	**rvar**	**dp**	**tbl**	**tsp**	**rvar**
Num pockets	25	14	12	20	20	16	11	24
Fraction of sample	0.35	0.32	0.24	0.28	0.32	0.33	0.34	0.25
Duration								
Min	37	20	4	25	33	87	24	66
Mean	315.9	331.0	254.0	307.7	353.8	296.8	383.8	231.9
Max	1,587	1,216	738	829	892	597	1,352	659
Integral R^2								
Min	-0.31	-0.08	-0.04	0.17	-0.99	0.15	0.03	0.04
Mean	5.74	5.63	3.78	6.31	4.14	3.46	5.31	3.13
Max	57.92	31.95	12.73	41.12	16.75	9.17	25.36	18.94

Table IA.XIV
Asset Pricing Model Statistics

This table reports reports various correlations and autocorrelations associated with financial predictors across the four different asset pricing models we consider in the paper. Panel A reports the estimated correlation between each model's true risk premium rp and the model's dividend-price ratio *dp*, risk-free rate r, and 60-day realized variance $rvar$. Panel B reports the true R^2 from predictive regressions of excess returns on *dp*, r, $rvar$, and *rp*. Panel C reports the estimated correlation between innovations to the excess return predictive regression and an AR(1) model estimated for *dp*, r, $rvar$, and *rp*. Panel D reports the annualized AR(1) coefficients for *dp*, r, $rvar$, and *rp*.

Variable	Bansal-Yaron	Campbell-Cochrane	Garleanu-Panageas	Wachter
		Panel A: Correlation with True Risk Premium		
dp	0.14	0.99	0.99	1.00
risk-free	-0.03	0.94	0.49	-1.00
rvar	0.75	0.84	0.92	0.18
		Panel B: True R^2 (in %)		
dp	1.28×10^{-4}	0.03	6.78×10^{-3}	0.04
risk-free	2.05×10^{-5}	0.03	1.66×10^{-3}	0.04
rvar	2.06×10^{-3}	0.02	5.81×10^{-3}	1.35×10^{-3}
rp	3.44×10^{-3}	0.03	6.92×10^{-3}	0.04
		Panel C: Stambaugh Correlation		
dp	-0.81	-1.00	-0.99	-0.82
risk-free	0.80	-0.97	-0.44	0.82
rvar	0.03	0.03	0.02	-0.33
rp	-0.05	-0.97	-0.95	-0.82
		Panel D: First-Order Autocorrelation (annualized)		
dp	0.78	0.89	0.97	0.93
risk-free	0.78	0.90	0.94	0.93
rvar	0.17	0.53	0.53	0.02
rp	0.85	0.88	0.97	0.93

Table IA.XV

OOS Asset Pricing Model Simulations (Unrestricted)

This table reports Monte Carlo simulation results of our one-sided kernel estimation applied to data simulated from six different asset pricing models. We report six statistics. The first three are Clark and West (2007) t-statistics relative to a prevailing mean benchmark in the full sample, in-pocket, and out-of-pocket. The second three are economic statistics associated with returns on a portfolio that uses the time-varying coefficient model forecast in-pocket and the prevailing mean forecast out-of-pocket to allocate between the risk-free asset and the market (portfolio weights are limited to be between zero and two): the annualized estimated alpha in percentage points, the HAC t-statistic associated with that alpha, and the annualized Sharpe ratio of the portfolio. The second column presents the corresponding statistics from the data for reference.

		Bansal-Yaron			**Campbell-Cochrane**			**DiTella**			**Garleanu-Panageas**			**Wachter**			**Wachter (no disasters)**		
Stats	**Sample**	**Avg.**	**Std. err.**	**p-val**	**Avg.**	**Std. err.**	**p-val**	**Avg.**	**Std. err.**	**p-val**	**Avg.**	**Std. err.**	**p-val**	**Avg.**	**Std. err.**	**p-val**	**Avg.**	**Std. err.**	**p-val**
Panel A: dp																			
CW_{FS}	−0.74	−0.10	0.98	0.74	0.09	0.98	0.81	−0.14	0.94	0.74	−0.03	0.96	0.78	0.31	1.02	0.84	0.65	1.05	0.90
CW_{IP}	3.00	−0.07	1.02	0.00	−0.08	0.99	0.00	−0.19	0.98	0.00	−0.09	0.97	0.00	0.15	1.00	0.00	0.34	1.00	0.00
CW_{OOP}	−1.62	−0.10	1.01	0.93	0.15	0.99	0.96	−0.07	0.97	0.94	−0.00	0.97	0.95	0.25	1.03	0.97	0.54	1.04	0.97
$\hat{\alpha}$	1.69	−0.38	1.67	0.11	0.05	0.95	0.04	−0.03	0.18	0.00	−0.02	0.86	0.02	0.21	1.82	0.19	0.31	1.34	0.15
$t_{\hat{\alpha}}$	2.10	−0.23	1.00	0.01	0.05	1.01	0.02	−0.19	0.98	0.01	−0.02	0.99	0.01	0.19	1.08	0.04	0.22	1.02	0.03
SR	0.47	0.44	0.13	0.40	0.47	0.07	0.54	0.25	0.11	0.04	0.33	0.11	0.08	0.46	0.12	0.46	0.58	0.10	0.89
Panel B: risk-free																			
CW_{FS}	0.68	−0.10	0.99	0.22	0.10	0.98	0.28	−0.12	0.93	0.20	−0.06	0.96	0.22	0.31	1.02	0.36	0.65	1.06	0.50
CW_{IP}	3.28	−0.09	1.03	0.00	−0.06	0.98	0.00	−0.18	0.98	0.00	−0.09	0.97	0.00	0.15	1.00	0.00	0.35	1.00	0.00
CW_{OOP}	−1.58	−0.09	1.01	0.93	0.14	0.99	0.96	−0.05	0.98	0.94	−0.04	0.96	0.94	0.25	1.03	0.96	0.54	1.04	0.97
$\hat{\alpha}$	3.57	−0.37	1.67	0.01	0.05	0.95	0.00	−0.03	0.18	0.00	−0.03	0.83	0.00	0.21	1.82	0.03	0.31	1.34	0.01
$t_{\hat{\alpha}}$	4.35	−0.23	1.00	0.00	0.05	1.01	0.00	−0.19	0.98	0.00	−0.02	0.99	0.00	0.19	1.08	0.00	0.22	1.02	0.00
SR	0.79	0.44	0.13	0.01	0.47	0.07	0.00	0.25	0.11	0.00	0.33	0.11	0.00	0.46	0.12	0.00	0.58	0.10	0.03
Panel C: rvar																			
CW_{FS}	−1.49	−0.08	1.03	0.91	−0.26	0.99	0.89	−0.25	0.98	0.89	−0.13	0.97	0.92	0.05	1.00	0.95	0.22	1.08	0.95
CW_{IP}	2.88	−0.08	1.02	0.00	−0.24	1.01	0.00	−0.25	0.99	0.00	−0.11	0.96	0.00	−0.01	0.97	0.00	0.08	1.00	0.00
CW_{OOP}	−1.77	−0.06	1.01	0.95	−0.15	0.99	0.95	−0.11	0.99	0.96	−0.10	0.98	0.95	0.10	1.01	0.97	0.22	1.02	0.98
$\hat{\alpha}$	2.31	−0.38	1.70	0.06	−0.06	0.96	0.01	−0.04	0.17	0.00	−0.04	0.85	0.01	0.05	1.29	0.04	0.17	1.33	0.06
$t_{\hat{\alpha}}$	3.63	−0.23	1.00	0.00	0.05	1.01	0.00	−0.19	0.98	0.00	−0.02	0.99	0.00	0.19	1.08	0.00	0.22	1.02	0.00
SR	0.68	0.44	0.13	0.03	0.47	0.07	0.01	0.25	0.11	0.00	0.33	0.11	0.00	0.45	0.12	0.03	0.58	0.10	0.12

Table IA.XVI

OOS Asset Pricing Model Simulations (+ Excess Return Forecasts)

This table reports Monte Carlo simulation results of our one-sided kernel estimation applied to data simulated from four different asset pricing models (this includes two specifications of Wachter's rare disasters model, one of which omits data from disaster episodes). We report six statistics. The first three are Clark and West (2007) t-statistics relative to a prevailing mean benchmark in the full sample, in-pocket, and out-of-pocket. The second three are economic statistics associated with returns on a portfolio that uses the time-varying coefficient model forecast in-pocket and the prevailing mean forecast out-of-pocket to allocate between the risk-free asset and the market (portfolio weights are limited to be between zero and two): the annualized estimated alpha in percentage points, the HAC t-statistic associated with that alpha, and the annualized Sharpe ratio of the portfolio. The second column presents the corresponding statistics from the data for reference. The time-varying coefficient model forecast is restricted to be greater than or equal to zero.

		Bansal-Yaron			Campbell-Cochrane			Garleanu-Panageas			Wachter			Wachter (no disasters)		
Statistic	**Sample**	**Avg.**	**Std. err.**	**p-val**	**Avg.**	**Std. err.**	**p-val**	**Avg.**	**Std. err.**	**p-val**	**Avg.**	**Std. err.**	**p-val**	**Avg.**	**Std. err.**	**p-val**
Panel A: dp																
CW_{FS}	0.40	−0.03	0.99	0.33	0.38	1.00	0.49	0.16	0.97	0.40	0.74	1.04	0.62	0.97	1.09	0.70
CW_{IP}	3.79	−0.03	1.00	0.00	0.11	0.98	0.00	0.06	0.98	0.00	0.43	1.04	0.00	0.59	1.08	0.00
CS_{OOP}	−1.94	−0.05	1.01	0.97	0.37	1.00	0.99	0.12	0.97	0.98	0.62	1.03	0.99	0.75	1.03	0.99
$\hat{\alpha}$	2.50	−0.40	1.83	0.06	0.05	1.06	0.01	−0.02	0.98	0.01	0.28	2.15	0.14	0.34	1.44	0.07
$t_{\hat{\alpha}}$	2.89	−0.22	1.00	0.00	0.04	1.01	0.00	−0.02	0.99	0.00	0.17	1.08	0.01	0.22	1.01	0.00
SR	0.54	0.44	0.13	0.23	0.47	0.07	0.18	0.33	0.11	0.03	0.46	0.12	0.24	0.58	0.10	0.65
Panel B: r																
CW_{FS}	1.98	−0.03	0.99	0.02	0.39	1.00	0.05	0.12	0.95	0.02	0.74	1.04	0.12	0.98	1.09	0.17
CW_{IP}	4.75	−0.05	1.02	0.00	0.13	0.98	0.00	0.05	0.95	0.00	0.43	1.04	0.00	0.60	1.07	0.00
CS_{OOP}	−1.33	−0.04	1.01	0.89	0.37	0.99	0.95	0.09	0.97	0.92	0.62	1.03	0.97	0.75	1.03	0.98
$\hat{\alpha}$	6.47	−0.39	1.83	0.00	0.06	1.06	0.00	−0.03	0.95	0.00	0.28	2.15	0.01	0.34	1.44	0.00
$t_{\hat{\alpha}}$	5.56	−0.22	1.00	0.00	0.05	1.01	0.00	−0.03	0.97	0.00	0.17	1.08	0.00	0.22	1.01	0.00
SR	0.94	0.44	0.13	0.00	0.47	0.07	0.00	0.33	0.11	0.00	0.46	0.12	0.00	0.58	0.10	0.00
Panel C: rvar																
CW_{FS}	−0.79	−0.02	1.02	0.78	0.02	0.99	0.80	0.04	0.97	0.81	0.33	1.00	0.83	0.53	1.10	0.90
CW_{IP}	3.93	−0.04	1.04	0.00	−0.03	1.03	0.00	−0.01	0.98	0.00	0.18	0.99	0.00	0.29	1.07	0.00
CS_{OOP}	−1.07	−0.02	1.00	0.86	0.03	0.99	0.87	0.04	0.96	0.89	0.33	1.02	0.92	0.46	1.02	0.93
$\hat{\alpha}$	2.88	−0.39	1.84	0.04	−0.06	1.04	0.00	−0.04	0.96	0.00	0.07	1.72	0.05	0.19	1.41	0.03
$t_{\hat{\alpha}}$	3.46	−0.23	1.02	0.00	−0.06	0.99	0.00	−0.05	0.98	0.00	0.09	1.06	0.00	0.10	1.02	0.00
SR	0.71	0.44	0.13	0.02	0.47	0.07	0.00	0.33	0.11	0.00	0.45	0.12	0.02	0.58	0.10	0.09

Table IA.XVII

OOS Asset Pricing Model Simulations (All Sign Restrictions)

This table reports Monte Carlo simulation results of our one-sided kernel estimation applied to data simulated from four different asset pricing models (this includes two specifications of Wachter's rare disasters model, one of which omits data from disaster episodes). We report six statistics. The first three are Clark and West (2007) t-statistics relative to a prevailing mean benchmark in the full sample, in-pocket, and out-of-pocket. The second three are economic statistics associated with returns on a portfolio that uses the time-varying coefficient model forecast in-pocket and the prevailing mean forecast out-of-pocket to allocate between the risk-free asset and the market (portfolio weights are limited to be between zero and two): the annualized estimated alpha in percentage points, the HAC t-statistic associated with that alpha, and the annualized Sharpe ratio of the portfolio. The second column presents the corresponding statistics from the data for reference. The time-varying coefficient model forecast is restricted to be greater than or equal to zero and the estimated coefficients are restricted to be of a particular sign in accordance with economic theory: + for dp, - for r, and + for rvar.

		Bansal-Yaron			**Campbell-Cochrane**			**Garleanu-Panageas**			**Wachter**			**Wachter (no disasters)**		
Statistics	**Sample**	**Avg.**	**Std. err.**	***p*-val**	**Avg.**	**Std. err.**	***p*-val**	**Avg.**	**Std. err.**	***p*-val**	**Avg.**	**Std. err.**	***p*-val**	**Avg.**	**Std. err.**	***p*-val**
Panel A: dp																
CW_{FS}	0.68	0.00	1.00	0.24	0.39	0.99	0.39	0.20	0.96	0.31	0.77	1.05	0.52	1.02	1.09	0.61
CW_{IP}	4.03	0.03	1.01	0.00	0.13	1.00	0.00	0.11	1.00	0.00	0.46	1.04	0.00	0.64	1.08	0.00
CS_{OOP}	−1.84	−0.04	1.02	0.96	0.36	1.00	0.98	0.15	0.97	0.97	0.64	1.04	0.99	0.78	1.03	0.99
$\hat{\alpha}$	2.95	−0.40	1.87	0.04	0.04	1.05	0.00	−0.03	0.98	0.00	0.24	2.16	0.10	0.34	1.43	0.04
$t_{\hat{\alpha}}$	3.26	−0.21	1.00	0.00	0.03	1.00	0.00	−0.02	0.98	0.00	0.15	1.07	0.00	0.22	1.01	0.00
SR	0.57	0.44	0.13	0.17	0.47	0.07	0.08	0.33	0.11	0.02	0.46	0.12	0.16	0.58	0.10	0.48
Panel B: r																
CW_{FS}	2.03	−0.01	1.00	0.02	0.13	0.66	0.00	−0.20	0.98	0.01	0.80	1.03	0.11	1.01	1.08	0.17
CW_{IP}	4.69	−0.00	1.01	0.00	−0.08	0.83	0.00	−0.12	1.00	0.00	0.46	1.03	0.00	0.62	1.07	0.00
CS_{OOP}	−1.21	−0.03	1.02	0.87	0.25	0.86	0.94	−0.17	0.96	0.86	0.68	1.02	0.97	0.78	1.02	0.98
$\hat{\alpha}$	6.06	−0.41	1.86	0.00	−0.01	0.41	0.00	−0.22	0.88	0.00	0.34	2.08	0.01	0.35	1.43	0.00
$t_{\hat{\alpha}}$	5.36	−0.22	1.00	0.00	−0.26	0.99	0.00	−0.23	0.98	0.00	0.19	1.06	0.00	0.23	1.01	0.00
SR	0.92	0.44	0.13	0.00	0.48	0.07	0.00	0.33	0.11	0.00	0.46	0.12	0.00	0.58	0.10	0.00
Panel C: rvar																
CW_{FS}	−0.44	0.06	0.99	0.69	0.62	0.90	0.88	0.08	1.00	0.69	0.47	0.96	0.81	0.80	0.97	0.91
CW_{IP}	3.10	−0.01	0.96	0.00	0.17	0.97	0.00	−0.04	0.98	0.00	0.24	0.97	0.00	0.40	1.06	0.00
CS_{OOP}	−0.97	0.05	0.98	0.85	0.59	0.90	0.96	0.07	1.01	0.85	0.45	0.97	0.92	0.66	0.95	0.96
$\hat{\alpha}$	2.69	−0.08	1.74	0.06	0.31	0.98	0.01	0.05	0.98	0.00	0.10	1.77	0.07	0.46	1.31	0.05
$t_{\hat{\alpha}}$	3.03	−0.05	0.99	0.00	0.29	0.96	0.00	0.05	1.01	0.00	0.18	1.03	0.00	0.31	0.96	0.00
SR	0.54	0.44	0.13	0.20	0.47	0.07	0.16	0.33	0.11	0.02	0.45	0.12	0.22	0.58	0.10	0.62

Table IA.XVIII

Calibrated Parameters and Moments for Sticky Expectations Model

This table reports the calibrated parameters and analytic moments of the sticky expectations VAR model. All parameters and moments are reported in annualized units. Panel A reports calibrated parameters and Panel B reports the implied moments of interest. The "Data" column in Panel B lists the annualized empirical targets used for calibration. The three rightmost columns refer to three separate calibrations. In order, "Baseline" refers to the standard calibration with sticky expectations, "Baseline ($\lambda = 0$)" refers to the "Baseline" calibration but with rational expectations (i.e., $\lambda = 0$), and "RE Recalibrated" refers to a recalibration of the rational expectations model to match the target moments. Calibrated parameters are chosen by minimizing the weighted sum of squared deviations of analytic moments from empirical targets.

Panel A: Parameters

Parameter	**Notation**	**Baseline**	**Baseline ($\lambda = 0$)**	**RE recalibrated**
Sticky expectations parameter	λ	0.98107	0.00000	0.00000
Persistence of cash flow state	ρ_{cf}^{252}	0.66854	0.66854	0.66854
Persistence of discount rate state	ρ_{dr}^{252}	0.62264	0.62264	0.66171
Persistence of time-preference state	ρ_{tp}^{252}	0.94772	0.94772	0.96139
Loading of risk-free rate on cash flows	$\beta_{rf,cf}$	-0.01707	-0.01707	-0.19009
Loading of risk-free rate on discount rates	$\beta_{rf,dr}$	-0.27597	-0.27597	-0.21651
Volatility of cash flow state	$\sigma_{z_{cf}}$	0.08584	0.08584	0.04000
Volatility of discount rate state	$\sigma_{z_{dr}}$	0.06263	0.06263	0.06461
Volatility of time-preference state	$\sigma_{z_{tp}}$	0.02668	0.02668	0.01903
Volatility of dividend growth	$\sigma_{\Delta d}$	0.07999	0.07999	0.06042
Volatility of subj. expected dividend growth	$\sigma_{F[\Delta d]}$	0.07638	0.08584	0.04000

Panel B: Moments

Moments	**Data**	**Baseline**	**Baseline ($\lambda = 0$)**	**RE recalibrated**
Volatility of log returns	0.15664	0.15706	0.23823	0.17180
Volatility of log dp ratio	0.37688	0.36889	0.38010	0.30874
Volatility of risk-free	0.03198	0.03182	0.03182	0.02482
AC1 of log pd ratio	0.89277	0.93531	0.83991	0.87304
AC1 of log rf ratio	0.89873	0.83734	0.83670	0.82521
Correlation(pd,rf)	-0.60772	-0.59254	-0.57996	-0.56010
OLS coef. of excess returns on log pd	-0.04317	-0.01638	-0.01543	-0.03061
OLS coef. of excess returns on log rf	-0.00616	-0.00424	-0.00424	-0.00582
Stambaugh correlation of log dp	-0.86622	-0.85846	-0.94193	-0.93608
Stambaugh correlation of log rf	-0.04100	0.19509	0.07516	0.09248

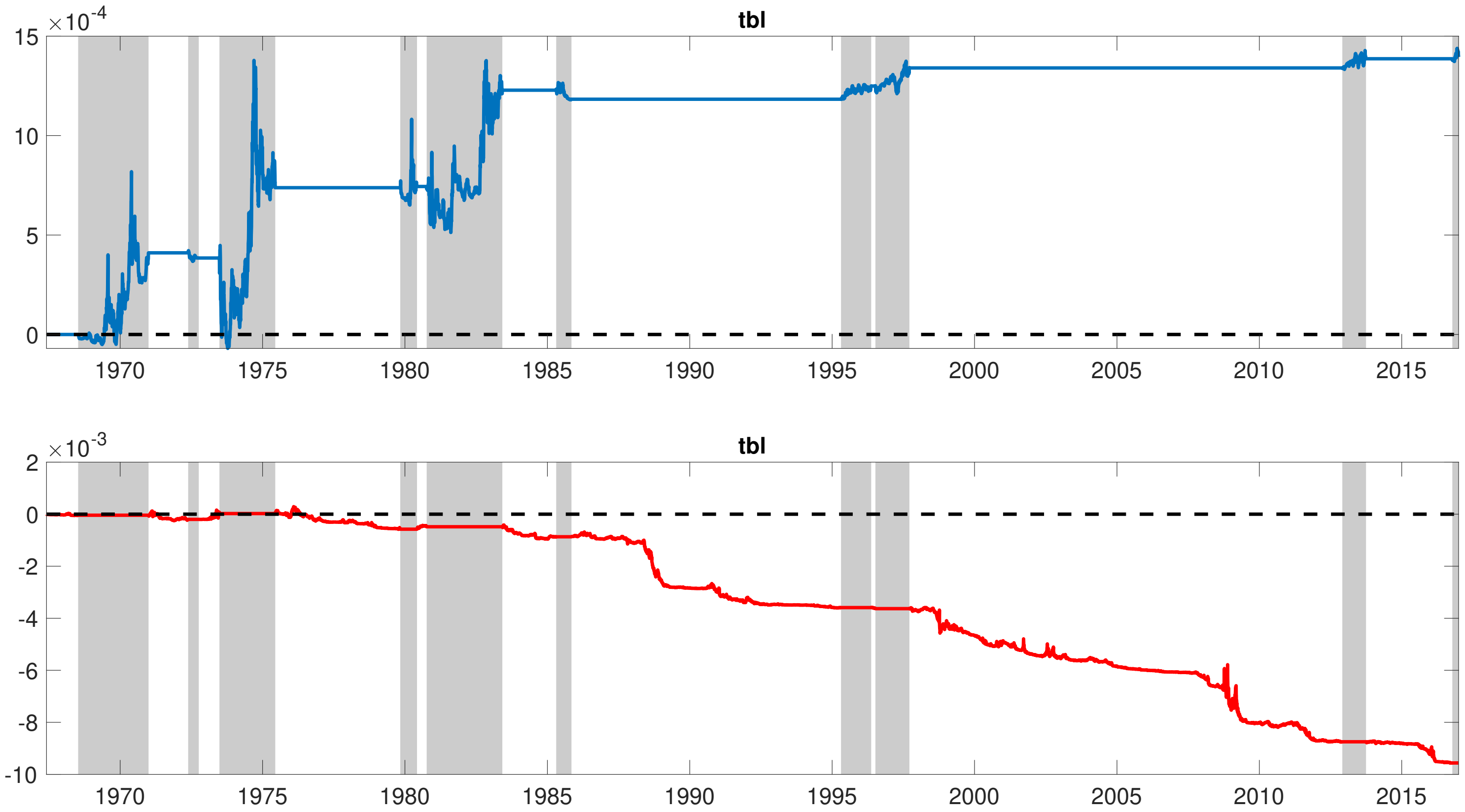

Figure IA.1: Cumsum of squared forecast error differentials for the tbl model. Each panel presents the cumulated sum of squared forecast errors between a time-varying coefficient model with a 2.5-effective sample size and the prevailing mean model. Areas shaded in gray are pocket periods identified in real-time when a fitted squared forecast error differential estimated using a one-year effective sample size is greater than zero in the preceding period. The top row shows a forecasting rule that uses the time-varying coefficient model in pockets and the prevailing mean model out of pockets. The bottom row shows a foreacsting rule that uses the prevailing mean model in pockets and the time-varying coefficient model out of pockets.

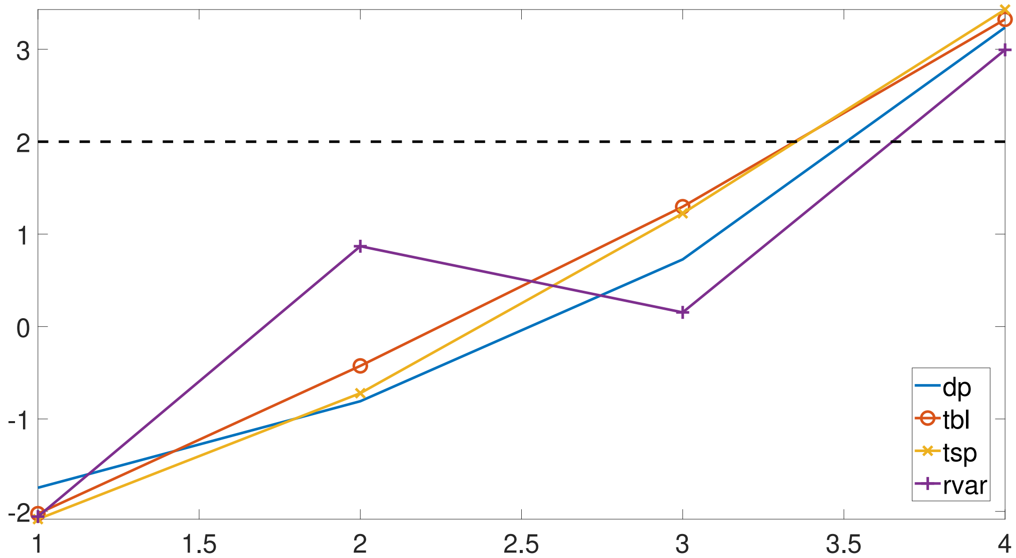

Figure IA.2: Clark-West statistics by quartile of fitted squared forecast error differential. For each univariate model's forecasts, we sort them into four bins according to the quartiles of our $\widehat{SED}_t$ measure. For each of these quartiles, we report the estimated Clark-West (2007) statistic. The dashed line corresponds to a rough 95% cutoff level of two for the t-statistics.

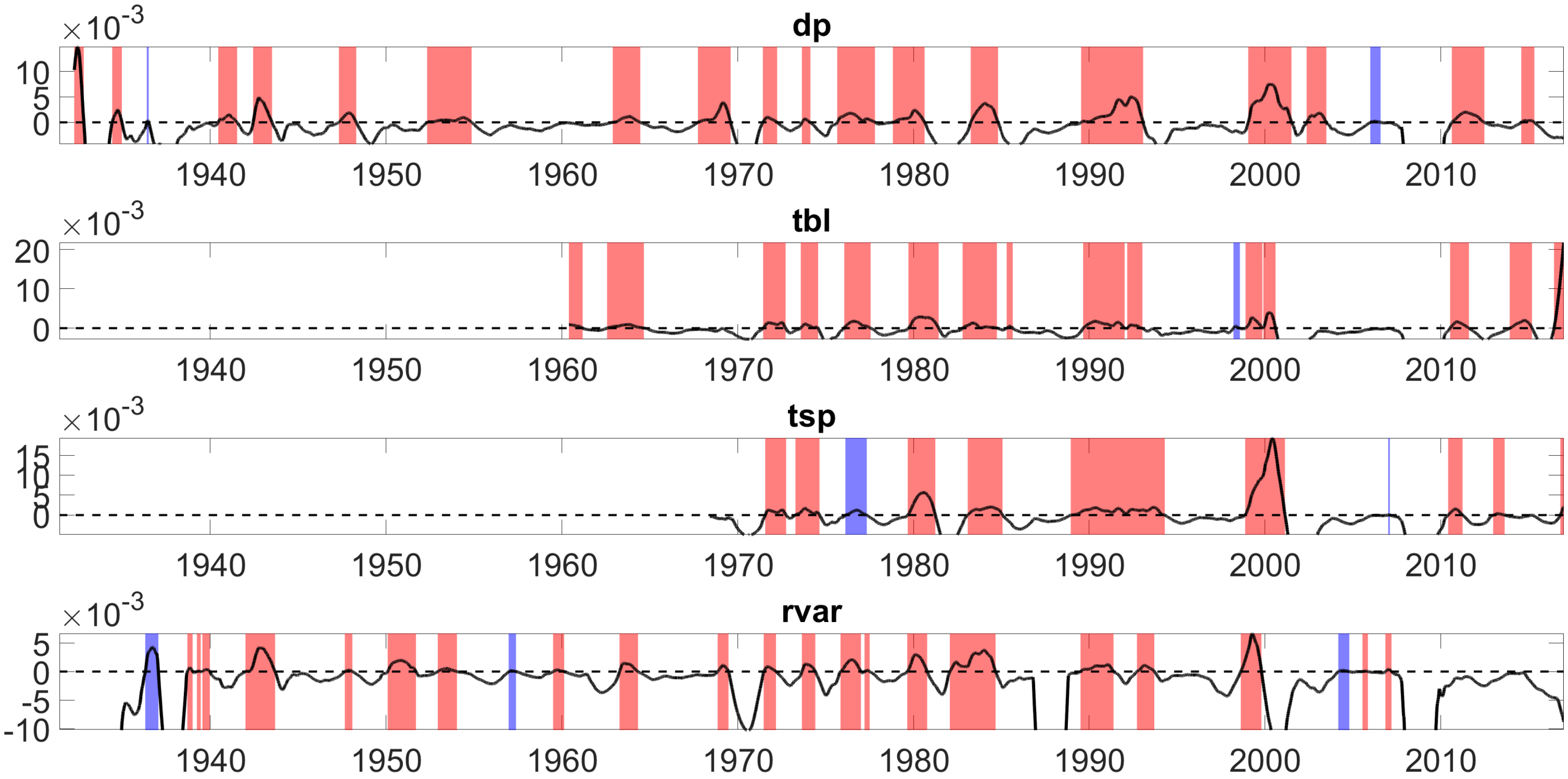

Figure IA.3: Local return predictability (HML). Each panel plots one-sided nonparametric kernel estimates of the local $\widehat{SED}_t$ (estimated using a one-year effective sample size) from a regression of daily returns on the Fama-French HML portfolio on each of the four predictor variables using an effective sample size of 2.5 years. The shaded areas represent periods when $\widehat{SED}_t > 0$, with areas in red representing pockets that have less than a 10% chance of being spurious, and areas in blue representing pockets that have more than a 10% chance of being spurious. The sampling distribution used to determine spuriousness comes from an EGARCH(1,1) residiual bootstrap design.

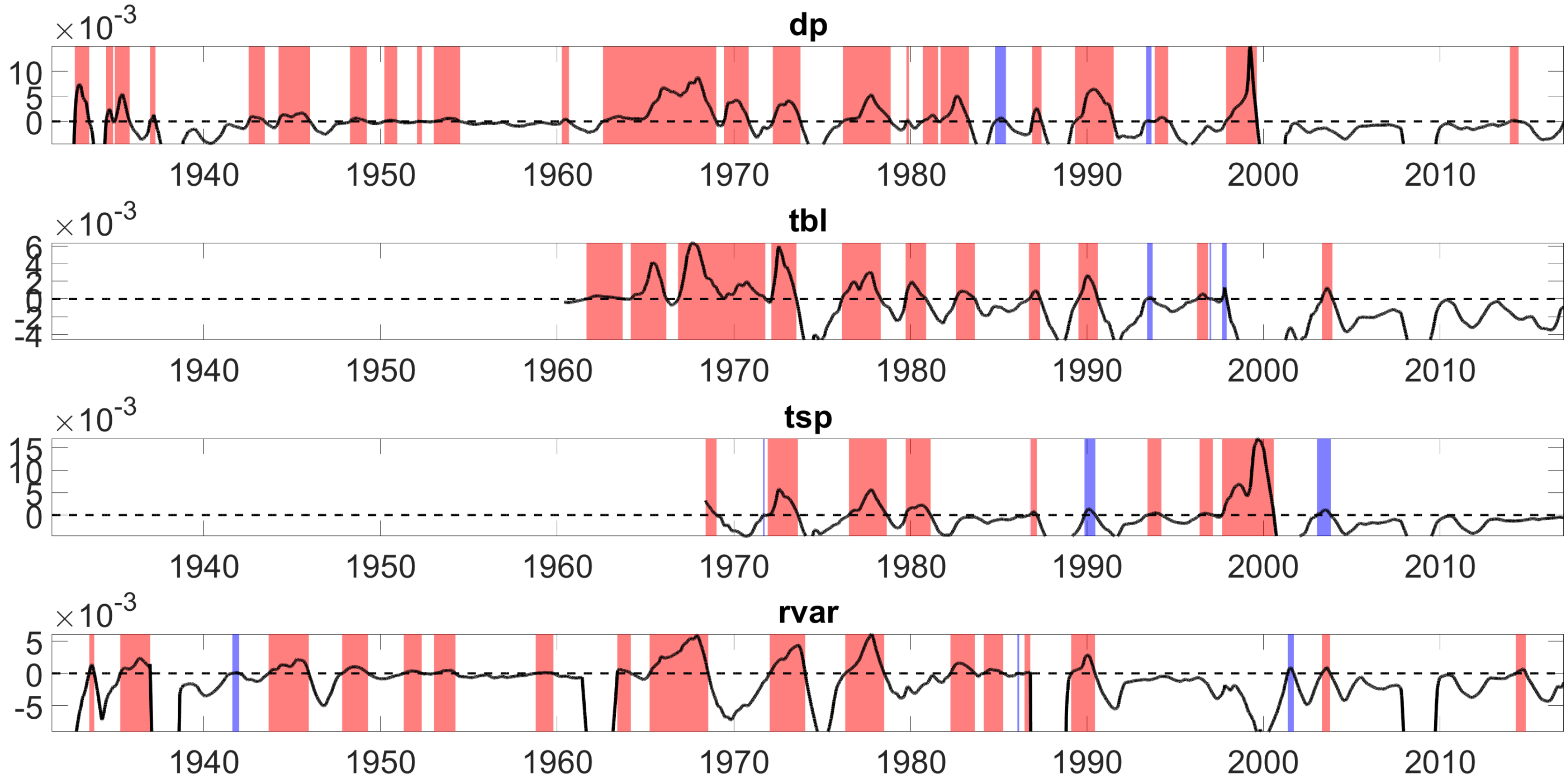

Figure IA.4: Local return predictability (SMB). Each panel plots one-sided nonparametric kernel estimates of the local $\widehat{SED}_t$ (estimated using a one-year effective sample size) from a regression of daily returns on the Fama-French SMB portfolio on each of the four predictor variables using an effective sample size of 2.5 years. The shaded areas represent periods when $\widehat{SED}_t > 0$, with areas in red representing pockets that have less than a 10% chance of being spurious, and areas in blue representing pockets that have more than a 10% chance of being spurious. The sampling distribution used to determine spuriousness comes from an EGARCH(1,1) residual bootstrap design.

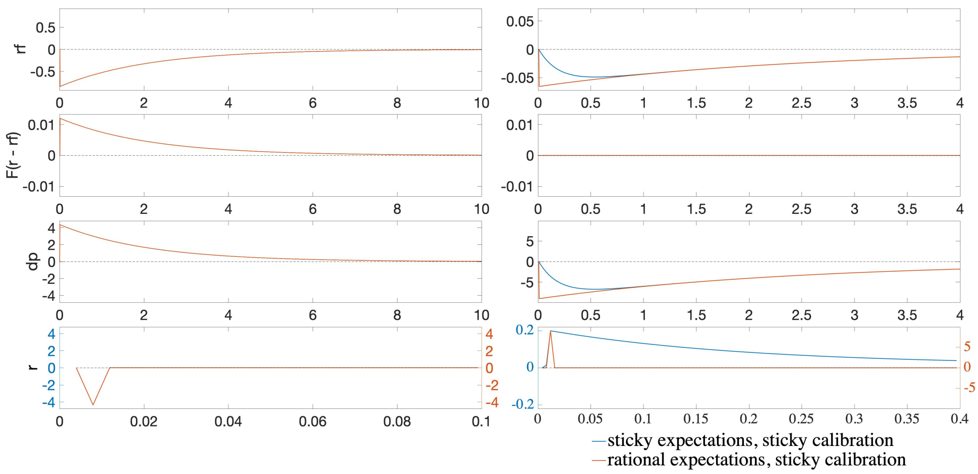

Figure IA.5: Analytic impulse response functions to one-quarter discount rate shock one-quarter cash-flow shocks. This figure displays the impulse response function of four variables: log risk-free rate, log subjective risk premium, log dividend-price ratio, and log returns. The left panel shows results for a one-quarter discount rate shock and the right panel shows results for a one-quarter cash-flow shock. The impulse response functions are calculated analytically according to baseline calibrated parameters. The blue line corresponds to sticky expectations ($\lambda \neq 0$) and the orange line to rational expectations ($\lambda = 0$).

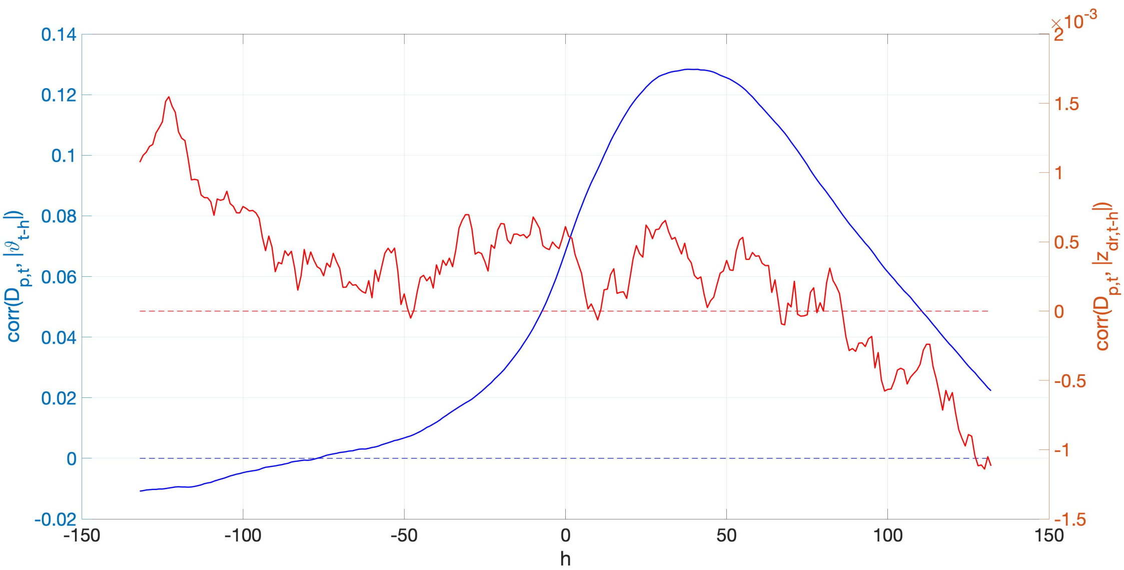

Figure IA.6: Cross-correlations between pocket indicator and measures of return predictability in sticky expectations model. The figure depicts the cross-correlations from simulated data from the sticky expectations between a pocket indicator, p_t, and two measures of the strength of predictability coming from the sticky expectations channel $|\vartheta_{t-h}|$ and the subjective risk premium $|z_{dr,t-h}|$, respectively.

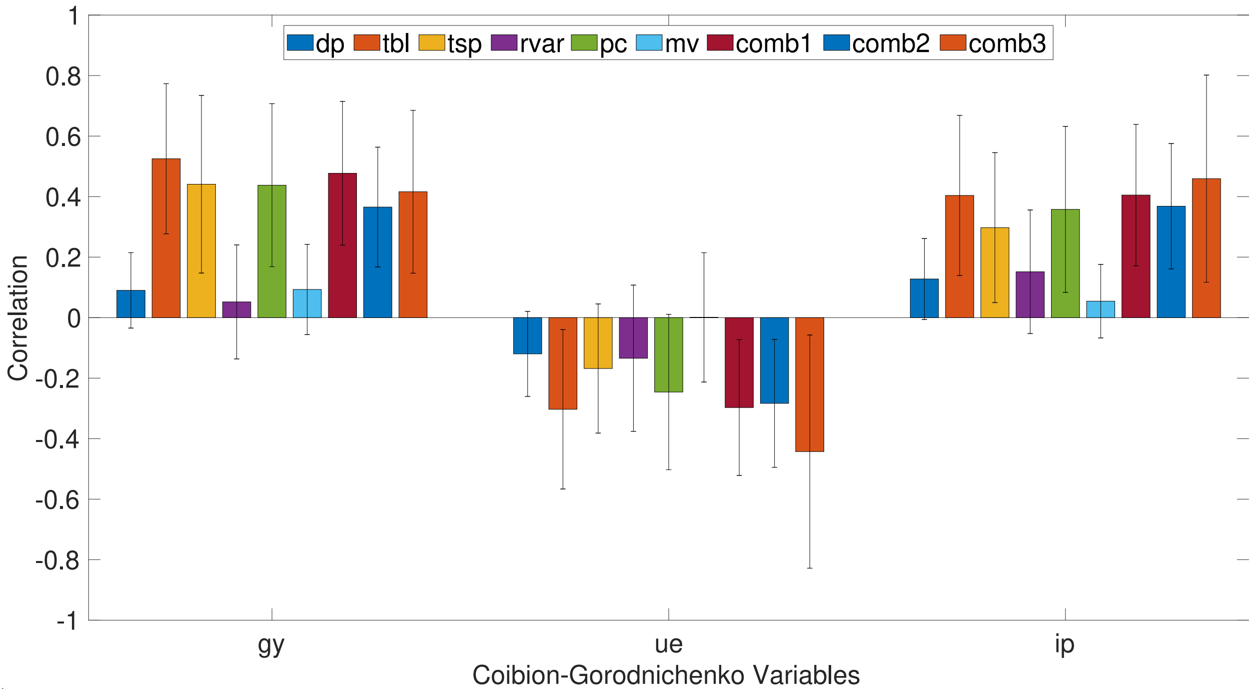

Figure IA.7: Correlation of Coibion-Gorodnichenko forecast errors with excess return forecasts (one-quarter lead). This figure shows correlations between forecast errors of three macroeconomic variables from the Survey of Professional Forecasters (SPF) with excess return forecasts from our time-varying coefficient models. The three sets of bar graphs correspond to forecast errors for real GDP growth (gy), the unemployment rate (ue), and real industrial production growth (ip). The height of the nine colored bars represents correlations of those forecast errors with the labeled excess return forecasts from our time-varying predictor models. Each bar is bracketed by a 95% confidence interval computed using HAC standard errors. Since the SPF respondents send in their forecasts in the middle of each quarter, we lead the SPF forecasts by one quarter to be conservative about information sets.

REFERENCES

Bansal, Ravi, and Amir Yaron, 2004, Risks for the long run: A potential resolution of asset pricing puzzles, *Journal of Finance* 59, 1481–1509.

Bekaert, Geert, and Eric Engstrom, 2017, Asset return dynamics under habits and bad environment good environment fundamentals, *Journal of Political Economy* 125, 713–760.

Bouchaud, Jean-Philippe, Philipp Krueger, Augustin Landier, and David Thesmar, 2019, Sticky expectations and the profitability anomaly, *Journal of Finance* 74, 639–674.

Cai, Zongwu, 2007, Trending time-varying coefficient time series models with serially correlated errors, *Journal of Econometrics* 136, 163–188.

Campbell, John Y., and John H. Cochrane, 1999, By force of habit: A consumption-based explanation of aggregate stock market behavior, *Journal of Political Economy* 107, 205–251.

Campbell, John Y., and Robert J. Shiller, 1988, The dividend-price ratio and expectations of future dividends and discount factors, *Review of Financial Studies* 1, 195–228.

Chen, Yu, Thomas Cosimano, Alex Himonas, and Peter Kelly, 2009, Asset pricing with long run risk and stochastic differential utility: An analytic approach, *SSRN Electronic Journal* .

Clark, Todd E., and Kenneth D. West, 2007, Approximately normal tests for equal predictive accuracy in nested models, *Journal of Econometrics* 138, 291–311.

Coibion, Olivier, and Yuriy Gorodnichenko, 2015, Information rigidity and the expectations formation process: A simple framework and new facts, *American Economic Review* 105, 2644–2678.

Constantinides, George M., and Darrell Duffie, 1996, Asset pricing with heterogeneous consumers, *Journal of Political Economy* 104, 219–240.

Constantinides, George. M., and Anisha Ghosh, 2017, Asset pricing with countercyclical household consumption risk, *Journal of Finance* 72, 415–460.

Creal, Drew D., and Jing C. Wu, 2016, Bond risk premia in consumption-based models, NBER Working Paper 22183.

Drechsler, Itamar, and Amir Yaron, 2011, What's vol got to do with it, *Review of Financial Studies* 24, 1–45.

Duffie, Darrell, and Larry G. Epstein, 1992, Stochastic differential utility, *Econometrica: Journal of the Econometric Society* 353–394.

Epstein, Larry G., and Stanley E. Zin, 1989, Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework, *Econometrica* 57, 937–969.

Eraker, Bjørn, and Ivan Shaliastovich, 2008, An equilibrium guide to designing affine asset pricing models, *Mathematical Finance* 18, 519–543.

Gârleanu, Nicolae, and Stavros Panageas, 2015, Young, old, conservative, and bold: The implications of heterogeneity and finite lives for asset pricing, *Journal of Political Economy* 123, 670–685.

Hansen, Lars Peter, Paymon Khorrami, and Fabrice Tourre, 2018, Comparative valuation dynamics in models with financing restrictions, Working paper, University of Chicago.

Hansen, Lars Peter, John C. Heaton, and Nan Li, 2008, Consumption strikes back? Measuring long-run risk, *Journal of Political Economy* 116, 260–302.

Herskovic, Bernard, Bryan Kelly, Hanno Lustig, and Stijn Nieuwerburgh, 2015, The common factor in idiosyncratic volatility: Quantitative asset pricing implications, *Journal of Financial Economics* 119.

Lettau, Martin, and Stijn Van Nieuwerburgh, 2007, Reconciling the return predictability evidence, *Review of Financial Studies* 21, 1607–1652.

Lustig, Hanno, Stijn van Nieuwerburgh, and Adrien Verdelhan, 2013, The wealth-consumption ratio, *Review of Asset Pricing Studies* 3, 38–94.

Pettenuzzo, Davide, Riccardo Sabbatucci, and Allan Timmermann, 2020, Cash flow news and stock price dynamics, *Journal of Finance* 75, 2221–2270.

Politis, Dimitris N., and Halbert White, 2004, Automatic block-length selection for the dependent bootstrap, *Econometric Review* 23, 53–70.

Robinson, Peter M., 1989, Nonparametric estimation of time-varying parameters, in Peter Hackl, ed., *Statistical Analysis and Forecasting of Economic Structural Change*, 253–264 (Springer).

Schmidt, Lawrence, 2020, Climbing and falling off the ladder: Asset pricing implications of labor market event risk, SSRN Working Paper, 2471342.

Schorfheide, Frank, Dongho Song, and Amir Yaron, 2018, Identifying long-run risks: A Bayesian mixed-frequency approach, *Econometrica* 86, 617–654.

Stambaugh, Robert F., 1999, Predictive regressions, *Journal of Financial Economics* 54, 375–421.

Wachter, Jessica A., 2005, Solving models with external habit, *Finance Research Letters* 2, 210–226.

Wachter, Jessica A., 2013, Can time-varying risk of rare disasters explain aggregate stock market volatility?, *Journal of Finance* 68, 987–1035.

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