Estimating beta: forecast adjustments and the impact of stock characteristics for a broad cross-section

Article

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Estimating Beta: Forecast Adjustments and the Impact of Stock Characteristics for a Broad Cross-Section[*]

Fabian Hollstein[†], Marcel Prokopczuk[†,‡], and Chardin Wese Simen[‡]

December 21, 2018

Abstract

Researchers and practitioners face many choices when estimating an asset's sensitivities toward risk factors, i.e., betas. Using the entire U.S. stock universe and a sample period of more than 50 years, we find that a historical estimator based on daily return data with an exponential weighting scheme as well as simple shrinkage adjustments yield the best predictions for future beta. Adjustments for asynchronous trading, macroeconomic conditions, or regression-based combinations, on the other hand, typically yield very high prediction errors and fail to create market-neutral anomaly portfolios. Finally, we document a robust link between stock characteristics and beta predictability.

JEL classification: G12, G11, G17

Keywords: Beta estimation, forecast combinations, forecast adjustments

I Introduction

Researchers and practitioners need estimates of betas for a wide variety of applications. Typically, historical data is used to estimate beta. Many researchers use the simple historical estimate. Others shrink the estimates toward the average beta of similar stocks. Some condition their estimates on macroeconomic state variables while others apply a weighting scheme on the historical data. Finally, some directly combine estimates obtained from different methods. Often, these decisions are made ad hoc without much guidance and discussion on how they impact the resulting estimates. The primary goal of this study is to deliver guidance for making the optimal choice among these and many more options one faces when estimating beta.[1]

To be more precise, we study the impact that these choices – e.g., different data sampling frequencies, estimation windows, forecast adjustments, and forecast combinations – have on estimates for beta. We analyze the impact of these choices both unconditionally and conditionally for stocks with different stock characteristics.

We use a large cross-section of stocks and more than 50 years of data to comprehensively study the estimation of beta. Relative to existing studies, we substantially expand the scope both in the asset space and in the time dimension. We also illuminate several aspects of the estimation of beta. We evaluate the predictability for realized beta by computing the average root mean squared error (RMSE) of all approaches, testing the significance in mean squared and median squared forecast errors. In addition, we examine the *economic value* of more accurate beta forecasts in portfolio formation. We test if the beta estimators are able to generate market-neutral anomaly portfolios.

We examine several estimation and adjustment approaches. First, we study the impact of different estimation windows and data sampling frequencies. Regarding the estimation

---

[1]While betas are generally estimated with respect to various possible state variables, we focus our analysis on market beta.

window, the researcher faces a trade-off between conditionality, i.e., using the most recent data, and a sufficient sample size that reduces measurement errors when predicting a time-varying beta using historical data. We find that a historical window of 1 year typically yields the lowest average prediction errors and performs well in generating market-neutral anomaly portfolios. Furthermore, consistent with the findings of Hollstein et al. (2018), we find that the data frequency should be as high as possible, i.e., estimators based on daily data outperform those based on monthly or quarterly data.

Second, we examine the impact of different weighting schemes. Conceptually, exponentially weighing past observations could deliver a possible solution to the conditionality vs. sample size trade-off because one can *"have it both ways"*, placing a higher weight on more recent observations to get a conditional estimate and using a long historical window to reduce measurement noise. Indeed, we find that exponentially weighting the observations yields significantly more precise estimates for beta and brings the realized betas of ex-ante market-neutral portfolios toward zero.

Third, we examine the impact of imposing priors for the beta estimates. The idea behind this approach is that the beta estimate of a stock should not be too dissimilar to that of other stocks with similar characteristics. We find that the shrinkage adjustments of Vasicek (1973), Karolyi (1992), and Cosemans et al. (2016) reduce average squared prediction errors. The simple shrinkage approach of Vasicek (1973) also performs better than the simple historical estimator in forming market-neutral anomaly portfolios while the more informative prior models of Karolyi (1992) and Cosemans et al. (2016) perform worse.

Fourth, we examine the effect of adjustments for asynchronous trading. Scholes & Williams (1977) and Dimson (1979) suggest that we can account for asynchronous trading by including betas with respect to lagged market returns. Arguing that it takes investors time to process and understand the impact of systematic news on opaque firms, Gilbert et al. (2014) suggest using quarterly instead of daily data to estimate beta. However, contrary to

these arguments, we find that the Dimson adjusted beta and, as indicated previously, estimators based on monthly and quarterly data, yield very high RMSEs and average realized betas in absolute terms in the portfolio formation test.

Fifth, following Shanken (1990) and Ferson & Schadt (1996), we also examine the impact of conditioning information from macroeconomic state variables for beta estimation and find that all estimators that build on such information underperform the simple historical model.

Sixth, we investigate forecast combinations. We examine simple, regression-based, and Bayesian combinations. We find that a simple forecast combination of an exponentially weighted and a prior-based historical estimator yields the lowest average prediction errors overall. However, more elaborated combination approaches perform considerably worse, especially if we combine many individual models.

We also analyze *why* certain estimators perform well while others perform very poorly. We find that weighting schemes and shrinking the estimates toward informative priors work well because both approaches correct for measurement errors, particularly in stocks with very high and very low betas. On the other hand, asynchronicity adjustments, macroeconomic conditioning variables, and model-based forecast combinations increase measurement errors. The RMSE difference of betas based on macroeconomic variables is typically strongly correlated with the conditioning variables. It thus seems that individual stock betas are not related to macroeconomic variables and imposing this structure therefore induces systematic measurement errors. Model-based forecast combinations appear to overfit the dynamics of beta in-sample and thus perform very poorly out-of-sample.

Finally, we analyze the link between stock characteristics and the performance of different approaches. We find that betas are particularly hard to predict for stocks with extreme (high or low) betas or momentum, small and illiquid stocks, as well as those with high idiosyncratic volatility. The betas for value stocks are harder to predict than those for growth stocks. Additionally, we detect a heterogeneous degree of predictability across different industries

with betas of stocks, e.g., of the manufacturing sector being markedly more predictable than those of the energy sector. For all the sorted portfolios, we obtain similar results in terms of ranking our main approaches. Those estimators that perform well unconditionally, typically also do so conditional on certain stock characteristics.

We test the robustness of our results, and find that these are largely similar for forecast horizons of 1, 3, 6, 12, and 60 months. Our results are also robust to computing hedging error ratios or estimators for realized beta that account for infrequent trading. Finally, we obtain qualitatively similar results for equally and value-weighted RMSEs, for an evaluation in the time series of individual firms, as well as for an alternative statistical loss function.

Our study is related to but materially different from Hollstein & Prokopczuk (2016). The authors examine the performance of different option-implied and time-series based estimators of beta and the historical estimator, finding that the option-implied estimator of Buss & Vilkov (2012) performs best. Our focus is considerably broader. While Hollstein & Prokopczuk (2016) face limitations imposed by option-implied estimators (only large-cap stocks of the S&P 500 and a relatively short sample period), we can analyze how to optimally estimate beta for *all* stocks, especially those for which there is no sufficient options data available. Furthermore, we use numerous estimators that are not included in Hollstein & Prokopczuk (2016) and link beta predictability to different stock characteristics.

More recently, Hollstein et al. (2018) make use of the results of Bollerslev & Zhang (2003), Barndorff-Nielsen & Shephard (2004), and Andersen et al. (2006) and show that, using high-frequency data, betas can be estimated more precisely for the firms of the S&P 500. However, the same shortcoming as for option-implied estimators applies for estimators relying on high-frequency data: they are only reliable for the subset of the most liquid stocks. Thus, our research directly complements these studies.

Our paper also connects to studies on the conditional capital asset pricing model (CCAPM). Shanken (1990), Ferson & Schadt (1996), Lettau & Ludvigson (2001), and Guo et al. (2017)

condition on macroeconomic variables to obtain time-varying betas. In contrast, Lewellen & Nagel (2006) use the simple historical estimator based on short windows for the same purpose. We complement these studies by examining the predictive accuracy of the estimators based on linear macroeconomic conditioning variables relative to the historical estimator and other models.

Our paper also adds to the literature on forecast combinations. Bates & Granger (1969), Clemen (1989), and Timmermann (2006) show that forecast combinations can be beneficial in many fields of financial forecasting. The authors show that forecast combinations are especially beneficial when the combined forecasts use data from different sources. We extend the forecast combinations literature in the context of beta estimation.

Lastly, we also connect to the literature on forecast adjustments for beta pioneered by Vasicek (1973). The author shrinks beta estimates toward the cross-sectional average beta. Recent developments turn towards more informative priors, as in Karolyi (1992) and Cosemans et al. (2016). We thoroughly examine the performance of prior-based combinations vis-à-vis single models and other possible forecast combinations. Levi & Welch (2017) test different shrinking approaches and suggest best practices to obtain cost-of-capital estimates.

The remainder of this paper is organized as follows. In Section II, we introduce the data and the methodology for the estimation of the different models. We present our main empirical results for estimating beta in Section III. In Section IV, we examine why some models work while others do not. Section V analyzes the impact of stock characteristics on beta predictability. In Section VI, we present additional analyses and test the robustness of our results. Section VII concludes.

II Data and Methodology

A Data

We obtain daily data on stock returns and shares outstanding from the Center for Research in Security Prices (CRSP). We use all stocks traded on the New York Stock Exchange (NYSE), the American Stock Exchange (AMEX), and the National Association of Securities Dealers Automated Quotations (NASDAQ). We start our sample period in January 1963 and end it in December 2015. Our sample period thus starts after the cross-section expansion of CRSP in the mid-1962 and spans more than 50 years. We obtain data on the risk-free (1-month Treasury Bill) rate from Kenneth French's data library. To proxy for the market return, we use the CRSP value-weighted index.

B Estimation Methodology

Historical Beta   We consider historical beta estimates (HIST) following, e.g., Fama & MacBeth (1973), regressing an asset's excess return on a constant and the market excess return:

$$ r_{j,\tau} - r_{f,\tau} \;=\; \alpha_{j,t} + \beta_{j,t}^{\text{HIST}}(r_{M,\tau} - r_{f,\tau}) + \epsilon_{j,\tau}, \tag{1} $$

where $\beta_{j,t}^{\text{HIST}}$ denotes the estimate for the historical beta of asset j at time t. We use data from time $t - k$ to t, observed at discrete intervals τ, where k is the length of the estimation window. $r_{j,\tau}$ is the return on asset j, $r_{M,\tau}$ denotes the return of the market portfolio, and $r_{f,\tau}$ is the risk-free rate, all observed at time τ.

EWMA Beta   We also examine a weighted version of the historical estimator with an exponentially weighted moving average structure. To be precise, we estimate Equation (1) with weighted least squares (WLS) using the weights $\frac{exp(-|t-\tau|h)}{\sum_{\tau=1}^{t-1} exp(-|t-\tau|h)}$ with $h = \frac{log(2)}{\iota}$. ι

characterizes the horizon, to which the half-life of the weights converges for large samples. We try two alternatives for ι: (i) one third and (ii) two thirds of the number of observations of the (initial) estimation window.[2]

This estimator is similar in spirit to that of Ang & Kristensen (2012), who estimate conditional betas using a kernel-weighted ordinary least squares (OLS) approach. The estimator used here can be regarded as a special case, only using past data and implementing an exponentially-weighted kernel. The Gaussian kernel used in Ang & Kristensen (2012) is not computable in real time since it hinges on future return data.

Shrinkage Beta   Following Vasicek (1973), we obtain a posterior belief of beta by combining the historical estimate ($\beta_{j,t}^{\text{HIST}}$) with one or multiple priors ($b_{i,j,t}$) in the following way:

$$\beta_{j,t}^{\text{Shr}} = \frac{\beta_{j,t}^{\text{HIST}} + \sum_i \frac{\sigma^2_{\beta_{j,t}^{\text{HIST}}}}{s^2_{b_{i,j,t}}} b_{i,j,t}}{1 + \sum_i \frac{\sigma^2_{\beta_{j,t}^{\text{HIST}}}}{s^2_{b_{i,j,t}}}}. \tag{2}$$

$\sigma^2_{\beta_{j,t}^{\text{HIST}}}$ and $s^2_{b_{i,j,t}}$ are the variances of the historical estimate and the prior(s), respectively. Hence, the degree of shrinkage depends on the relative precision of the historical estimate and the priors. We use as priors (i) the cross-sectional average beta (Vasicek, 1973) (β^{V}), (ii) a multiple shrinkage adjustment using the cross-sectional average beta, the cross-sectional average beta of firms in the same Global Industry Classification Standard (GICS) industry sector, and the cross-sectional average beta of firms in the same size decile (Karolyi, 1992) (β^{K}), and (iii) the fundamentals-based prior of Cosemans et al. (2016) (β^{I}).[3]

Dimson Beta   Following Dimson (1979) and Lewellen & Nagel (2006), we account for

---

[2]We try both a rolling window estimation using the same window as for HIST and an expanding window. To reduce the computational burden, we limit the maximum amount of daily returns used to 10 years.

[3]Cosemans et al. (2016) use the firms' size, book-to-market ratio, operating leverage, financial leverage, momentum, and industry classification, as well as the default yield spread to estimate the prior. For further information, we refer to the original article. We obtain the balance sheet data necessary to compute the above ratios from Compustat.

potential infrequent trading effects. If stocks trade less frequently than the market index, stock prices adjust gradually to new information. Therefore, Dimson (1979) adds lagged market returns in the regression:

$$
\begin{aligned}
r_{j,\tau} - r_{f,\tau} &= \alpha_{j,t} + \beta_{j,t}^{(0)}(r_{M,\tau} - r_{f,\tau}) + \beta_{j,t}^{(1)}(r_{M,\tau-1} - r_{f,\tau-1}) \\
&\quad + \beta_{j,t}^{(2)}\left(\sum_{n=2}^{N} r_{M,\tau-n} - r_{f,\tau-n}\right) + \epsilon_{j,\tau}.
\end{aligned} \tag{3}
$$

We incorporate $N = 1$ up to $N = 5$ lagged returns. In the case $N = 1$, the term associated with $\beta_{j,t}^{(2)}$ drops. The estimator for beta is then $\beta_{j,t}^{\mathrm{Dim(N)}} = \sum_{i=0}^{\min(2,N)} \beta_{j,t}^{(i)}$, where $\min(\cdot)$ is the minimum operator.

Scholes–Williams Beta   We also examine the beta estimator of Scholes & Williams (1977). That is, we estimate three separate regressions as in Equation (1), once using the contemporaneous market return, exactly as in Equation (1). Once, we estimate a similar regression using the lagged market excess return, that is $r_{j,\tau} - r_{f,\tau} = \alpha_{j,t} + \beta_{j,t}^{-}(r_{M,\tau-1} - r_{f,\tau-1}) + \epsilon_{j,\tau}$ and once we use the leaded market excess return $r_{j,\tau} - r_{f,\tau} = \alpha_{j,t} + \beta_{j,t}^{+}(r_{M,\tau+1} - r_{f,\tau+1}) + \epsilon_{j,\tau}$. The final estimator for beta is:

$$
\beta_{j,t}^{\mathrm{SW}} = \frac{\beta_{j,t}^{-} + \beta_{j,t}^{\mathrm{HIST}} + \beta_{j,t}^{+}}{1 + 2\rho}, \tag{4}
$$

where ρ is the first order autocorrelation of the market excess return.

Correlation-separated Beta   Following Frazzini & Pedersen (2014), we use an estimator that separates the estimation of volatilities and correlations. The authors estimate volatilities from daily return data and correlations from overlapping 3-day returns to account for asynchronous trading. Additionally, Frazzini & Pedersen (2014) argue that correlations move more slowly over time, allowing for different estimation windows used to compute

volatilities and correlations. We thus obtain the beta as:

$$\beta_{j,t}^{\mathrm{FP^{hor}}} = \rho_{j,M,t}^{\mathrm{hor}} \frac{\sigma_{j,t}}{\sigma_{M,t}}, \tag{5}$$

where $\rho_{j,M,t}^{\mathrm{hor}}$ is the correlation between the return of asset j with the market during the past *hor* months and $\sigma_{j,t}$ and $\sigma_{M,t}$ are the volatilities of the return of asset j and the market, respectively.

Macro Beta   We follow Shanken (1990) and Ferson & Schadt (1996) in assuming that $\beta_{j,t}^{Mac}$ is a linear function of state variable(s):

$$\beta_{j,t}^{\mathrm{Mac}} = b_{0,j} + B_j' z_t. \tag{6}$$

We define z_t as the vector of deviations of the state variables from their average up to time t, so that $b_{0,j}$ can be interpreted as the average beta while the elements in the matrix B_j determine the sensitivity of beta to the state variable(s). We estimate the parameters for Equation (6) using the time series of past (quarterly) macroeconomic variables and estimates for historical beta as on the left-hand side. We use a rolling estimation window of 20 quarters.[4]

We use the variables examined by Goyal & Welch (2008). The dataset is available from Amit Goyal's webpage. Specifically, we examine the book-to-market ratio of the Dow Jones Industrial Average (bm), the consumption–wealth–income ratio (Lettau & Ludvigson, 2001, cay), the default yield spread (dfy), the dividend–price (dp) and earnings–price (ep) ratios of the S&P 500, the investment–capital ratio (ic), inflation (inf), the long-term government bond yield (lty), and Treasury Bill rates (tbl).[5] We also use the 1-month macroeconomic uncertainty (unc) of Jurado et al. (2015) from Sydney Ludvigson's webpage and the unem-

---

[4]We also try an expanding window and find that the results are qualitatively similar, while the prediction errors for the expanding window are typically slightly higher.

[5]For further description of the construction of the variables, we refer to Goyal & Welch (2008).

ployment rate (une) from the Federal Reserve Economic Database.

We follow Goyal & Welch (2008) and also estimate a "kitchen-sink" (all) regression using all these variables. In a recent study, Guo et al. (2017) find that the earnings–price ratio, inflation, and the unemployment rate are the best predictors for the beta of the value premium. The authors cannot reject the null hypothesis of a linear relationship of the state variables and beta, which supports our choice of a simple linear specification.

Forecast Combinations Bates & Granger (1969) note that the combination of estimation techniques may prove worthwhile, especially when the combined estimates exploit (at least partially) different information sets. To investigate whether combinations are worthwhile for estimating beta, we try several approaches. The first is a simple equally weighted combination of different estimates. However, while such a simple ad hoc combination is easy to implement, the procedure might not provide the optimal result.

Second, we estimate weights by performing multivariate regressions for each stock.[6] We employ an expanding window to make use of a maximum length of history to be able to estimate the parameters with greater precision.[7] The regression equation takes the following form:

$$\beta_{j,\tau}^{\mathrm{R}} \quad = \quad a_{j,t} + \sum_{m=1}^{M} b_{j,t}^{(m)} \beta_{j,\tau}^{(\mathrm{m})} + \epsilon_{j,\tau}. \tag{7}$$

$\beta_{j,\tau}^{(\mathrm{m})}$ is the beta estimate for asset j of approach m at time τ. We combine the estimates of M different models. $\beta_{j,\tau}^{\mathrm{R}}$ denotes the corresponding realized beta of that asset. At every point in time the estimation moves forward, one additional observation is added to each of these vectors. After obtaining the time-t regression coefficients, we adjust the beta estimates,

---

[6]We use the first 100 months as our initial training sample. At each point in time t, we use estimates of beta up to $t - k$, since realized beta with a k-month window is only available up to the period $t - k$ until t at time t.

[7]We also try a rolling window approach. The results indicate that the expanding window approach indeed yields superior results.

using the following equation:

$$\beta_{j,t}^{\mathrm{C}} \;\; = \;\; \hat{a}_{j,t} + \sum_{m=1}^{M} \hat{b}_{j,t}^{(m)} \beta_{j,t}^{(\mathrm{m})}. \tag{8}$$

$\beta_{j,t}^{\mathrm{C}}$ is the combined beta forecast for asset j at time t and $\hat{a}_{j,t}$, $\hat{b}_{j,t}^{(m)}$ are the respective regression coefficients, i.e., weights.[8]

We also consider the Bayesian shrinkage approach proposed by Diebold & Pauly (1990). This approach shrinks regression coefficients towards a prior of equal weights for each forecast and an intercept of zero. To obtain $\beta_{j,t}^{\mathrm{shr}}$, we use Equations (7) and (8) with the empirical Bayes estimator.

Bayesian Model Averaging   Finally, we examine optimal forecast combinations using Bayesian model averaging. The basic idea of this approach is that there are K different possibilities to combine M different forecasts. To be more precise, one can use one single forecast only (M different possibilities), there are various possible combinations of two forecasts, three forecasts, and so forth. The models thus differ in the subset of predictors used. Under the uninformative prior specification of Fernandez et al. (2001), assuming that all variables are equally likely to enter the model, and that the likelihood that a variable enters the model is independent of that of another variable, the optimal combinations are (Stock & Watson, 2006):

$$\beta_{j,t}^{\mathrm{BMA}} = \sum_{k=1}^{K} \omega_k \beta_{j,t}^{(\mathrm{k})}, \tag{9}$$

where $\beta_{j,t}^{(\mathrm{k})}$ is the OLS combination (as of Equation (8)) of forecast models for one possible

---

[8]Note that now the $\beta_{j,t}^{(\mathrm{k})}$ have a t-subscript. This is because we only use the current beta estimates instead of the vector of all previous beta estimates.

way k to combine the M forecasts. The weights ω_k are:

$$\omega_k = \frac{a(g)^{\frac{1}{2}P_k}[1 + g^{-1}SSR_k^U/SSR^R]^{-\frac{1}{2}df^R}}{\sum_{i=1}^{K} a(g)^{\frac{1}{2}P_i}[1 + g^{-1}SSR_i^U/SSR^R]^{-\frac{1}{2}df^R}}. \tag{10}$$

Essentially, we first estimate a restricted forecasting model as in Equation (7) with OLS using only the variables that ought to be included in each model.[9] From this, we get the sum of squared residuals (SSR^R). Second, we estimate a forecasting model as of Equation (7) for each of the K possible combinations of predictors and get the forecast $\beta_{j,t}^{(k)}$ and the sum of squared residuals (SSR_k^U). P_k is the number of parameters in the kth regression combination, df^R is the number of the degrees of freedom of the restricted model, and $a(g) = g/(1 + g)$ with $g = 1/min(T, M^2)$ following Fernandez et al. (2001). T is the number of time periods in the estimation window.

C   Evaluation Methodology

Realized Beta   To evaluate predictions for beta, we follow Andersen et al. (2006) and use the realized beta (RB). Andersen et al. (2006) show that under weak regulatory conditions, RB is a consistent estimator of the underlying beta. We use daily (log-)returns during the prediction window t until T to estimate:[10]

$$\beta_{j,t}^{R} = \frac{\sum_{\tau=t+1}^{T} r_{j,\tau} r_{M,\tau}}{\sum_{\tau=t+1}^{T} r_{M,\tau}^2}, \tag{11}$$

where $r_{j,\tau}$ and $r_{M,\tau}$ refer to the return of asset j and the market return at time τ, respectively. Throughout our main empirical analysis, we follow Chang et al. (2012) and Hollstein &

---

[9]When empirically implementing the approach, specifying variables that are included in each model can substantially reduce the computational effort.

[10]Note that the formula for realized beta makes use of the expanded formula for the variance, neglecting the drift term. Andersen et al. (2006) note that the effect of the drift term vanishes as the sampling frequency is reduced, which effectively "annihilates" the mean. However, the average daily excess-return of the CRSP value-weighted index amounts to only 2.37 basis points. Thus, it is unlikely that this simplification induces a material bias.

Prokopczuk (2016) and focus on a prediction horizon for realized beta of 6 months.[11]

Root Mean Squared Error (RMSE) To examine the out-of-sample forecast accuracy of the different approaches, we perform the analysis using the RMSE, a loss function commonly applied in the literature:[12]

$$\text{RMSE}_j \;=\; \sqrt{\frac{1}{o}\sum_{j=1}^{o}(\beta_{j,t}^{\text{R}} - \beta_{j,t})^2}, \tag{12}$$

where o is the number of out-of-sample observations of realized and predicted betas at one point in time. $\beta_{j,t}^{R}$ is the realized beta in the period ranging from t to T, and $\beta_{j,t}$ denotes an estimate for beta. We rely on the RMSE criterion since it is robust to the presence of (mean zero) noise in the evaluation proxy while other commonly employed loss functions are not (Patton, 2011). We test for significance in RMSE differences using the modified Diebold–Mariano test proposed by Harvey et al. (1997). We use Newey & West (1987) standard errors with 4 lags. To test for significance in root median squared error (RMedSE) differences, we employ the non-parametric Wilcoxon signed rank test.[13] In general, the results for the RMedSE and its significance are similar to those for the RMSE. Hence, when discussing our results, we mainly focus on the RMSE results.

We believe that the combination of realized beta and the RMSE criterion is appropriate for evaluating forecasts for beta. Given that realized beta is a consistent estimator, in finite samples, the main concern with using this to evaluate beta forecasts is that it might be a noisy proxy of the true underlying beta. However, the RMSE criterion is robust to this noise (if it is zero on average). Thus, this combination appears to be well-suited.

---

[11]For all estimators, we require at least half of the data to be available in order to obtain an estimate.

[12]In Section VI.H, we also examine the Mean Absolute Error (MAE) loss function as an alternative, and obtain largely similar results as for the RMSE.

[13]Strictly speaking, the Wilcoxon signed rank test incorporates the joint null hypothesis of zero median in the loss differentials as well as a symmetric distribution. We stick to this test instead of an alternative only testing on zero median, like the simple sign test, since the Wilcoxon signed rank test turns out to be more powerful in many applications (Conover, 1999).

Market-Neutral Anomaly Portfolios   For estimating a firm's cost of capital and other applications, individual betas are very relevant. However, the ability of different beta estimation methods to create market-neutral portfolios is an interesting economic criterion on which to assess different beta estimation approaches that is at least as important. Therefore, along with the RMSE, we evaluate the average ex-post realized betas of ex-ante market-neutral long–short anomaly portfolios. We use several anomaly variables, which we describe in Section A of the Appendix.

Each month, we sort the stocks to form the different anomaly portfolios based on NYSE breakpoints. For size and value, as in Lewellen & Nagel (2006), we build 25 independently sorted portfolios. S is the average of the 5 low market-cap portfolios and B is the average of the 5 high market-cap portfolios while SMB is their difference. Similarly, L is the average of the 5 low book-to-market portfolios and H is the average of the 5 high book-to-market portfolios. HML is the difference between the value and growth portfolios. For momentum, we sort the stocks into 10 portfolios based on their return over the past 12 months while skipping the most recent month (Jegadeesh & Titman, 1993). W and L are the top and bottom deciles while WML is their difference. For idiosyncratic volatility (Ivol) and illiquidity (Illiq), we sort the stocks into 5 portfolios and form low-minus-high and high-minus-low portfolios, respectively. For leverage, we follow the same approach as for the book-to-market ratio, that is, we build 25 portfolios, independently sorted on size and leverage and form the high-minus-low portfolio.

For each beta estimator, we compute the long portfolio beta and the short portfolio beta predictions. We set the weight $v_{j,t}$ so that it fulfills the equation $v_{j,t}\beta_{j,t}^{\text{long}} - \beta_{j,t}^{\text{short}} = 0$.[14] We thus create anomaly portfolios that are ex-ante market-neutral. We then test whether the ex-post realized beta of the anomaly portfolios is indeed 0 on average.

---

[14]The results are qualitatively similar when keeping the weight of the long side at 1 and instead weighing the short side to make the portfolios market neutral.

III Estimating Beta

A Optimal Window Length and Sampling Frequency

We start the main analysis looking for the optimal sampling frequency and window length for the simple historical estimator. For the historical estimator, we consider windows of 1, 3, 6, 12, 24, 36, and 60 months when using daily data. Additionally, we consider the historical estimator based on monthly data ($HIST_{mon}$) using windows of 12, 36, and 60 months, as well as an estimator based on quarterly data using the returns over the previous 10 years.[15]

In Table 1, we present the summary statistics of these estimators. Several properties of the different estimators are worth mentioning. First, the value-weighed average beta, which should be equal to 1 when examining a complete market, is close to that value for most approaches. Values below 1 provide some indication that stocks are traded infrequently or that opacity prevents market participants from fully understanding the impact of systematic news during the chosen return interval. Values above 1 indicate that an estimator overestimates the systematic risk on average.

Second, we examine the average cross-sectional standard deviation of the approaches. A high standard deviation might be an indication of high measurement errors, whereas a very low standard deviation might indicate that an approach fails to sufficiently capture the heterogeneity in the estimates. Naturally, the average cross-sectional standard deviation is larger and the quantiles are wider for shorter estimation windows. Thus, the short-window historical estimators likely suffer from high measurement errors.

Third, we examine the average value-weighted correlation among the estimates. We find that the correlations are far from perfect even though we use exactly the same estimator for all approaches, and only change the historical window size and sampling frequency. For

---

[15]The subscript of the HIST estimators denotes the return frequency. This is left blank for daily data. We use the subscript "mon" for monthly and "q" for quarterly data. The superscript of the estimators indicates the period included in the estimation window (expressed in months).

example, the correlation of the historical estimator based on daily data and a 1-month historical window with that using a 60-month window is as low as 0.40. Additionally, even when using the same data window, the correlation of the 60-month historical estimator based on daily data with that based on monthly data is only 0.65.

To find the optimal combination of window length and sampling frequency, in Table 2 we present the average out-of-sample prediction errors of different historical estimators. We detect the typical trade-off between conditionality and sample size. On the one hand, beta changes over time. Hence, an estimate based on a short historical window delivers a more timely conditional estimate. On the other, estimates based on a small sample are prone to measurement error. Starting with daily data, we find that the average value-weighted RMSE is highest for the 1-month horizon. It falls gradually up to the 12-month horizon and begins to rise again for longer estimation windows. The average RMSE of the 12-month historical estimator (HIST^{12}) is significantly lower than that of the 1-month horizon 59% of the time, compared to the 3-month horizon estimator 42% of the time, and relative to the 60-month estimator 17% of the time.

Additionally, we find that low-frequency estimators, i.e., those based on monthly and quarterly data, yield very high average RMSEs, which are significantly higher than the RMSE of HIST^{12} about 80% of the time. This result is also in line with the finding of Hollstein et al. (2018), who examine the stocks of the S&P 500 and show that estimators based on higher-frequency data outperform those based on lower-frequency data. It thus seems that estimators based on higher-frequency data are generally preferable whenever reliable data are available.

In Table 3, we further test to what extent the estimators help create market-neutral anomaly portfolios. We find results that are broadly consistent with those when using the RMSE. For the 12-month estimation window using daily data, the average ex-post realized betas of the anomaly portfolios are small and significant for only 2 out of the 6 anomaly

portfolios. For very long or short horizons, the magnitude of the average realized betas is typically larger and the average ex-post realized betas are statistically significant more frequently. A very good alternative is also the 24-month estimation window. For this estimation window, only 1 anomaly portfolio has a statistically significant realized beta. However, the point estimates and standard errors are typically higher for the 24-month window compared to the 12-month window.

Overall, the daily historical estimator using a 12-month window both yields the most accurate predictions and fares well in creating market-neutral anomaly portfolios. In the following sections, we therefore concentrate on the 12-month estimation window, indicate $HIST^{12}$ simply by HIST, and examine if we can further improve its predictive accuracy by imposing different adjustments on the estimator.

B   Different Weighting Schemes

In the previous section, we address the trade-off between conditionality and sample size by searching for an optimal window that balances both arguments. However, it may also be possible to resolve this trade-off in an alternative manner. While, thus far, we weigh all observations equally, independently of whether the returns occur 11 months or 1 week before the date of the estimation, one could also implement an exponentially decaying weighting scheme. This way, we can use a large sample to estimate the parameters precisely and, at the same time, give a higher weight to more recent observations that likely carry better information on the current conditional beta. We use two different half-lifes for the exponential weighting, one that has a higher level of conditionality, where the half-life corresponds to 84 trading days (indicated by the additional subscript "s" for "short") and one where it is 168 trading days.[16] Additionally, we use each of the two half-lifes together with an expand-

---

[16]We compute this as 12 (months) times 21 (average daily return observations per month) times $\frac{1}{3}$ in the former and $\frac{2}{3}$ in the latter case.

ing window ($\text{HIST}_{\text{ewma,s,ex}}$ and $\text{HIST}_{\text{ewma,ex}}$), where we have an even larger sample size that might further increase the precision of the estimates.[17]

In Table 1, we present summary statistics for the exponentially weighted historical estimator.[18] We find that the overall properties of $\text{HIST}_{\text{ewma}}$ and $\text{HIST}_{\text{ewma,ex}}$ are similar to those of HIST^{12} and the correlation is high with the 12-month historical estimator employing equal weights. Thus, we expect that the differences might not be very large.

We present the results on prediction errors when using an exponential weighting scheme in Table 4. We find that, independently of the specification, the exponential weighting reduces the average value-weighted RMSE. We obtain the lowest average value-weighted RMSE for $\text{HIST}_{\text{ewma,ex}}$. The value-weighted RMSE is significantly lower for $\text{HIST}_{\text{ewma,ex}}$ compared to HIST 30% of the time. Thus, the exponential weighting, especially combined with an expanding estimation window, can reduce prediction errors in beta.

In Table 5, we present the portfolio results. We find that most approaches perform similarly well as HIST. The best approach, just like for the RMSE, is $\text{HIST}_{\text{ewma,ex}}$, which is able to create ex-post market-neutral portfolios on average for all anomalies except for momentum.

C   Imposing Priors

Another way to correct for potential measurement errors is to shrink potentially noisy estimates toward an informative prior. Estimates that have higher standard errors are thus shrunk more heavily toward their prior than estimates with lower standard errors. We use three different shrinkage estimators, HIST^{V}, HIST^{K}, and HIST^{I}.

Summary statistics of these estimators are presented in Table 1. Naturally, we find that

---

[17]One might wonder how much of the weight is assigned to observations more than 1 year past, when using an expanding window. For $\iota = 84$, this is roughly 12% and for $\iota = 168$, about 35% of the weight is placed on observations further back.

[18]To enhance the exposition, we only present the summary statistics for the estimator with $\iota = 168$. Those with a shorter half-life of the weights ($\iota = 84$) are qualitatively similar.

the distributions of HIST^V, HIST^K, and HIST^I are narrower than that of the unadjusted 12-month historical estimator; also, quite naturally, since HIST^V, HIST^K, and HIST^I are directly derived from HIST^{12} the estimators are highly cross-sectionally correlated with it and among each other. The correlation with HIST^{12} decreases with the amount of information imposed for the prior from 0.98 for HIST^V to 0.83 for HIST^I. On the other hand, the value-weighted average for HIST^V and HIST^K is slightly below 1 because HIST^V and HIST^K shrink the beta estimates toward an equally weighted average, which is typically below 1.[19] For HIST^I, we have far fewer observations compared to the simple historical approaches. This is because the approach requires accounting data that is not as widely available as stock data and because we need an initial window to estimate the parameters. The value-weighted average of HIST^I of 1.02 slightly exceeds 1.

We present the prediction errors of the different prior-adjusted betas in Table 6.[20] We find that HIST^V, HIST^K, and HIST^I yield lower average value-weighted RMSEs compared to HIST. The differences are significant 8%, 14%, and 4% of the time, respectively. HIST^K, which shrinks the beta estimates three times, is slightly better than the less informative HIST^V, which shrinks estimates toward the overall average beta only. The individual-priors-based estimator, HIST^I yields a slightly lower average RMSE compared to HIST^V (which becomes visible in the decimal places not tabulated) but a higher average RMSE than HIST^K. The differences are generally stronger in the median. HIST^I yields a relatively high RMedSE.

Our findings are in line with Karolyi (1992), who shows that the Vasicek (1973) beta adjustment yields a substantially lower RMSE compared to HIST and, furthermore, the multiple priors approach by Karolyi (1992) performs even better than HIST and than HIST^V.

---

[19]We also try HIST^V and HIST^K shrinking the beta estimates toward a value-weighted average. We find that in that case the value-weighted averages are closer to 1. The overall performance of the two estimators is qualitatively similar and generally even slightly better when shrinking towards the value-weighted average.

[20]Note that the average RMSE for HIST is different from that of Table 2 because both the sample period and stock universe differ slightly. We only include stock–month observations, for which all approaches in the table yield an estimate. The number of firm–month observations reduces because we need an in-sample period to first estimate the prior for HIST^I and because many firms lack accounting data.

Our findings are also consistent with those in Cosemans et al. (2016). The authors find that their individual-priors-based approach yields lower average RMSEs compared to HIST, HIST^V, and a version of the Karolyi (1992) estimator. We confirm these results in that we find that HIST^I yields a lower average RMSE than HIST and HIST^V, but the average RMSE is higher than that of HIST^K. The reason for this finding is that Cosemans et al. (2016) use a weaker specification of the Karolyi (1992) estimator, only shrinking the beta estimates to the industry average instead of the three-step shrinkage adjustment which Karolyi (1992) uses in his initial paper. When we use this weaker version of the Karolyi (1992) estimator used by Cosemans et al. (2016), we are able to confirm the findings of the authors that the Cosemans et al. (2016) estimator yields a slightly lower average RMSE.

Table 7 presents the results for market-neutral anomaly portfolios. We find that the more informative the prior, the worse the performance in generating ex-post market-neutral portfolios. While HIST^V only fails in creating market-neutral momentum portfolios, HIST^K fails for 3 portfolios, and HIST^I even fails for 5 out of the 6 portfolios. Since the prior for HIST^I is based on a characteristics-augmented monthly regression, the poor result for HIST^I in creating market-neutral anomaly portfolios is consistent with poor results for estimators based on monthly data. Thus, shrinkage toward more informative priors helps reduce mean squared errors but worsens the performance in portfolio formation.[21]

D Asynchronicity Adjustments

A possible concern when estimating betas is that some stocks might be traded less frequently than the market portfolio. If the stock price reacts days after the arrival of systematic news, the usual historical beta estimator will be biased downward. The usual approach to

---

[21]Cosemans et al. (2016) find that HIST^I performs well in creating a market-neutral minimum variance portfolio (MVP). For their MVP approach, we obtain similar results as they do. However, their approach has to make structural assumptions on the covariance matrix, which makes the test a joint test of the performance of a beta estimator and these structural assumptions. Furthermore, we feel that anomaly portfolios are a more relevant application, especially in practice.

handle this are the Scholes & Williams (1977) and Dimson (1979) adjustments, which also account for betas with respect to the leaded and lagged market return(s). Additionally, we use the correlation-separated beta of Frazzini & Pedersen (FP) (2014).

We present summary statistics for Dimson-betas with 1, 3, and 5 lags, SW, as well as FP betas with 12-month and 60-month correlations in Table 1. We find that the overall value-weighted averages of the Dimson-betas and SW are similar to that of HIST^{12}. Hence, the estimators do not appear to be systematically biased. Interestingly however, the value-weighted averages of FP^{12} and FP^{60} amount to 1.06 and 1.09, respectively, which indicates some upward bias for the estimator. We find that the standard deviation as well as the quantile range rise the more lags we use. Additionally, the correlations with HIST^{12} fall with an increasing number of lags. The average value-weighted cross-sectional correlation between HIST^{12} and $\text{Dim}^{(5)}$ is 0.70. Thus, adding betas with respect to lagged market returns materially affects the properties of the historical estimator.

We present the results for prediction errors when using up to 5 lags for the Dimson beta, SW, and 3 different horizons for the FP beta in Table 8. We find that the asynchronicity-adjustment does not improve the beta estimates on average. The more lags we use, the higher the average value-weighted RMSE. Those of $\text{Dim}^{(1)}$, $\text{Dim}^{(5)}$, and SW are significantly higher than that of HIST 48%, 77%, and 50% of the time, respectively. Hence, there is very little evidence to warrant a lag adjustment.[22] For the FP beta, we find a similar performance of the estimator whose correlation is based on a 12-month window as for $\text{Dim}^{(1)}$, while the average RMSE of FP is substantially smaller than that of, e.g., $\text{Dim}^{(3)}$ and $\text{Dim}^{(5)}$. Hence, the FP beta may be a bit better suited for beta estimation than the Dimson adjustment. The longer the estimation window for correlations, the worse is the performance of the FP

---

[22]Since we evaluate the predictions using realized beta without an adjustment for infrequent trading in the measurement of this quantity, we might fail to capture infrequent trading effects ex-post. We account for this possibility in Section VI.F and show that even under an evaluation that accounts for potential infrequent trading, the Dimson-adjusted estimator still falls short of the simple historical estimator.

estimator. Thus, it seems that correlations also exhibit substantial time-variation.

In Table 9, we examine the ability of the estimators to create ex-post market-neutral anomaly portfolios. Consistent with the RMSE results, we find that all estimators perform poorly and create a significant market exposure for at least 4 out of the 6 anomaly portfolios.

E Macroeconomic Conditioning Information

If betas change over the business cycle, one could make use of information on macroeconomic state variables to obtain better estimates for conditional betas. Thus, we examine the predictions of several potential state variables.

To enhance the exposition, in Table 1, we only present the summary statistics on one of the betas combined with macroeconomic state variables, Beta$^{\text{cay}}$ (Lettau & Ludvigson, 2001). The results of the other estimators are qualitatively similar. Since we first need initial data to estimate Equation (6), we have fewer overall observations. We find that the value-weighted average beta is very close to 1. The cross-sectional standard deviation and quantile range of Beta$^{\text{cay}}$ are neither very large nor very small and correlations to other approaches are rather low in comparison to those of most estimators.

In Table 10, we present the prediction error results for different macroeconomic conditioning variables. Because information on some of these is issued only on a quarterly basis, we sample the betas at the end of each quarter instead of at the end of each month.[23] We find that all of the estimators based on macroeconomic conditioning variables substantially and significantly underperform HIST. The performance of the "kitchen-sink" approach Beta$^{\text{all}}$ is particularly poor. The portfolio formation results of Table 11 are similarly poor. Thus, it appears to be much more favorable to roughly follow Lewellen & Nagel (2006), and use a (short) historical 12-month window to estimate conditional betas instead of using macroeconomic

---

[23]The results when sampling monthly for all variables that are available at that frequency are qualitatively similar.

conditioning variables as in, e.g., Lettau & Ludvigson (2001) or Guo et al. (2017).

F Forecast Combinations

Finally, we examine whether one can improve upon HIST by combining different estimates. We use two different sets of models to be combined: (i) only the estimators that performed best during the previous sections, $\text{HIST}_{\text{ewma,ex}}$ and HIST^{K} (Best), and (ii) a much larger subset of the different possible adjustments (All). For the latter, we combine HIST, $\text{HIST}_{\text{ewma,ex}}$, HIST^{K}, HIST^{I}, $\text{Dim}^{(5)}$, and Beta^{cay}.[24] For both model sets, we use four combination possibilities: (i) a simple combination, (ii) a model-based combination as of Equation (8), (iii) a model-based combination as in (ii) with the shrinkage approach of Diebold & Pauly (1990), and (iv) Bayesian model averaging.

Table A1 of the Online Appendix presents summary statistics on these combinations. We find that the properties of Best^{sim} are overall very similar to those of HIST and the average value-weighted cross-sectional correlation is 0.98. For the model-based combinations, we typically have far fewer observations. This is because we first need observations for each of the models we combine. Additionally, we need an initial window to perform the estimation of the weights. This further reduces the number of observations available when combining many models in All. Overall, we find that the value-weighted average, especially of the model combinations, slightly exceeds 1, which indicates that these combinations yield a bias on average.

We present the prediction error results in Table 12. We find that the simple combination Best^{sim} yields a significantly lower average value-weighted RMSE compared to HIST 32% of the time. The model-based combinations Best^{C}, Best^{shr}, and Best^{BMA} perform similarly to HIST, while the Bayesian approaches perform slightly better than the non-Bayesian com-

---

[24]We choose to only use a subset of all adjustments in the paper, since using too many highly correlated approaches creates problems of multicollinearity and yields extreme weights for the OLS-based combinations. Our overall conclusions are not sensitive to different choices of the models from the respective subsets.

bination Best$^{\text{C}}$. When combining all approaches, independently of whether they work or not individually, only the simple combination All$^{\text{sim}}$ performs better than HIST, but not as well as the simple combination of the best models. The model-based combinations of all approaches work clearly less well. These underperform HIST about 50% of the time and more often. Interestingly, we find that the Bayesian combinations perform even worse than All$^{\text{C}}$. Hence, simple equally weighted combinations typically yield lower prediction errors than more elaborated regression-based combinations, even when these use a Bayesian approach.

The portfolio formation results in Table 13 reveal that the combinations generally do not outperform HIST in making anomaly portfolios market-neutral. Best$^{\text{sim}}$ and All$^{\text{C}}$ work reasonably well, but overall not better than HIST. Thus, combining estimators appears to reduce the RMSE, but the evidence for portfolio formation is mixed.

G Which is the Best Approach?

Thus far, we examine which of the approaches yields an improvement relative to HIST. However, it is of course of high practical interest to know which of the adjustments and combinations yields the lowest prediction errors overall. We present the results for the best models of the previous sections, HIST$_{\text{ewma,ex}}$, HIST$^{\text{K}}$, an approach that directly imposes the industry-based prior on the EWMA Beta, HIST$^{\text{K}}_{\text{ewma,ex}}$, and the simple combination of the 2 best models, Best$^{\text{sim}}$ in Table 14.[25]

We find that, individually, HIST$_{\text{ewma,ex}}$ yields a slightly lower average value-weighted RMSE compared to HIST$^{\text{K}}$. Directly applying the prior suggested by Karolyi (1992) yields another small improvement for HIST$_{\text{ewma,ex}}$. However, the simple combination Best$^{\text{sim}}$ yields the overall lowest average value-weighted RMSE. While the differences in RMSE between

---

[25]One might wonder why the prediction errors of Table 14 are partially higher than those in Tables 2, 6, and 12 for the same models. As already indicated in footnote 20, for each table, we use only firm–month observations that are available for all the approaches presented. This yields a substantial reduction of firm–month observations in the earlier tables.

Best$^{\text{sim}}$ and HIST$_{\text{ewma,ex}}$ as well as HIST$^{\text{K}}_{\text{ewma,ex}}$ are only rarely significant, Best$^{\text{sim}}$ significantly outperforms HIST 52% of the time.

For portfolio formation, the overall picture of Section III indicates that the best approaches are HIST, HIST^{24}, HIST$_{\text{ewma,ex}}$, and HIST$^{\text{V}}$.

IV   Why do the Adjustments "Work"?

Given that some of the adjustments and combinations improve the predictability for beta while others yield substantially higher prediction errors, one may wonder what the reason for these different results is. We address this by performing a decomposition of the mean squared errors (MSE). To do so, we follow Mincer & Zarnowitz (1969) and decompose the MSE in the following fashion:

$$\text{MSE}_j = \underbrace{(\bar{\beta}_j^{\text{R}} - \bar{\beta}_j)^2}_{\text{bias}} + \underbrace{(1 - b_j)^2\sigma^2(\beta_j)}_{\text{inefficiency}} + \underbrace{(1 - \rho_j^2)\sigma^2(\beta_j^{\text{R}})}_{\text{random error}}. \tag{13}$$

b_j is the slope coefficient of the regression $\beta_j^{\text{R}} = a_j + b_j\beta_j + \epsilon_j$ and ρ_j^2 is the coefficient of determination of this regression. A bias indicates that the prediction is, on average, different from the realization. Inefficiency represents a tendency of an estimator to systematically yield positive forecast errors for low values and negative forecast errors for high values or vice versa. The remaining *random* forecast errors are unrelated to the predictions and realizations.

We present the results of the MSE decomposition for different adjustments that "work" and for others that "do not work" in Table 15. Furthermore, we present key statistics about forecast errors in Table 16. We choose to report a subset of the estimators in order to study the main mechanisms while keeping the presentation manageable. The results for other estimators of the same group are generally very similar.

Weighting Schemes   Comparing $HIST_{ewma,ex}$ to HIST, Table 15 shows that for the weighted estimator, the bias and random error parts are of similar magnitude, but $HIST_{ewma,ex}$ substantially reduces the inefficiency compared to HIST. Thus, the larger estimation window appears to reduce the likelihood of high measurement error in low-beta and high-beta stocks. The weighting scheme ensures that the estimates are conditional. Table 16 confirms this interpretation. The share of very high squared forecast errors is substantially reduced. Particularly for very high and very small realized betas, $HIST_{ewma,ex}$ works better than HIST.

Imposing Priors   We find that $HIST^K$, $HIST^I$, and $HIST^K_{ewma,ex}$ reduce the inefficiency relative to HIST. Thus, it seems that the approaches partially correct the tendency of HIST of underestimating the betas of low-beta stocks and overestimating the betas of high-beta stocks. The correction mechanism is especially intuitive for the shrinkage approaches: some betas will mechanically be estimated with high measurement error. It is particularly likely that very low beta estimates contain negative measurement errors while very high beta estimates contain positive measurement errors. The Bayesian prior approaches take a step toward detecting and correcting these measurement errors. If the standard error in a beta estimate is high, it is shrunk more strongly toward an informative prior.

The forecast error statistics reveal further interesting patterns: compared to HIST, $HIST^K$ reduces the share of very high squared forecast errors and works better particularly for stocks with very high or very low realized betas. For $HIST^I$, the results are somewhat different. $HIST^I$ yields the lowest RMSE among all approaches for stocks with high realized betas. However, the approach performs very poorly for stocks with low realized betas and has an overall higher share than HIST with very high squared forecast errors. This pattern also delivers an explanation for the poor results of $HIST^I$ in creating market-neutral anomaly portfolios. These extreme long–short portfolios typically contain disproportionate numbers of low-beta stocks.

The portfolio formation results indicate that $HIST^K$ and $HIST^I$ place too much weight

on their respective priors. In particular, the prior beta of Cosemans et al. (2016) changes mechanically when new accounting data becomes available, which might not be very informative about a stock's beta. Our findings are thus consistent with recent evidence in Dittmar & Lundblad (2017), who find that market betas are only weakly related to stock characteristics.

Asynchronicity Adjustments The estimators $\text{Dim}^{(1)}$, $\text{Dim}^{(3)}$, and $\text{Dim}^{(5)}$ do not strongly increase the bias, but both the inefficiency and the random error increase with increasing number of lags. In particular the increase in random errors indicates that adding lagged betas introduces a lot of noise on the estimators, amplifying measurement errors. FP^{12} exhibits a large bias part, which is more than double that of HIST, a high inefficiency when compared to the best models, but a random error that is only slightly above that of HIST. Thus, the correlation-adjustment to three-day correlations induces substantially less noise than the Dimson adjustment. The forecast error statistics corroborate these findings. The Dimson estimators yield a very high share of stocks with high squared forecast errors, the share increasing with the number of lags.

Macroeconomic Conditioning Information The beta augmented by macroeconomic conditioning variables, Beta^{cay}, yields a higher bias component compared to HIST, a high inefficiency, and a very high random error. Thus, the approach using macroeconomic conditioning variables appears to yield systematic measurement errors for high-beta and low-beta stocks. More importantly, the approach seems to add a lot of noise to beta estimates. Since the reasons for the failure of the macroeconomic conditioning variables might be diverse, we present the forecast error statistics for all estimators in Table A2 of the Online Appendix. All approaches yield a large share of high squared forecast errors and perform poorly for particularly for stocks with high realized betas. Part of the motivation to use macroeconomic conditioning variables is that betas might adapt more quickly to changing economic conditions. However, all approaches work poorly in both expansions and recessions,

28

as well as in the transition between expansion and recession and vice versa.

Finally, a common scheme seems to be that the average RMSEs and the difference in RMSE to HIST are both significantly correlated with the underlying conditioning variables. This correlation results when a majority of the slope estimates of Equation (6) share the same sign but the realized beta does not follow the imposed dynamics. For few variables, the RMSE is largely uncorrelated to the underlying conditioning variables, indicating that the stocks have heterogeneous slope estimates. However, these variables also do not work better than those correlated with the RMSEs. Overall, it seems that individual stock betas are not systematically related to macroeconomic variables and imposing macroeconomic conditioning variables induces random noise by assuming a spurious relation between macroeconomic variables and betas.

Combinations   For the combinations, we find that Best$^{\text{sim}}$ yields a similar bias as HIST, but a very low inefficiency part and the lowest overall random error part. Thus, by simply averaging two of the overall best approaches, the estimator diversifies and reduces random measurement errors. This is also visible when looking at the forecast error statistics: Best$^{\text{sim}}$ yields the lowest share of high squared forecast errors.

The model-based combinations typically have a higher inefficiency part and the combinations including a large number of predictors yield a very high random error part. Particularly All$^{\text{C}}$ and All$^{\text{BMA}}$ yield huge shares of very high squared forecast errors. These large errors are likely caused by in-sample overfitting of the beta dynamics during the estimation window. Because the beta estimates are typically strongly correlated, we often observe that slope coefficients turn negative. This is also the case for the Bayesian approaches. That in-sample overfitted models generate high prediction errors is a common result in financial economics (e.g., Stock & Watson, 2006; Goyal & Welch, 2008).

V   Beta Estimation and Stock Characteristics

In the previous sections, we examine the performance of different beta estimation techniques *unconditionally*, i.e., without regard to the specific characteristics of the stocks as well as by their ability to create market-neutral anomaly portfolios. In this section, we complement our previous analysis with a deeper examination of how stock characteristics affect the optimality of different beta estimators. In the following, we explicate the results for each sorting characteristic, presented in Table 17.[26]

Past Beta   First, we sort on the observation of HIST, obtained using data from $t-24$ until $t-12$.[27] Overall, we find that the squared prediction errors are highest in the extreme portfolios, i.e., those with the lowest and highest past betas. Thus, it seems to be particularly difficult to accurately predict the betas of the stocks with very high or very low systematic risk. Comparing the performance of the different estimators, we find that those that perform best unconditionally also perform well for each of the quintiles. In total, HIST and the best adjustments perform almost similarly well. Thus, in general it seems that building portfolios helps reduce the inefficiency of HIST. $\text{Dim}^{(5)}$ performs particularly badly for the portfolio with the lowest past betas, where the asynchronicity correction might be most relevant. Thus, one should interpret this result with caution. However, also for the stocks with high past beta, where we can compare the performance more accurately since an upward-bias due to infrequent trading is highly unlikely, $\text{Dim}^{(5)}$ falls short of the best approaches. FP^{12} performs poorly across all quintiles.

Size   Sorting the stocks according to the market capitalizations, we find that for P5,

---

[26]An alternative way to study such a question are panel regressions, where we can include several variables at the same time. However, it is likely that the relationship between the prediction errors and the characteristics is highly non-linear. Therefore, we decided to use portfolio sorts, where we do not have to take a stance on the parametric form of the relation. The results of multivariate panel regressions, when applied nevertheless, are largely similar to those of the portfolio sorts.

[27]Thus, the periods we use for sorting and the estimation of HIST are strictly non-overlapping. For all stocks, for which we cannot obtain a past beta due to lack of sufficient data during the period $t-24$ until $t-12$, we set it to 1.

the RMSE is markedly low, amounting to less than half of that of the other portfolios. This result holds for each beta estimation approach. Hence, for large stocks it is considerably less difficult to accurately estimate betas than for the majority of stocks. Among the different estimators, we find that the unconditionally best adjustments and HIST perform similarly well. Best$^{\text{sim}}$ yields the lowest average RMSE over the quintiles. Dim$^{(5)}$ and FP^{12} also yield higher RMSEs than the best approaches for the portfolios of the largest stocks (P4 and P5). Thus, at a minimum, the asynchronicity adjustment is harmful for estimating the betas of large stocks, for which the realized beta should not be affected by non-synchronous trading.

Book-to-market   For value stocks (high book-to-market ratios), betas appear to be harder to estimate than for growth stocks. For all approaches, the RMSE is higher for the value portfolio (P5) compared to the growth portfolio (P1). Overall, the RMSEs are lowest for P2 and P3. The pattern across the different approaches is similar as for the previous sorts.

Momentum   When sorting on past performance, we observe a distinct pattern for all approaches: the highest RMSEs are in the portfolios of the stocks with the worst (P1) and best (P5) past performance. This pattern is consistent with the evidence found in Kothari & Shanken (1992) and Daniel & Moskowitz (2016), who show that the extreme momentum portfolios exhibit substantial time-variation in beta. Furthermore, Chen et al. (2016) argue that past winners in general exhibit high betas and vice versa. Thus, these extreme betas in general do not persist in future periods. Consistent with this, in Section III, we find that most approaches fail in creating market-neutral momentum portfolios. Among the approaches, the relation is largely similar as before. However, interestingly, for the momentum winner portfolio (P5), the RMSEs of Dim$^{(5)}$ and FP^{12} are comparably low. Thus, in part, the lag and correlation adjustments of these approaches appear to be helpful in correcting for the measurement error typically present in high-momentum stocks. On the other hand, the simple historical estimator performs better than Dim$^{(5)}$ and FP^{12} for all other portfolios.

Idiosyncratic volatility   Sorting on idiosyncratic volatility, we detect a natural pattern. For the stocks with the lowest idiosyncratic volatilities (P1 and P2), the betas are most accurately predictable while for the remaining portfolios, the RMSEs rise with idiosyncratic volatility.

Illiquidity   For illiquidity, the results are largely inverse to those for size. The cross-sectional correlation among those two variables exceeds -90%. Thus, the conclusions we can draw are similar.

Leverage   In case of sorting on the stocks' leverage, we find that the betas of the stocks in the extreme portfolios are least accurately predictable for all our approaches. For the intermediate leverage portfolios, the RMSEs for all approaches are lower than for the extremes.

Industries   For the industry portfolios, we observe the lowest RMSEs across all approaches for Other, Manufacturing, Shops, and Non-Durables. In contrast, the betas for stocks of Energy, High-Tech, Health, Telecommunication, and Utilities stocks are substantially harder to predict by the approaches we study.[28] Comparing the different approaches, as observed throughout all portfolio sorts, we find that Best$^{\text{sim}}$ yields the lowest RMSE for 6 out of the 10 industry portfolios, while Dim$^{(5)}$ and FP^{12} always yield higher RMSEs than HIST.

---

[28]One might wonder whether this result is mechanical, driven by different levels of idiosyncratic volatility among the different industries. However, we find that there is essentially no relation detectable. Other and Shops are among the sectors with the highest average idiosyncratic volatilities while Telecommunication and Utilities belong to the industries with the lowest average idiosyncratic volatilities.

VI   Additional Analyses and Robustness

A   Different Horizons

In this section, we examine the results for different forecast horizons of 1, 3, 12, and 60 months. Table A3 of the Online Appendix presents these results. To enhance the exposition, we only report the results on the best models and an estimation horizon of 12 months. The results for the remaining specifications are qualitatively similar as those for the 6-month forecast horizon.[29]

We start the analysis by examining 1-month forecasts. We present these results in Panel A of Table A3. We find that for all approaches, the average value-weighted RMSE is higher than for the 6-month horizon. This is most likely due to higher measurement errors in the estimator for realized beta which suffers from a reduced evaluation window.[30] We find that the adjustments that "work" for the 6-month horizon also yield lower average value-weighted RMSEs than the simple historical model. Similar to the 6-month horizon, the simple combination Best[sim] yields the lowest overall average value-weighted RMSE.

The results for the 3-month horizon are presented in Panel B of Table A3. With the longer evaluation horizon, for all approaches the average value-weighted RMSEs are substantially lower than for the 1-month horizon. All other results are qualitatively similar to the 1- and 6-month horizons.

Panel C of Table A3 presents the results for the 12-month forecast horizon. We find that for all approaches, the average value-weighted RMSEs are lower than for the 6-month horizon. This pattern indicates that 12-month betas are slightly more predictable than betas of shorter horizons. All adjustments that perform better than the simple historical model

---

[29]For the 1- and 3-month horizons, the 12-month historical window also turns out optimal. For the 12- and 60-month forecast horizons, longer historical windows yield slightly lower average value-weighted RMSEs.

[30]However, as indicated previously, the RMSE criterion is still a robust evaluation criterion if the sampling error is zero on average (Patton, 2011).

for the 6-month horizon also do so for the 12-month horizon.

Finally, we present the results for the 60-month forecast horizon, relevant for long-term investors, in Panel D of Table A3. We find that the average value-weighted RMSEs for all approaches are slightly higher than for the 12-month horizon. Thus, it appears that time-variation in beta renders 60-month betas slightly harder to predict than 12-month betas. However, the average value-weighted RMSEs are still slightly lower than for the 6-month horizon. Apart from that, the results for the 60-month horizon are qualitatively similar to those for other horizons. Overall, Best$^{\text{sim}}$ yields the lowest average value-weighted RMSE.

B Hedging Errors

The RMSE results show that the approaches HIST$_{\text{ewma,ex}}$, HIST$^{\text{K}}$, HIST$^{\text{K}}_{\text{ewma,ex}}$, and Best$^{\text{sim}}$ yield the best results, while, e.g., Dim$^{(5)}$ performs very poorly. To account for the possibility that our ex-post realized betas are measured with error, we follow Liu et al. (2018) and examine the out-of-sample hedging errors of our main approaches. If realized beta estimates are biased, we may falsely conclude that an approach is superior simply because it is biased in a similar fashion. We thus compute the hedging error for each stock as

$$h_{j,t,T} = (r_{j,t,T} - r_{f,t,T}) - \beta_{j,t}(r_{M,t,T} - r_{f,t,T}). \tag{14}$$

$r_{j,t,T}$ is the return of stock j between t and T. $r_{f,t,T}$ and $r_{M,t,T}$ are the risk-free rate and the return on the market portfolio over the same horizon. We use 1-month returns. β_t is the estimate for beta, using data up to time t. Liu et al. (2018) show that under certain assumptions the hedging error variance ratio $\frac{\text{var}(h_{j,t,T})}{\text{var}(r_{M,t,T} - r_{f,t,T})}$ is approximately equal to the mean squared error relative to the *true* realized beta plus a term that is unrelated to the beta estimation, i.e., constant across all estimation approaches. We follow Liu et al. (2018) and estimate the variance ratios using rolling 5-year windows to account for the possibility that

the variances in the numerator and denominator change over time. We report the average ratio over time.

We present the results in Table A4 of the Online Appendix. These results are consistent with our previous results relying on the RMSE and realized beta computations. We find that $\text{HIST}_{\text{ewma,ex}}$, HIST^{K}, $\text{HIST}^{\text{K}}_{\text{ewma,ex}}$, as well as Best^{sim} yield significantly lower mean average hedging error ratios than HIST. $\text{Dim}^{(5)}$ yields a substantially and significantly higher mean average hedging error ratio than HIST. $\text{HIST}^{\text{K}}_{\text{ewma,ex}}$ achieves the lowest mean average hedging error ratio. Thus, our main results appear to be robust to the specification of forecast error measurement.

C   Mincer–Zarnowitz Regressions

As an alternative way to evaluate the performance of the beta estimators, we use simple Mincer & Zarnowitz (1969) regressions. We regress the 6-month (ex-post) realized beta on the different predictions for beta:

$$\beta_{j,t}^{\text{R}} \;\; = \;\; a + b\beta_{j,t} + \epsilon_{j,t}, \tag{15}$$

where all variables are as previously defined.

The regression model in Equation (15) is designed to test for unbiasedness of different estimators. We use univariate regressions to test for unbiasedness using a Wald test, which imposes the joint hypothesis that $a = 0$ and $b = 1$. If the model is unbiased, the joint hypothesis of the Wald test cannot be rejected.[31] We stick to level Mincer–Zarnowitz regressions instead of logarithmically transforming our variables since beta is theoretically unbounded and can also take on negative values. Hansen & Lunde (2006) show that level

---

[31]Mincer–Zarnowitz regressions are also designed to test for informational efficiency in multivariate regressions. However, since the estimates of our models are partially very highly cross-sectionally correlated, we consider the significance tests for differences in RMSE as more sensible to directly rank the models. We thus refrain from performing multivariate regressions.

Mincer–Zarnowitz regressions are robust to (mean zero) errors in the evaluation proxy.

We present the results in Table A5 of the Online Appendix. Consistent with our previous results, we find that the models performing best thus far are also less biased than HIST. For all adjustment models, the average intercept coefficient is closer to 0 and the slope is closer to 1. However, in the vast majority of the cases, we still reject unbiasedness of the models. Also consistent with our previous results, $\text{Dim}^{(5)}$ and FP^{12} yield heavily biased forecasts for realized beta.

D Equally Weighted Results

Thus far, we present primarily value-weighted results. We regard this as the most relevant case, since for investment decisions the stocks provide investment opportunities relative to their total market capitalizations. However, small stocks make up a very large fraction of the total number of stocks and, thus, it is also interesting to examine to what extent the adjustments are beneficial for these. Therefore, in this section, we examine the robustness of our main findings when weighing all stocks equally.

We present the equally weighted prediction error results in Table A6 of the Online Appendix.[32] We find that all the average RMSEs are higher for all approaches than for the value-weighted examination. This is consistent with previous results showing that it is considerably more difficult to estimate the betas of small stocks than it is for large stocks. Apart from that, the adjustment approaches that work best when value-weighting also significantly outperform HIST when weighing equally. Typically, the difference in the equally weighted RMSE is significant considerably more often than that in the value-weighted RMSE. Thus, the adjustments appear to be even more beneficial for small stocks compared to large stocks.

---

[32]One might wonder whether the asynchronicity adjustment performs better for small stocks. However, we find that the Dimson beta estimators are even more clearly inferior compared to HIST when weighing all stocks equally. The average RMSE is significantly higher than that of HIST nearly all the time, independent of the number of lags used.

Overall, $\text{HIST}^{\text{K}}_{\text{ewma,ex}}$ yields the lowest average RMSE.

E   Firm-Level Evaluation

In the main part of this paper we evaluate the forecasts cross-sectionally. That is, each month we examine how well one approach predicts future betas of all stocks in the cross-section. However, it may also be of interest to see how the adjustment approaches perform for different stocks on average in the time series dimension. To perform this analysis, and in order to assess the statistical significance and to prevent stocks which are only available over short intervals during our sample period from biasing our results, we use only stocks with more than 100 observations. Essentially, this approach implies that we potentially lose information from stocks available for a shorter sample period.

We present the results in Table A7 of the Online Appendix. These are qualitatively similar to those for the cross-sectional evaluation. The best adjustments also yield lower value-weighted average RMSEs compared to HIST. Best$^{\text{sim}}$ yields the overall lowest value-weighted average RMSE.

F   Dimson Evaluation

To further test the robustness of our main results to infrequent trading effects, in this section, we present the results when evaluating estimates with respect to a Dimson realized beta. We estimate the realized beta as the sum of the realized beta as of Equation (11) with 0 up to 5 lags.[33],[34]

We present the results in Table A8 of the Online Appendix. First, we find that the average value-weighted RMSEs are higher for all approaches. Thus, it seems to be very hard to predict future Dimson realized betas. This is most likely due to higher measurement

---

[33]E.g., for 1 lag, in the numerator we multiply $r_{j,\tau}$ by $r_{M,\tau-1}$ instead of $r_{M,\tau}$, etc.

[34]Using the ex-post historical Dimson estimator of Equation (4) instead yields results that are qualitatively similar.

error caused by adding betas with respect to lagged market returns. Second, we find the same patterns as when using realized beta without lags. The best approaches also yield improvements over HIST under the Dimson realized beta. Finally, we find that $\text{Dim}^{(5)}$ yields a higher average value-weighted RMSE than HIST even under the Dimson realized beta evaluation. However, the difference is considerably smaller and significant less frequent than under the regular realized beta.[35] Interestingly, however, under the Dimson realized beta evaluation, FP^{12} yields the lowest average RMSE. Thus, the infrequent trading adjustment of Frazzini & Pedersen (2014) might have some benefit in selected cases.

G   Monthly Evaluation

In the previous section, we account for infrequent trading effects by evaluating the forecasts with a Dimson realized beta. However, as we show in Section IV, the Dimson estimators yield a substantial increase in both the inefficiency and the random error compared to HIST. Therefore, in this section, we examine an alternative possibility to account for asynchronous trading in realized beta: the use of monthly data. We estimate the realized beta using monthly data over a 5-year horizon. Since this approach mixes a change in the data frequency and estimation window, it is probably most suitable to compare the results to those in Panel D of Table A3 of the Online Appendix.

Table A9 of the Online Appendix presents the results. These are very similar to those when using daily returns for realized beta. For the monthly evaluation, the average value-weighted RMSEs of all approaches are larger than for daily evaluation over 60 months. This indicates that realized beta with monthly data is also difficult to predict, likely because it is more prone to measurement errors than realized beta with daily data. Additionally, we

---

[35]Untabulated results reveal that, e.g., when we use 3 lags only both for the estimation ($\text{Dim}^{(3)}$) and the realized beta, $\text{Dim}^{(3)}$ also yields a slightly higher average value-weighted RMSE than HIST. Interestingly, contrary to what one might expect, for all numbers of lags, HIST is even more strongly favorable when weighting all stocks equally. Thus, the Dimson adjustment seems to work even less well for the stocks it was initially designed for.

find that Best$^{\text{sim}}$ yields the lowest overall average value-weighted RMSE. Finally, Dim$^{(5)}$ and FP^{12} still yield substantially and often significantly higher average value-weighted RMSEs compared to HIST and even more so compared to Best$^{\text{sim}}$. Hence, in general, estimators that do not account for infrequent trading appear to be superior to those that use a Dimson or FP adjustment.

H   Mean Absolute Error

Finally, we examine the robustness of our results to the loss function employed. As an alternative to the RMSE, in this section, we use the mean absolute error (MAE):

$$\text{MAE} \;\; = \;\; \frac{1}{o} \sum_{t=1}^{o} \mid \beta_{j,t}^{R} - \beta_{j,t} \mid, \tag{16}$$

where all variables are as previously defined. The MAE penalizes all forecast errors in the same way. Hence large forecast errors are less influential under the MAE than under the RMSE. We present the results in Table A10 of the Online Appendix. These are very similar to those using the RMSE. The best models under the RMSE also yield improvements over HIST under the MAE, the differences are significant for similar shares of the time, and Best$^{\text{sim}}$ yields the lowest value-weighted average MAE.

VII   Conclusion

We examine the effects of different historical windows, sampling frequencies, and various forecast adjustments on beta estimation. We find that using daily data over a 12-month horizon generally yields lower prediction errors than alternative historical windows and estimators based on low-frequency data. Furthermore, exponential weighting schemes, simple shrinkage adjustments toward a prior, as well as simple combinations, yield improvements.

For portfolio formation, the results are generally similar. However, elaborated shrinkage methods and forecast combinations perform worse.

In contrast, adjusting for asynchronous trading, conditioning beta on the deviations of macroeconomic state variables from their historical averages, and regression-based as well as Bayesian model averaging combinations, typically yield high prediction errors and fail to create market-neutral portfolios.

Analyzing the cross-section of beta predictability with respect to different stock characteristics, we find that in particular stocks with high and low betas, low size, high and low momentum, and high idiosyncratic volatility, are relatively difficult to predict.

Appendix

A   Cross-Sectional Variables

- **Book-to-market** (Fama & French, 1992) is the most current observation for book equity divided by the current market capitalization. Following the standard literature, we assume that the book equity of the previous year's balance sheet statement becomes available at the end of June. Book equity is defined as stockholders' equity, plus balance sheet deferred taxes and investment tax credit, plus post-retirement benefit liabilities, minus the book value of preferred stock. We obtain the data necessary from Compustat.

- **Idiosyncratic volatility** (Ang et al., 2006, "Ivol") is the standard deviation of the residuals $\epsilon_{j,\tau}$ in the Fama & French (1993) 3-factor model $r_{j,\tau} - r_{f,\tau} = \alpha_{j,t} + \beta_{j,t}^M(r_{M,\tau} - r_{f,\tau}) + \beta_{j,t}^S SMB_\tau + \beta_{j,t}^H HML_\tau + \epsilon_{j,\tau}$, using daily returns over the previous year.

- **Illiquidity** (Amihud, 2002, "Illiq") is the absolute value of the stock's return divided by the daily dollar volume, averaged over the previous year. Specifically, it is $Illiq_t = \frac{1}{n}\sum_{\tau=1}^{n}\frac{|r_{j,\tau}|}{Volume\$_\tau}$, with the daily dollar volume ($Volume\$_\tau$, in thousand dollars) being calculated as last trade price times shares traded on day τ, while the summation is taken over all n trading days during the examination period.

- **Industry Classifications** employ the definition for 10 industry portfolios applied by Kenneth French. "NoDur" is Consumer Non-Durables, "Durbl" is Consumer Durables, "Manuf" is Manufacturing, "Enrgy" is the oil, gas, and coal extraction industry, "HiTec" is Business Equipment, "Telcm" is Telephone and Television Transmission, "Shops" are Wholesale, Retail, and Services, "Hlth" is Healthcare, Medical Equipment, and Drugs, "Utils" is Utilities, and "Other" contains Mines, Construction, Construction Materials, Transport, Hotels, Bus Services, Entertainment, as well as Finance.

- **Leverage** (Bhandari, 1988) is defined as one minus book equity (see "Book-to-market") divided by total assets (Compustat: AT). Book equity and total assets are updated every 12 months at the end of June. We obtain the data necessary from Compustat.

- **Momentum** (Jegadeesh & Titman, 1993) is the cumulative stock return over the period from $t - 12$ until $t - 1$.

- **Size** (Banz, 1981) is the current market capitalization of a firm. Market capitalization is computed as the product of the stock price and the number of shares outstanding.

References

Amihud, Y. (2002). Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets*, *5*(1), 31–56.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Wu, J. G. (2006). Realized beta: Persistence and predictability. In T. B. Fomby, & D. Terrell (Eds.) *Advances in Econometrics: Econometric Analysis of Economic and Financial Time Series*, (pp. 1–40). New York: Elsevier.

Ang, A., Hodrick, R. J., Xing, Y., & Zhang, X. (2006). The cross-section of volatility and expected returns. *Journal of Finance*, *61*(1), 259–299.

Ang, A., & Kristensen, D. (2012). Testing conditional factor models. *Journal of Financial Economics*, *106*(1), 132–156.

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics*, *9*(1), 3–18.

Barndorff-Nielsen, O. E., & Shephard, N. (2004). Econometric analysis of realized covariation: High frequency based covariance, regression, and correlation in financial economics. *Econometrica*, *72*(3), 885–925.

Bates, J. M., & Granger, C. W. (1969). The combination of forecasts. *Operations Research Quarterly*, *20*(4), 451–468.

Bhandari, L. C. (1988). Debt/equity ratio and expected common stock returns: Empirical evidence. *Journal of Finance*, *43*(2), 507–528.

Bollerslev, T., & Zhang, B. Y. (2003). Measuring and modeling systematic risk in factor pricing models using high-frequency data. *Journal of Empirical Finance*, *10*(5), 533–558.

Buss, A., & Vilkov, G. (2012). Measuring equity risk with option-implied correlations. *Review of Financial Studies*, *25*(10), 3113–3140.

Chang, B.-Y., Christoffersen, P., Jacobs, K., & Vainberg, G. (2012). Option-implied measures of equity risk. *Review of Finance*, *16*(2), 385–428.

Chen, H., Singal, V., & Whitelaw, R. F. (2016). Comovement revisited. *Journal of Financial Economics*, *121*(3), 624–644.

Clemen, R. T. (1989). Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting*, *5*(4), 559–583.

Conover, W. (1999). *Practical Nonparametric Statistics*. New York: Wiley, 3 ed.

Cosemans, M., Frehen, R., Schotman, P. C., & Bauer, R. (2016). Estimating security betas using prior information based on firm fundamentals. *Review of Financial Studies*, *29*(4), 1072–1112.

Daniel, K., & Moskowitz, T. J. (2016). Momentum crashes. *Journal of Financial Economics*, *122*(2), 221–247.

Diebold, F. X., & Pauly, P. (1990). The use of prior information in forecast combination. *International Journal of Forecasting*, *6*(4), 503–508.

Dimson, E. (1979). Risk measurement when shares are subject to infrequent trading. *Journal of Financial Economics*, *7*(2), 197–226.

Dittmar, R. F., & Lundblad, C. T. (2017). Firm characteristics, consumption risk, and firm-level risk exposures. *Journal of Financial Economics*, *125*(2), 326–343.

Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *Journal of Finance*, *47*(2), 427–465.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3–56.

Fama, E. F., & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, *81*(3), 607–636.

Fernandez, C., Ley, E., & Steel, M. F. (2001). Benchmark priors for Bayesian model averaging. *Journal of Econometrics*, *100*(2), 381–427.

Ferson, W. E., & Schadt, R. W. (1996). Measuring fund strategy and performance in changing economic conditions. *Journal of Finance*, *51*(2), 425–461.

Frazzini, A., & Pedersen, L. H. (2014). Betting against beta. *Journal of Financial Economics*, *111*(1), 1–25.

Gilbert, T., Hrdlicka, C., Kalodimos, J., & Siegel, S. (2014). Daily data is bad for beta: Opacity and frequency-dependent betas. *Review of Asset Pricing Studies*, *4*(1), 78–117.

Goyal, A., & Welch, I. (2008). A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies*, *21*(4), 1455–1508.

Guo, H., Wu, C., & Yu, Y. (2017). Time-varying beta and the value premium. *Journal of Financial and Quantitative Analysis*, *52*(4), 1551–1576.

Hansen, P. R., & Lunde, A. (2006). Consistent ranking of volatility models. *Journal of Econometrics*, *131*(1), 97–121.

Harvey, D., Leybourne, S., & Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting*, *13*(2), 281–291.

Hollstein, F., & Prokopczuk, M. (2016). Estimating beta. *Journal of Financial and Quantitative Analysis*, *51*(4), 1437–1466.

Hollstein, F., Prokopczuk, M., & Wese Simen, C. (2018). The conditional CAPM revisited: Evidence from high-frequency betas. *Leibniz University Hannover Working Paper*.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance*, *48*(1), 65–91.

Jurado, K., Ludvigson, S. C., & Ng, S. (2015). Measuring uncertainty. *American Economic Review*, *105*(3), 1177–1216.

Karolyi, G. A. (1992). Predicting risk: Some new generalizations. *Management Science*, *38*(1), 57–74.

Kothari, S. P., & Shanken, J. (1992). Stock return variation and expected dividends: A time-series and cross-sectional analysis. *Journal of Financial Economics*, *31*(2), 177–210.

Lettau, M., & Ludvigson, S. (2001). Resurrecting the (C) CAPM: A cross-sectional test when risk premia are time-varying. *Journal of Political Economy*, *109*(6), 1238–1287.

Levi, Y., & Welch, I. (2017). Best practice for cost-of-capital estimates. *Journal of Financial and Quantitative Analysis*, *52*(2), 427–463.

Lewellen, J., & Nagel, S. (2006). The conditional CAPM does not explain asset-pricing anomalies. *Journal of Financial Economics*, *82*(2), 289–314.

Liu, J., Stambaugh, R. F., & Yuan, Y. (2018). Absolving beta of volatility's effects. *Journal of Financial Economics*, *128*(1), 1–15.

Mincer, J. A., & Zarnowitz, V. (1969). The evaluation of economic forecasts. In J. Mincer (Ed.) *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance*, (pp. 1–46). Cambridge, MA: Elsevier.

Newey, W. K., & West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, *55*(3), 703–708.

Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, *160*(1), 246–256.

Scholes, M., & Williams, J. (1977). Estimating betas from nonsynchronous data. *Journal of Financial Economics*, *5*(3), 309–327.

Shanken, J. (1990). Intertemporal asset pricing: An empirical investigation. *Journal of Econometrics*, *45*(1), 99–120.

Stock, J. H., & Watson, M. W. (2006). Forecasting with many predictors. In G. Elliott, C. W. J. Granger, & A. Timmermann (Eds.) *Handbook of Economic Forecasting*, (pp. 515–554). Amsterdam: Elsevier.

Timmermann, A. (2006). Forecast combinations. *Handbook of Economic Forecasting*, *1*, 135–196.

Vasicek, O. A. (1973). A note on using cross-sectional information in Bayesian estimation of security betas. *Journal of Finance*, *28*(5), 1233–1239.

Table 1: Summary Statistics

This table presents value-weighted summary statistics for the main estimators as well as averages of value-weighted cross-sectional correlations. *Nobs* indicates the total number of firm–month observations for which we have estimates. $Mean^{vw}$ is the time-series average of the value-weighted mean of the estimates. *Std. dev.* presents the average cross-sectional standard deviation. $q^{0.05}$ and $q^{0.95}$ indicate the averages of the cross-sectional 5% and 95% quantiles, respectively. The sample period runs from January 1963 until December 2015.

	Nobs	$Mean^{vw}$	Std. dev	$q^{0.05}$	$q^{0.95}$	$HIST^3$	$HIST^6$	$HIST^{12}$	$HIST^{24}$	$HIST^{36}$	$HIST^{60}$	$HIST^{60}_{mon}$	$HIST_q$	$HIST_{ewma}$	$HIST_{ewma,ex}$	$HIST^V$	$HIST^K$	$HIST^I$	$Dim^{(1)}$	$Dim^{(3)}$	$Dim^{(5)}$	SW	FP12	FP60	$Beta^{cay}$
$HIST^1$	3,482,306	0.99	1.32	−1.03	2.81	0.68	0.57	0.50	0.45	0.43	0.40	0.26	0.20	0.56	0.52	0.49	0.48	0.42	0.44	0.38	0.35	0.43	0.44	0.39	0.31
$HIST^3$	3,458,916	1.00	0.86	−0.39	2.20	*	0.84	0.73	0.65	0.62	0.57	0.38	0.29	0.80	0.75	0.71	0.69	0.61	0.65	0.57	0.51	0.62	0.65	0.56	0.45
$HIST^6$	3,424,766	1.00	0.72	−0.18	2.01		*	0.88	0.78	0.74	0.69	0.46	0.35	0.93	0.87	0.85	0.83	0.73	0.78	0.68	0.61	0.74	0.79	0.67	0.53
$HIST^{12}$	3,354,313	1.00	0.63	−0.05	1.89			*	0.90	0.85	0.79	0.53	0.40	0.99	0.95	0.98	0.94	0.83	0.88	0.77	0.70	0.84	0.90	0.77	0.61
$HIST^{24}$	3,213,995	1.00	0.57	0.02	1.80				*	0.95	0.89	0.59	0.45	0.89	0.97	0.90	0.89	0.80	0.80	0.70	0.64	0.78	0.82	0.82	0.74
$HIST^{36}$	3,054,285	1.00	0.54	0.06	1.75					*	0.95	0.62	0.48	0.84	0.95	0.86	0.86	0.78	0.76	0.66	0.60	0.74	0.78	0.85	0.80
$HIST^{60}$	2,737,036	1.00	0.51	0.09	1.68						*	0.65	0.51	0.78	0.90	0.81	0.82	0.74	0.71	0.62	0.57	0.69	0.73	0.87	0.85
$HIST^{60}_{mon}$	2,770,429	1.00	0.69	0.16	2.33							*	0.70	0.52	0.60	0.54	0.54	0.65	0.55	0.54	0.54	0.55	0.56	0.66	0.58
$HIST_q$	2,133,299	1.01	0.75	0.26	2.53								*	0.40	0.46	0.41	0.41	0.50	0.43	0.42	0.42	0.42	0.43	0.52	0.46
$HIST_{ewma}$	3,354,313	1.00	0.64	−0.06	1.89									*	0.95	0.96	0.93	0.82	0.87	0.76	0.69	0.83	0.89	0.76	0.60
$HIST_{ewma,ex}$	3,354,313	1.00	0.57	0.02	1.78										*	0.95	0.93	0.83	0.85	0.74	0.67	0.81	0.86	0.83	0.73
$HIST^V$	3,354,313	0.98	0.52	0.05	1.68											*	0.99	0.85	0.87	0.76	0.68	0.84	0.88	0.78	0.64
$HIST^K$	3,274,082	0.96	0.44	0.12	1.51												*	0.84	0.85	0.74	0.66	0.82	0.86	0.77	0.65
$HIST^I$	1,414,615	1.02	0.50	0.16	1.76													*	0.77	0.69	0.63	0.74	0.79	0.77	0.60
$Dim^{(1)}$	3,355,001	1.00	0.67	−0.03	2.05														*	0.86	0.78	0.93	0.92	0.75	0.55
$Dim^{(3)}$	3,355,001	1.00	0.75	−0.06	2.25															*	0.90	0.80	0.83	0.68	0.48
$Dim^{(5)}$	3,355,001	1.00	0.83	−0.10	2.47																*	0.72	0.74	0.63	0.44
SW	3,355,001	1.00	0.69	−0.07	2.08																	*	0.95	0.76	0.54
FP12	3,354,295	1.06	0.70	−0.04	2.15																		*	0.80	0.57
FP60	3,354,295	1.09	0.63	0.12	2.10																			*	0.73
$Beta^{cay}$	2,113,109	0.99	0.58	−0.02	1.81																				*

46

Table 2: Prediction Errors – Window Length and Sampling Frequency

This table presents the out-of-sample prediction errors of the simple historical estimator implemented for different estimation windows and data sampling frequencies. We measure the realized beta with daily returns over the horizon of 6 months. Each month, we compute the value-weighted RMSE for the approaches considered using the entire cross-section of stocks. The first row reports the average value-weighted RMSE over our sample period. We indicate the lowest average RMSE with *italic* font. The remainder of the table reports the differences in prediction errors. The upper triangular matrix reports the average differences in RMSE and the lower triangular matrix reports the average differences in RMedSE. We report the error loss differential between the model *[name in row]* and the model *[name in column]*. The absolute values of the numbers in parentheses indicate the share of time periods for which the difference is significant at 5% (e.g., 0.4 indicates that the difference is significant 40% of the time). We test the significance using the modified Diebold–Mariano and the Wilcoxon signed rank tests for the upper and lower triangular matrix, respectively. The sign of the number in parentheses indicates the direction of the significant differences.

	HIST^1	HIST^3	HIST^6	HIST^{12}	HIST^{24}	HIST^{36}	HIST^{60}	$\text{HIST}^{12}_{\text{mon}}$	$\text{HIST}^{36}_{\text{mon}}$	$\text{HIST}^{60}_{\text{mon}}$	HIST_q
avg. RMSE	0.564	0.388	0.339	*0.322*	0.325	0.332	0.347	0.640	0.472	0.449	0.511
HIST^1		0.175 (0.53)	0.225 (0.60)	0.242 (0.59)	0.239 (0.54)	0.232 (0.54)	0.216 (0.50)	−0.076 (−0.22)	0.091 (0.11)	0.115 (0.16)	0.053 (−0.00)
HIST^3	−0.151 (−0.99)		0.050 (0.39)	0.067 (0.42)	0.064 (0.34)	0.056 (0.25)	0.041 (0.14)	−0.252 (−0.67)	−0.084 (−0.43)	−0.060 (−0.38)	−0.123 (−0.48)
HIST^6	−0.199 (−1.00)	−0.048 (−0.93)		0.017 (0.20)	0.014 (0.08)	0.007 (−0.01)	−0.009 (−0.17)	−0.301 (−0.80)	−0.134 (−0.73)	−0.110 (−0.65)	−0.172 (−0.66)
HIST^{12}	−0.222 (−1.00)	−0.071 (−0.92)	−0.023 (−0.78)		−0.003 (−0.05)	−0.010 (−0.14)	−0.026 (−0.29)	−0.319 (−0.83)	−0.151 (−0.81)	−0.127 (−0.78)	−0.190 (−0.76)
HIST^{24}	−0.226 (−0.99)	−0.075 (−0.89)	−0.027 (−0.67)	−0.004 (−0.36)		−0.007 (−0.15)	−0.023 (−0.32)	−0.315 (−0.82)	−0.148 (−0.87)	−0.124 (−0.79)	−0.186 (−0.79)
HIST^{36}	−0.222 (−0.98)	−0.071 (−0.85)	−0.023 (−0.56)	0.000 (−0.17)	0.004 (0.02)		−0.015 (−0.29)	−0.308 (−0.82)	−0.140 (−0.85)	−0.117 (−0.83)	−0.179 (−0.81)
HIST^{60}	−0.211 (−0.95)	−0.060 (−0.69)	−0.012 (−0.34)	0.011 (0.13)	0.015 (0.26)	0.011 (0.23)		−0.293 (−0.79)	−0.125 (−0.77)	−0.101 (−0.80)	−0.164 (−0.79)
$\text{HIST}^{12}_{\text{mon}}$	0.093 (0.45)	0.244 (0.97)	0.292 (1.00)	0.315 (1.00)	0.319 (1.00)	0.315 (1.00)	0.304 (1.00)		0.168 (0.57)	0.191 (0.61)	0.129 (0.29)
$\text{HIST}^{36}_{\text{mon}}$	−0.034 (−0.16)	0.117 (0.77)	0.165 (0.98)	0.188 (1.00)	0.192 (1.00)	0.188 (1.00)	0.177 (1.00)	−0.127 (−0.92)		0.024 (0.27)	−0.039 (−0.19)
$\text{HIST}^{60}_{\text{mon}}$	−0.049 (−0.22)	0.102 (0.70)	0.150 (0.93)	0.173 (1.00)	0.177 (1.00)	0.173 (0.99)	0.162 (0.99)	−0.142 (−0.94)	−0.015 (−0.41)		−0.062 (−0.30)
HIST_q	0.051 (0.28)	0.201 (0.88)	0.250 (0.99)	0.272 (1.00)	0.276 (1.00)	0.273 (1.00)	0.262 (1.00)	−0.042 (−0.25)	0.084 (0.77)	0.100 (0.92)	

Table 3: Market–Neutral Anomaly Portfolios – Window Length and Sampling Frequency

This table presents average ex-post realized betas of ex-ante beta-neutral anomaly portfolios based on the simple historical estimator implemented for different estimation windows and data sampling frequencies. Each month, we sort the stocks to form the different anomaly portfolios based on NYSE breakpoints. For size and value, we build 25 independently sorted portfolios. S is the average of the 5 low-market-cap portfolios and B is the average of the 5 high-market-cap portfolios while SMB is their difference. Similarly, L is the average of the 5 low-book-to-market portfolios and H is the average of the 5 high-book-to-market portfolios. HML is the difference between the Value and Growth portfolios. For momentum, we sort the stocks into 10 portfolios based on their return over the past 12 months (skipping 1 month). W and L are the top and bottom deciles while WML is their difference. For idiosyncratic volatility (Ivol) and illiquidity (Illiq), we sort the stocks into 5 portfolios and form low-minus-high and high-minus-low portfolios, respectively. For leverage, we follow the same approach as for the book-to-market ratio, that is, we build 25 portfolios, independently sorted on size and leverage and form the high-minus-low portfolio. We make the anomaly portfolios ex-ante beta-neutral by solving the equation $v_{j,t}\beta_{j,t}^{\text{long}} - \beta_{j,t}^{\text{short}} = 0$ and applying the resulting weight $v_{j,t}$ to the long side of the anomaly. In parentheses, we present robust Newey & West (1987) standard errors, using 6 lags. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

	HIST^1	HIST^3	HIST^6	HIST^{12}	HIST^{24}	HIST^{36}	HIST^{60}	$\text{HIST}^{12}_{\text{mon}}$	$\text{HIST}^{36}_{\text{mon}}$	$\text{HIST}^{60}_{\text{mon}}$	HIST_{q}
SMB	0.129***	0.040***	0.009	−0.007	−0.012	−0.010	−0.002	−0.328***	−0.368***	−0.360***	−0.413***
(s.e.)	(0.015)	(0.012)	(0.015)	(0.019)	(0.023)	(0.024)	(0.027)	(0.026)	(0.023)	(0.023)	(0.027)
HML	−0.002	0.025***	0.023*	0.011	−0.027	−0.057***	−0.085***	0.017	−0.062***	−0.073***	−0.014
(s.e.)	(0.034)	(0.009)	(0.013)	(0.016)	(0.017)	(0.017)	(0.017)	(0.030)	(0.019)	(0.018)	(0.020)
WML	0.038***	0.009	0.016	0.052***	0.107***	0.119***	0.113***	0.211***	0.205***	0.139***	0.124***
(s.e.)	(0.009)	(0.011)	(0.015)	(0.020)	(0.023)	(0.026)	(0.031)	(0.051)	(0.034)	(0.028)	(0.034)
Ivol (1–5)	0.039***	0.056***	0.069***	0.064***	−0.006	−0.059*	−0.114***	0.761***	0.523***	0.331***	0.345***
(s.e.)	(0.012)	(0.015)	(0.019)	(0.025)	(0.029)	(0.031)	(0.033)	(0.105)	(0.077)	(0.048)	(0.053)
Illiq (5–1)	0.148***	0.060***	0.034*	0.034	0.035	0.032	0.031	−0.402***	−0.432***	−0.423***	−0.476***
(s.e.)	(0.016)	(0.016)	(0.018)	(0.022)	(0.025)	(0.027)	(0.031)	(0.025)	(0.022)	(0.023)	(0.026)
Lev (5–1)	0.017***	0.007	0.006	0.007	0.012	0.017	0.028*	0.044*	0.045*	0.045**	0.096***
(s.e.)	(0.004)	(0.005)	(0.007)	(0.009)	(0.012)	(0.014)	(0.015)	(0.023)	(0.026)	(0.023)	(0.030)

Table 4: Prediction Errors – Different Weighting Schemes

This table presents the out-of-sample prediction errors for the exponentially weighted estimators. We measure the realized beta with daily returns over the horizon of 6 months. Each month, we compute the value-weighted RMSE for the approaches considered using the entire cross-section of stocks. The first row reports the average value-weighted RMSE over our sample period. We indicate the lowest average RMSE with *italic* font. The remainder of the table reports the differences in prediction errors. The upper triangular matrix reports the average differences in RMSE and the lower triangular matrix reports the average differences in RMedSE. We report the error loss differential between the model *[name in row]* and the model *[name in column]*. The absolute values of the numbers in parentheses indicate the share of time periods for which the difference is significant at 5% (e.g., 0.4 indicates that the difference is significant 40% of the time). We test the significance using the modified Diebold–Mariano and the Wilcoxon signed rank tests for the upper and lower triangular matrix, respectively. The sign of the number in parentheses indicates the direction of the significant differences.

	HIST	HIST$_{ewma,s}$	HIST$_{ewma,s,ex}$	HIST$_{ewma}$	HIST$_{ewma,ex}$
avg. RMSE	0.322	0.317	0.309	0.317	*0.308*
HIST		0.004	0.013	0.005	0.013
		(0.15)	(0.32)	(0.29)	(0.30)
HIST$_{ewma,s}$	0.001		0.009	0.001	0.009
	(0.09)		(0.43)	(−0.02)	(0.14)
HIST$_{ewma,s,ex}$	−0.008	−0.009		−0.008	0.001
	(−0.53)	(−0.85)		(−0.32)	(−0.03)
HIST$_{ewma}$	−0.002	−0.003	0.006		0.009
	(−0.23)	(−0.35)	(0.59)		(0.19)
HIST$_{ewma,ex}$	−0.014	−0.015	−0.006	−0.012	
	(−0.74)	(−0.63)	(−0.34)	(−0.67)	

Table 5: Market-Neutral Anomaly Portfolios – Different Weighting Schemes

This table presents average ex-post realized betas of ex-ante beta-neutral anomaly portfolios when using an exponential weighting scheme. Each month, we sort the stocks to form the different anomaly portfolios based on NYSE breakpoints. For size and value, we build 25 independently sorted portfolios. S is the average of the 5 low-market-cap portfolios and B is the average of the 5 high-market-cap portfolios while SMB is their difference. Similarly, L is the average of the 5 low-book-to-market portfolios and H is the average of the 5 high-book-to-market portfolios. HML is the difference between the Value and Growth portfolios. For momentum, we sort the stocks into 10 portfolios based on their return over the past 12 months (skipping 1 month). W and L are the top and bottom deciles while WML is their difference. For idiosyncratic volatility (Ivol) and illiquidity (Illiq), we sort the stocks into 5 portfolios and form low-minus-high and high-minus-low portfolios, respectively. For leverage, we follow the same approach as for the book-to-market ratio, that is, we build 25 portfolios, independently sorted on size and leverage and form the high-minus-low portfolio. We make the anomaly portfolios ex-ante beta-neutral by solving the equation $\upsilon_{j,t}\beta_{j,t}^{\text{long}} - \beta_{j,t}^{\text{short}} = 0$ and applying the resulting weight $\upsilon_{j,t}$ to the long side of the anomaly. In parentheses, we present robust Newey & West (1987) standard errors, using 6 lags. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

	HIST	HIST$_{\text{ewma,s}}$	HIST$_{\text{ewma,s,ex}}$	HIST$_{\text{ewma}}$	HIST$_{\text{ewma,ex}}$
SMB	−0.007	−0.006	−0.012	−0.008	−0.013
(s.e.)	(0.019)	(0.016)	(0.017)	(0.018)	(0.020)
HML	0.011	0.019	0.006	0.015	−0.022
(s.e.)	(0.016)	(0.013)	(0.014)	(0.015)	(0.015)
WML	0.052***	0.030*	0.050***	0.041**	0.089***
(s.e.)	(0.020)	(0.016)	(0.017)	(0.018)	(0.021)
Ivol (1–5)	0.064***	0.065***	0.043**	0.065***	−0.007
(s.e.)	(0.025)	(0.020)	(0.021)	(0.022)	(0.024)
Illiq (5–1)	0.034	0.025	0.022	0.028	0.024
(s.e.)	(0.022)	(0.018)	(0.019)	(0.020)	(0.022)
Lev (5–1)	0.007	0.004	0.003	0.005	0.006
(s.e.)	(0.009)	(0.008)	(0.008)	(0.009)	(0.010)

Table 6: Prediction Errors – Imposing Priors

This table presents the out-of-sample prediction errors for the shrinkage estimators. We measure the realized beta with daily returns over the horizon of 6 months. Each month, we compute the value-weighted RMSE for the approaches considered using the entire cross-section of stocks. The first row reports the average value-weighted RMSE over our sample period. We indicate the lowest average RMSE with *italic* font. The remainder of the table reports the differences in prediction errors. The upper triangular matrix reports the average differences in RMSE and the lower triangular matrix reports the average differences in RMedSE. We report the error loss differential between the model *[name in row]* and the model *[name in column]*. The absolute values of the numbers in parentheses indicate the share of time periods for which the difference is significant at 5% (e.g., 0.4 indicates that the difference is significant 40% of the time). We test the significance using the modified Diebold–Mariano and the Wilcoxon signed rank tests for the upper and lower triangular matrix, respectively. The sign of the number in parentheses indicates the direction of the significant differences.

	HIST	HIST^V	HIST^K	HIST^I
avg. RMSE	0.295	0.289	*0.285*	0.289
HIST		0.006	0.009	0.006
		(0.08)	(0.14)	(0.04)
HIST^V	−0.008		0.004	0.000
	(−0.83)		(0.16)	(0.02)
HIST^K	−0.014	−0.006		−0.004
	(−0.77)	(−0.63)		(0.02)
HIST^I	0.012	0.020	0.026	
	(0.25)	(0.61)	(0.71)	

Table 7: Market-Neutral Anomaly Portfolios – Imposing Priors

This table presents average ex-post realized betas of ex-ante beta-neutral anomaly portfolios when using the shrinkage estimators. Each month, we sort the stocks to form the different anomaly portfolios based on NYSE breakpoints. For size and value, we build 25 independently sorted portfolios. S is the average of the 5 low-market-cap portfolios and B is the average of the 5 high-market-cap portfolios while SMB is their difference. Similarly, L is the average of the 5 low-book-to-market portfolios and H is the average of the 5 high-book-to-market portfolios. HML is the difference between the Value and Growth portfolios. For momentum, we sort the stocks into 10 portfolios based on their return over the past 12 months (skipping 1 month). W and L are the top and bottom deciles while WML is their difference. For idiosyncratic volatility (Ivol) and illiquidity (Illiq), we sort the stocks into 5 portfolios and form low-minus-high and high-minus-low portfolios, respectively. For leverage, we follow the same approach as for the book-to-market ratio, that is, we build 25 portfolios, independently sorted on size and leverage and form the high-minus-low portfolio. We make the anomaly portfolios ex-ante beta-neutral by solving the equation $v_{j,t}\beta_{j,t}^{\text{long}} - \beta_{j,t}^{\text{short}} = 0$ and applying the resulting weight $v_{j,t}$ to the long side of the anomaly. In parentheses, we present robust Newey & West (1987) standard errors, using 6 lags. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

	HIST	HIST$^{\text{V}}$	HIST$^{\text{K}}$	HIST$^{\text{I}}$
SMB	−0.007	0.009	0.020	−0.179***
(s.e.)	(0.019)	(0.019)	(0.019)	(0.023)
HML	0.011	−0.023	−0.052***	−0.018
(s.e.)	(0.016)	(0.016)	(0.016)	(0.013)
WML	0.052***	0.046**	0.040*	0.056**
(s.e.)	(0.020)	(0.020)	(0.022)	(0.024)
Ivol (1–5)	0.064***	−0.031	−0.130***	0.069**
(s.e.)	(0.025)	(0.024)	(0.024)	(0.035)
Illiq (5–1)	0.034	0.032	0.003	−0.162***
(s.e.)	(0.022)	(0.022)	(0.022)	(0.025)
Lev (5–1)	0.007	0.002	0.007	0.025**
(s.e.)	(0.009)	(0.009)	(0.010)	(0.011)

Table 8: Prediction Errors − Asynchronicity

This table presents the out-of-sample prediction errors for the estimators with asynchronicity adjustment. We measure the realized beta with daily returns over the horizon of 6 months. Each month, we compute the value-weighted RMSE for the approaches considered using the entire cross-section of stocks. The first row reports the average value-weighted RMSE over our sample period. We indicate the lowest average RMSE with *italic* font. The remainder of the table reports the differences in prediction errors. The upper triangular matrix reports the average differences in RMSE and the lower triangular matrix reports the average differences in RMedSE. We report the error loss differential between the model *[name in row]* and the model *[name in column]*. The absolute values of the numbers in parentheses indicate the share of time periods for which the difference is significant at 5% (e.g., 0.4 indicates that the difference is significant 40% of the time). We test the significance using the modified Diebold–Mariano and the Wilcoxon signed rank tests for the upper and lower triangular matrix, respectively. The sign of the number in parentheses indicates the direction of the significant differences.

	HIST	Dim$^{(1)}$	Dim$^{(2)}$	Dim$^{(3)}$	Dim$^{(4)}$	Dim$^{(5)}$	SW	FP^{12}	FP^{36}	FP^{60}
avg. RMSE	*0.322*	0.356	0.372	0.396	0.418	0.440	0.360	0.354	0.366	0.376
HIST		−0.034	−0.051	−0.074	−0.096	−0.119	−0.039	−0.032	−0.044	−0.054
		(−0.48)	(−0.55)	(−0.63)	(−0.69)	(−0.77)	(−0.50)	(−0.36)	(−0.39)	(−0.44)
Dim$^{(1)}$	0.044		−0.016	−0.040	−0.062	−0.084	−0.005	0.002	−0.010	−0.020
	(0.90)		(−0.21)	(−0.44)	(−0.61)	(−0.72)	(0.04)	(0.11)	(0.06)	(−0.01)
Dim$^{(2)}$	0.068	0.023		−0.023	−0.046	−0.068	0.012	0.018	0.007	−0.003
	(0.96)	(0.95)		(−0.41)	(−0.54)	(−0.64)	(0.23)	(0.24)	(0.20)	(0.09)
Dim$^{(3)}$	0.094	0.050	0.026		−0.022	−0.045	0.035	0.042	0.030	0.020
	(0.98)	(0.99)	(0.96)		(−0.37)	(−0.51)	(0.38)	(0.38)	(0.33)	(0.27)
Dim$^{(4)}$	0.117	0.073	0.050	0.023		−0.022	0.057	0.064	0.052	0.042
	(0.99)	(1.00)	(0.99)	(0.95)		(−0.37)	(0.50)	(0.48)	(0.41)	(0.36)
Dim$^{(5)}$	0.142	0.097	0.074	0.048	0.024		0.080	0.086	0.074	0.065
	(1.00)	(1.00)	(0.99)	(0.98)	(0.95)		(0.58)	(0.62)	(0.52)	(0.45)
SW	0.046	0.001	−0.022	−0.048	−0.072	−0.096		0.006	−0.005	−0.015
	(0.93)	(0.15)	(−0.68)	(−0.91)	(−0.94)	(−0.97)		(0.20)	(0.09)	(−0.02)
FP^{12}	0.055	0.010	−0.013	−0.039	−0.063	−0.087	0.009		−0.012	−0.022
	(0.86)	(0.14)	(−0.42)	(−0.75)	(−0.86)	(−0.90)	(0.04)		(−0.04)	(−0.13)
FP^{36}	0.060	0.016	−0.007	−0.034	−0.057	−0.081	0.015	0.006		−0.010
	(0.60)	(−0.14)	(−0.43)	(−0.60)	(−0.66)	(−0.75)	(−0.18)	(−0.13)		(−0.13)
FP^{60}	0.067	0.022	−0.001	−0.027	−0.051	−0.075	0.021	0.012	0.006	
	(0.63)	(0.11)	(−0.18)	(−0.40)	(−0.57)	(−0.71)	(0.04)	(0.05)	(0.11)	

Table 9: Market-Neutral Anomaly Portfolios – Asynchronicity

This table presents average ex-post realized betas of ex-ante beta-neutral anomaly portfolios when using estimators with asynchronicity adjustment. Each month, we sort the stocks to form the different anomaly portfolios based on NYSE breakpoints. For size and value, we build 25 independently sorted portfolios. S is the average of the 5 low-market-cap portfolios and B is the average of the 5 high-market-cap portfolios while SMB is their difference. Similarly, L is the average of the 5 low-book-to-market portfolios and H is the average of the 5 high-book-to-market portfolios. HML is the difference between the Value and Growth portfolios. For momentum, we sort the stocks into 10 portfolios based on their return over the past 12 months (skipping 1 month). W and L are the top and bottom deciles while WML is their difference. For idiosyncratic volatility (Ivol) and illiquidity (Illiq), we sort the stocks into 5 portfolios and form low-minus-high and high-minus-low portfolios, respectively. For leverage, we follow the same approach as for the book-to-market ratio, that is, we build 25 portfolios, independently sorted on size and leverage and form the high-minus-low portfolio. We make the anomaly portfolios ex-ante beta-neutral by solving the equation $v_{j,t}\beta_{j,t}^{\text{long}} - \beta_{j,t}^{\text{short}} = 0$ and applying the resulting weight $v_{j,t}$ to the long side of the anomaly. In parentheses, we present robust Newey & West (1987) standard errors, using 6 lags. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

	HIST	$\text{Dim}^{(1)}$	$\text{Dim}^{(2)}$	$\text{Dim}^{(3)}$	$\text{Dim}^{(4)}$	$\text{Dim}^{(5)}$	SW	FP^{12}	FP^{36}	FP^{60}
SMB	−0.007	−0.171***	−0.208***	−0.253***	−0.281***	−0.312***	−0.144***	−0.156***	−0.177***	−0.176***
(s.e.)	(0.019)	(0.020)	(0.019)	(0.020)	(0.021)	(0.020)	(0.019)	(0.019)	(0.022)	(0.023)
HML	0.011	0.028*	0.016	0.012	0.011	0.002	0.035**	0.003	−0.084***	−0.122***
(s.e.)	(0.016)	(0.015)	(0.015)	(0.016)	(0.017)	(0.017)	(0.015)	(0.015)	(0.015)	(0.015)
WML	0.052***	0.037*	0.051**	0.068***	0.089***	0.110***	0.057***	0.064***	0.171***	0.189***
(s.e.)	(0.020)	(0.020)	(0.021)	(0.023)	(0.026)	(0.029)	(0.019)	(0.019)	(0.023)	(0.024)
Ivol (1–5)	0.064***	0.186***	0.210***	0.262***	0.323***	0.380***	0.214***	0.180***	0.197***	0.191***
(s.e.)	(0.025)	(0.029)	(0.029)	(0.031)	(0.033)	(0.034)	(0.029)	(0.027)	(0.034)	(0.036)
Illiq (5–1)	0.034	−0.205***	−0.248***	−0.304***	−0.335***	−0.370***	−0.164***	−0.178***	−0.203***	−0.209***
(s.e.)	(0.022)	(0.021)	(0.020)	(0.021)	(0.021)	(0.022)	(0.021)	(0.021)	(0.022)	(0.023)
Lev (5–1)	0.007	0.021**	0.021*	0.021*	0.021	0.026*	0.018*	0.005	0.016	0.018
(s.e.)	(0.009)	(0.010)	(0.012)	(0.013)	(0.013)	(0.015)	(0.010)	(0.010)	(0.011)	(0.011)

Table 10: Prediction Errors – Macroeconomic Conditioning Variables

This table presents the out-of-sample prediction errors for the estimators that condition on macroeconomic information. We measure the realized beta with daily returns over the horizon of 6 months. Each month, we compute the value-weighted RMSE for the approaches considered using the entire cross-section of stocks. The first row reports the average value-weighted RMSE over our sample period. We indicate the lowest average RMSE with *italic* font. The remainder of the table reports the differences in prediction errors. The upper triangular matrix reports the average differences in RMSE and the lower triangular matrix reports the average differences in RMedSE. We report the error loss differential between the model *[name in row]* and the model *[name in column]*. The absolute values of the numbers in parentheses indicate the share of time periods for which the difference is significant at 5% (e.g., 0.4 indicates that the difference is significant 40% of the time). We test the significance using the modified Diebold–Mariano and the Wilcoxon signed rank tests for the upper and lower triangular matrix, respectively. The sign of the number in parentheses indicates the direction of the significant differences.

	HIST	Beta^{bm}	Beta^{cay}	Beta^{def}	Beta^{dp}	Beta^{ep}	Beta^{ic}	Beta^{inf}	Beta^{lty}	Beta^{tbl}	Beta^{unc}	Beta^{une}	Beta^{all}
avg. RMSE	*0.342*	0.448	0.442	0.448	0.440	0.461	0.472	0.421	0.445	0.449	0.433	0.486	0.885
HIST		−0.106 (−0.65)	−0.100 (−0.59)	−0.106 (−0.63)	−0.098 (−0.63)	−0.119 (−0.62)	−0.130 (−0.73)	−0.079 (−0.63)	−0.103 (−0.61)	−0.107 (−0.63)	−0.091 (−0.60)	−0.144 (−0.74)	−0.543 (−0.94)
Beta^{bm}	0.060 (0.65)		0.005 (−0.00)	0.000 (−0.01)	0.007 (−0.06)	−0.013 (−0.07)	−0.024 (−0.26)	0.027 (0.01)	0.003 (0.02)	−0.002 (−0.12)	0.015 (0.04)	−0.038 (−0.16)	−0.437 (−0.90)
Beta^{cay}	0.067 (0.60)	0.007 (0.04)		−0.006 (−0.01)	0.002 (0.02)	−0.019 (−0.06)	−0.030 (−0.13)	0.021 (0.07)	−0.003 (−0.01)	−0.007 (−0.14)	0.010 (−0.04)	−0.044 (−0.16)	−0.443 (−0.93)
Beta^{def}	0.070 (0.61)	0.010 (0.04)	0.003 (−0.05)		0.008 (0.07)	−0.013 (−0.05)	−0.024 (−0.17)	0.027 (0.05)	0.003 (0.09)	−0.002 (−0.13)	0.015 (0.01)	−0.038 (−0.19)	−0.437 (−0.93)
Beta^{dp}	0.064 (0.60)	0.005 (0.14)	−0.003 (−0.07)	−0.005 (−0.04)		−0.021 (−0.05)	−0.032 (−0.09)	0.019 (−0.02)	−0.005 (−0.02)	−0.009 (−0.14)	0.008 (0.02)	−0.046 (−0.17)	−0.445 (−0.91)
Beta^{ep}	0.076 (0.59)	0.016 (0.05)	0.009 (−0.03)	0.006 (0.06)	0.012 (−0.06)		−0.011 (−0.08)	0.040 (0.06)	0.016 (0.07)	0.012 (−0.07)	0.028 (0.11)	−0.025 (−0.15)	−0.413 (−0.88)
Beta^{ic}	0.076 (0.72)	0.016 (0.19)	0.009 (0.08)	0.006 (0.11)	0.012 (−0.05)	0.000 (0.05)		0.051 (0.22)	0.027 (0.26)	0.023 (0.08)	0.039 (0.19)	−0.014 (−0.05)	−0.413 (−0.88)
Beta^{inf}	0.044 (0.56)	−0.016 (−0.10)	−0.024 (−0.16)	−0.026 (−0.21)	−0.021 (−0.13)	−0.032 (−0.16)	−0.033 (−0.13)		−0.024 (−0.03)	−0.029 (−0.16)	−0.012 (−0.05)	−0.065 (−0.27)	−0.464 (−0.90)
Beta^{lty}	0.068 (0.61)	0.008 (−0.03)	0.001 (−0.06)	−0.002 (−0.11)	0.004 (−0.02)	−0.008 (−0.06)	−0.008 (−0.16)	0.024 (0.07)		−0.004 (−0.08)	0.012 (−0.08)	−0.041 (−0.25)	−0.440 (−0.82)
Beta^{tbl}	0.069 (0.70)	0.009 (0.10)	0.002 (0.11)	−0.001 (0.12)	0.005 (0.11)	−0.007 (0.03)	−0.007 (0.08)	0.026 (0.20)	0.001 (0.08)		0.017 (0.11)	−0.037 (−0.07)	−0.436 (−0.87)
Beta^{unc}	0.059 (0.64)	−0.001 (−0.01)	−0.008 (0.04)	−0.010 (−0.01)	−0.005 (0.08)	−0.017 (−0.13)	−0.017 (−0.07)	0.016 (0.12)	−0.009 (0.02)	−0.010 (−0.11)		−0.053 (−0.21)	−0.452 (−0.94)
Beta^{une}	0.091 (0.71)	0.031 (0.12)	0.024 (0.07)	0.022 (0.12)	0.027 (0.08)	0.015 (0.14)	0.015 (0.05)	0.048 (0.26)	0.023 (0.22)	0.022 (0.01)	0.032 (0.13)		−0.399 (−0.84)
Beta^{all}	0.436 (0.99)	0.376 (0.97)	0.369 (0.95)	0.366 (0.98)	0.372 (0.98)	0.360 (0.94)	0.360 (0.94)	0.392 (0.96)	0.368 (0.89)	0.367 (0.96)	0.377 (0.99)	0.345 (0.93)	

Table 11: Market-Neutral Anomaly Portfolios – Macroeconomic Conditioning Variables

This table presents average ex-post realized betas of ex-ante beta-neutral anomaly portfolios when using estimators that condition on macroeconomic information. Each month, we sort the stocks to form the different anomaly portfolios based on NYSE breakpoints. For size and value, we build 25 independently sorted portfolios. S is the average of the 5 low-market-cap portfolios and B is the average of the 5 high-market-cap portfolios while SMB is their difference. Similarly, L is the average of the 5 low-book-to-market portfolios and H is the average of the 5 high-book-to-market portfolios. HML is the difference between the Value and Growth portfolios. For momentum, we sort the stocks into 10 portfolios based on their return over the past 12 months (skipping 1 month). W and L are the top and bottom deciles while WML is their difference. For idiosyncratic volatility (Ivol) and illiquidity (Illiq), we sort the stocks into 5 portfolios and form low-minus-high and high-minus-low portfolios, respectively. For leverage, we follow the same approach as for the book-to-market ratio, that is, we build 25 portfolios, independently sorted on size and leverage and form the high-minus-low portfolio. We make the anomaly portfolios ex-ante beta-neutral by solving the equation $v_{j,t}\beta_{j,t}^{\text{long}} - \beta_{j,t}^{\text{short}} = 0$ and applying the resulting weight $v_{j,t}$ to the long side of the anomaly. In parentheses, we present robust Newey & West (1987) standard errors, using 6 lags. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

	HIST	Beta$^{\text{bm}}$	Beta$^{\text{cay}}$	Beta$^{\text{def}}$	Beta$^{\text{dp}}$	Beta$^{\text{ep}}$	Beta$^{\text{ic}}$	Beta$^{\text{inf}}$	Beta$^{\text{lty}}$	Beta$^{\text{tbl}}$	Beta$^{\text{unc}}$	Beta$^{\text{une}}$	Beta$^{\text{all}}$
SMB	−0.007	0.013	0.013	0.013	−0.005	−0.058**	−0.004	−0.016	−0.056*	0.565	−0.018	0.074*	−0.464
(s.e.)	(0.019)	(0.037)	(0.032)	(0.033)	(0.035)	(0.027)	(0.035)	(0.029)	(0.034)	(0.525)	(0.032)	(0.040)	(0.439)
HML	0.011	−0.070***	−0.062**	−0.082***	−0.085***	−0.007	−0.029	−0.073***	−0.053**	−0.078***	−0.076***	−0.049	0.173
(s.e.)	(0.016)	(0.023)	(0.024)	(0.021)	(0.022)	(0.054)	(0.047)	(0.020)	(0.024)	(0.024)	(0.020)	(0.046)	(0.136)
WML	0.052***	0.112***	0.102**	0.167***	0.129***	0.114***	0.107**	0.103**	0.133***	0.172***	0.143***	0.101**	−0.448
(s.e.)	(0.020)	(0.040)	(0.040)	(0.041)	(0.037)	(0.042)	(0.045)	(0.040)	(0.039)	(0.054)	(0.040)	(0.044)	(0.603)
Ivol (1–5)	0.064***	−0.164***	−0.116***	−0.096**	−0.134***	−0.101**	−0.201***	−0.139***	−0.019	−0.110**	−0.019	−0.207***	−0.115
(s.e.)	(0.025)	(0.045)	(0.041)	(0.043)	(0.044)	(0.045)	(0.050)	(0.041)	(0.051)	(0.044)	(0.068)	(0.046)	(0.083)
Illiq (5–1)	0.034	0.052	0.069*	0.102*	0.038	0.000	0.047	0.042	−0.005	−0.264	0.036	0.131***	−1.058
(s.e.)	(0.022)	(0.036)	(0.038)	(0.053)	(0.037)	(0.030)	(0.038)	(0.034)	(0.037)	(0.249)	(0.036)	(0.048)	(1.204)
Lev (5–1)	0.007	0.053**	0.062***	0.087***	0.039**	0.104***	0.094**	0.068***	0.026	0.052**	0.077***	0.076**	0.299**
(s.e.)	(0.009)	(0.026)	(0.023)	(0.023)	(0.017)	(0.033)	(0.046)	(0.020)	(0.021)	(0.023)	(0.021)	(0.034)	(0.134)

Table 12: Prediction Errors – Combinations

This table presents the out-of-sample prediction errors for forecast combinations. We measure the realized beta with daily returns over the horizon of 6 months. Each month, we compute the value-weighted RMSE for the approaches considered using the entire cross-section of stocks. The first row reports the average value-weighted RMSE over our sample period. We indicate the lowest average RMSE with *italic* font. The remainder of the table reports the differences in prediction errors. The upper triangular matrix reports the average differences in RMSE and the lower triangular matrix reports the average differences in RMedSE. We report the error loss differential between the model *[name in row]* and the model *[name in column]*. The absolute values of the numbers in parentheses indicate the share of time periods for which the difference is significant at 5% (e.g., 0.4 indicates that the difference is significant 40% of the time). We test the significance using the modified Diebold–Mariano and the Wilcoxon signed rank tests for the upper and lower triangular matrix, respectively. The sign of the number in parentheses indicates the direction of the significant differences.

	HIST	Best^{sim}	Best^{C}	Best^{shr}	Best^{BMA}	All^{sim}	All^{C}	All^{shr}	All^{BMA}
avg. RMSE	0.344	*0.326*	0.350	0.347	0.344	0.333	0.469	0.475	0.650
HIST		0.018	−0.006	−0.004	−0.001	0.010	−0.126	−0.132	−0.306
		(0.32)	(−0.12)	(−0.07)	(−0.08)	(0.09)	(−0.56)	(−0.49)	(−0.78)
Best^{sim}	−0.018		−0.024	−0.022	−0.019	−0.008	−0.144	−0.150	−0.324
	(−0.90)		(−0.42)	(−0.32)	(−0.50)	(−0.20)	(−0.62)	(−0.56)	(−0.79)
Best^{C}	0.002	0.020		0.002	0.005	0.016	−0.120	−0.126	−0.300
	(−0.12)	(0.72)		(0.16)	(0.14)	(0.30)	(−0.55)	(−0.47)	(−0.78)
Best^{shr}	−0.003	0.015	−0.005		0.003	0.014	−0.122	−0.128	−0.302
	(−0.27)	(0.68)	(−0.68)		(0.06)	(0.18)	(−0.54)	(−0.47)	(−0.79)
Best^{BMA}	−0.009	0.009	−0.011	−0.006		0.011	−0.125	−0.131	−0.305
	(−0.45)	(0.64)	(−0.28)	(−0.09)		(0.19)	(−0.55)	(−0.48)	(−0.79)
All^{sim}	−0.005	0.013	−0.007	−0.002	0.004		−0.136	−0.142	−0.316
	(−0.37)	(0.57)	(−0.36)	(−0.16)	(0.03)		(−0.61)	(−0.55)	(−0.78)
All^{C}	0.077	0.095	0.075	0.080	0.086	0.082		−0.006	−0.180
	(0.97)	(0.98)	(1.00)	(0.99)	(0.99)	(0.97)		(0.23)	(−0.46)
All^{shr}	0.052	0.071	0.051	0.055	0.061	0.058	−0.025		−0.174
	(0.91)	(0.99)	(0.99)	(1.00)	(0.97)	(0.97)	(−0.98)		(−0.50)
All^{BMA}	0.102	0.120	0.100	0.105	0.111	0.107	0.025	0.050	
	(0.98)	(0.99)	(0.98)	(0.99)	(1.00)	(0.97)	(0.42)	(0.85)	

Table 13: Market-Neutral Anomaly Portfolios – Combinations

This table presents average ex-post realized betas of ex-ante beta-neutral anomaly portfolios when using forecast combinations. Each month, we sort the stocks to form the different anomaly portfolios based on NYSE breakpoints. For size and value, we build 25 independently sorted portfolios. S is the average of the 5 low-market-cap portfolios and B is the average of the 5 high-market-cap portfolios while SMB is their difference. Similarly, L is the average of the 5 low-book-to-market portfolios and H is the average of the 5 high-book-to-market portfolios. HML is the difference between the Value and Growth portfolios. For momentum, we sort the stocks into 10 portfolios based on their return over the past 12 months (skipping 1 month). W and L are the top and bottom deciles while WML is their difference. For idiosyncratic volatility (Ivol) and illiquidity (Illiq), we sort the stocks into 5 portfolios and form low-minus-high and high-minus-low portfolios, respectively. For leverage, we follow the same approach as for the book-to-market ratio, that is, we build 25 portfolios, independently sorted on size and leverage and form the high-minus-low portfolio. We make the anomaly portfolios ex-ante beta-neutral by solving the equation $v_{j,t}\beta_{j,t}^{\text{long}} - \beta_{j,t}^{\text{short}} = 0$ and applying the resulting weight $v_{j,t}$ to the long side of the anomaly. In parentheses, we present robust Newey & West (1987) standard errors, using 6 lags. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

	HIST	Best$^{\text{sim}}$	Best$^{\text{C}}$	Best$^{\text{shr}}$	Best$^{\text{BMA}}$	All$^{\text{sim}}$	All$^{\text{C}}$	All$^{\text{shr}}$	All$^{\text{BMA}}$
SMB	−0.007	0.000	−0.049**	−0.042*	0.044*	−0.120***	−0.001	−0.039	0.106***
(s.e.)	(0.019)	(0.019)	(0.024)	(0.023)	(0.025)	(0.022)	(0.027)	(0.026)	(0.038)
HML	0.011	−0.036**	−0.055***	−0.051***	−0.013	−0.043***	−0.020	−0.027*	−0.017
(s.e.)	(0.016)	(0.015)	(0.016)	(0.016)	(0.015)	(0.014)	(0.017)	(0.016)	(0.023)
WML	0.052***	0.063***	0.064**	0.061**	0.054***	0.071***	0.085***	0.076**	0.122***
(s.e.)	(0.020)	(0.021)	(0.025)	(0.024)	(0.021)	(0.027)	(0.030)	(0.029)	(0.037)
Ivol (1–5)	0.064***	−0.065***	−0.028	−0.038	0.050*	0.029	−0.044	−0.012	0.064
(s.e.)	(0.025)	(0.024)	(0.030)	(0.028)	(0.027)	(0.031)	(0.038)	(0.035)	(0.048)
Illiq (5–1)	0.034	0.010	−0.012	−0.006	0.120***	−0.108***	0.041	−0.006	0.278***
(s.e.)	(0.022)	(0.021)	(0.026)	(0.026)	(0.032)	(0.023)	(0.029)	(0.027)	(0.075)
Lev (5–1)	0.007	0.006	0.035***	0.047***	0.016	0.023*	0.030*	0.039**	0.091**
(s.e.)	(0.009)	(0.009)	(0.012)	(0.018)	(0.011)	(0.013)	(0.016)	(0.015)	(0.037)

Table 14: Prediction Errors – Best Models

This table presents the out-of-sample prediction errors of the best estimators. We measure the realized beta with daily returns over the horizon of 6 months. Each month, we compute the value-weighted RMSE for the approaches considered using the entire cross-section of stocks. The first row reports the average value-weighted RMSE over our sample period. We indicate the lowest average RMSE with *italic* font. The remainder of the table reports the differences in prediction errors. The upper triangular matrix reports the average differences in RMSE and the lower triangular matrix reports the average differences in RMedSE. We report the error loss differential between the model *[name in row]* and the model *[name in column]*. The absolute values of the numbers in parentheses indicate the share of time periods for which the difference is significant at 5% (e.g., 0.4 indicates that the difference is significant 40% of the time). We test the significance using the modified Diebold–Mariano and the Wilcoxon signed rank tests for the upper and lower triangular matrix, respectively. The sign of the number in parentheses indicates the direction of the significant differences.

	HIST	$\text{HIST}_{\text{ewma,ex}}$	HIST^{K}	$\text{HIST}^{\text{K}}_{\text{ewma,ex}}$	Best^{sim}
avg. RMSE	0.368	0.352	0.355	0.349	*0.346*
HIST		0.016	0.013	0.019	0.022
		(0.44)	(0.24)	(0.33)	(0.52)
$\text{HIST}_{\text{ewma,ex}}$	−0.015		−0.003	0.003	0.006
	(−0.79)		(−0.04)	(0.07)	(0.21)
HIST^{K}	−0.017	−0.002		0.006	0.009
	(−0.80)	(−0.11)		(0.26)	(0.37)
$\text{HIST}^{\text{K}}_{\text{ewma,ex}}$	−0.021	−0.005	−0.004		0.003
	(−0.85)	(−0.73)	(−0.30)		(0.12)
Best^{sim}	−0.021	−0.006	−0.004	−0.001	
	(−0.94)	(−0.56)	(−0.45)	(0.05)	

Table 15: Forecast Error Decomposition

This table presents the decomposition of mean squared forecast errors for selected estimators. We measure the realized beta with daily returns over the horizon of 6 months. We decompose the MSE for all approaches into a bias, inefficiency, and a random error part. We perform the decomposition for each stock and weight the results by the stocks' average market capitalizations.

	Bias	Inefficiency	Random Error
HIST	0.003	0.017	0.102
$\mathrm{HIST}_{\mathrm{ewma,ex}}$	0.003	0.011	0.100
$\mathrm{HIST}^{\mathrm{K}}$	0.005	0.008	0.102
$\mathrm{HIST}^{\mathrm{I}}$	0.003	0.013	0.103
$\mathrm{HIST}^{\mathrm{K}}_{\mathrm{ewma,ex}}$	0.003	0.010	0.100
$\mathrm{Dim}^{(1)}$	0.004	0.026	0.122
$\mathrm{Dim}^{(3)}$	0.005	0.045	0.140
$\mathrm{Dim}^{(5)}$	0.005	0.079	0.154
FP^{12}	0.008	0.026	0.111
$\mathrm{Beta}^{\mathrm{cay}}$	0.005	0.060	0.151
$\mathrm{Best}^{\mathrm{sim}}$	0.003	0.008	0.099
$\mathrm{Best}^{\mathrm{C}}$	0.004	0.013	0.110
$\mathrm{Best}^{\mathrm{BMA}}$	0.003	0.017	0.102
$\mathrm{All}^{\mathrm{C}}$	0.006	0.078	0.147
$\mathrm{All}^{\mathrm{BMA}}$	0.007	0.296	0.163

Table 16: Forecast Error Statistics

This table presents the key statistics about squared forecast errors (SFEs) for selected estimators. For each estimator, we present the square-root of different quantiles of the estimators' SFEs ($q^{0.01}$, $q^{0.05}$, $q^{0.95}$, $q^{0.99}$). In addition, we present the frequency of square-root SFEs lower than 0.1 (Share $\leq$ 0.1) of at least 1 (Share $\geq$ 1). RMSE(RB> 1.5) and RMSE (RB< 0.5) indicate the RMSE of stock-month observations with realized beta greater than 1.5 and smaller than 0.5, respectively. Expansion, Recession, Exp. $\Rightarrow$ Rec., and Rec. $\Rightarrow$ Exp. present the average RMSE in expansions, recessions, the first period after the transition from expansion to recession and the first period after the transition from recession to expansion, respectively.

	$q^{0.01}$	$q^{0.05}$	$q^{0.95}$	$q^{0.99}$	Share $\leq$ 0.1	Share $\geq$ 1	RMSE(RB> 1.5)	RMSE(RB< 0.5)	Expansion	Recession	Exp. $\Rightarrow$ Rec.	Rec. $\Rightarrow$ Exp.
HIST	0.005	0.023	1.125	1.953	0.211	0.067	0.962	0.605	0.345	0.335	0.344	0.288
HIST$_{\text{ewma,ex}}$	0.004	0.022	1.048	1.804	0.220	0.057	0.938	0.564	0.333	0.325	0.339	0.271
HIST$^{\text{K}}$	0.004	0.022	1.025	1.750	0.221	0.053	0.958	0.552	0.332	0.331	0.347	0.278
HIST$^{\text{I}}$	0.005	0.024	1.198	2.156	0.203	0.077	0.904	0.762	0.340	0.331	0.347	0.279
HIST$^{\text{K}}_{\text{ewma,ex}}$	0.004	0.022	1.031	1.760	0.221	0.054	0.922	0.555	0.331	0.324	0.339	0.271
Dim$^{(1)}$	0.005	0.027	1.258	2.124	0.181	0.089	0.929	0.722	0.383	0.379	0.388	0.341
Dim$^{(3)}$	0.006	0.031	1.441	2.437	0.156	0.122	0.979	0.852	0.430	0.423	0.442	0.374
Dim$^{(5)}$	0.007	0.036	1.631	2.765	0.140	0.156	1.061	0.965	0.480	0.464	0.484	0.394
FP^{12}	0.005	0.027	1.282	2.201	0.181	0.092	0.931	0.754	0.376	0.359	0.366	0.314
Beta$^{\text{cay}}$	0.005	0.026	1.409	2.357	0.186	0.113	1.179	0.685	0.446	0.433	0.447	0.423
Best$^{\text{sim}}$	0.004	0.022	1.018	1.749	0.224	0.052	0.937	0.550	0.326	0.325	0.339	0.272
Best$^{\text{C}}$	0.005	0.024	1.074	1.824	0.207	0.061	0.942	0.601	0.348	0.358	0.373	0.306
Best$^{\text{BMA}}$	0.005	0.022	1.066	1.806	0.217	0.059	0.945	0.556	0.345	0.337	0.347	0.285
All$^{\text{C}}$	0.006	0.029	1.564	2.883	0.168	0.128	1.264	0.899	0.469	0.471	0.522	0.407
All$^{\text{BMA}}$	0.006	0.031	1.781	3.440	0.160	0.152	1.845	1.053	0.669	0.515	0.549	0.441

Table 17: Characteristics Sorts

At the end of each month, we sort the stocks into 5 portfolios according to the characteristic indicated in the panel headings. P1–P5 indicate the 5 portfolios, where P1 contains the 20% of the stocks with the lowest values of the sorting characteristic and P5 contains the 20% with the highest values. Within the portfolios, the stocks are value-weighted. avg. RMSE reports the average RMSE over our sample period. In each column, we indicate the lowest (average) RMSE with *italic* font. (share) presents the relative frequency with which the model *[name in row]* achieves the lowest RMSE for the different portfolios.

Panel A. Past HIST

	P1	P2	P3	P4	P5	avg. RMSE	(share)
HIST	0.138	0.107	*0.084*	*0.065*	0.149	0.109	(0.40)
HIST$_{\text{ewma,ex}}$	0.152	0.115	0.088	0.070	0.159	0.117	(0.00)
HIST$^{\text{K}}$	*0.129*	*0.100*	0.085	0.073	0.136	0.105	(0.40)
HIST$^{\text{K}}_{\text{ewma,ex}}$	0.143	0.113	0.088	0.068	0.140	0.111	(0.00)
Best$^{\text{sim}}$	0.130	0.105	0.085	0.067	*0.123*	*0.102*	(0.20)
Dim$^{(5)}$	0.233	0.175	0.121	0.094	0.169	0.159	(0.00)
FP^{12}	0.160	0.141	0.121	0.119	0.208	0.150	(0.00)

Panel B. Size

	P1	P2	P3	P4	P5	avg. RMSE	(share)
HIST	0.151	0.149	*0.154*	*0.131*	0.054	0.128	(0.40)
HIST$_{\text{ewma,ex}}$	0.151	0.143	0.156	0.132	0.055	0.128	(0.00)
HIST$^{\text{K}}$	0.149	0.146	0.163	0.144	0.064	0.133	(0.00)
HIST$^{\text{K}}_{\text{ewma,ex}}$	0.149	0.142	0.159	0.136	*0.054*	0.128	(0.20)
Best$^{\text{sim}}$	*0.147*	*0.142*	0.156	0.133	0.055	*0.126*	(0.40)
Dim$^{(5)}$	0.436	0.394	0.339	0.277	0.063	0.302	(0.00)
FP^{12}	0.275	0.268	0.249	0.220	0.102	0.223	(0.00)

Panel C. Book-to-Market

	P1	P2	P3	P4	P5	avg. RMSE	(share)
HIST	0.097	0.072	0.077	*0.097*	0.120	*0.093*	(0.20)
HIST$_{\text{ewma,ex}}$	*0.096*	0.074	0.078	0.100	*0.114*	0.093	(0.40)
HIST$^{\text{K}}$	0.125	0.075	0.076	0.102	0.123	0.100	(0.00)
HIST$^{\text{K}}_{\text{ewma,ex}}$	0.100	0.073	0.078	0.102	0.114	0.093	(0.00)
Best$^{\text{sim}}$	0.105	*0.071*	*0.075*	0.098	0.115	0.093	(0.40)
Dim$^{(5)}$	0.116	0.088	0.103	0.123	0.157	0.117	(0.00)
FP^{12}	0.141	0.130	0.125	0.132	0.153	0.136	(0.00)

Table 17: Characteristics Sorts (continued)

Panel D. Momentum

	P1	P2	P3	P4	P5	avg. RMSE	(share)
HIST	0.185	0.117	0.086	0.078	*0.147*	*0.123*	(0.20)
$\text{HIST}_{\text{ewma,ex}}$	0.175	0.123	0.088	0.077	0.165	0.126	(0.00)
HIST^{K}	0.181	0.117	0.084	0.082	0.212	0.135	(0.00)
$\text{HIST}^{\text{K}}_{\text{ewma,ex}}$	*0.165*	0.121	0.087	0.077	0.185	0.127	(0.20)
Best^{sim}	0.166	*0.116*	*0.083*	*0.077*	0.186	0.126	(0.60)
$\text{Dim}^{(5)}$	0.300	0.160	0.107	0.109	0.166	0.168	(0.00)
FP^{12}	0.253	0.171	0.138	0.124	0.154	0.168	(0.00)

Panel E. Idiosyncratic Volatility

	P1	P2	P3	P4	P5	avg. RMSE	(share)
HIST	0.083	0.075	0.147	0.227	0.304	0.167	(0.00)
$\text{HIST}_{\text{ewma,ex}}$	0.082	0.072	*0.136*	*0.205*	*0.289*	*0.157*	(0.60)
HIST^{K}	0.081	0.077	0.169	0.273	0.381	0.196	(0.00)
$\text{HIST}^{\text{K}}_{\text{ewma,ex}}$	0.081	0.071	0.138	0.222	0.335	0.169	(0.00)
Best^{sim}	*0.080*	*0.070*	0.142	0.223	0.314	0.166	(0.40)
$\text{Dim}^{(5)}$	0.104	0.125	0.218	0.338	0.591	0.275	(0.00)
FP^{12}	0.105	0.161	0.240	0.320	0.401	0.245	(0.00)

Panel F. Illiquidity

	P1	P2	P3	P4	P5	avg. RMSE	(share)
HIST	0.057	0.112	0.134	0.148	0.174	0.125	(0.00)
$\text{HIST}_{\text{ewma,ex}}$	0.058	*0.110*	*0.134*	0.148	0.170	0.124	(0.40)
HIST^{K}	0.067	0.122	0.142	0.151	*0.161*	0.128	(0.20)
$\text{HIST}^{\text{K}}_{\text{ewma,ex}}$	*0.057*	0.112	0.137	0.150	0.163	0.124	(0.20)
Best^{sim}	0.058	0.111	0.135	*0.147*	0.161	*0.123*	(0.20)
$\text{Dim}^{(5)}$	0.076	0.251	0.303	0.341	0.385	0.271	(0.00)
FP^{12}	0.100	0.202	0.221	0.238	0.276	0.207	(0.00)

Panel G. Leverage

	P1	P2	P3	P4	P5	avg. RMSE	(share)
HIST	0.141	0.073	0.072	0.076	0.128	0.098	(0.00)
$\text{HIST}_{\text{ewma,ex}}$	*0.132*	*0.069*	0.076	0.078	0.127	*0.096*	(0.40)
HIST^{K}	0.161	0.084	0.073	0.082	0.133	0.107	(0.00)
$\text{HIST}^{\text{K}}_{\text{ewma,ex}}$	0.136	0.069	0.075	0.078	0.126	0.097	(0.00)
Best^{sim}	0.138	0.072	*0.071*	*0.076*	*0.124*	0.096	(0.60)
$\text{Dim}^{(5)}$	0.164	0.094	0.085	0.100	0.174	0.124	(0.00)
FP^{12}	0.179	0.110	0.119	0.130	0.185	0.145	(0.00)

Table 17: Characteristics Sorts (continued 2)

Panel H. Industries

	NoDur	Durbl	Manuf	Enrgy	HiTec	Telcm	Shops	Hlth	Utils	Other	avg. RMSE	(share)
HIST	0.117	0.146	*0.096*	0.210	0.180	0.141	0.132	0.159	0.166	0.095	0.144	(0.10)
HIST$_{\text{ewma,ex}}$	0.120	0.152	0.101	0.215	*0.176*	0.144	0.126	0.157	0.161	0.096	0.145	(0.10)
HIST$^{\text{K}}$	0.116	0.145	0.099	*0.203*	0.206	0.142	0.132	0.152	0.163	0.100	0.146	(0.10)
HIST$^{\text{K}}_{\text{ewma,ex}}$	0.120	0.152	0.101	0.214	0.178	0.144	*0.125*	0.155	0.161	0.095	0.144	(0.10)
Best$^{\text{sim}}$	*0.115*	*0.145*	0.097	0.206	0.178	*0.140*	0.126	*0.151*	*0.159*	*0.094*	*0.141*	(0.60)
Dim$^{(5)}$	0.155	0.184	0.137	0.260	0.219	0.198	0.160	0.198	0.201	0.150	0.186	(0.00)
FP^{12}	0.158	0.178	0.149	0.221	0.215	0.164	0.184	0.197	0.187	0.164	0.182	(0.00)

64

Estimating Beta: Forecast Adjustments and the Impact of Stock Characteristics for a Broad Cross-Section

Online Appendix

Table A1: Summary Statistics – Combinations

This table presents value-weighted summary statistics for estimators based on forecast combinations as well as value-weighted averages of firm-level correlations. *Nobs* indicates the total number of firm–month observations for which we have estimates. *Mean^{vw}* is the overall value-weighted average of the estimates over the entire sample period. *Std. dev.* presents the average cross-sectional standard deviation. $q^{0.05}$ and $q^{0.95}$ indicate the averages of the cross-sectional 5% and 95% quantiles, respectively. The sample period runs from January 1963 until December 2015.

	Nobs	Mean$_{vw}$	Std. dev	$q^{0.05}$	$q^{0.95}$	Best$^{\text{sim}}$	Best$^{\text{C}}$	Best$^{\text{shr}}$	Best$^{\text{BMA}}$	All$^{\text{sim}}$	All$^{\text{C}}$	All$^{\text{shr}}$	All$^{\text{BMA}}$
HIST	3,354,313	1.00	0.63	−0.05	1.89	0.98	0.89	0.86	0.96	0.96	0.76	0.77	0.62
Best$^{\text{sim}}$	3,274,082	0.98	0.49	0.08	1.63	*	0.91	0.88	0.97	0.98	0.78	0.79	0.64
Best$^{\text{C}}$	2,385,441	1.01	0.57	0.04	1.73		*	0.93	0.92	0.92	0.80	0.80	0.60
Best$^{\text{shr}}$	2,385,441	1.01	0.61	0.04	1.71			*	0.87	0.91	0.79	0.79	0.60
Best$^{\text{BMA}}$	2,206,006	1.02	0.58	0.02	1.75				*	0.95	0.76	0.76	0.63
All$^{\text{sim}}$	1,403,858	1.00	0.45	0.16	1.64					*	0.78	0.79	0.64
All$^{\text{C}}$	1,047,550	1.03	0.72	−0.14	1.91						*	0.92	0.58
All$^{\text{shr}}$	1,047,550	1.02	0.68	−0.04	1.83							*	0.56
All$^{\text{BMA}}$	912,912	1.05	0.94	−0.34	2.09								*

1

	$q^{0.01}$	$q^{0.05}$	$q^{0.95}$	$q^{0.99}$	Share $\leq$ 0.1	Share $\geq$ 1	RMSE (RB> 1.5)	RMSE (RB< 1.5)	Expansion	Recession	Exp. $\Rightarrow$ Rec.	Rec. $\Rightarrow$ Exp.	Corr(ΔRMSE,CV)	Corr(RMSE,CV)
Beta$^{\mathrm{bm}}$	0.005	0.026	1.393	2.322	0.186	0.110	1.196	0.665	0.443	0.491	0.447	0.494	−0.08	−0.14*
Beta$^{\mathrm{cay}}$	0.005	0.026	1.409	2.357	0.186	0.113	1.179	0.685	0.446	0.433	0.447	0.423	−0.02	0.08
Beta$^{\mathrm{def}}$	0.005	0.026	1.409	2.433	0.184	0.112	1.229	0.673	0.433	0.544	0.449	0.608	0.52***	0.47***
Beta$^{\mathrm{dp}}$	0.005	0.026	1.428	2.489	0.185	0.114	1.195	0.694	0.431	0.509	0.431	0.585	−0.09	−0.13*
Beta$^{\mathrm{ep}}$	0.005	0.026	1.463	2.687	0.183	0.117	1.248	0.705	0.447	0.558	0.438	0.634	−0.51***	−0.54***
Beta$^{\mathrm{ic}}$	0.005	0.027	1.493	2.588	0.181	0.122	1.248	0.693	0.475	0.469	0.441	0.506	−0.29***	−0.23***
Beta$^{\mathrm{inf}}$	0.005	0.025	1.289	2.144	0.193	0.095	1.142	0.626	0.414	0.465	0.449	0.413	0.08	0.03
Beta$^{\mathrm{lty}}$	0.005	0.026	1.397	2.373	0.184	0.110	1.178	0.669	0.433	0.524	0.551	0.427	0.18**	0.19***
Beta$^{\mathrm{tbl}}$	0.005	0.027	1.426	2.348	0.180	0.117	1.215	0.665	0.445	0.489	0.514	0.476	−0.03	0.00
Beta$^{\mathrm{unc}}$	0.005	0.026	1.382	2.312	0.188	0.109	1.186	0.651	0.417	0.537	0.453	0.566	0.60***	0.51***
Beta$^{\mathrm{une}}$	0.005	0.027	1.513	2.664	0.178	0.125	1.303	0.675	0.486	0.496	0.446	0.543	0.22***	0.17**
Beta$^{\mathrm{all}}$	0.009	0.046	3.070	5.457	0.107	0.316	2.060	1.337	0.873	0.970	0.755	1.116		

Table A3: Prediction Errors – Different Horizons

This table presents the out-of-sample prediction errors for selected estimators and various forecast horizons. We measure the realized beta with daily returns over the horizons of 1 month (Panel A), 3 Months (Panel B), 12 Months (Panel C), and 60 Months (Panel D). Each month, we compute the value-weighted RMSE for the approaches considered using the entire cross-section of stocks. The first row reports the average value-weighted RMSE over our sample period. We indicate the lowest average RMSE with *italic* font. The remainders of the panels report the differences in prediction errors. The upper triangular matrices report the average differences in RMSE and the lower triangular matrices report the average differences in RMedSE. We report the error loss differential between the model *[name in row]* and the model *[name in column]*. The absolute values of the numbers in parentheses indicate the share of time periods for which the difference is significant at 5% (e.g., 0.4 indicates that the difference is significant 40% of the time). We test the significance using the modified Diebold–Mariano and the Wilcoxon signed rank tests for the upper and lower triangular matrix, respectively. The sign of the number in parentheses indicates the direction of the significant differences.

Panel A. 1-Month Forecast Horizon

	HIST	$\text{HIST}_{\text{ewma,ex}}$	HIST^{K}	$\text{HIST}^{\text{K}}_{\text{ewma,ex}}$	Best^{sim}
avg. RMSE	0.616	0.609	0.612	0.608	*0.606*
HIST		0.007	0.004	0.008	0.010
		(0.19)	(0.04)	(0.11)	(0.21)
$\text{HIST}_{\text{ewma,ex}}$	−0.007		−0.003	0.001	0.003
	(−0.48)		(−0.03)	(−0.00)	(0.10)
HIST^{K}	−0.009	−0.001		0.004	0.006
	(−0.51)	(−0.08)		(0.16)	(0.20)
$\text{HIST}^{\text{K}}_{\text{ewma,ex}}$	−0.010	−0.003	−0.001		0.002
	(−0.58)	(−0.39)	(−0.15)		(0.08)
Best^{sim}	−0.011	−0.004	−0.002	−0.001	
	(−0.74)	(−0.38)	(−0.28)	(−0.04)	

Panel B. 3-Month Forecast Horizon

	HIST	$\text{HIST}_{\text{ewma,ex}}$	HIST^{K}	$\text{HIST}^{\text{K}}_{\text{ewma,ex}}$	Best^{sim}
avg. RMSE	0.427	0.415	0.419	0.414	*0.411*
HIST		0.012	0.008	0.013	0.016
		(0.33)	(0.12)	(0.24)	(0.37)
$\text{HIST}_{\text{ewma,ex}}$	−0.012		−0.003	0.002	0.005
	(−0.68)		(−0.04)	(0.02)	(0.16)
HIST^{K}	−0.013	−0.001		0.005	0.008
	(−0.71)	(−0.08)		(0.22)	(0.32)
$\text{HIST}^{\text{K}}_{\text{ewma,ex}}$	−0.016	−0.004	−0.003		0.003
	(−0.75)	(−0.58)	(−0.31)		(0.12)
Best^{sim}	−0.017	−0.005	−0.004	−0.001	
	(−0.88)	(−0.50)	(−0.44)	(0.03)	

3

Panel C. 12-Month Forecast Horizon

	HIST	HIST$_{ewma,ex}$	HIST^K	HIST$^K_{ewma,ex}$	Best^{sim}
avg. RMSE	0.343	0.321	0.323	0.316	*0.314*
HIST		0.023	0.020	0.027	0.030
		(0.59)	(0.39)	(0.51)	(0.67)
HIST$_{ewma,ex}$	−0.020		−0.003	0.005	0.007
	(−0.83)		(−0.03)	(0.17)	(0.29)
HIST^K	−0.022	−0.002		0.007	0.009
	(−0.91)	(−0.17)		(0.23)	(0.34)
HIST$^K_{ewma,ex}$	−0.027	−0.007	−0.005		0.002
	(−0.90)	(−0.86)	(−0.27)		(0.12)
Best^{sim}	−0.027	−0.007	−0.005	0.001	
	(−0.97)	(−0.61)	(−0.39)	(0.11)	

Panel D. 60-Month Forecast Horizon

	HIST	HIST$_{ewma,ex}$	HIST^K	HIST$^K_{ewma,ex}$	Best^{sim}
avg. RMSE	0.369	0.334	0.322	0.321	*0.320*
HIST		0.035	0.047	0.047	0.049
		(0.78)	(0.85)	(0.82)	(0.95)
HIST$_{ewma,ex}$	−0.028		0.012	0.013	0.014
	(−0.89)		(0.31)	(0.49)	(0.58)
HIST^K	−0.039	−0.011		0.001	0.002
	(−1.00)	(−0.57)		(−0.06)	(0.07)
HIST$^K_{ewma,ex}$	−0.041	−0.013	−0.002		0.001
	(−0.95)	(−0.99)	(−0.05)		(0.24)
Best^{sim}	−0.039	−0.012	−0.001	0.002	
	(−1.00)	(−0.81)	(0.15)	(0.16)	

Table A4: Hedging Errors

This table presents the ratio of hedging error variances to the market variance for different approaches. For each stock, estimator, and month, we obtain the hedging error over the next month as $h_{i,t} = R_{j,t} - \beta_{j,t} R_{M,t}$. We estimate the ratio of the hedging error variance to the market variance. We estimate the variances using rolling 5-year windows and use the average ratio over time. We present the mean hedging error ratios across all stocks. Additionally, we report differences of the mean to that of HIST (Diff). The lowest mean average hedging error ratio is indicated by *italic* font. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

	HIST	HIST$_{\text{ewma,ex}}$	HIST$^{\text{K}}$	HIST$^{\text{K}}_{\text{ewma,ex}}$	Best$^{\text{sim}}$	Dim$^{(5)}$	FP^{12}
Mean	14.086	14.033	14.031	*14.016*	14.021	14.137	14.056
Diff		−0.0524***	−0.0541***	−0.0698***	−0.0648***	0.0515***	−0.0292***

Table A5: Mincer–Zarnowitz Regressions

This table reports the results of univariate Mincer–Zarnowitz regressions for selected estimators. Each month, we regress realized beta ($\beta^{\mathrm{R}}_{j,t}$) on a constant and an ex-ante estimate $\beta_{j,t}$. The regression equation is $\beta^{\mathrm{R}}_{j,t} = a + b\beta_{j,t} + \epsilon_t$. For each regression coefficient, we report in parentheses (sig) the share of times at which the coefficient is significantly different from 0 (for a) and 1 (for b) at 5% using Newey & West (1987) standard errors. R^2_{adj} indicates the adjusted R^2 of the regressions. The Wald test imposes the joint hypothesis that $a = 0$ and $b = 1$ and the column sig($Wald$) presents the results of the percentage of times for which the test statistic is significant at 5%.

	a	sig	b_{HIST}	sig	$b_{HIST_{ewma,ex}}$	sig	$b_{HIST^{K}}$	sig	$b_{HIST^{K}_{ewma,ex}}$	sig	$b_{Best^{sim}}$	sig	$b_{Dim^{(5)}}$	sig	b_{FP12}	sig	R^2_{adj}	sig($Wald$)
HIST	0.25	(0.98)	0.66	(0.99)													0.37	(1.00)
HIST$_{ewma,ex}$	0.18	(0.94)			0.75	(0.96)											0.39	(1.00)
HIST^{K}	0.03	(0.78)					1.00	(0.82)									0.40	(0.95)
HIST$^{K}_{ewma,ex}$	0.09	(0.81)							0.90	(0.85)							0.41	(0.96)
Best^{sim}	0.10	(0.81)									0.89	(0.83)					0.41	(0.97)
Dim$^{(5)}$	0.37	(1.00)											0.38	(1.00)			0.21	(1.00)
FP^{12}	0.25	(0.96)													0.56	(0.99)	0.32	(1.00)

6

Table A6: Prediction Errors – Equally Weighted

This table presents the out-of-sample prediction errors when weighting the results equally. We measure the realized beta with daily returns over the horizon of 6 months. Each month, we compute the equally weighted RMSE for the approaches considered using the entire cross-section of stocks. The first row reports the average equally weighted RMSE over our sample period. We indicate the lowest average RMSE with *italic* font. The remainder of the table reports the differences in prediction errors. The upper triangular matrix reports the average differences in RMSE and the lower triangular matrix reports the average differences in RMedSE. We report the error loss differential between the model *[name in row]* and the model *[name in column]*. The absolute values of the numbers in parentheses indicate the share of time periods for which the difference is significant at 5% (e.g., 0.4 indicates that the difference is significant 40% of the time). We test the significance using the modified Diebold–Mariano and the Wilcoxon signed rank tests for the upper and lower triangular matrix, respectively. The sign of the number in parentheses indicates the direction of the significant differences.

	HIST	$\text{HIST}_{\text{ewma,ex}}$	HIST^{K}	$\text{HIST}^{\text{K}}_{\text{ewma,ex}}$	Best^{sim}
avg. RMSE	0.627	0.591	0.572	*0.569*	0.570
HIST		0.036	0.055	0.058	0.057
		(0.90)	(0.84)	(0.92)	(0.97)
$\text{HIST}_{\text{ewma,ex}}$	−0.015		0.020	0.023	0.021
	(−0.79)		(0.46)	(0.82)	(0.81)
HIST^{K}	−0.017	−0.002		0.003	0.002
	(−0.80)	(−0.11)		(0.15)	(0.11)
$\text{HIST}^{\text{K}}_{\text{ewma,ex}}$	−0.021	−0.005	−0.004		−0.001
	(−0.85)	(−0.73)	(−0.30)		(−0.11)
Best^{sim}	−0.021	−0.006	−0.004	−0.001	
	(−0.94)	(−0.56)	(−0.45)	(0.05)	

Table A7: Prediction Errors – Firm-Level Evaluation

This table presents the out-of-sample prediction errors when evaluating the forecasts at the firm-level. We measure the realized beta with daily returns over the horizon of 6 months. For each stock, we compute the RMSE for the approaches considered using the entire time series. The first row reports the value-weighted average RMSE. We indicate the lowest average RMSE with *italic* font. The remainder of the table reports the differences in prediction errors. The upper triangular matrix reports the average differences in RMSE, averaged over all stocks. Similarly, the lower triangular matrix reports the average differences in RMedSE. We report the error loss differential between the model *[name in row]* and the model *[name in column]*. The absolute values of the numbers in parentheses indicate the share of time periods for which the difference is significant at 5% (e.g., 0.4 indicates that the difference is significant for firms representing 40% of the average total market capitalization). We test the significance using the modified Diebold–Mariano and the Wilcoxon signed rank tests for the upper and lower triangular matrix, respectively. The sign of the number in parentheses indicates the direction of the significant differences.

	HIST	$\text{HIST}_{\text{ewma,ex}}$	HIST^{K}	$\text{HIST}^{\text{K}}_{\text{ewma,ex}}$	Best^{sim}
avg. RMSE	0.381	0.367	0.370	0.365	*0.360*
HIST		0.014	0.011	0.016	0.020
		(0.12)	(0.24)	(0.12)	(0.27)
$\text{HIST}_{\text{ewma,ex}}$	−0.007		−0.003	0.003	0.007
	(−0.09)		(−0.01)	(0.06)	(0.06)
HIST^{K}	−0.008	−0.001		0.006	0.010
	(−0.16)	(0.01)		(0.02)	(0.09)
$\text{HIST}^{\text{K}}_{\text{ewma,ex}}$	−0.008	−0.001	0.000		0.004
	(−0.11)	(−0.06)	(−0.02)		(0.02)
Best^{sim}	−0.013	−0.006	−0.004	−0.004	
	(−0.23)	(−0.04)	(−0.11)	(−0.02)	

Table A8: Prediction Errors – Dimson Evaluation

This table presents the out-of-sample prediction errors when estimates are evaluated relative to the Dimson beta. We measure the realized beta using the Dimson approach and 5 lags with daily returns over the horizon of 6 months. Each month, we compute the value-weighted RMSE for the approaches considered using the entire cross-section of stocks. The first row reports the average value-weighted RMSE over our sample period. We indicate the lowest average RMSE with *italic* font. The remainder of the table reports the differences in prediction errors. The upper triangular matrix reports the average differences in RMSE and the lower triangular matrix reports the average differences in RMedSE. We report the error loss differential between the model *[name in row]* and the model *[name in column]*. The absolute values of the numbers in parentheses indicate the share of time periods for which the difference is significant at 5% (e.g., 0.4 indicates that the difference is significant 40% of the time). We test the significance using the modified Diebold–Mariano and the Wilcoxon signed rank tests for the upper and lower triangular matrix, respectively. The sign of the number in parentheses indicates the direction of the significant differences.

	HIST	$\text{HIST}_{\text{ewma,ex}}$	HIST^{K}	$\text{HIST}^{\text{K}}_{\text{ewma,ex}}$	Best^{sim}	$\text{Dim}^{(5)}$	FP^{12}
avg. RMSE	0.682	0.671	0.677	0.672	0.670	0.699	*0.661*
HIST		0.011	0.005	0.010	0.012	−0.017	0.021
		(0.34)	(0.12)	(0.21)	(0.32)	(−0.14)	(0.37)
$\text{HIST}_{\text{ewma,ex}}$	−0.013		−0.005	−0.001	0.001	−0.028	0.010
	(−0.62)		(−0.09)	(0.01)	(0.07)	(−0.22)	(0.21)
HIST^{K}	−0.009	0.003		0.005	0.007	−0.023	0.016
	(−0.48)	(0.05)		(0.17)	(0.28)	(−0.19)	(0.21)
$\text{HIST}^{\text{K}}_{\text{ewma,ex}}$	−0.015	−0.003	−0.006		0.002	−0.027	0.011
	(−0.63)	(−0.42)	(−0.24)		(0.08)	(−0.23)	(0.21)
Best^{sim}	−0.014	−0.002	−0.005	0.001		−0.029	0.009
	(−0.75)	(−0.20)	(−0.28)	(0.08)		(−0.25)	(0.17)
$\text{Dim}^{(5)}$	−0.008	0.004	0.001	0.007	0.006		0.038
	(0.03)	(0.13)	(0.06)	(0.14)	(0.12)		(0.43)
FP^{12}	−0.035	−0.022	−0.026	−0.020	−0.021	−0.027	
	(−0.37)	(−0.13)	(−0.15)	(−0.09)	(−0.11)	(−0.52)	

Table A9: Prediction Errors – Monthly Evaluation

This table presents the out-of-sample prediction errors when using monthly data for to evaluate the estimates. We measure the realized beta with monthly returns over the horizon of 60 months. Each month, we compute the value-weighted RMSE for the approaches considered using the entire cross-section of stocks. The first row reports the average value-weighted RMSE over our sample period. We indicate the lowest average RMSE with *italic* font. The remainder of the table reports the differences in prediction errors. The upper triangular matrix reports the average differences in RMSE and the lower triangular matrix reports the average differences in RMedSE. We report the error loss differential between the model *[name in row]* and the model *[name in column]*. The absolute values of the numbers in parentheses indicate the share of time periods for which the difference is significant at 5% (e.g., 0.4 indicates that the difference is significant 40% of the time). We test the significance using the modified Diebold–Mariano and the Wilcoxon signed rank tests for the upper and lower triangular matrix, respectively. The sign of the number in parentheses indicates the direction of the significant differences.

	HIST	$\text{HIST}_{\text{ewma,ex}}$	HIST^{K}	$\text{HIST}^{\text{K}}_{\text{ewma,ex}}$	Best^{sim}	$\text{Dim}^{(5)}$	FP^{12}
avg. RMSE	0.524	0.496	0.490	0.489	*0.488*	0.544	0.537
HIST		0.029	0.034	0.035	0.037	−0.020	−0.013
		(0.77)	(0.67)	(0.71)	(0.80)	(−0.16)	(−0.16)
$\text{HIST}_{\text{ewma,ex}}$	−0.025		0.005	0.007	0.008	−0.049	−0.042
	(−0.82)		(0.24)	(0.29)	(0.41)	(−0.41)	(−0.59)
HIST^{K}	−0.035	−0.010		0.002	0.003	−0.054	−0.047
	(−0.87)	(−0.46)		(−0.10)	(0.03)	(−0.48)	(−0.58)
$\text{HIST}^{\text{K}}_{\text{ewma,ex}}$	−0.036	−0.011	−0.001		0.001	−0.056	−0.049
	(−0.85)	(−0.78)	(0.12)		(0.19)	(−0.46)	(−0.61)
Best^{sim}	−0.034	−0.010	0.000	0.001		−0.057	−0.050
	(−0.92)	(−0.56)	(0.22)	(0.08)		(−0.49)	(−0.65)
$\text{Dim}^{(5)}$	0.004	0.029	0.039	0.040	0.039		0.007
	(0.08)	(0.40)	(0.49)	(0.51)	(0.50)		(0.03)
FP^{12}	−0.018	0.007	0.017	0.018	0.016	−0.022	
	(−0.42)	(0.04)	(0.16)	(0.20)	(0.18)	(−0.48)	

Table A10: Prediction Errors – MAE

This table presents the out-of-sample mean absolute prediction errors. We measure the realized beta with daily returns over the horizon of 6 months. Each month, we compute the value-weighted MAE for the approaches considered using the entire cross-section of stocks. The first row reports the average value-weighted MAE over our sample period. We indicate the lowest average MAE with *italic* font. The remainder of the table reports the differences in prediction errors. The upper triangular matrix reports the average differences in RMSE and the lower triangular matrix reports the average differences in MedAE. We report the error loss differential between the model *[name in row]* and the model *[name in column]*. The absolute values of the numbers in parentheses indicate the share of time periods for which the difference is significant at 5% (e.g., 0.4 indicates that the difference is significant 40% of the time). We test the significance using the modified Diebold–Mariano and the Wilcoxon signed rank tests for the upper and lower triangular matrix, respectively. The sign of the number in parentheses indicates the direction of the significant differences.

	HIST	$\text{HIST}_{\text{ewma,ex}}$	HIST^{K}	$\text{HIST}^{\text{K}}_{\text{ewma,ex}}$	Best^{sim}
avg. RMSE	0.270	0.259	0.261	0.257	*0.255*
HIST		0.011	0.010	0.013	0.016
		(0.56)	(0.52)	(0.56)	(0.77)
$\text{HIST}_{\text{ewma,ex}}$	−0.015		−0.002	0.002	0.004
	(−0.79)		(−0.17)	(0.24)	(0.25)
HIST^{K}	−0.017	−0.002		0.004	0.006
	(−0.80)	(−0.10)		(0.29)	(0.48)
$\text{HIST}^{\text{K}}_{\text{ewma,ex}}$	−0.021	−0.005	−0.004		0.003
	(−0.84)	(−0.74)	(−0.30)		(0.08)
Best^{sim}	−0.021	−0.006	−0.004	−0.001	
	(−0.94)	(−0.56)	(−0.44)	(0.03)	