Shrinking beta

Published in:
Journal of Risk

Publication status and date:
Published: 01/09/2022

DOI (link to publisher):
10.21314/jor.2022.037

Document Version
Publisher's PDF, also known as Version of record

Document License/Available under:
Article 25fa Dutch Copyright Act

Citation for the published version (APA):
Blitz, D., Swinkels, L., Ūsaitė, K., & van Vliet, WN. (2022). Shrinking beta. *Journal of Risk*, *24*(6), 25-40. https://doi.org/10.21314/jor.2022.037

Journal of Risk **24**(6), 1–20
DOI: 10.21314/JOR.2022.037

Research Paper

Shrinking beta

David Blitz,[1] Laurens Swinkels,[1,2] Kristina Ūsaitė[1] and Pim van Vliet[1]

[1]Robeco Institutional Asset Management, Weena 850, 3014 DA Rotterdam, Netherlands; emails: d.c.blitz@robeco.com, l.swinkels@robeco.com, k.usaite@robeco.com, p.van.vliet@robeco.com

[2]Erasmus University Rotterdam, Burgemeester Oudlaan 50, 3062 PA Rotterdam, Netherlands; email: lswinkels@ese.eur.nl

(Received April 17, 2020; revised May 15, 2022; accepted May 25, 2022)

ABSTRACT

Beta is used in many applications, ranging from asset pricing tests to cost of capital estimation, investment management and risk management. Beta needs to be estimated, and shrinkage to its cross-sectional average value of 1 is often applied to reduce estimation error. Since beta is the product of the return correlation of a security with the market and its return volatility relative to that of the market, we shrink correlation and volatility separately and evaluate the predictive power of this approach. We find economically and statistically significant gains from applying more shrinkage to correlations than to volatilities.

Keywords: beta; correlation; investing; low risk; shrinkage; volatility.

1 INTRODUCTION

The capital asset pricing model (CAPM) developed by Sharpe (1964) presumes that the expected return on a security is related solely, and positively linearly, to its systematic risk; this relationship is measured by the covariance of the security's returns

Corresponding author: L. Swinkels

Print ISSN 1465-1211 | Online ISSN 1755-2842

with those of the market, also known as the security's beta. Although beta is not good at predicting returns, it can still be useful for a wide range of purposes, such as security selection, risk management, investment fund performance evaluation and capital budgeting decisions. All things being equal, risk-averse investors should demand a higher expected return for stocks and investment projects with a higher beta. In any case, beta is a characteristic that cannot be observed directly and needs to be estimated.

Hence, an important question for academics and investors is how to best estimate the beta of a security. It is especially important to know how to deal with estimation error. Fama and MacBeth (1973) sort stocks into beta-based portfolios, and they use the estimated betas for these portfolios, which contain less noise, as proxies for the betas of individual stocks. Vasicek (1973) generates a Bayesian estimate of beta that shrinks the empirical estimate toward 1. Blume (1975) suggests a practical approach that combines the historical beta estimate with a prior of 1 in a ratio of 2 to 1. These solutions reduce large biases resulting from more extreme beta estimates.

In this study, we focus on beta estimation using shrinkage techniques. We compare methods that shrink the beta in its entirety to methods that reduce estimation error in correlation and in volatility separately, and we examine which best forecasts beta. We then apply these insights to investment management. In particular, we focus on the construction of low-beta portfolios.

The main contribution of our study is that we empirically examine the effect of using different shrinkage parameters for correlation and for volatility when predicting equity betas. We then apply these insights to a specific application in order to construct low-risk portfolios, and we compare the ex post portfolio characteristics of these methods. Our empirical findings can be summarized as follows. First, shrinking correlations more than relative volatilities significantly improves beta forecasts compared with shrinking betas in their entirety. Second, investors can efficiently construct low-risk portfolios by combining regular unshrunk betas and volatilities, as this implicitly resembles portfolio construction based on shrunk beta estimates in which correlations are shrunk more than volatilities.

The remainder of the paper is structured as follows. In Section 2 we briefly discuss the data. In Section 3 we describe the methodology. Section 4 contains the empirical results for predicting betas. In Section 5 we discuss our application to low-risk strategies. Finally, Section 6 presents our conclusions.

2 DATA

The data used in this research consists of US stock data from the Center for Research in Security Prices (CRSP), covering the period from January 1, 1963 to December 31, 2017. The weekly data is based on an aggregate of the returns from Monday to Friday. Only shares that are traded on the New York Stock Exchange (NYSE), the NYSE American Exchange (formerly the AMEX) or the Nasdaq are included. Further, at each point in time, penny stocks and micro-cap stocks are excluded; penny stocks are defined as stocks with a price less than or equal to US$1, and, following Fama and French (2008), micro-cap stocks are defined as stocks (from any exchange) that fall below the value of the 20th percentile of the market capitalization of all NYSE-listed stocks. On average, for any point in the studied period, about 1500 large- and mid-cap stocks are included in the analysis. Figure A1 in the online appendix shows the number of included stocks over time. The Treasury bill rate is from the online data library of Kenneth French,[1] and it is used to calculate returns in excess of the risk-free rate.

3 METHODOLOGY

The market beta, which is the central measure of a security's expected return in the CAPM, is defined as the product of stock i's correlation with the market portfolio m and the ratio of its volatility to the volatility of the market:

$$\beta_{i,t} = \frac{\operatorname{cov}(R^{\mathrm{e}}_{i,t}, R^{\mathrm{e}}_{m,t})}{\operatorname{var}(R^{\mathrm{e}}_{m,t})} = \frac{\rho(R^{\mathrm{e}}_{i,t}, R^{\mathrm{e}}_{m,t})\sigma(R^{\mathrm{e}}_{i,t})\sigma(R^{\mathrm{e}}_{m,t})}{\sigma^2(R^{\mathrm{e}}_{m,t})} = \rho(R^{\mathrm{e}}_{i,t}, R^{\mathrm{e}}_{m,t})\frac{\sigma(R^{\mathrm{e}}_{i,t})}{\sigma(R^{\mathrm{e}}_{m,t})}. \tag{3.1}$$

We assume that the correlations and the volatility ratios are independent, such that the expectation of beta is equal to the product of the expectations of the two components, and for convenience we ignore Jensen's inequality in the volatility ratio term.

Our approach to forecasting betas makes use of returns only.[2] Estimating beta is commonly done using a rolling window with a certain lookback period and data frequency. The decomposition in (3.1) allows us to estimate each of the three components separately to obtain an estimator for a security's market beta.

Since the stocks in our data set form an unbalanced panel, we require that at least 50% of a stock's return information should be available within the estimation window, otherwise we set the beta estimate to represent a missing value and do not use the stock in portfolio construction for that particular point in time.

The obtained estimates of beta, correlation and volatility may contain a substantial estimation error, which can be reduced by implementing shrinkage, although it is not

[1] URL: https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

[2] Theoretical models such as those in Ehling and Heyerdahl-Larsen (2017) link time-varying stock correlations across good and bad states to preference heterogeneity and aggregate risk aversion, which we ignore here. We also do not include more sophisticated methods or other forms of security information to estimate a security's beta, such as those found in Cosemans *et al* (2016).

a priori clear how much these estimates should be shrunk and to what target value.[3] However, instead of shrinking the entire beta estimate to 1 as in Blume (1975), we propose to shrink the correlation and the relative volatility separately to their respective cross-sectional means, since both components may include a different degree of estimation error:

$$R^{\text{shrink}}_{i,m,t} = (1-c)R_{i,m,t} + cR^{\text{target}}_{i,m,t}, \tag{3.2}$$

$$V^{\text{shrink}}_{i,m,t} = (1-v)V_{i,m,t} + vV^{\text{target}}_{i,m,t}, \tag{3.3}$$

where $R^{\text{shrink}}_{i,m,t}$ is the shrunk correlation between stock i's returns and the market's returns based on the correlation estimate $R_{i,m,t}$. Further, $V^{\text{shrink}}_{i,m,t}$ is the shrunk volatility ratio of the stock i's returns relative to the market's returns, with $V_{i,m,t}$ being its estimate. The target values of the correlation component, $R^{\text{target}}_{i,m,t}$, and the relative volatility component, $V^{\text{target}}_{i,m,t}$, are set equal to the cross-sectional mean over all securities at time t. It is important to note here that, in the case of correlations being shrunk more than relative volatilities, this implicitly means that the spread in stock correlations is reduced more than the spread in volatilities. Such a difference in shrinkage parameters implies that the estimated correlations become less important than estimated volatilities.

4 PREDICTING BETAS

A security's market beta is a characteristic that cannot be observed directly and thus has to be estimated. There are a range of ways of estimating these betas. To ensure our results are not dependent on one particular choice, we use a variety of lengths of the out-of-sample period, which we label the forward period to distinguish it from the lookback period (ie, the historical sample period used for the prediction of the betas), and we use a variety of data frequencies ranging from daily to monthly. We start by examining which rolling-window lookback periods and data frequencies predict these differently estimated future betas best.

Table 1 shows the mean squared error (MSE) between the rolling-window estimates and their subsequent realized values for a range of forward periods, lookback periods and data frequencies.[4] Part (a) contains the MSE values for the estimates of correlation, part (b) shows the relative volatility and part (c) gives the beta. The minimum values per forward period and frequency (ie, per column) are highlighted

[3] Welch (2019) proposes an alternative method to calculate robust security betas by capping outliers in individual stock returns.

[4] In the main text we use the MSE, which penalizes large differences more strongly than other error measures, but we show in Tables A1 and A2 in the online appendix that the mean absolute deviation would lead to qualitatively similar results.

TABLE 1 Predicting correlation, relative volatility and beta. [Table continues on next two pages.]

(a) MSE of correlation estimates

Look-back period	Look-back freq.	Forward period and frequency										
		1M D	**6M D**	**1Y D**	**3Y D**	**5Y D**	**1Y W**	**3Y W**	**5Y W**	**1Y M**	**3Y M**	**5Y M**
1M	**D**	0.0820	0.0545	0.0524	0.0539	0.0558	0.0740	0.0697	0.0699	0.1407	0.1072	0.1013
6M	**D**	0.0560	0.0247	0.0213	0.0219	0.0225	0.0388	0.0333	0.0332	0.1027	0.0677	0.0616
1Y	**D**	**0.0547**	**0.0224**	**0.0192**	0.0190	0.0195	0.0356	0.0293	0.0290	0.0991	0.0628	0.0564
3Y	**D**	0.0562	0.0231	0.0194	**0.0174**	**0.0175**	0.0342	0.0257	0.0251	0.0955	0.0570	0.0497
5Y	**D**	0.0572	0.0239	0.0202	0.0178	0.0184	0.0347	0.0260	0.0256	0.0950	0.0558	0.0489
1Y	**W**	0.0732	0.0382	0.0341	0.0324	0.0320	0.0414	0.0334	0.0322	0.0941	0.0560	0.0489
3Y	**W**	0.0683	0.0325	0.0281	0.0241	0.0239	0.0335	0.0228	0.0220	0.0844	0.0431	0.0361
5Y	**W**	0.0674	0.0314	0.0269	0.0232	0.0232	**0.0322**	**0.0219**	**0.0211**	**0.0820**	**0.0411**	**0.0341**
1Y	**M**	0.1395	0.1018	0.0974	0.0967	0.0969	0.0939	0.0861	0.0856	0.1324	0.0942	0.0875
3Y	**M**	0.1084	0.0700	0.0654	0.0610	0.0613	0.0586	0.0469	0.0468	0.0949	0.0519	0.0456
5Y	**M**	0.1028	0.0640	0.0591	0.0550	0.0557	0.0523	0.0412	0.0411	0.0876	0.0450	0.0389

TABLE 1 Continued.

(b) MSE of relative volatility estimates

Look-back period	Look-back freq.	Forward period and frequency										
		1M D	**6M D**	**1Y D**	**3Y D**	**5Y D**	**1Y W**	**3Y W**	**5Y W**	**1Y M**	**3Y M**	**5Y M**
1M	**D**	3.8839	2.5870	2.4854	2.6733	2.8259	3.2655	3.4107	3.5528	4.2146	4.1976	4.2800
6M	**D**	2.8172	1.1733	0.9959	1.0312	1.1031	1.3994	1.4350	1.5389	2.1362	2.0210	2.0724
1Y	**D**	**2.7354**	**1.0620**	**0.8515**	0.8424	0.9013	1.1434	1.1557	1.2505	1.8621	1.6856	1.7272
3Y	**D**	2.8790	1.0928	0.8598	0.7450	0.7982	1.0069	0.9456	1.0319	1.6175	1.3965	1.4134
5Y	**D**	2.8742	1.1284	0.9002	0.7743	0.7983	1.0208	0.9554	1.0115	1.6022	1.3701	1.3664
1Y	**W**	3.3964	1.3833	1.0698	0.8592	0.8260	0.8951	0.7726	0.7916	1.2995	1.0131	0.9975
3Y	**W**	3.4565	1.3584	1.0343	0.7509	0.7253	0.7548	0.5759	0.6049	1.0689	0.7557	0.7398
5Y	**W**	3.3723	1.3202	1.0053	**0.7406**	**0.6928**	**0.7222**	**0.5640**	**0.5705**	1.0264	0.7269	0.6984
1Y	**M**	4.3927	2.1324	1.7491	1.4406	1.3730	1.2624	1.0715	1.0716	1.4301	1.0809	1.0554
3Y	**M**	4.2907	1.9667	1.5725	1.1998	1.1328	0.9843	0.7366	0.7370	1.0513	0.6885	0.6486
5Y	**M**	4.1615	1.8934	1.5123	1.1530	1.0513	0.9208	0.6875	0.6538	**0.9900**	**0.6198**	**0.5606**

TABLE 1 Continued.

(c) MSE of beta estimates

Look-back period	Look-back freq.	Forward period and frequency										
		1M D	**6M D**	**1Y D**	**3Y D**	**5Y D**	**1Y W**	**3Y W**	**5Y W**	**1Y M**	**3Y M**	**5Y M**
1M	**D**	1.2180	0.7071	0.6691	0.6642	0.6846	0.7959	0.7347	0.7383	1.1541	0.8469	0.8180
6M	**D**	0.7421	0.2116	0.1771	0.1725	0.1820	0.2759	0.2196	0.2221	0.6229	0.3259	0.2955
1Y	**D**	0.7162	0.1847	0.1483	0.1360	0.1425	0.2409	0.1778	0.1779	0.5783	0.2805	0.2489
3Y	**D**	0.7046	**0.1823**	**0.1402**	**0.1138**	0.1152	0.2193	0.1463	0.1431	0.5415	0.2414	0.2093
5Y	**D**	**0.6983**	0.1882	0.1451	0.1139	**0.1133**	0.2174	0.1429	0.1388	0.5261	0.2324	0.2014
1Y	**W**	0.8118	0.2687	0.2276	0.2070	0.2068	0.2953	0.2264	0.2214	0.6158	0.3138	0.2794
3Y	**W**	0.7404	0.2104	0.1663	0.1348	0.1326	0.2245	0.1475	0.1419	0.5311	0.2294	0.1965
5Y	**W**	0.7239	0.2062	0.1614	0.1265	0.1220	**0.2148**	**0.1374**	**0.1305**	**0.5099**	**0.2151**	**0.1829**
1Y	**M**	1.2155	0.6486	0.5875	0.5408	0.5241	0.6384	0.5412	0.5228	0.9397	0.6074	0.5587
3Y	**M**	0.8570	0.3216	0.2747	0.2376	0.2338	0.3163	0.2337	0.2283	0.6023	0.2947	0.2638
5Y	**M**	0.8055	0.2836	0.2376	0.2021	0.1980	0.2771	0.1986	0.1932	0.5518	0.2578	0.2286

Forward and lookback periods range from one month (1M) to five years (5Y) and use daily (D), weekly (W) or monthly (M) data frequencies. Numbers in bold have the lowest MSE for a particular combination of forward period and frequency.

in bold. We see that, in most cases, the data frequency used for the forward period is also the optimal frequency for the lookback period, but a weekly lookback period frequency is also sometimes optimal when forward period frequencies are daily or monthly. In most cases, the optimal lookback period is at least as long as the forward period. A five-year lookback period with a weekly frequency gives the lowest MSE in most cases, especially when we are interested in predicting one- to three-year weekly and monthly betas. This is consistent with Gilbert *et al* (2014), who show that daily beta estimates are not necessarily better risk measures.

Parts (a) and (b) of Table 1 show that this longer-term lookback period is the most efficient choice for both components of beta: correlations and volatilities. Intuitively, it could be argued that correlations move more slowly than relative volatilities. However, we find no empirical support for differentiating the lookback periods for correlations and volatilities.[5] For ease of presentation, in the remainder of the paper we use a lookback period of five years with weekly data for both correlations and volatilities. This setting is robust and can predict betas on a one- to five-year basis. Consequently, for the evaluation of betas we use a three-year forward horizon with a weekly frequency. This horizon connects well with our application, as Van Vliet (2018) finds evidence that the optimal holding time for a stock in a low-risk portfolio when transaction costs are taken into account is close to three years.[6]

Table 2 shows the effect of shrinkage on the prediction error. Part (a) contains two-parameter shrinkage results, with volatility shrinkage varying between columns and correlation shrinkage varying between rows. Part (b) shows the prediction error when only one shrinkage factor is applied; no shrinkage is denoted by a zero, and the no-shrinkage value of 0.1374 was also given in Table 1(c). The optimal shrinkage factors for the sample estimates are $c = 0.5$ (correlation shrinkage) and $v = 0.2$ (volatility shrinkage). These optimal factors result in an MSE value of 0.1121.

Table 2(b) shows the results of four important statistical tests. The difference between the two-parameter shrinkage estimate and the no-shrinkage estimate is statistically significant according to the Diebold and Mariano (1995) test statistic (DM = 5.92).[7] The MSE of the optimal two-parameter shrinkage estimate is statistically significantly lower than the estimate for one-parameter shrinkage of either 0.2 or 0.5, as shown by the shaded values on the diagonal (DM = 3.68 and DM = 2.17, respectively). Unlike the optimal lookback period, which is similar for correlations

[5] Frazzini and Pedersen (2014), for example, use a one-year lookback period for volatility and a five-year lookback period for correlation.

[6] With a longer holding period, we also circumvent possible “beta bubbles”, ie, mean reversion in betas, as documented by Jylhä *et al* (2018). For more on estimating high-frequency betas, see Hansen *et al* (2014).

[7] We use the Newey–West estimator with a 36-month lag to deal with the overlapping nature of our estimates.

TABLE 2 Comparison of two-parameter and one-parameter shrinkage.

(a) MSE of rolling-window beta estimates with two shrinkage factors

c	v: 0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0	0.1374	0.1327	0.1304	0.1304	0.1327	0.1374	0.1443	0.1536	0.1652	0.1791	0.1954
0.1	0.1304	0.1259	0.1237	0.1239	0.1263	0.1311	0.1382	0.1475	0.1592	0.1733	0.1896
0.2	0.1250	0.1207	0.1186	0.1188	0.1214	0.1262	0.1334	0.1429	0.1547	0.1688	0.1852
0.3	0.1212	0.1169	0.1149	0.1152	0.1178	0.1228	0.1300	0.1396	0.1515	0.1657	0.1822
0.4	0.1190	0.1147	0.1127	0.1131	0.1158	0.1208	0.1281	0.1377	0.1497	0.1640	0.1806
0.5	0.1184	0.1141	**0.1121**	0.1124	0.1151	0.1202	0.1275	0.1373	0.1493	0.1637	0.1805
0.6	0.1193	0.1149	0.1129	0.1132	0.1159	0.1210	0.1284	0.1382	0.1503	0.1648	0.1817
0.7	0.1218	0.1173	0.1152	0.1155	0.1182	0.1232	0.1307	0.1405	0.1527	0.1673	0.1843
0.8	0.1259	0.1213	0.1191	0.1193	0.1219	0.1269	0.1344	0.1442	0.1565	0.1713	0.1884
0.9	0.1316	0.1268	0.1244	0.1245	0.1270	0.1320	0.1395	0.1494	0.1617	0.1766	0.1939
1.0	0.1388	0.1338	0.1312	0.1312	0.1336	0.1385	0.1460	0.1559	0.1684	0.1833	0.2007

(b) Statistical tests

	DM	***p*-value**
No shrinkage versus optimal two-parameter shrinkage	5.92	0.000
Optimal two-parameter versus symmetric 0.2/0.2 parameter shrinkage	3.68	0.000
Optimal two-parameter versus symmetric 0.5/0.5 parameter shrinkage	2.17	0.015
Optimal two-parameter versus optimal one-parameter shrinkage	1.63	0.052

The lookback period is five years with weekly data. The forward period is three years with weekly data. Part (a) shows the two-parameter shrinkage, with correlation varying between rows and volatility varying between columns (the diagonal is shaded), the minimum MSE is given in bold and no shrinkage is denoted by a zero. Part (b) shows the results of the Diebold and Mariano (1995) test for differences in prediction accuracy.

and relative volatilities, optimal shrinkage is different for these two components of the beta. The difference between optimal one- and two-parameter shrinkage is statistically significant at the 10% level according to the Diebold–Mariano test statistic (DM $=$ 1.63). Table 2(a) gives an optimal shrinkage factor of 0.3 for the case of a single shrinkage factor for correlation and volatility alike, similar to the one-third that Blume (1975) suggested.[8]

To further test if two-parameter shrinkage improves beta estimates, we perform an additional analysis that zooms in on stocks for which the beta estimates differ most between the two methods. Beta estimates that are similar for both methods are not particularly relevant when comparing which model leads to better forecasting performance. We therefore rank stocks on the difference in the beta estimates between the two methods and form 10 portfolios based on this ranking. To examine whether there is an asymmetric forecasting performance relationship, we consider the sign of the differences when ranking, rather than using their absolute value. We then compare the average ex ante beta estimates with the realized betas over the full sample period. As we do not differentiate forecasting performance between under- or overestimation, the average squared forecast error is used to gauge the empirical performance.

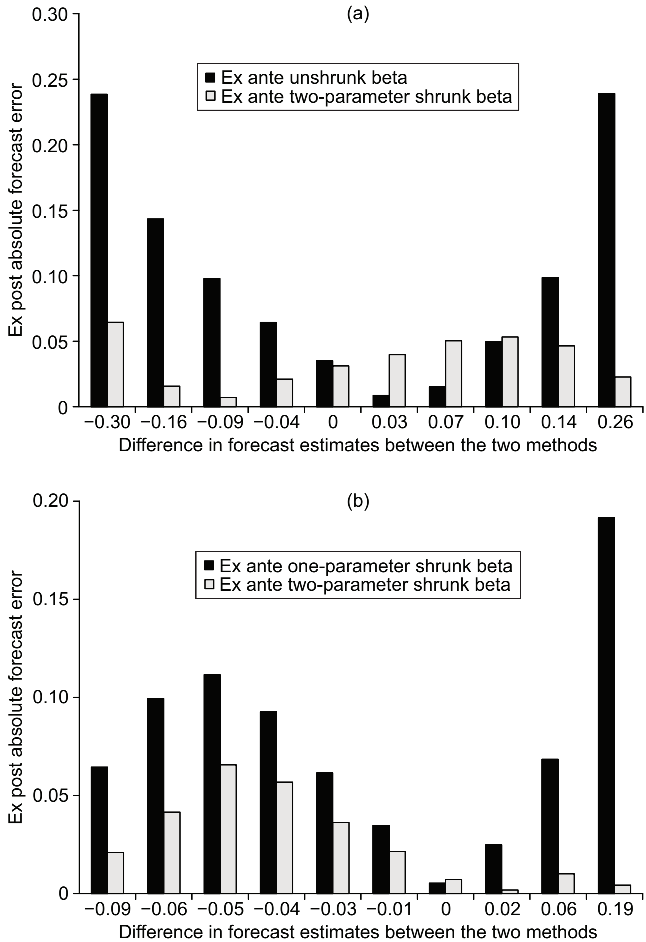
Figure 1 displays these forecasting errors. Part (a) compares the rolling-window estimator without shrinkage against the estimator using the two-parameter shrinkage that we propose in this paper, while part (b) compares the one- and two-parameter shrinkage methods against each other. The horizontal axis shows the differences in average beta estimates for each of the 10 decile portfolios based on these differences. Both figure parts clearly show that the forecasting errors are lower for the two-parameter shrinkage method.

It also becomes clear that the two-parameter shrinkage method is much better in the cases where the choice of method matters most. For example, Figure 1(a) shows that for the 10th-decile (D10) portfolio (for which the ex ante difference in beta estimates is 0.26) the ex post difference from the realized beta is 0.24 for the rolling-window estimator without shrinkage and 0.02 when the two-parameter shrinkage method is used.

In Figure 1(b) the ex ante difference for the D10 portfolio is 0.19, and the one-parameter shrinkage method has a forecast error of 0.19, while this error is only 0.01 for the two-parameter shrinkage method. For portfolios where the methods show

[8] This is also close to the optimal value reported by Frazzini and Pedersen (2014), who propose shrinking beta (with a mix of lookback periods for volatility and correlation) to the cross-sectional mean (of 1) with a shrinkage factor of 0.4. For this sample we also replicate the Frazzini and Pedersen (2014) parameter settings: a one-year lookback period for volatility, a five-year lookback period for correlation and a shrinkage factor for beta of 0.4. This results in a mean absolute deviation value of 0.1241.

FIGURE 1 Comparing prediction errors where it matters most.



(a) Portfolios sorted on the difference between no shrinkage and two-parameter shrinkage. (b) Portfolios sorted on the difference between one- and two-parameter shrinkage. The figure shows the ex post absolute forecast error of 10 portfolios, each containing 10% of stocks ranked on differences in estimated betas, with D1 containing stocks for which the unshrunk betas (part (a)) or one-parameter shrunk betas (part (b)) most exceed the two-parameter shrunk betas. The vertical axis shows the absolute value of the difference between the average ex ante estimate and the realization over the entire sample period. The numbers on the horizontal axis are the differences between ex ante estimates for each of the 10 portfolios.

little ex ante difference, the forecast errors are low, and forecasting betas for those cases does not seem to be that challenging to begin with.

5 PRACTICAL APPLICATION TO LOW-RISK PORTFOLIOS

We could stop our analysis after the previous section, as we have established that betas can be more accurately estimated by two-parameter shrinkage than one-parameter shrinkage. However, we also want to show how practitioners may use these superior beta estimates in portfolio management of low-risk portfolios. Empirically, the CAPM beta is unable to explain returns, since the risk–return relationship is too flat or even negative.[9] This has led academics and investors to focus on using beta to construct low-risk portfolios. In our application, we use the portfolio ranking methodology: at the end of each month we separate all stocks into 10 portfolios based on their betas, and we calculate the 10 equally weighted portfolio returns for the next month.[10] Note that, when this sorting algorithm is based on unshrunk betas or on betas that are shrunk to 1 with the same shrinkage factor, it leads to exactly the same portfolios as it would without shrinkage, as the order of betas will not be affected by such transformations. However, when we use different shrinkage factors for the correlation and the volatility, the order of the betas may change, and hence the portfolios are likely contain different stocks than portfolios based on unshrunk betas. In this section, we aim to improve low-risk portfolio management by applying the insights from the previous sections.

5.1 Comparing beta-sorted portfolios

Black *et al* (1972) show that low-beta stocks tend to have a positive CAPM alpha. As shown above, the two-parameter shrunk beta estimate is a better prediction of the forward beta than an estimate with no shrinkage or with one-parameter shrinkage. We now turn to asking whether the improved ex ante beta estimates at the individual stock level also lead to lower ex post portfolio betas.

Parts (a) and (b) of Table 3 show the statistics for decile portfolios constructed by ranking stocks on their historical beta estimates and on their two-parameter shrinkage beta estimates, respectively. In addition to the decile portfolios, we also include a column with a long position in the D1 portfolio and a short position in the D10 portfolio, which combines the positive alpha of D1 with the negative alpha of D10.

Table 3(a) shows that the decile with the lowest historical volatility, D1, has an average excess return of 7.26%, a volatility of 11.04%, a Sharpe ratio of 0.66 and a

---

[9] See Black *et al* (1972), Haugen and Baker (1991), Clarke *et al* (2006), Blitz and Van Vliet (2007), Baker *et al* (2011), Chow *et al* (2014), Frazzini and Pedersen (2014) and Blitz *et al* (2020). Blitz *et al* (2014) review a wide range of explanations for this low-risk effect.

[10] Soe (2012) finds that minimum-variance portfolio optimization leads to a similar degree of risk reduction, even though the ranking approach typically ignores correlations between stock returns. For more details on estimating high-frequency covariance matrixes that can accommodate trade asynchronicities, see, for example, Boudt *et al* (2017). Callot *et al* (2017) use lasso-type estimators to reduce the dimensionality of estimating large covariance matrixes.

TABLE 3 Comparing the characteristics of beta-ranked portfolios.

(a) Portfolios based on historical rolling-window betas

	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D1–D10
Excess return (%)	7.26	9.00	9.89	9.97	10.17	9.68	9.53	9.12	8.53	7.26	0.00
Standard dev. (%)	11.04	12.70	14.30	15.49	16.70	18.17	19.70	22.03	25.55	32.59	28.05
Sharpe ratio	0.66	0.71	0.69	0.64	0.61	0.53	0.48	0.41	0.33	0.22	0.00
CAPM alpha (%)	3.41	3.81	3.74	3.20	2.78	1.56	0.61	−0.90	−2.89	−6.74	10.14
t-statistic	3.45	4.73	5.26	4.63	4.24	2.48	1.14	−1.68	−3.34	−4.15	4.29
CAPM beta	0.45	0.61	0.72	0.80	0.87	0.96	1.05	1.18	1.34	1.65	−1.19

(b) Portfolios based on two-parameter shrunk betas

	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D1–D10
Excess return (%)	7.48	9.03	9.58	9.72	10.37	9.68	9.69	9.95	8.71	6.22	1.26
Standard dev. (%)	10.79	12.64	14.18	15.65	16.71	18.17	19.95	22.20	26.00	33.18	29.30
Sharpe ratio	0.69	0.71	0.68	0.62	0.62	0.53	0.49	0.45	0.33	0.19	0.04
CAPM alpha (%)	3.89	3.93	3.55	2.92	3.03	1.57	0.65	−0.13	−2.87	−7.98	11.87
t-statistic	3.81	4.69	4.70	4.00	4.26	2.44	1.22	−0.24	−3.08	−4.70	4.82
CAPM beta	0.42	0.60	0.71	0.80	0.86	0.95	1.06	1.19	1.36	1.67	−1.25

Ex post portfolio characteristics of 10 portfolios each containing 10% of stocks ranked on their estimated historical beta, with D1 having the lowest betas and D10 the highest. The excess returns are relative to the risk-free rate. D1–D10 is a long position in portfolio D1 and a short position in D10. Part (a) contains the statistics for sorting on historical rolling-window betas (with a lookback of five years with weekly returns). In part (b) we apply fixed two-parameter shrinkage ($v = 0.2$ and $c = 0.5$).

market beta of 0.45. Note that the beta here is the ex post beta of the decile portfolios, and not the ex ante beta on which the sorting is based. The decile with the highest historical volatility, D10, also has an average return of 7.26%, but has a volatility of 32.59%, a Sharpe ratio of 0.22 and a market beta of 1.65. The D1–D10 portfolio gives an alpha of 10.14% relative to the CAPM. This is the low-beta effect originally documented by Black *et al* (1972).

Table 3(b) uses the two-parameter shrinkage method (ie, shrinking correlation and relative volatility separately), which, as we have shown above, predicts future betas better than methods without shrinkage or shrinking beta in its entirety. Note that one-parameter shrinkage for beta (with the same shrinkage parameter for each stock) does not change the ordering of betas, and hence the results of the portfolio sorting would be the same as they are for the beta estimates without shrinkage, presented in Table 3(a). The shrinkage parameters are set equal to $c = 0.5$ and $v = 0.2$, as this is the combination that can best forecast betas.

We see that this two-parameter shrinkage leads to the D1 portfolio being less risky and to the D10 portfolio being riskier in terms of both volatility and beta. The volatility of the D1 portfolio declines from 11.04% to 10.79%, and its beta from 0.45 to 0.42, while for D10 its volatility increases from 32.59% to 33.18%, and its beta from 1.65 to 1.67. However, these differences are not statistically significant. The alpha in the D1–D10 portfolio increases from 10.14% to 11.87%. This increase is statistically significant, with a p-value of 0.00, and the increase of the Sharpe ratio from 0.55 to 0.60 is also statistically significant, with a p-value of 0.01 based on the Jobson and Korkie (1983) test with Memmel (2003) correction.

5.2 Comparing low-beta and low-volatility portfolios

Blitz and Van Vliet (2007) document a positive alpha for both low-volatility and low-beta portfolios, with alphas for low-volatility generally being higher. In Section 4 we found that beta predictions are more accurate when correlations receive greater shrinkage to their cross-sectional average than relative volatilities do. One interpretation of our result is that estimated correlations receive less weight in the portfolio construction than estimated relative volatilities do, because individual estimated correlations are shrunk more to their cross-sectional mean than are relative volatilities. We now take a different approach and form three sets of portfolios: portfolios based on two-parameter shrunk betas; portfolios based on volatility only; and portfolios based on a combination of these two variables. Because the beta and volatility estimates have a different mean and variance, we standardize them to obtain a joint ranking:

$$Z_{ms,t}^{\text{combined}} = (1 - q)Z(\beta_t) + qZ(\sigma_t), \tag{5.1}$$

TABLE 4 The impact of including correlations on volatility.

Forward period	Forward freq.	q: 0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	p-value ($q = 0$)	p-value ($q = 1$)
1M	D	10.84	10.79	10.69	10.68	10.64	10.64	10.62	**10.62**	10.65	10.68	10.74	0.178	0.178
6M	D	10.70	10.70	10.66	10.62	10.63	10.63	**10.62**	10.63	10.64	10.69	10.74	0.538	0.212
1Y	D	10.67	10.63	10.63	10.62	10.63	10.62	**10.61**	10.65	10.68	10.71	10.74	0.608	0.125
3Y	D	10.94	10.86	10.86	**10.81**	10.82	10.83	10.84	10.84	10.91	11.00	11.06	0.128	0.155
5Y	D	11.33	11.26	11.18	11.09	10.99	10.96	10.92	10.86	**10.83**	10.89	10.98	**0.094**	**0.034**
1Y	W	10.79	**10.75**	10.79	10.79	10.80	10.77	10.77	10.82	10.83	10.89	10.98	0.158	**0.097**
3Y	W	10.79	**10.75**	10.79	10.79	10.80	10.77	10.77	10.82	10.83	10.89	10.98	0.158	**0.097**
5Y	W	10.86	10.80	10.78	10.78	10.79	10.80	**10.78**	10.79	10.81	10.88	10.98	0.595	**0.001**
1Y	M	10.87	10.86	10.85	10.86	10.86	**10.83**	10.85	10.91	10.89	10.97	11.01	0.643	**0.038**
3Y	M	10.85	10.84	10.85	10.87	**10.83**	10.85	10.87	10.88	10.92	10.96	11.01	0.822	**0.040**
5Y	M	10.86	10.85	10.84	10.85	10.86	**10.84**	10.87	10.91	10.91	10.97	11.01	0.826	0.051

Annualized volatility of the D1 portfolio for different forward and frequency combinations using the optimal lookback periods and frequencies displayed in Table 1. For $q = 0$ we have the standard two-parameter shrunk beta portfolio, while for $q = 1$ we have a ranking based purely on stocks' volatilities. The lowest volatility per row is highlighted in bold. The last two columns contain the p-values of the statistical test of whether the minimum volatility of the row is significantly below the leftmost column ($q = 0$) or the rightmost column ($q = 1$). A p-value below 0.1 is marked in bold.

where β_t are the (shrunk) betas, σ_t are the volatility estimates and $Z(\cdot)$ denotes cross-sectional robust z-scores.[11] For $q = 0$ the beta estimates are the same as in the previous subsection, whereas for $q = 1$ we have a pure volatility strategy in which correlations are completely ignored. This analysis gives insight into the effect of including correlations in addition to volatilities, and it therefore connects the literature on low-beta and low-volatility investing.

Table 4 shows the volatilities for the D1 portfolio for a varying parameter q, based on the shrunk beta and a pure volatility estimate. For the forward period of three years with a weekly frequency (with the optimal lookback period of five years with a weekly frequency), the volatility of 10.79% in Table 4 corresponds to the volatility of the D1 portfolio reported in Table 3(b). Note that the forward period and forward frequency vary between rows, and this may lead to different rows sharing the same historical estimation period and frequency (see Table 1), resulting in duplicate rows, given that the shrinkage parameters remain the same. Table 4 shows that most of the lowest D1 volatilities are obtained for a parameter q ranging between 0.3 and 0.7.

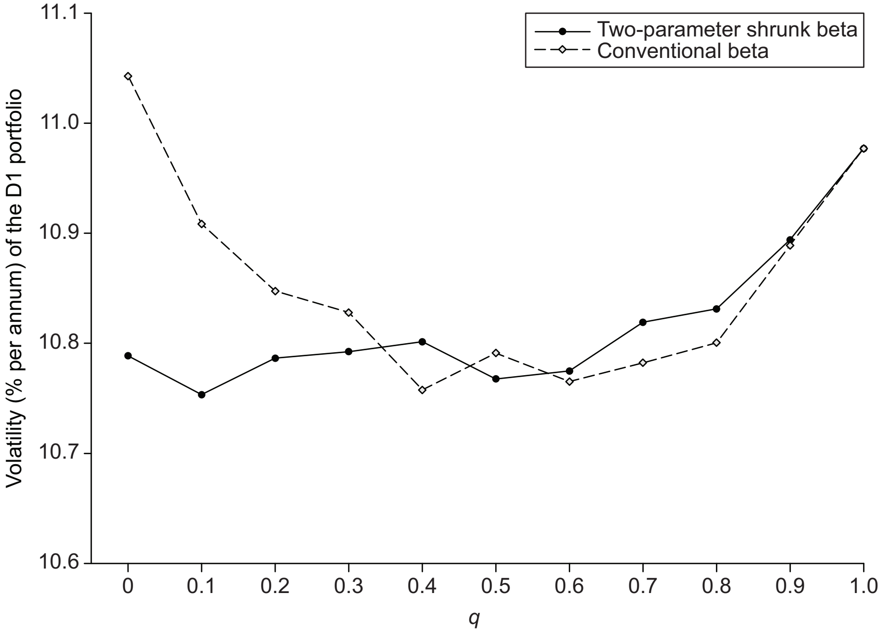
For the weekly and monthly forward frequencies, the optimal volatilities are significantly lower than those for the border case $q = 1$ at a 10% confidence level, which indicates that including correlations significantly reduces the volatility of a low-risk investment portfolio. However, in all cases except one, the optimal volatilities are not significantly lower than those for the $q = 0$ border case. This suggests that adding more volatility does not help to significantly reduce the volatility of a low-beta strategy, provided that the correlations are shrunk more than the volatilities when estimating a security's beta.

The above results are based on combining the shrunk beta estimate with volatility according to a varying q parameter. Figure 2 shows the results from Tables 4 and A3 for the rows corresponding to the three-year weekly forward period, for varying q and using either betas with two-parameter shrinkage (Table 4) or conventionally estimated betas (Table A3).[12] It shows that for the two-parameter shrunk beta the volatility is below 10.8% per annum for $q = 0$ to $q = 0.6$, with the lowest volatility at $q = 0.1$, as shown in Table 4. For the conventional beta the range is narrower and starts later ($q = 0.4$–0.7). In practical terms, this means that a half–half combination of a conventional beta and pure volatility leads to the lowest portfolio volatility, just

[11] Compared with regular z-scores, the z-scores we use here are based on replacing the (cross-sectional) average with the median and the (cross-sectional) standard deviation with a constant (1.483) times the median absolute deviation. The normalized values are capped at -3 and 3 to further reduce the impact of outliers. See Rousseeuw and Croux (1993) for more on robust z-scores.

[12] To examine whether this combined strategy benefits from the shrunk beta introduced in this paper, Table A3 in the online appendix includes the D1 portfolio statistics for the combined strategy (as in Table 4) but based on a regular beta estimate that is not shrunk.

FIGURE 2 Comparing beta- and volatility-ranked portfolios.

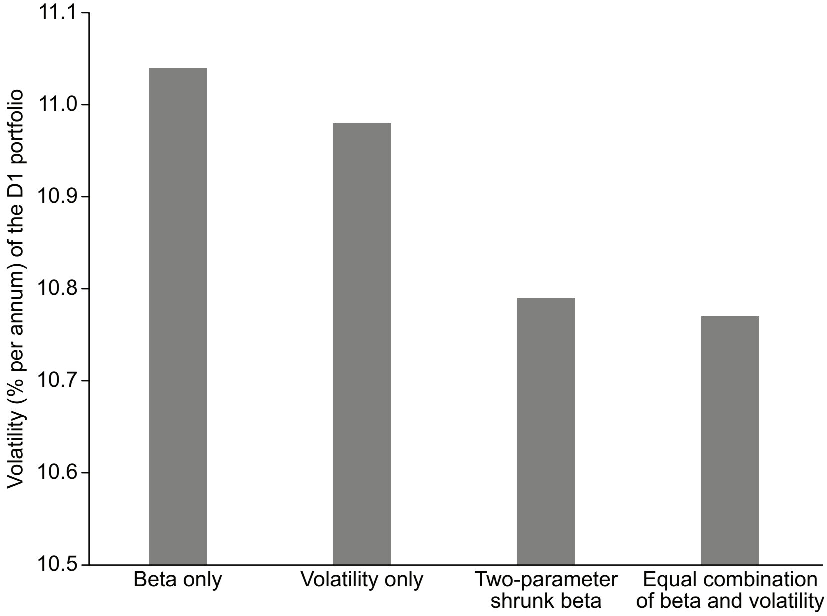


Annualized volatility of the D1 portfolio for a three-year weekly forward period and frequency with an optimal five-year weekly lookback period and frequency. For $q = 0$ the ranking is based on the (shrunk or unshrunk) beta, while for $q = 1$ the ranking is based purely on a stock's volatility.

below 10.8% per annum, indicating that correlation is indeed an important element for creating a low-volatility portfolio. However, sorting stocks into portfolios based on their conventional betas gives too much weight to the correlation, which is less accurately forecast than the volatility. Reducing the weight of the correlation can be achieved in two different ways that lead to similar results: shrinking correlations more to their cross-sectional average compared with volatilities, or combining the conventional unshrunk beta with a pure volatility estimate. Both approaches lead to similar portfolio volatilities, just below 10.8% in our specific application.

This result is illustrated in Figure 3, in which the first two columns are the portfolio volatilities when the ranking is based on only conventional beta and when it is based on only volatility; both are close to 11.0% per annum. The last two columns are the portfolio volatilities when ranking is based on two-parameter shrunk betas and when it is based on an equally weighted combination of the z-scores of unshrunk betas and volatilities, leading to similar outcomes of just below 10.8% per annum.

FIGURE 3 Comparing portfolio volatilities of different methods.



6 CONCLUSION

We challenged the common approach of simply shrinking security betas to their sample mean of 1. We disentangled correlations from relative volatilities and allowed different shrinkage parameters for each in order to better predict betas. Our empirical results confirmed that beta prediction errors are significantly lower when correlations are shrunk more than relative volatilities, especially in cases where the betas are difficult to forecast.

In our application to low-risk portfolios, we found that portfolios ranked on betas for which correlations have been shrunk more than relative volatilities have lower ex post betas, albeit not to a statistically significant extent. In the debate over whether low-beta or low-volatility stocks result in portfolios with lower risk, we found that an equally weighted combination performs best. This combination illustrates how our results on two-parameter beta shrinkage should be interpreted: using correlation helps for low-risk portfolio management, but since it is less accurately estimated than volatilities, it has to be handled with care.

DECLARATION OF INTEREST

All the authors are affiliated with Robeco Institutional Asset Management, an investment company that may use the insights from this paper for managing investment portfolios for their clients. The views expressed in this paper are not necessarily shared by Robeco Institutional Asset Management. The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

ACKNOWLEDGEMENTS

We thank Dixie de Klerk for excellent research assistance and Mathijs Cosemans, Andrea Frazzini, Winfried Hallerbach, Rutger-Jan Lange, Frederik Muskens and Karolina Scholtus for valuable comments and suggestions.

REFERENCES

Baker, M., Bradley, B., and Wurgler, J. (2011). Benchmarks as limits to arbitrage: understanding the low-volatility anomaly. *Financial Analysts Journal* **67**(1), 40–54 (https://doi.org/10.2469/faj.v67.n1.4).

Black, F., Jensen, M. C., and Scholes, M. (1972). The capital asset pricing model: some empirical tests. In *Studies in the Theory of Capital Markets*, Jensen, M. C. (ed), pp. 79–121. Praeger, New York.

Blitz, D., and Van Vliet, P. (2007). The volatility effect: lower risk without lower return. *Journal of Portfolio Management* **34**(1), 102–113 (https://doi.org/10.3905/jpm.2007.698039).

Blitz, D., Falkenstein, E., and Van Vliet, P. (2014). Explanations for the volatility effect: an overview based on the CAPM assumptions. *Journal of Portfolio Management* **40**(3), 61–76 (https://doi.org/10.3905/jpm.2014.40.3.061).

Blitz, D., Van Vliet, P., and Baltussen, G. (2020). The volatility effect revisited. *Journal of Portfolio Management* **46**(2), 45–63 (https://doi.org/10.3905/jpm.2019.1.114).

Blume, M. (1975). Betas and their regression tendencies. *Journal of Finance* **30**(3), 785–795 (https://doi.org/10.1111/j.1540-6261.1975.tb01850.x).

Boudt, K., Laurent, S., Lunde, A., Quaedvlieg, R., and Sauri, O. (2017). Positive semidefinite integrated covariance estimation, factorizations and asynchronicity. *Journal of Econometrics* **196**(2), 347–367 (https://doi.org/10.1016/j.jeconom.2016.09.016).

Callot, L., Kock, A., and Medeiros, M. (2017). Modeling and forecasting large realized covariance matrices and portfolio choice. *Journal of Applied Econometrics* **32**(1), 140–158 (https://doi.org/10.1002/jae.2512).

Chow, T. M., Hsu, J. C., Kuo, L. L., and Li, F. (2014). A study of low-volatility portfolio construction methods. *Journal of Portfolio Management* **40**(4), 89–105 (https://doi.org/10.3905/jpm.2014.40.4.089).

Clarke, R., De Silva, H., and Thorley, S. (2006). Minimum-variance portfolios in the US equity market. *Journal of Portfolio Management* **33**(1), 1–13 (https://doi.org/10.3905/jpm.2006.661366).

Cosemans, M., Frehen, R., Schotman, P., and Bauer, R. (2016). Estimating security betas using prior information based on firm fundamentals. *Review of Financial Studies* **29**(4), 1072–1112 (https://doi.org/10.1093/rfs/hhv131).

Diebold, F., and Mariano, R. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics* **13**(3), 253–263 (https://doi.org/10.1080/07350015.1995.10524599).

Ehling, P., and Heyerdahl-Larsen, C. (2017). Correlations. *Management Science* **63**(6), 1657–2048 (https://doi.org/10.1287/mnsc.2015.2413).

Fama, E., and French, K. (2008). Dissecting anomalies. *Journal of Finance* **63**(4), 1653–1678 (https://doi.org/10.1111/j.1540-6261.2008.01371.x).

Fama, E., and MacBeth, J. (1973). Risk, return, and equilibrium: empirical tests. *Journal of Political Economy* **81**(3), 607–636 (https://doi.org/10.1086/260061).

Frazzini, A., and Pedersen, L. (2014). Betting against beta. *Journal of Financial Economics* **111**(1), 1–25 (https://doi.org/10.1016/j.jfineco.2013.10.005).

Gilbert, T., Hrdlicka, C., Kalodimos, J., and Siegel, S. (2014). Daily data is bad for beta: opacity and frequency-dependent betas. *Review of Asset Pricing Studies* **4**(1), 78–117 (https://doi.org/10.1093/rapstu/rau001).

Hansen, P., Lunde, A., and Voev, V. (2014). Realized beta GARCH: a multivariate GARCH model with realized measures of volatility. *Journal of Applied Econometrics* **29**(5), 774–799 (https://doi.org/10.1002/jae.2389).

Haugen, R., and Baker, N. (1991). The efficient market inefficiency of capitalization-weighted stock portfolios. *Journal of Portfolio Management* **17**(3), 35–40 (https://doi.org/10.3905/jpm.1991.409335).

Jobson, J., and Korkie, B. (1983). Statistical inference in two-parameter portfolio theory with multiple regression software. *Journal of Financial and Quantitative Analysis* **18**(2), 189–197 (https://doi.org/10.2307/2330917).

Jylhä, P., Suominen, M., and Tomunen, T. (2018). Beta bubbles. *Review of Asset Pricing Studies* **8**(1), 1–35 (https://doi.org/10.1093/rapstu/rax014).

Memmel, C. (2003). Performance hypothesis testing with the Sharpe ratio. *Finance Letters*, **1**, 21–23.

Rousseeuw, P., and Croux, C. (1993). Alternatives to the median absolute deviation. *Journal of the American Statistical Association* **88**(424), 1273–1283 (https://doi.org/10.1080/01621459.1993.10476408).

Sharpe, W. (1964). Capital asset prices: a theory of market equilibrium under conditions of risk. *Journal of Finance* **19**(3), 425–442 (https://doi.org/10.1111/j.1540-6261.1964.tb02865.x).

Soe, A. (2012). Low-volatility portfolio construction: ranking versus optimization. *Journal of Index Investing* **3**(3), 63–73 (https://doi.org/10.3905/jii.2012.3.3.063).

Van Vliet, P. (2018). Low-volatility needs little trading. *Journal of Portfolio Management* **44**(3), 33–42 (https://doi.org/10.3905/jpm.2018.44.3.033).

Vasicek, O. (1973). A note on using cross-sectional information in Bayesian estimation of security betas. *Journal of Finance* **28**(5), 1233–1239 (https://doi.org/10.1111/j.1540-6261.1973.tb01452.x).

Welch, I. (2019). Simpler better market betas. Working Paper 26015, National Bureau of Economic Research, Cambridge, MA (https://doi.org/10.3386/w26105).