

# How to Estimate Beta?\*

Fabian Hollstein<sup>†</sup>, Marcel Prokopczuk<sup>†,‡</sup>, and Chardin Wese Simen<sup>‡</sup>

November 11, 2017

## Abstract

Researchers and practitioners face many choices when estimating an asset's sensitivities toward risk factors, i.e., betas. We study the effect of different data sampling frequencies, forecast adjustments, and model combinations for beta estimation. 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 a shrinkage toward the industry average 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.

**JEL classification:** G12, G11, G17

**Keywords:** Beta estimation, forecast combinations, forecast adjustments

---

\*Contact: hollstein@fmt.uni-hannover.de (F. Hollstein), prokopczuk@fmt.uni-hannover.de (M. Prokopczuk), and C.Wese-Simen@icmacentre.ac.uk (C. Wese Simen).

<sup>†</sup>School of Economics and Management, Leibniz University Hannover, Koenigsworther Platz 1, 30167 Hannover, Germany.

<sup>‡</sup>ICMA Centre, Henley Business School, University of Reading, Reading, RG6 6BA, UK.

# I Introduction

Researchers and practitioners need estimates of betas for a wide variety of applications. Typically, historical data is used to estimate beta. Often the simple historical estimate is used. Others shrink the estimates toward the average beta of similar stocks. Some condition their estimates on macroeconomic state variables while others fit some kind of 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 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.<sup>1</sup>

More precisely, 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 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.

We examine several estimation and adjustment approaches. First, we study the impact of different estimation windows and data sampling frequencies. Regarding the estimation 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. Furthermore, consistent with the findings of [Hollstein et al. \(2017\)](#), 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.

---

<sup>1</sup>In this study, we concentrate on market beta. While betas are generally estimated with respect to various possible state variables, market beta is the most important and we therefore focus our analysis on it.

Second, we examine the impact of different weighting schemes. Conceptually, exponentially weighting past observations could deliver a possible solution to the conditionality vs. sample size trade-off because one can “*have it both ways*”, place higher weight on more recent observations to get a conditional estimate and use a long historical window to reduce measurement noise. Indeed, we find that exponentially weighting the observations yields significantly more precise estimates for beta.

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 simple shrinkage adjustments of [Vasicek \(1973\)](#) and [Karolyi \(1992\)](#) yield improvements for the simple historical estimator while the more elaborated individuals’ prior model of [Cosemans et al. \(2016\)](#) works considerably less well. With the simple shrinkage, one can tackle the extreme estimates that are likely associated with high measurement errors.

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.

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.

Finally, 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 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 contributes to the literature on beta estimation. [Buss & Vilkov \(2012\)](#), [Chang et al. \(2012\)](#), and [Baule et al. \(2016\)](#) develop option-implied beta estimators. [Hollstein & Prokopczuk \(2016\)](#) show that the [Buss & Vilkov \(2012\)](#) approach works particularly well in predicting future betas. While the intrinsically forward-looking nature of option-based estimators seems to be favorable, the estimators face one important shortcoming. They are only applicable for a subset of large stocks with active options markets.

More recently, [Hollstein et al. \(2017\)](#) 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. We complement these studies, first, by examining the whole stock universe available at the Center for Research in Security Prices (CRSP) and, second, by studying the impact of lower sampling frequencies, as well as different adjustments and forecast combinations.

Our paper 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.

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. In Section V, we present additional analyses and test the robustness of our results. Section VI concludes.

## II Data and Methodology

### A Data

We obtain daily data on stock returns, prices, 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 well 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 the market excess return:

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

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  $\tau$ , 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  $\tau$ .

**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  $h = \frac{\log(2)}{\iota}$ .  $\iota$  characterizes the horizon, to which the half-life of the weights converges for large samples. We try two alternatives for  $\iota$ : (i) one third and (ii) two thirds of the number of observations of the (initial) estimation window.<sup>2</sup>

**Shrinkage Estimators** Following [Vasicek \(1973\)](#), we obtain a posterior belief of beta by combining the historical estimate ( $\beta_{j,t}^{\text{HIST}}$ ) with a prior ( $b_{j,t}$ ) in the following way:

$$\beta_{j,t}^{\text{Shr}} = \frac{s_{b_{j,t}}^2}{\sigma_{\beta_{j,t}^{\text{HIST}}}^2 + s_{b_{j,t}}^2} \beta_{j,t}^{\text{HIST}} + \frac{\sigma_{\beta_{j,t}^{\text{HIST}}}^2}{\sigma_{\beta_{j,t}^{\text{HIST}}}^2 + s_{b_{j,t}}^2} b_{j,t}. \quad (2)$$

$\sigma_{\beta_{j,t}^{\text{HIST}}}^2$  and  $s_{b_{j,t}}^2$  are the variances of the historical estimate and the prior, respectively. Hence, the degree of shrinkage depends on the relative precisions of the historical estimate and the prior. We use as priors (i) the cross-sectional average beta ([Vasicek, 1973](#)) ( $\beta^{\text{V}}$ ), (ii) the cross-sectional average beta of firms in the same Global Industry Classification Standard (GICS) industry sector ([Karolyi, 1992](#)) ( $\beta^{\text{K}}$ ), and (iii) the fundamentals-based prior of [Cosemans et al. \(2016\)](#) ( $\beta^{\text{I}}$ ).<sup>3</sup>

**Dimson Beta** Following [Dimson \(1979\)](#) and [Lewellen & Nagel \(2006\)](#), we estimate a beta that ought to account for potential infrequent trading effects. If stocks trade less frequently than the market index, stock prices adjust gradually to new information. Therefore,

---

<sup>2</sup>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.

<sup>3</sup>[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 the interested reader to the original article. We obtain the balance sheet data necessary to compute the key ratios from Compustat.

Dimson (1979) adds lagged market returns in the regression:

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

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  $\beta_{j,t}^{\text{Dim}^{(N)}} = \sum_{i=0}^{\min(2,N)} \beta_{j,t}^{(i)}$ , where  $\min(\cdot)$  is the minimum operator.

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

$$\beta_{j,t}^{\text{Mac}} = b_{0,j} + B_j' z_t. \quad (4)$$

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 (4) 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.<sup>4</sup>

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).<sup>5</sup> We also use the 1-month macroeconomic uncertainty (unc) of Jurado et al. (2015) from Sydney Ludvigson’s webpage and the unemployment rate (une) from the Federal Reserve Economic Database.

We follow Goyal & Welch (2008) and also estimate a “kitchen-sink” (all) regression using

---

<sup>4</sup>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.

<sup>5</sup>For further description of the construction of the variables, we refer to Goyal & Welch (2008).

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 function.

**Forecast Combinations** [Bates & Granger \(1969\)](#) note that the combination of estimation techniques may prove worthwhile, especially when the estimates combined 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.<sup>6</sup> We employ an expanding window to make use of a maximum length of history to be able to estimate the parameters with greater precision.<sup>7</sup> The regression equation takes the following form:

$$\beta_{j,\tau}^R = a_{j,t} + \sum_{m=1}^M b_{j,t}^{(m)} \beta_{j,\tau}^{(m)} + \epsilon_{j,\tau}. \quad (5)$$

$\beta_{j,\tau}^{(m)}$  is the beta estimate for asset  $j$  of approach  $m$  at time  $\tau$ . We combine the estimates of  $M$  different models.  $\beta_{j,\tau}^R$  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, using the following equation:

$$\beta_{j,t}^C = \hat{a}_{j,t} + \sum_{m=1}^M \hat{b}_{j,t}^{(m)} \beta_{j,t}^{(m)}. \quad (6)$$

$\beta_{j,t}^C$  is the combined beta forecast for asset  $j$  at time  $t$  and  $\hat{a}_{j,t}$ ,  $\hat{b}_{j,t}^{(m)}$  are the respective

---

<sup>6</sup>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  $t$ .

<sup>7</sup>We also try a rolling window approach. The results indicate that the expanding window approach indeed yields superior results.



regression coefficients, i.e. weights.<sup>8</sup>

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  $\beta_{j,t}^{\text{shr}}$ , we use Equations (5) and (6) 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  $k = 1, \dots, K$  different possible ways to combine  $M$  different forecasts. The models differ in the subset of predictors used. Under the uninformative prior specification of [Fernandez et al. \(2001\)](#), assuming 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](#)):

$$\beta_{j,t}^{\text{BMA}} = \sum_{k=1}^K \omega_k \beta_{j,t}^{(k)}, \quad (7)$$

where  $\beta_{j,t}^{(k)}$  is the OLS combination (as of Equation (6)) of forecast models for one possible way  $k$  to combine the  $M$  forecasts. The weights  $\omega_k$  are:

$$\omega_k = \frac{a(g)^{\frac{1}{2}P_k} [1 + g^{-1}SSR_k^U/SSR^R]^{-\frac{1}{2}df^R}}{\sum_{i=1}^K a(g)^{\frac{1}{2}P_i} [1 + g^{-1}SSR_i^U/SSR^R]^{-\frac{1}{2}df^R}}. \quad (8)$$

Essentially, we first estimate a restricted forecasting model as in Equation (5) with OLS using only the variables that ought to be included in each model.<sup>9</sup> From this, we get the sum of squared residuals ( $SSR^R$ ). Second, we estimate a forecasting model as of Equation (5) 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  $k$ th 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

---

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

<sup>9</sup>When empirically implementing the approach, specifying variables that are included in each model can substantially reduce the computational effort.

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). We use daily (log-)returns during the prediction window  $t$  until  $T$  to estimate:<sup>10</sup>

$$\beta_{j,t}^R = \frac{\sum_{\tau=t+1}^T r_{j,\tau} r_{M,\tau}}{\sum_{\tau=t+1}^T r_{M,\tau}^2}, \quad (9)$$

where  $r_{j,\tau}$  and  $r_{M,\tau}$  refer to the return of asset  $j$  and the market return at time  $\tau$ , respectively.

**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:<sup>11</sup>

$$\text{RMSE}_j = \sqrt{\frac{1}{o} \sum_{t=1}^o (\beta_{j,t}^R - \beta_{j,t})^2}, \quad (10)$$

where  $o$  is the number of out-of-sample observations.  $\beta_{j,t}^R$  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. \(1997\)](#). For the cross-sectional evaluation, we use [Newey & West \(1987\)](#) standard errors with 4 lags. To test for significance in root median squared error (RMedSE) differences, we employ the

---

<sup>10</sup>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.

<sup>11</sup>In Section V.G, we also examine the Mean Absolute Error (MAE) loss function as an alternative, and obtain largely similar results as for the RMSE.

non-parametric Wilcoxon signed rank test.<sup>12</sup> 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.

### 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. Throughout our main empirical analysis, we follow [Chang et al. \(2012\)](#) and [Hollstein & Prokopczuk \(2016\)](#) and focus on a prediction horizon for realized beta of 6 months. 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 ( $\text{HIST}_{\text{mon}}$ ) 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.<sup>13</sup>

In Table 1, we present summary statistics on these estimators. Several properties of the estimators are worth mentioning. First, the value-weighted 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 hinders 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. Interestingly, while the value-weighted average of daily estimators with short windows is close to 1, it is clearly below 1 for estimators that use very long historical windows. For the 60-month estimator, the value-weighted average is 0.93. On the other hand, when using monthly or quarterly data with a long window, the value-weighted average is close to 1 again.

---

<sup>12</sup>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](#)).

<sup>13</sup>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).

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, but change the historical window size and sampling frequency, for all approaches. For 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 0.39. 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 0.73.

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<sup>12</sup>) 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 29% 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 each significantly higher than that of HIST<sup>12</sup> about 80% of the time. This result is also in line with the finding of [Hollstein et al. \(2017\)](#), 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.

Overall, the historical estimator using a 12-month window yields the most accurate predictions. In the following sections, we therefore concentrate on the 12-month estimation window, indicate HIST<sup>12</sup> simply by HIST, and examine if we can further improve the predictive accuracy by imposing different adjustments on the estimator.

## B Different Weighting Schemes

In the previous section, we analyzed the conditionality vs. sample size trade-off 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 weight all observations equally, independently of whether the returns occur 11 months or 1 week before the date of the estimation, one could also impose 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-lives 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.<sup>14</sup> Additionally, we use each of the two  $\iota$ s together with an expanding window (HIST<sub>ewma,s,ex</sub> and HIST<sub>ewma,ex</sub>), where we have an even larger sample size that might further increase the precision of the estimates.<sup>15</sup>

In Table 1, we present summary statistics for the exponentially weighted historical estimator.<sup>16</sup> We find that the overall properties of HIST<sub>ewma</sub> and HIST<sub>ewma,ex</sub> are similar to those of HIST<sup>12</sup> 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

---

<sup>14</sup>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.

<sup>15</sup>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.

<sup>16</sup>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.

in Table 3. 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 52% of the time. Thus, the exponential weighting, especially combined with an expanding estimation window, can materially reduce prediction errors in beta.

## 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,  $\text{HIST}^{\text{V}}$ ,  $\text{HIST}^{\text{K}}$ , and  $\text{HIST}^{\text{I}}$ .

Summary statistics of these estimators are presented in Table 1. Naturally, we find that the distributions of  $\text{HIST}^{\text{V}}$  and  $\text{HIST}^{\text{K}}$  are more narrow than that of the unadjusted 12-month historical estimator; also, quite naturally, since  $\text{HIST}^{\text{V}}$  and  $\text{HIST}^{\text{K}}$  are directly derived from  $\text{HIST}^{12}$  and (almost) perfectly cross-sectionally correlated with it. On the other hand, the value-weighted average is, with 0.97, slightly below 1 because  $\text{HIST}^{\text{V}}$  and  $\text{HIST}^{\text{K}}$  shrink the beta estimates toward an equally weighted average, which is typically below 1.<sup>17</sup> When imposing the fundamentals-based prior in  $\text{HIST}^{\text{I}}$ , things look quite different. First, the approach requires accounting data that is not as widely available as stock data. Because of this, and because we need an initial window to estimate the parameters, we have far fewer observations for  $\text{HIST}^{\text{I}}$  compared to the simple historical approaches. Furthermore, the value-weighted average of  $\text{HIST}^{\text{I}}$  is far below 1 with 0.87. While this low average could be due to wider availability of accounting data for low-beta stocks, it does deliver some indication that the approach is biased on average. The correlations of  $\text{HIST}^{\text{I}}$  with other approaches are moderate, e.g., the average cross-sectional correlation of  $\text{HIST}^{\text{I}}$  with  $\text{HIST}^{12}$  is 0.77.

---

<sup>17</sup>We also try  $\text{HIST}^{\text{V}}$  and  $\text{HIST}^{\text{K}}$  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.

We present the prediction errors of the different prior-adjusted betas in Table 4.<sup>18</sup> We find that both HIST<sup>V</sup> and HIST<sup>K</sup> yield lower average value-weighted RMSEs compared to HIST. The differences are significant 21% of the time. HIST<sup>K</sup>, which shrinks the beta estimates toward the average of the same industry, is slightly better than the less informative HIST<sup>V</sup>, which shrinks estimates toward the overall average beta. The picture looks quite different in the case of the individuals-based prior estimator, HIST<sup>I</sup>. The average value-weighted RMSE is substantially higher even compared to HIST. The differences are especially strong in the median. HIST<sup>I</sup> yields a higher root median squared error practically all the time. Thus, overall HIST<sup>I</sup> performs poorly. This finding is consistent with recent evidence in [Dittmar & Lundblad \(2017\)](#), who find that market betas are only weakly related to stock characteristics.

## 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 handle this is the [Dimson \(1979\)](#) adjustment, which sums up betas with respect to the current and lagged market return(s).

We present summary statistics for Dimson-betas with 1, 3, and 5 lags in Table 1. We find that the overall value-weighted average beta is similar to that of HIST<sup>12</sup>. Hence, neither of the estimators appears to be systematically biased. We find that the standard deviation as well as the quantile range rise the more lags we use. Additionally, the correlations with HIST<sup>12</sup> fall with an increasing number of lags. The average value-weighted cross-sectional correlation between HIST<sup>12</sup> and Dim<sup>(5)</sup> is 0.74. 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 in Table 5. 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 Dim<sup>(1)</sup> and Dim<sup>(5)</sup>

---

<sup>18</sup>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<sup>I</sup> and because many firms lack accounting data.

are significantly higher than those of HIST 35% and 73% of the time, respectively. Hence, there is very little evidence to warrant a Dimson adjustment.<sup>19</sup>

## E Macroeconomic Conditioning Information

If beta changes 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,  $\text{Beta}^{\text{cay}}$  (Lettau & Ludvigson, 2001). The results of the other estimators are qualitatively similar. Since we first need initial data to estimate Equation (4), we have fewer overall observations. We find that the value-weighted average beta is quite low with 0.91, which indicates that the approach may suffer from a systematic downward-bias. The cross-sectional standard deviation and quantile range of  $\text{Beta}^{\text{cay}}$  are neither very large nor very small and correlations to other approaches are rather low on average.

In Table 6, 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.<sup>20</sup> We find that all of the estimators based on macroeconomic conditioning variables substantially and significantly underperform HIST. The performance of the “kitchen-sink” approach  $\text{Beta}^{\text{all}}$  is especially poor. Thus, it appears to be much more favorable to broadly follow Lewellen & Nagel (2006), who use a (short) historical 12-month window to estimate conditional betas instead of using macroeconomic conditioning variables as in, e.g., Lettau & Ludvigson (2001) or Guo et al. (2017).

---

<sup>19</sup>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 V.E 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.

<sup>20</sup>The results when sampling monthly for all variables that are available at that frequency are qualitatively similar.



## 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  $\text{HIST}^{\text{K}}$  (Best), and (ii) a much larger subset of the different possible adjustments (All). For the latter, we combine HIST,  $\text{HIST}_{\text{ewma,ex}}$ ,  $\text{HIST}^{\text{K}}$ ,  $\text{HIST}^{\text{I}}$ ,  $\text{Dim}^{(5)}$ , and  $\text{Beta}^{\text{cay}}$ .<sup>21</sup> For both model sets, we use four combination possibilities: (i) a simple combination, (ii) a model-based combination as of Equation (6), (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  $\text{Best}^{\text{sim}}$  are overall very similar to those of HIST and the average value-weighted cross-sectional correlation is 0.99. 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, is far below 1, which indicates that these combinations yield a bias on average.

We present the prediction error results in Table 7. We find that the simple combination  $\text{Best}^{\text{sim}}$  yields a significantly lower average value-weighted RMSE compared to HIST 48% of the time. The model-based combinations  $\text{Best}^{\text{C}}$ ,  $\text{Best}^{\text{shr}}$ , and  $\text{Best}^{\text{BMA}}$  perform similarly to HIST, while the Bayesian approaches perform slightly better than the non-Bayesian combination  $\text{Best}^{\text{C}}$ . When combining all approaches, independently of whether they work or not individually, only the simple combination  $\text{All}^{\text{sim}}$  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 more than 50% of the time. Additionally, as in the case of just combining the best 2 models, we find that the Bayesian

---

<sup>21</sup>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.

combinations perform better than All<sup>C</sup>.

Thus, it appears worthwhile to combine estimators, but one should concentrate on those models that “work” individually and simple equally weighted combinations typically yield lower prediction errors than more elaborated regression-based combinations even when these use a Bayesian approach.

## G Which is the Best Approach?

Thus far, we examine which of the approaches yields an improvement relative to HIST. However, it is probably of high practical interest which of the adjustments and combinations overall yields the lowest prediction errors. We present the results for the models HIST<sub>ewma,ex</sub>, HIST<sup>K</sup>, an approach that directly imposes the industry-based prior on the EWMA Beta, HIST<sup>K</sup><sub>ewma,ex</sub>, and the simple combination of the 2 best models, Best<sup>sim</sup> in Table 8.<sup>22</sup>

We find that individually, HIST<sub>ewma,ex</sub> yields a significantly lower value-weighted RMSE compared to HIST<sup>K</sup> 21% of the time. Directly applying the prior suggested by [Karolyi \(1992\)](#) yields a small further improvement for HIST<sub>ewma,ex</sub>. However, the simple combination Best<sup>sim</sup> yields the overall lowest average value-weighted RMSE. While the differences in RMSE between Best<sup>sim</sup> and HIST<sub>ewma,ex</sub> as well as HIST<sup>K</sup><sub>ewma,ex</sub> are only rarely significant, Best<sup>sim</sup> significantly outperforms HIST 46% of the time.

## 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 try to address this by performing a decomposition of the mean squared errors (MSE). To do so, we follow [Mincer & Zarnowitz \(1969\)](#) and decompose

---

<sup>22</sup>One might wonder why the prediction errors of Table 8 are partially higher than those in Tables 4 and 7 for the same models. As already indicated in footnote 18, 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 two earlier tables.

the MSE in the following fashion:

$$\text{MSE}_j = \underbrace{(\bar{\beta}_j^{\text{R}} - \bar{\beta}_j)^2}_{\text{bias}} + \underbrace{(1 - b_j)^2 \sigma^2(\beta_j)}_{\text{inefficiency}} + \underbrace{(1 - \rho_j^2) \sigma^2(\beta_j^{\text{R}})}_{\text{random error}}. \quad (11)$$

$b_j$  is the slope coefficient of the regression  $\beta_j^{\text{R}} = a_j + b_j \beta_j + \epsilon_j$  and  $\rho_j^2$  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 9. We find that the best approaches,  $\text{HIST}_{\text{ewma,ex}}$ ,  $\text{HIST}^{\text{K}}$ ,  $\text{HIST}_{\text{ewma,ex}}^{\text{K}}$ , and  $\text{Best}^{\text{sim}}$ , yield improvements in all 3 dimensions. They slightly reduce the bias relative to HIST from 0.4% to 0.3%, they reduce the inefficiency of HIST by up to one third, and they also slightly reduce the random error. Hence, these adjustments create betas that are more accurate on average, for low and high beta stocks, and they reduce random measurement errors.

On the other hand, the reasons for why other models do not work are diverse.  $\text{HIST}^{\text{I}}$  exhibits a substantial bias of 2.4%. Additionally, both the inefficiency and the random error increase relative to HIST. The estimators  $\text{Dim}^{(1)}$ ,  $\text{Dim}^{(3)}$ , and  $\text{Dim}^{(5)}$ , on the other hand, do not strongly increase the bias, but both the inefficiency and the random error increase with increasing number of lags. Hence, the Dimson adjustment appears to systematically increase the forecast errors for high and low beta stocks. The beta augmented by macroeconomic conditioning variables,  $\text{Beta}^{\text{cay}}$ , also strongly increases the inefficiency relative to HIST and yields one of the highest overall random errors among all approaches.

Regarding the combinations of the two best individual models, we find that the regression-based combination  $\text{Best}^{\text{C}}$  slightly reduces the inefficiency relative to HIST, but increases the random error.  $\text{Best}^{\text{BMA}}$ , on the other hand, slightly reduces the bias and random error, but yields a marginally larger inefficiency. Finally, when combining more models, among them also models from which the adjustments “do not work” individually, creates very large

random errors. Thereby,  $All^{BMA}$  is superior to  $All^C$ , mainly because the inefficiency is not as high. However, both the inefficiency and the random measurement error of  $All^{BMA}$  and  $All^C$  are dramatically higher than that of HIST.

## V 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 10 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.<sup>23</sup>

We start the analysis by examining 1-month forecasts. We present these results in Panel A of Table 10. 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.<sup>24</sup> 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 10. 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 10 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 for the 6-month horizon continue to do so for the 12-month horizon. However, the best approach

---

<sup>23</sup>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 horizons yield slightly lower average value-weighted RMSEs.

<sup>24</sup>However, as indicated previously, the RMSE criterion is still a robust evaluation criterion if the sampling error is zero on average (Patton, 2011).

is the direct shrinkage adjustment  $\text{HIST}_{\text{ewma,ex}}^{\text{K}}$  and not the simple combination  $\text{Best}^{\text{sim}}$  that works best for shorter horizons.  $\text{HIST}_{\text{ewma,ex}}^{\text{K}}$  often yields significantly lower prediction errors than HIST,  $\text{HIST}^{\text{K}}$ , and  $\text{Best}^{\text{sim}}$ , both in RMSE and especially in RmedSE.

Finally, we present the results for the 60-month forecast horizon, relevant for very long-term investors, in Panel D of Table 10. 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 the 12-month horizon. Overall,  $\text{HIST}_{\text{ewma,ex}}^{\text{K}}$  yields the lowest average value-weighted RMSE.

## B Hedging Errors

The RMSE results show that the approaches  $\text{HIST}_{\text{ewma,ex}}$ ,  $\text{HIST}^{\text{K}}$ ,  $\text{HIST}_{\text{ewma,ex}}^{\text{K}}$ , and  $\text{Best}^{\text{sim}}$  yield the best results, while, e.g.,  $\text{Dim}^{(5)}$  performs very poorly. To account for the possibility that our ex-post realized betas are measured with error, we follow [Liu et al. \(2017\)](#) 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 compute the hedging error for each stock as

$$h_{j,t,T} = (r_{j,t,T} - r_{f,t,T}) - \beta_{j,t}(r_{M,t,T} - r_{f,t,T}). \quad (12)$$

$r_{j,t,T}$  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.  $\beta_t$  is the estimate for beta, using data up to time  $t$ . [Liu et al. \(2017\)](#) 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. \(2017\)](#) and estimate the variances 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 A2 of the online appendix. These results are consistent with our previous results relying on the RMSE and realized beta computations. We find that  $\text{HIST}_{\text{ewma,ex}}$ ,  $\text{HIST}^{\text{K}}$ ,  $\text{HIST}_{\text{ewma,ex}}^{\text{K}}$ , as well as  $\text{Best}^{\text{sim}}$  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.  $\text{HIST}_{\text{ewma,ex}}^{\text{K}}$  achieves the lowest mean average hedging error ratio. Thus, our main results appear to be robust to the specification of realized beta.

## C 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 weighting all stocks equally.

We present the equally weighted prediction error results in Table A3 of the online appendix.<sup>25</sup> We find that all the average RMSEs are higher for all approaches than for the value-weighted examination. Thus, it seems to be considerably more difficult to estimate the betas for 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 weighting 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. Overall,  $\text{HIST}_{\text{ewma,ex}}^{\text{K}}$  yields the lowest average RMSE.

---

<sup>25</sup>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 weighting all stocks equally. The average RMSE is significantly higher than that of HIST nearly all the time, independent of the number of lags used.

## D Firm-Level Evaluation

In the main part of this paper we evaluate the forecasts cross-sectionally. However, it may also be of interest to see how the adjustment approaches perform for different stocks separately in the time series. To perform this analysis, 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 A4 of the online appendix. These are qualitatively similar as those for the cross-sectional evaluation. The best adjustments also yield lower value-weighted average RMSEs compared to HIST. Best<sup>sim</sup> yields the overall lowest value-weighted average RMSE.

## E 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 Dimson realized beta. We estimate the realized beta as the sum of the realized beta as of Equation (9) with 0 up to 5 lags.<sup>26,27</sup>

We present the results in Table A5 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 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 Dim<sup>(5)</sup> 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

---

<sup>26</sup>E.g., for 1 lag, in the numerator we multiply  $r_{j,\tau}$  by  $r_{M,\tau-1}$  instead of  $r_{M,\tau}$ , etc.

<sup>27</sup>Using the ex post historical Dimson estimator of Equation (4) instead yields results that are qualitatively similar.

under the regular realized beta.<sup>28</sup>

## F 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 10.

Table A6 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 find that  $\text{HIST}_{\text{ewma,ex}}^{\text{K}}$  yields the lowest overall average value-weighted RMSE. Finally,  $\text{Dim}^{(5)}$  still yields a substantially and often significantly higher average value-weighted RMSE compared to HIST and even more so compared to  $\text{HIST}_{\text{ewma,ex}}^{\text{K}}$ . Hence, in general, estimators that do not account for infrequent trading appear to be superior to those that use a Dimson adjustment.

---

<sup>28</sup>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.



## G 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):

$$\text{MAE} = \frac{1}{o} \sum_{t=1}^o |\beta_{j,t}^R - \beta_{j,t}|, \quad (13)$$

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 A7 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  $\text{HIST}_{\text{ewma,ex}}^K$  and  $\text{Best}^{\text{sim}}$  yield the lowest value-weighted average MAE.

## VI 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.

On the other hand, an adjustment 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.

## References

- 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.
- 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.
- Baule, R., Korn, O., & Sakning, S. (2016). Which beta is best? On the information content of option-implied betas. *European Financial Management*, *22*(3), 450–483.
- 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.
- 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.
- 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*, *forthcoming*.
- Fama, E. F., & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, *81*(4), 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.
- 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*, *forthcoming*.
- 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. (2017). The value of high-frequency data for beta estimation. *Leibniz University Working Paper*.
- 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.
- 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.
- 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. (2017). Absolving beta of volatility’s effects. *Journal of Financial Economics*, *forthcoming*.
- 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. In G. Elliot, C. W. J. Granger, & A. Timmermann (Eds.) *Handbook of Economic Forecasting*, (pp. 135–196). Amsterdam: Elsevier.
- 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.

	<i>Nobs</i>	$Mean^{vw}$	<i>Std. dev</i>	$q^{0.05}$	$q^{0.95}$	HIST <sup>3</sup>	HIST <sup>6</sup>	HIST <sup>12</sup>	HIST <sup>24</sup>	HIST <sup>36</sup>	HIST <sup>60</sup>	HIST <sup>60mon</sup>	HIST <sup>q</sup>	HIST <sup>ewma</sup>	HIST <sup>ewma,ex</sup>	HIST <sup>V</sup>	HIST <sup>K</sup>	HIST <sup>I</sup>	Dim <sup>(1)</sup>	Dim <sup>(3)</sup>	Dim <sup>(5)</sup>	Beta <sup>cav</sup>	
HIST <sup>1</sup>	3,482,306	0.99	1.32	-1.03	2.81	0.78	0.70	0.64	0.59	0.57	0.53	0.39	0.31	0.68	0.64	0.64	0.64	0.50	0.59	0.53	0.48	0.43	
HIST <sup>3</sup>	3,458,916	0.99	0.86	-0.39	2.20	*	0.90	0.82	0.76	0.73	0.68	0.50	0.40	0.87	0.82	0.82	0.82	0.64	0.76	0.69	0.62	0.54	
HIST <sup>6</sup>	3,424,766	0.99	0.72	-0.18	2.01	*	*	0.93	0.85	0.81	0.75	0.56	0.45	0.96	0.90	0.92	0.92	0.72	0.85	0.77	0.69	0.60	
HIST <sup>12</sup>	3,354,313	0.99	0.63	-0.05	1.89	*	*	*	0.93	0.89	0.83	0.61	0.49	0.99	0.96	1.00	1.00	0.77	0.91	0.82	0.74	0.67	
HIST <sup>24</sup>	3,213,995	0.99	0.57	0.02	1.80	*	*	*	*	0.96	0.90	0.67	0.53	0.92	0.98	0.93	0.93	0.75	0.84	0.75	0.69	0.77	
HIST <sup>36</sup>	3,054,285	0.97	0.54	0.06	1.75	*	*	*	*	*	0.95	0.70	0.56	0.87	0.97	0.89	0.89	0.72	0.79	0.71	0.65	0.82	
HIST <sup>60</sup>	2,737,036	0.93	0.51	0.09	1.68	*	*	*	*	*	*	0.73	0.58	0.81	0.92	0.83	0.83	0.70	0.73	0.66	0.60	0.85	
HIST <sup>60mon</sup>	2,770,429	0.99	0.69	0.16	2.33	*	*	*	*	*	*	*	0.77	0.61	0.68	0.60	0.60	0.55	0.64	0.64	0.62	0.60	
HIST <sup>q</sup>	2,133,299	1.00	0.75	0.26	2.53	*	*	*	*	*	*	*	*	0.49	0.54	0.48	0.48	0.54	0.54	0.52	0.51	0.51	
HIST <sup>ewma</sup>	3,354,313	0.99	0.64	-0.06	1.89	*	*	*	*	*	*	*	*	*	0.96	0.99	0.99	0.77	0.91	0.82	0.74	0.66	
HIST <sup>ewma,ex</sup>	3,354,313	0.99	0.57	0.02	1.78	*	*	*	*	*	*	*	*	*	*	0.96	0.96	0.77	0.87	0.78	0.71	0.77	
HIST <sup>V</sup>	3,354,313	0.97	0.52	0.05	1.68	*	*	*	*	*	*	*	*	*	*	*	1.00	0.78	0.91	0.81	0.73	0.67	
HIST <sup>K</sup>	3,277,719	0.97	0.52	0.06	1.70	*	*	*	*	*	*	*	*	*	*	*	*	0.79	0.90	0.81	0.73	0.67	
HIST <sup>I</sup>	1,397,454	0.87	0.38	0.41	1.60	*	*	*	*	*	*	*	*	*	*	*	*	*	0.71	0.64	0.59	0.59	
Dim <sup>(1)</sup>	3,355,001	0.99	0.67	-0.03	2.05	*	*	*	*	*	*	*	*	*	*	*	*	*	*	0.91	0.84	0.57	
Dim <sup>(3)</sup>	3,355,001	0.99	0.75	-0.06	2.25	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	0.93	0.51	
Dim <sup>(5)</sup>	3,355,001	0.99	0.83	-0.10	2.47	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Beta <sup>cav</sup>	2,113,109	0.91	0.58	-0.02	1.81	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	0.46	

**Table 2: Prediction Errors – Window Length and Sampling Frequency**

This table presents the out-of-sample prediction errors. We measure the realized beta with daily returns over the horizon of 6 months. 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 <sup>1</sup>	HIST <sup>3</sup>	HIST <sup>6</sup>	HIST <sup>12</sup>	HIST <sup>24</sup>	HIST <sup>36</sup>	HIST <sup>60</sup>	HIST <sup>12</sup> <sub>mon</sub>	HIST <sup>36</sup> <sub>mon</sub>	HIST <sup>60</sup> <sub>mon</sub>	HIST <sub>q</sub>
avg. RMSE	0.567	0.391	0.341	<i>0.323</i>	0.327	0.334	0.349	0.644	0.475	0.451	0.514
HIST <sup>1</sup>		0.176 (0.54)	0.226 (0.61)	0.244 (0.59)	0.240 (0.54)	0.233 (0.55)	0.218 (0.50)	-0.077 (-0.23)	0.092 (0.11)	0.116 (0.16)	0.053 (0.00)
HIST <sup>3</sup>	-0.151 (-0.99)		0.050 (0.40)	0.067 (0.42)	0.064 (0.34)	0.057 (0.26)	0.041 (0.14)	-0.253 (-0.67)	-0.084 (-0.43)	-0.061 (-0.38)	-0.123 (-0.49)
HIST <sup>6</sup>	-0.199 (-1.00)	-0.048 (-0.93)		0.017 (0.21)	0.014 (0.09)	0.007 (-0.01)	-0.009 (-0.17)	-0.303 (-0.80)	-0.134 (-0.73)	-0.111 (-0.66)	-0.173 (-0.67)
HIST <sup>12</sup>	-0.222 (-1.00)	-0.071 (-0.92)	-0.023 (-0.78)		-0.003 (-0.05)	-0.011 (-0.14)	-0.026 (-0.29)	-0.321 (-0.84)	-0.152 (-0.82)	-0.128 (-0.79)	-0.191 (-0.77)
HIST <sup>24</sup>	-0.226 (-0.99)	-0.075 (-0.89)	-0.027 (-0.67)	-0.004 (-0.36)		-0.007 (-0.15)	-0.023 (-0.32)	-0.317 (-0.83)	-0.148 (-0.87)	-0.125 (-0.82)	-0.187 (-0.80)
HIST <sup>36</sup>	-0.222 (-0.98)	-0.071 (-0.85)	-0.023 (-0.56)	0.000 (-0.17)	0.004 (0.02)		-0.016 (-0.29)	-0.310 (-0.82)	-0.141 (-0.86)	-0.117 (-0.83)	-0.180 (-0.82)
HIST <sup>60</sup>	-0.211 (-0.95)	-0.060 (-0.69)	-0.012 (-0.34)	0.011 (0.13)	0.015 (0.26)	0.011 (0.23)		-0.295 (-0.80)	-0.126 (-0.78)	-0.102 (-0.80)	-0.165 (-0.80)
HIST <sup>12</sup> <sub>mon</sub>	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)		0.169 (0.57)	0.193 (0.61)	0.130 (0.30)
HIST <sup>36</sup> <sub>mon</sub>	-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 (-0.20)
HIST <sup>60</sup> <sub>mon</sub>	-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)		-0.063 (-0.31)
HIST <sub>q</sub>	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: Prediction Errors – Different Weighting Schemes**

This table presents the out-of-sample prediction errors. We measure the realized beta with daily returns over the horizon of 6 months. 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 <sub>ewma,s</sub>	HIST <sub>ewma,s,ex</sub>	HIST <sub>ewma</sub>	HIST <sub>ewma,ex</sub>
avg. RMSE	0.327	0.319	0.311	0.318	<i>0.310</i>
HIST		0.007 (0.13)	0.016 (0.30)	0.008 (0.13)	0.017 (0.52)
HIST <sub>ewma,s</sub>	0.005 (0.24)		0.009 (0.44)	0.001 (-0.02)	0.009 (0.14)
HIST <sub>ewma,s,ex</sub>	-0.004 (-0.07)	-0.009 (-0.85)		-0.008 (-0.32)	0.001 (-0.02)
HIST <sub>ewma</sub>	0.002 (0.23)	-0.003 (-0.35)	0.006 (0.59)		0.009 (0.19)
HIST <sub>ewma,ex</sub>	-0.010 (-0.68)	-0.015 (-0.63)	-0.006 (-0.34)	-0.012 (-0.67)	

**Table 4: Prediction Errors – Imposing Priors**

This table presents the out-of-sample prediction errors. We measure the realized beta with daily returns over the horizon of 6 months. 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 <sup>V</sup>	HIST <sup>K</sup>	HIST <sup>I</sup>
avg. RMSE	0.302	0.292	<i>0.292</i>	0.353
HIST		0.010 (0.21)	0.010 (0.21)	-0.051 (-0.30)
HIST <sup>V</sup>	-0.005 (-0.14)		0.000 (0.03)	-0.060 (-0.40)
HIST <sup>K</sup>	-0.005 (-0.18)	0.000 (-0.11)		-0.060 (-0.41)
HIST <sup>I</sup>	0.115 (0.97)	0.120 (0.98)	0.120 (0.99)	



**Table 5: Prediction Errors – Asynchronicity**

This table presents the out-of-sample prediction errors. We measure the realized beta with daily returns over the horizon of 6 months. 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 <sup>(1)</sup>	Dim <sup>(2)</sup>	Dim <sup>(3)</sup>	Dim <sup>(4)</sup>	Dim <sup>(5)</sup>
avg. RMSE	<i>0.327</i>	0.358	0.374	0.398	0.420	0.443
HIST		-0.031 (-0.35)	-0.048 (-0.46)	-0.071 (-0.56)	-0.094 (-0.65)	-0.116 (-0.73)
Dim <sup>(1)</sup>	0.048 (0.86)		-0.017 (-0.22)	-0.040 (-0.45)	-0.062 (-0.62)	-0.085 (-0.72)
Dim <sup>(2)</sup>	0.072 (0.93)	0.023 (0.95)		-0.024 (-0.43)	-0.046 (-0.55)	-0.068 (-0.65)
Dim <sup>(3)</sup>	0.098 (0.96)	0.050 (0.99)	0.026 (0.96)		-0.022 (-0.38)	-0.045 (-0.53)
Dim <sup>(4)</sup>	0.121 (0.97)	0.073 (1.00)	0.050 (0.99)	0.023 (0.95)		-0.023 (-0.38)
Dim <sup>(5)</sup>	0.146 (0.98)	0.097 (1.00)	0.074 (0.99)	0.048 (0.98)	0.024 (0.95)	

**Table 6: Prediction Errors – Macroeconomic Conditioning Variables**

This table presents the out-of-sample prediction errors. We measure the realized beta with daily returns over the horizon of 6 months. 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 <sup>bm</sup>	Beta <sup>cav</sup>	Beta <sup>def</sup>	Beta <sup>dp</sup>	Beta <sup>ep</sup>	Beta <sup>ic</sup>	Beta <sup>inf</sup>	Beta <sup>ly</sup>	Beta <sup>tbl</sup>	Beta <sup>unc</sup>	Beta <sup>une</sup>	Beta <sup>all</sup>
avg. RMSE	<i>0.345</i>	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		-0.103 (-0.74)	-0.097 (-0.68)	-0.103 (-0.73)	-0.095 (-0.68)	-0.116 (-0.73)	-0.127 (-0.83)	-0.076 (-0.72)	-0.100 (-0.70)	-0.104 (-0.80)	-0.088 (-0.72)	-0.141 (-0.84)	-0.540 (-0.98)
Beta <sup>bm</sup>	0.064 (0.79)		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)	-0.438 (-0.90)
Beta <sup>cav</sup>	0.071 (0.74)	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 <sup>def</sup>	0.074 (0.79)	0.010 (0.03)	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 <sup>dp</sup>	0.069 (0.77)	0.005 (0.14)	-0.003 (-0.07)	-0.005 (-0.04)		-0.021 (-0.02)	-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 <sup>ep</sup>	0.080 (0.77)	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)	-0.42 (-0.88)
Beta <sup>ic</sup>	0.080 (0.82)	0.016 (0.19)	0.009 (0.08)	0.006 (0.11)	0.012 (-0.05)	0.000 (0.04)		0.051 (0.23)	0.027 (0.26)	0.023 (0.08)	0.039 (0.19)	-0.014 (-0.05)	-0.413 (-0.88)
Beta <sup>inf</sup>	0.048 (0.73)	-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 <sup>ly</sup>	0.072 (0.75)	0.008 (-0.03)	0.001 (-0.06)	-0.002 (-0.12)	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 <sup>tbl</sup>	0.073 (0.82)	0.009 (0.12)	0.002 (0.11)	-0.001 (0.09)	0.005 (0.11)	-0.007 (0.03)	-0.007 (0.08)	0.026 (0.21)	0.001 (0.09)		0.017 (0.11)	-0.037 (-0.07)	-0.436 (-0.87)
Beta <sup>unc</sup>	0.063 (0.79)	-0.001 (-0.01)	-0.008 (0.03)	-0.010 (-0.01)	-0.005 (0.08)	-0.017 (-0.13)	-0.017 (-0.07)	0.016 (0.12)	-0.009 (0.03)	-0.010 (-0.11)		-0.053 (-0.21)	-0.452 (-0.94)
Beta <sup>une</sup>	0.095 (0.83)	0.031 (0.12)	0.024 (0.07)	0.022 (0.13)	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 (0.14)		-0.399 (-0.84)
Beta <sup>all</sup>	0.440 (1.00)	0.376 (0.97)	0.369 (0.95)	0.366 (0.98)	0.372 (0.98)	0.360 (0.94)	0.360 (0.94)	0.393 (0.96)	0.368 (0.89)	0.367 (0.96)	0.377 (1.00)	0.345 (0.93)	

**Table 7: Prediction Errors – Combinations**

This table presents the out-of-sample prediction errors. We measure the realized beta with daily returns over the horizon of 6 months. 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 <sup>sim</sup>	Best <sup>C</sup>	Best <sup>shr</sup>	Best <sup>BMA</sup>	All <sup>sim</sup>	All <sup>C</sup>	All <sup>shr</sup>	All <sup>BMA</sup>
avg. RMSE	0.346	<i>0.325</i>	0.352	0.349	0.343	0.331	0.543	0.501	0.503
HIST		0.021 (0.48)	-0.006 (-0.13)	-0.002 (-0.04)	0.003 (0.08)	0.016 (0.24)	-0.196 (-0.64)	-0.155 (-0.56)	-0.156 (-0.58)
Best <sup>sim</sup>	-0.013 (-0.62)		-0.027 (-0.40)	-0.023 (-0.33)	-0.018 (-0.52)	-0.005 (-0.15)	-0.217 (-0.68)	-0.176 (-0.63)	-0.177 (-0.63)
Best <sup>C</sup>	0.009 (0.23)	0.021 (0.61)		0.003 (0.32)	0.009 (0.14)	0.021 (0.35)	-0.191 (-0.69)	-0.149 (-0.58)	-0.151 (-0.59)
Best <sup>shr</sup>	0.002 (0.01)	0.015 (0.55)	-0.006 (-0.73)		0.006 (0.09)	0.018 (0.24)	-0.194 (-0.68)	-0.153 (-0.59)	-0.154 (-0.59)
Best <sup>BMA</sup>	-0.005 (-0.23)	0.008 (0.48)	-0.013 (-0.32)	-0.007 (-0.12)		0.013 (0.21)	-0.199 (-0.66)	-0.158 (-0.57)	-0.159 (-0.59)
All <sup>sim</sup>	0.003 (-0.03)	0.016 (0.43)	-0.006 (-0.27)	0.001 (-0.07)	0.008 (0.10)		-0.212 (-0.69)	-0.171 (-0.64)	-0.172 (-0.62)
All <sup>C</sup>	0.121 (1.00)	0.134 (1.00)	0.112 (1.00)	0.118 (1.00)	0.125 (1.00)	0.118 (0.99)		0.041 (0.51)	0.040 (0.16)
All <sup>shr</sup>	0.088 (1.00)	0.101 (1.00)	0.079 (1.00)	0.086 (1.00)	0.093 (1.00)	0.085 (0.97)	-0.033 (-1.00)		-0.001 (0.02)
All <sup>BMA</sup>	0.084 (0.99)	0.097 (0.98)	0.076 (0.97)	0.082 (0.98)	0.089 (1.00)	0.081 (0.93)	-0.037 (-0.73)	-0.004 (0.08)	

**Table 8: Prediction Errors – Best Models**

This table presents the out-of-sample prediction errors. We measure the realized beta with daily returns over the horizon of 6 months. 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 <sub>ewma,ex</sub>	HIST <sup>K</sup>	HIST <sup>K</sup> <sub>ewma,ex</sub>	Best <sup>sim</sup>
avg. RMSE	0.364	0.345	0.353	0.344	<i>0.343</i>
HIST		0.018 (0.57)	0.011 (0.18)	0.020 (0.59)	0.020 (0.46)
HIST <sub>ewma,ex</sub>	-0.011 (-0.70)		-0.008 (-0.21)	0.001 (0.01)	0.002 (0.06)
HIST <sup>K</sup>	-0.006 (-0.24)	0.005 (0.29)		0.009 (0.36)	0.010 (0.52)
HIST <sup>K</sup> <sub>ewma,ex</sub>	-0.013 (-0.79)	-0.002 (-0.64)	-0.007 (-0.50)		0.000 (-0.03)
Best <sup>sim</sup>	-0.012 (-0.64)	-0.002 (-0.18)	-0.007 (-0.65)	0.001 (0.23)	

**Table 9: Forecast Error Decomposition**

This table presents the decomposition of mean squared forecast errors. 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.

*Panel A. Adjustments that “Work”*

	Bias	Inefficiency	Random Error
HIST	0.004	0.015	0.105
HIST <sub>ewma,ex</sub>	0.003	0.011	0.098
HIST <sup>K</sup>	0.003	0.012	0.100
HIST <sub>ewma,ex</sub> <sup>K</sup>	0.003	0.011	0.098
Best <sup>sim</sup>	0.003	0.010	0.097

*Panel B. Adjustments that “do not Work”*

	Bias	Inefficiency	Random Error
HIST <sup>I</sup>	0.024	0.018	0.136
Dim <sup>(1)</sup>	0.004	0.025	0.120
Dim <sup>(3)</sup>	0.005	0.044	0.138
Dim <sup>(5)</sup>	0.005	0.078	0.151
Beta <sup>cay</sup>	0.004	0.050	0.162
Best <sup>C</sup>	0.004	0.013	0.112
Best <sup>BMA</sup>	0.003	0.016	0.102
All <sup>C</sup>	0.004	0.139	0.168
All <sup>BMA</sup>	0.004	0.107	0.152

**Table 10: Prediction Errors – Different Horizons**

This table presents the out-of-sample prediction errors. 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). 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	HIST <sub>ewma,ex</sub>	HIST <sup>K</sup>	HIST <sup>K</sup> <sub>ewma,ex</sub>	Best <sup>sim</sup>
avg. RMSE	0.612	0.599	0.602	0.598	<i>0.597</i>
HIST		0.014 (0.46)	0.011 (0.23)	0.014 (0.47)	0.016 (0.42)
HIST <sub>ewma,ex</sub>	-0.009 (-0.65)		-0.003 (-0.09)	0.001 (-0.01)	0.002 (0.09)
HIST <sup>K</sup>	-0.007 (-0.32)	0.002 (0.15)		0.004 (0.14)	0.005 (0.24)
HIST <sup>K</sup> <sub>ewma,ex</sub>	-0.010 (-0.73)	-0.001 (-0.35)	-0.003 (-0.26)		0.001 (0.06)
Best <sup>sim</sup>	-0.011 (-0.58)	-0.002 (-0.15)	-0.004 (-0.40)	-0.001 (0.04)	

*Panel B. 3-Month Forecast Horizon*

	HIST	HIST <sub>ewma,ex</sub>	HIST <sup>K</sup>	HIST <sup>K</sup> <sub>ewma,ex</sub>	Best <sup>sim</sup>
avg. RMSE	0.425	0.408	0.413	0.407	<i>0.406</i>
HIST		0.017 (0.55)	0.012 (0.21)	0.018 (0.56)	0.019 (0.48)
HIST <sub>ewma,ex</sub>	-0.010 (-0.69)		-0.005 (-0.14)	0.001 (-0.00)	0.002 (0.08)
HIST <sup>K</sup>	-0.007 (-0.29)	0.004 (0.25)		0.006 (0.26)	0.007 (0.41)
HIST <sup>K</sup> <sub>ewma,ex</sub>	-0.012 (-0.79)	-0.002 (-0.55)	-0.006 (-0.45)		0.001 (0.02)
Best <sup>sim</sup>	-0.012 (-0.64)	-0.002 (-0.17)	-0.006 (-0.62)	0.000 (0.11)	

Table 10: Prediction Errors – Different Horizons (continued)

*Panel C. 12-Month Forecast Horizon*

	HIST	HIST <sub>ewma,ex</sub>	HIST <sup>K</sup>	HIST <sup>K</sup> <sub>ewma,ex</sub>	Best <sup>sim</sup>
avg. RMSE	0.332	0.315	0.326	<i>0.313</i>	0.314
HIST		0.018 (0.58)	0.006 (0.10)	0.020 (0.59)	0.018 (0.43)
HIST <sub>ewma,ex</sub>	-0.010 (-0.66)		-0.011 (-0.28)	0.002 (0.06)	0.001 (0.05)
HIST <sup>K</sup>	-0.003 (-0.17)	0.007 (0.36)		0.013 (0.45)	0.012 (0.57)
HIST <sup>K</sup> <sub>ewma,ex</sub>	-0.014 (-0.80)	-0.003 (-0.79)	-0.010 (-0.57)		-0.002 (-0.09)
Best <sup>sim</sup>	-0.012 (-0.61)	-0.001 (-0.13)	-0.008 (-0.72)	0.002 (0.32)	

*Panel D. 60-Month Forecast Horizon*

	HIST	HIST <sub>ewma,ex</sub>	HIST <sup>K</sup>	HIST <sup>K</sup> <sub>ewma,ex</sub>	Best <sup>sim</sup>
avg. RMSE	0.341	0.327	0.341	<i>0.323</i>	0.329
HIST		0.014 (0.43)	0.000 (0.02)	0.019 (0.54)	0.013 (0.31)
HIST <sub>ewma,ex</sub>	-0.009 (-0.49)		-0.014 (-0.35)	0.005 (0.28)	-0.001 (-0.03)
HIST <sup>K</sup>	-0.002 (-0.13)	0.007 (0.22)		0.018 (0.55)	0.013 (0.63)
HIST <sup>K</sup> <sub>ewma,ex</sub>	-0.014 (-0.71)	-0.006 (-0.97)	-0.013 (-0.50)		-0.006 (-0.35)
Best <sup>sim</sup>	-0.010 (-0.50)	-0.001 (-0.16)	-0.009 (-0.56)	0.004 (0.41)	

# How to Estimate Beta?

## Online Appendix

**JEL classification:** G11, G12, G17

**Keywords:** Beta estimation, forecast combinations, forecast adjustments



**Table A1: Summary Statistics – Combinations**

This table presents value-weighted summary statistics for our main estimators 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.

	<i>Nobs</i>	$Mean^{vw}$	<i>Std. dev.</i>	$q^{0.05}$	$q^{0.95}$	$Best^{sim}$	$Best^C$	$Best^{shr}$	$Best^{BMA}$	$All^{sim}$	$All^C$	$All^{shr}$	$All^{BMA}$
HIST	3,354,313	0.99	0.63	-0.05	1.89	0.99	0.85	0.84	0.94	0.95	0.63	0.68	0.67
$Best^{sim}$	3,277,719	0.98	0.54	0.05	1.73	*	0.89	0.87	0.96	0.97	0.67	0.71	0.71
$Best^C$	2,226,506	0.86	0.55	0.04	1.70		*	0.95	0.89	0.92	0.72	0.75	0.71
$Best^{shr}$	2,226,506	0.86	0.57	0.04	1.70			*	0.86	0.91	0.71	0.74	0.71
$Best^{BMA}$	2,226,506	0.88	0.58	0.00	1.74				*	0.94	0.66	0.69	0.72
$All^{sim}$	1,396,958	0.85	0.43	0.23	1.62					*	0.68	0.73	0.71
$All^C$	988,922	0.77	0.79	-0.23	1.99						*	0.96	0.72
$All^{shr}$	988,922	0.77	0.72	-0.09	1.89							*	0.72
$All^{BMA}$	863,603	0.82	0.81	-0.19	1.98								*

**Table A2: 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. To account for the possibility that the variances in the numerator and denominator change over time, 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 to of the mean to those 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	$HIST_{ewma,ex}$	$HIST^K$	$HIST_{ewma,ex}^K$	$Best^{sim}$	$Dim^{(5)}$
Mean	13.810	13.801	13.795	<i>13.787</i>	13.790	13.902
Diff	0.0000	-0.0097***	-0.0156***	-0.0239***	-0.0204***	0.0915***

**Table A3: Prediction Errors – Equally Weighted**

This table presents the out-of-sample prediction errors. We measure the realized beta with daily returns over the horizon of 6 months. 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	HIST <sub>ewma,ex</sub>	HIST <sup>K</sup>	HIST <sup>K</sup> <sub>ewma,ex</sub>	Best <sup>sim</sup>
avg. RMSE	0.604	0.588	0.582	<i>0.575</i>	0.577
HIST		0.016 (0.59)	0.022 (0.55)	0.029 (0.86)	0.027 (0.78)
HIST <sub>ewma,ex</sub>	-0.011 (-0.70)		0.006 (0.14)	0.013 (0.90)	0.011 (0.59)
HIST <sup>K</sup>	-0.006 (-0.24)	0.005 (0.29)		0.007 (0.35)	0.005 (0.37)
HIST <sup>K</sup> <sub>ewma,ex</sub>	-0.013 (-0.79)	-0.002 (-0.64)	-0.007 (-0.50)		-0.002 (-0.25)
Best <sup>sim</sup>	-0.012 (-0.64)	-0.002 (-0.18)	-0.007 (-0.65)	0.001 (0.23)	

**Table A4: Prediction Errors – Firm-Level Evaluation**

This table presents the out-of-sample prediction errors. We measure the realized beta with daily returns over the horizon of 6 months. 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	HIST <sub>ewma,ex</sub>	HIST <sup>K</sup>	HIST <sup>K</sup> <sub>ewma,ex</sub>	Best <sup>sim</sup>
avg. RMSE	0.382	0.363	0.369	0.362	<i>0.360</i>
HIST		0.019 (0.25)	0.013 (0.04)	0.020 (0.26)	0.022 (0.19)
HIST <sub>ewma,ex</sub>	-0.013 (-0.12)		-0.006 (-0.08)	0.001 (0.09)	0.003 (0.00)
HIST <sup>K</sup>	-0.010 (-0.01)	0.002 (0.02)		0.007 (0.10)	0.009 (0.20)
HIST <sup>K</sup> <sub>ewma,ex</sub>	-0.013 (-0.13)	-0.001 (-0.02)	-0.003 (-0.03)		0.002 (-0.02)
Best <sup>sim</sup>	-0.016 (-0.08)	-0.003 (-0.00)	-0.005 (-0.07)	-0.002 (0.00)	

**Table A5: Prediction Errors – Dimson Evaluation**

This table presents the out-of-sample prediction errors. We measure the realized beta using the Dimson approach and 5 lags with daily returns over the horizon of 6 months. 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 <sub>ewma,ex</sub>	HIST <sup>K</sup>	HIST <sup>K</sup> <sub>ewma,ex</sub>	Best <sup>sim</sup>	Dim <sup>(5)</sup>
avg. RMSE	0.669	0.659	0.666	<i>0.659</i>	0.660	0.686
HIST		0.010 (0.35)	0.002 (0.05)	0.010 (0.34)	0.009 (0.25)	-0.017 (-0.13)
HIST <sub>ewma,ex</sub>	-0.006 (-0.38)		-0.007 (-0.23)	0.000 (-0.01)	0.000 (-0.05)	-0.027 (-0.21)
HIST <sup>K</sup>	0.000 (-0.06)	0.006 (0.26)		0.007 (0.28)	0.007 (0.40)	-0.020 (-0.16)
HIST <sup>K</sup> <sub>ewma,ex</sub>	-0.008 (-0.50)	-0.002 (-0.42)	-0.008 (-0.38)		-0.001 (-0.08)	-0.027 (-0.20)
Best <sup>sim</sup>	-0.005 (-0.34)	0.001 (0.00)	-0.006 (-0.45)	0.002 (0.20)		-0.027 (-0.20)
Dim <sup>(5)</sup>	-0.003 (0.06)	0.004 (0.12)	-0.003 (0.06)	0.005 (0.13)	0.003 (0.10)	

**Table A6: Prediction Errors – Monthly Evaluation**

This table presents the out-of-sample prediction errors. We measure the realized beta with monthly returns over the horizon of 60 months. 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 <sub>ewma,ex</sub>	HIST <sup>K</sup>	HIST <sup>K</sup> <sub>ewma,ex</sub>	Best <sup>sim</sup>	Dim <sup>(5)</sup>
avg. RMSE	0.501	0.491	0.504	<i>0.488</i>	0.494	0.534
HIST		0.011 (0.42)	-0.003 (-0.08)	0.013 (0.49)	0.008 (0.27)	-0.032 (-0.23)
HIST <sub>ewma,ex</sub>	-0.007 (-0.30)		-0.014 (-0.38)	0.002 (0.14)	-0.003 (-0.17)	-0.043 (-0.35)
HIST <sup>K</sup>	-0.002 (-0.20)	0.006 (0.13)		0.016 (0.54)	0.011 (0.60)	-0.029 (-0.21)
HIST <sup>K</sup> <sub>ewma,ex</sub>	-0.012 (-0.52)	-0.005 (-0.81)	-0.011 (-0.41)		-0.005 (-0.37)	-0.045 (-0.36)
Best <sup>sim</sup>	-0.008 (-0.36)	0.000 (-0.16)	-0.006 (-0.38)	0.005 (0.40)		-0.040 (-0.34)
Dim <sup>(5)</sup>	0.021 (0.30)	0.028 (0.39)	0.023 (0.32)	0.033 (0.47)	0.029 (0.41)	

**Table A7: Prediction Errors – MAE**

This table presents the out-of-sample prediction errors. We measure the realized beta with daily returns over the horizon of 6 months. 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	HIST <sub>ewma,ex</sub>	HIST <sup>K</sup>	HIST <sup>K</sup> <sub>ewma,ex</sub>	Best <sup>sim</sup>
avg. RMSE	0.269	0.255	0.261	0.254	<i>0.254</i>
HIST		0.014 (0.70)	0.008 (0.16)	0.015 (0.71)	0.016 (0.52)
HIST <sub>ewma,ex</sub>	-0.011 (-0.71)		-0.006 (-0.37)	0.001 (0.15)	0.001 (-0.01)
HIST <sup>K</sup>	-0.006 (-0.25)	0.005 (0.29)		0.007 (0.40)	0.007 (0.63)
HIST <sup>K</sup> <sub>ewma,ex</sub>	-0.013 (-0.79)	-0.002 (-0.64)	-0.007 (-0.50)		0.001 (-0.10)
Best <sup>sim</sup>	-0.012 (-0.63)	-0.002 (-0.18)	-0.007 (-0.65)	0.001 (0.23)	