Better Betas

Lisa R. Goldberg[*] Alex Papanicolaou[†] Alex Shkolnik[‡] Simge Ulucam[§]

September 13, 2019[*]

We provide a data-driven adjustment for estimated betas that leads to material improvements in the accuracy of weights and risk forecasts for minimum variance portfolios. Like the widely used Blume 2/3 rule and Vasicek correction developed in the 1970s, our beta adjustment operates by shrinking raw beta estimates toward one. Unlike its antecedents, our adjustment adapts dynamically to volatility regimes, which matters since betas have tended to become more dispersed in stressed markets. Our analysis underscores the negative relationship between beta concentration and diversification, forcing a re-examination of the rule of thumb that diversification disappears in a crisis.

[*]Departments of Economics, Statistics and Consortium for Data Analytics in Risk, University of California, Berkeley, CA. Aperio Group, Sausalito, CA (lrg@berkeley.edu).

[†]Intelligent Financial Machines, Palo Alto, CA and Consortium for Data Analytics in Risk, University of California, Berkeley, CA (apapanicolaou@berkeley.edu).

[‡]Department of Statistics and Applied Probability, University of California, Santa Barbara, CA and Consortium for Data Analytics in Risk, University of California, Berkeley, CA (shkolnik@ucsb.edu).

[§]Aperio Group, Sausalito, CA (sulucam@aperiogroup.com).

[*] Acknowledgement. We wish to thank Torben Andersen, Ravi Jagannathan, Alec Kercheval, Bob Korajczyk, Olivier Ledoit and Viktor Todorov and Matti von Turk for insightful dialogs about the material in this article. We also thank the participants at the 9th Western Conference in Mathematical Finance and seminar attendees at Aperio Group, the San Francisco CFA Society and Oxford University Frontiers in Quantitative Finance for the comments and questions that led to improvements of a previous version of this article.

Market betas play a central role in quantitative portfolio construction.[1] For example, the weights of a minimum variance portfolio are inversely related to market betas and the portfolio tends to be long low-beta stocks and short high-beta stocks. This general, theoretically grounded phenomenon is illustrated in Figure 1, which shows scatter plots of minimum variance portfolio weights against market betas. The portfolios are constructed from securities in the S&P 500 at the end of December, 2015, and we use two different models.[2] The first is a multi-factor model, and the second is a trimmed version that retains only the market betas and specific risk. The example shows that, absent interference by a portfolio manager, minimum variance weights tend to have extreme positions corresponding to extreme values of beta. The effect is dramatic in a one-factor model, and it persists even when the model has many more factors.[3]

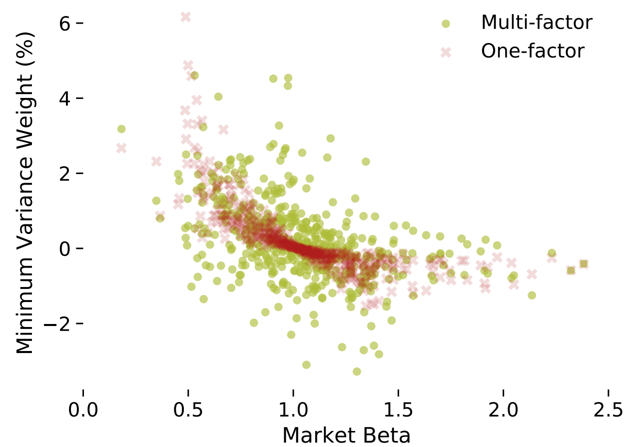

Figure 1. *Scatter plots of minimum variance portfolio weights versus betas in December 2015. The universe is the S&P 500 index and portfolios were constructed with the Barra Aegis optimizer using the 77-factor Barra USSLOW Model and a trimmed version that retains only the market factor and specific risk. Source: Aperio Group.*

Beta adjustments. The betas we use in practice are, of course, estimated from data. A series of influential articles beginning with Blume (1971), Levy (1971), Vasicek (1973) and Blume (1975) argues that time-series estimates of market betas are routinely overestimated for riskier securities and underestimated for securities that are less risky. Blume, Vasicek and their disciples remedied this shortcoming using an adjusted beta

$$\beta^{\text{adj}} = c\beta^{\text{raw}} + (1 - c), \tag{1}$$

where β^{raw} is an empirical (or, *raw*) estimate and the parameter c lies between zero and one. The Blume adjustment sets c to $2/3$ and the formula

$$\beta^{\text{Blume}} = \frac{2}{3}\beta^{\text{raw}} + \frac{1}{3} \tag{2}$$

is widely used in practice to adjust raw betas. It is noteworthy that the precise value of $2/3$ never appears in Blume's papers. Nevertheless, Blume's rule, as implemented by Bloomberg and others, uses the static value of $2/3$, and formula (2) is a staple of the CFA program curriculum. Blume's rule is fascinating in that it attempts to account for more than estimation error. It also adjusts for the fact that true betas tend to revert toward one for economic reasons (Blume 1975, Section IV).[4] In contrast, the proposal in Vasicek (1973), which is used by commercial models includ-

[1] Insightful articles on how market betas influence weights of optimized portfolios include Green & Hollifield (1992), Jagannathan & Ma (2003), Clarke, De Silva & Thorley (2006) and Clarke, De Silva & Thorley (2011).

[2] In practical situations, constraints are used to disallow concentrated positions and other undesirable features, such as sector overweights.

[3] The persistence of the relationship between betas and weights of minimum variance portfolios in the face of many factors can be explained by the fact that the market betas concentrate around one. In particular, the contribution of any factor to the minimum variance portfolio is precisely the projection of that factor onto the vector of all ones.

[4] Estimation error, a statistical artifact Blume referred to as "*order bias*", played a secondary role in Blume (1975). Indeed, it had to be removed to reveal the more significant "regression to the mean" tendency of the betas. See Klemkosky & Martin (1975), Blume (1979), Elgers, Haltiner & Hawthorne (1979), Eubank & Zumwalt (1979) and Mantripragada (1980) for a further discussion.

ing Barra, is focused solely on estimation error. In Vasicek's adjustment, the parameter c of formula (1) relies on a Bayesian prior for the true market betas.[5]

Dispersion shrinkage. Adjusted betas in statistics go by the name of "*shrinkage*" estimators, and the archetype is the James-Stein estimator.[6] The word shrinkage refers to the fact that β^{adj} is simply the β^{raw} estimate shrunk towards its mean. To see this, observe that (1) may be written as

$$\beta^{\text{adj}} = \mu_{\beta^{\text{raw}}} + c(\beta^{\text{raw}} - \mu_{\beta^{\text{raw}}}), \tag{3}$$

whenever the mean entry $\mu_{\beta^{\text{raw}}}$ of β^{raw} is taken to be one. Formula (3) shrinks β^{raw} by subtracting the mean from this estimate, scaling the entries by c and then distributing them around the mean once more. To our knowledge, the first application of the James-Stein estimator for beta adjustment is by Lavely, Wakefield & Barrett (1980). Neither Blume nor Vasicek normalize the raw betas. However, as their estimates cluster near one, doing so would not make a big difference.

In our discussion of beta adjustments, more emphasis is placed on the concept of *dispersion*, or coefficient of variation, a measure of deviation of a quantity from its mean. We define the dispersion of betas as the standard deviation taken after normalizing the mean to one. Our focus on dispersion facilitates a financial interpretation, and it allows us to compare the dispersion of market betas across various time periods.

PCA betas. Like the James Stein estimator, principal component analysis is an essential element of data science. A tool for exploring high-dimensional data sets, PCA is used in finance to identify factors that drive return and risk. PCA betas are the exposures to the dominant factor extracted from a security returns covariance matrix. A statistical analog to the market factor, the dominant principal component is the factor that maximizes explained variance. PCA betas, are generally all of the same sign, which suggests a parallel to market betas. The advantage of PCA lies in its ability to adapt to changes in market conditions.

Like time-series estimates of market betas, estimated PCA betas also tend to be overly dispersed. This general tendency is identified in Goldberg, Papanicolaou & Shkolnik (2018), who develop a market-regime dependent version of formula (1) to mitigate this "*dispersion bias.*" The proposed adjustment is tailored to minimum variance portfolio weights, like those in Figure 1. To this end, Goldberg et al. (2018) characterize the amount of excess dispersion to be removed to reveal the better betas.

Minimum variance. Arguments that the Blume, Vasicek and similar adjustments lead to better betas typically rely on tests of predictive power. Specifically, raw and adjusted beta estimates are compared out-of-sample with future realized betas, and the adjusted betas tend to be more accurate predictors than their raw counterparts.[7]

An alternative approach to evaluating betas builds on their tight relationship with minimum variance portfolio, which is carefully considered in Clarke et al. (2006), Clarke et al. (2011) and Goldberg et al. (2018), and illustrated in Figure 1. Indeed, estimation errors in betas have a profound impact on optimized portfolios.[8] As pointed out in **?**, minimum variance optimizers are "estimation-error maximizers." Goldberg et al. (2018) find a value for the parameter c in the adjustment (1) that maximizes the accuracy of weights and risk forecasts for minimum variance portfolios. We argue that it is this adjustment that generates better betas.

[5] Elton, Gruber, Brown & Goetzmann (2009, Chap. 7) and Sparks (2010) provide surveys of beta adjustments.

[6] An application of the James Stein estimator to beta shrinkage is Efron & Hastie (2016, Formula 7.43).

[7] See Klemkosky & Martin (1975), Elton, Gruber & Urich (1978), Elgers et al. (1979), Eubank & Zumwalt (1979), Mantripragada (1980), Hawawini & Vora (1983), Lally (1998) and Gray, Hall, Diamond & Brooks (2013).

[8] The literature on how estimation error corrupts optimized portfolios includes Klein & Bawa (1976), Brown (1979), Jobson & Korkie (1980), Michaud (1989), Best & Grauer (1991), Britten-Jones (1999), Michaud & Michaud (2008), Bender, Lee, Stefek & Yao (2009) and Bianchi, Goldberg & Rosenberg (2017) and references therein.

Diversification and volatility regimes. A challenge to the development of an optimal beta adjustment is the dynamics of betas over time. Figure 2 shows the means, standard deviations and dispersions of S&P 500 security betas at year end over the 21-year period from 1998 to 2018. The beta estimates therein are adjusted using Vasicek's rule. They provide support for Blume's finding that that true betas are non-stationary and move toward one over time. Empirically, we see that betas have been more dispersed in times of stress, such as the aftermath of the internet bubble and the great financial crisis, than in calmer periods.[9] Figure 2 suggests that a static rule like the Blume 2/3 adjustment may not be optimal.

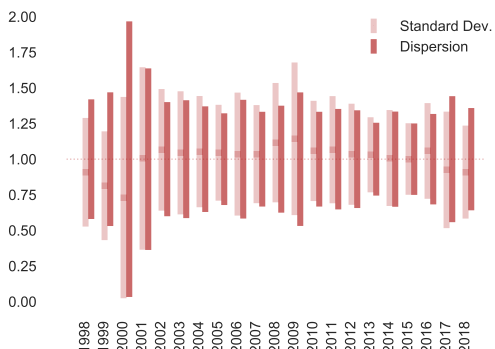

Figure 2. *Means, standard deviations and dispersions of Barra USSLOW historical S&P* 500 *betas at year end. Source: Aperio Group.*

Greater dispersion of betas tends to diversify a minimum variance portfolio by dampening extreme positions.[10] The relationship is illustrated in Figure 3, which shows scatter plots of minimum variance portfolio weights versus market betas at year end in 2000, when betas were more dispersed, and in 2015, when betas were more concentrated.[11]

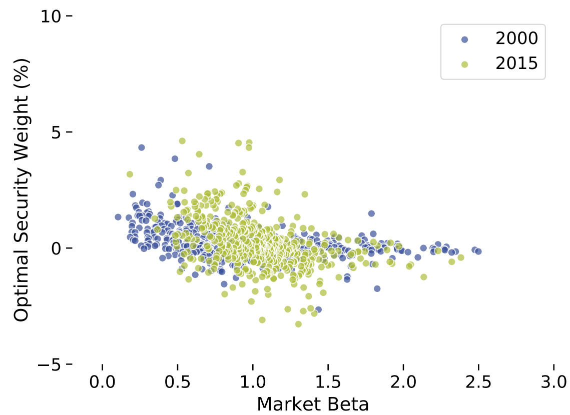

Figure 3. *Minimum variance portfolio weights versus betas at year end* 2000 *and* 2015. *The universe is the S&P* 500 *Index and the portfolios were constructed with the Barra USSLOW Model and the Barra Aegis optimizer. Source: Aperio Group.*

Since greater beta dispersion means greater diversification in a minimum variance portfolio, beta shrinkage must lead to greater concentration. Beta shrinkage is the enemy of diversification, and a portfolio manager must choose between greater accuracy and greater concentration. Goldberg et al. (2018) show that the optimal value of c in adjustment (1) is greater in stressed markets than in calm markets. In other words, all else equal, we should be shrinking less in stressed markets. There is a danger that the Blume 2/3 adjustment will lead to minimum variance portfolios that are both inaccurate and needlessly concentrated, especially in stressed markets.

Betas in the investment process. The analysis of the impact of beta dispersion on minimum variance portfolios was the starting point for this article. As outlined above, this led us to consider the subtle, regime-dependent linkages across a constellation of investment fundamentals, including beta dis-

[9] The tendency for betas to be more dispersed in turbulent periods is documented in Sefton, Jessop, De Rossi, Jones & Zhang (2011) for European markets.

[10] This relationship can be mathematically verified in a single-factor model for "reasonable" calibrations, but it does not hold in all circumstances. For example, in a single-factor homogenous specific variance model with all betas equal to one, any sufficiently small perturbation of the betas will concentrate a minimum variance portfolio. However, this special case seems to be atypical.

[11] The minimum variance portfolio was more diversified in 2000, with an effective security count (or inverse Herfindahl index) of 39, than it was in 2015, with an effective security count of 16.

persion and shrinkage, exposure normalization, diversification, estimation error, security correlations, and portfolio optimization. These fundamentals are tied together in the GPS adjustment, developed in Goldberg et al. (2018), which shrinks PCA betas in a way that improves the accuracy of weights and risk forecasts of minimum variance portfolios.[12]

We show that the GPS adjustment can be viewed as the PCA analog of the Vasicek correction for time series estimates of betas, and we outline some of the benefits of PCA-based factor models. We provide a simple formula for the GPS adjustment that can be easily implemented by practitioners, and we review metrics for assessing the impact of estimation error on optimized portfolios. We evaluate the GPS adjustment on minimum variance portfolios in a simulation based on a 4-factor model, which is carefully tuned to both calm and stressed regimes. The tuning is a delicate and interesting process, given the complex interactions of beta dispersion, market volatility and specific volatility, and the impact of these interactions on optimized portfolios. The results of our simulation are followed by a discussion of the limitations of the current study along with numerous suggestions for improvement and extension.

What are PCA betas?

The role of betas in the investment process dates back at least to the introduction of the Capital Asset Pricing Model (CAPM).[13] The CAPM identifies betas with the market factor governing security expected returns, volatilities and correlations. A more general factor-based approach than the CAPM is the arbitrage pricing theory (APT), as introduced in Ross (1973) and Ross (1976). The APT makes no upfront assumptions about the nature of the factors that drive risk, so it is natural to identify APT factors by applying statistical methods. This dynamic approach to factor identification can uncover emerging drivers of correlation, like the internet factor in the late 1990s, or the real estate factor that drove the financial crisis in 2008, or the climate factor that may soon be discovered.

The use of principal component analysis (PCA) to identify APT factors dates back at least to Connor & Korajczyk (1986).[14] In that article, PCA is applied to a sample covariance matrix of security returns to extract factors that drive equity return and risk. Of special interest are the PCA betas, which are the exposures to the dominant factor. As mentioned in the introduction, the dominant PCA factor in US equities is *market-like*, meaning that it has relatively large variance and mostly positive exposures.[15]

A scatter plot of Barra betas versus PCA-betas for the S&P 500 securities is in Figure 4.[16] It demonstrates a significant degree of correlation between the two estimates. That is no accident, but rather, an indication that human analysts, who identify market betas as the dominant factor, and statistical analysts, who identify PCA betas as the dominant factor, agree on the the most important source of risk in the US equity market.

The normalization of betas plays an important role in the adjustment given by for-

[12]The label GPS indicates that our adjustment navigates to the right destination. We note with humility that the label is composed of the initials of the first three authors' last names.

[13]See Treynor (1962), Sharpe (1964), Lintner (1965*b*), Lintner (1965*a*), Mossin (1966). The assumptions underlying the CAPM as well as empirical assessments of its predictions indicate it is likely misspecified. A few of the many studies that take issue with the CAPM are Blume & Friend (1973) Fama & French (2004) and Markowitz (2005).

[14]The authors build on this work further in Connor & Korajczyk (1988). Theoretical justifications for PCA as a tool for factor analysis were developped by Chamberlain & Rothschild (1983). Connor (1995) compares PCA models to the fundamental models that are typically used by financial practitioners (also see, Connor & Korajczyk (2010)).

[15]The market-like qualities of the dominant PCA factor in sample covariance matrices of US equity returns amount to an empirical argument for the existence of a large common factor. The assertion that the best instance of this factor is the market, itself, requires additional argument.

[16]More information about Barra beta estimation may be found in Bayraktar, Mashtaler, Meng & Radchenko (2014). It is possible that the differences between the PCA and Barra betas may be accounted for by adjustments like (1).

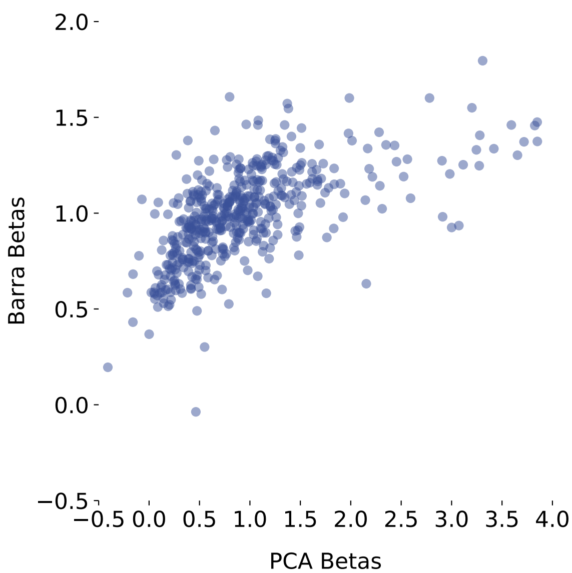

Figure 4. Barra USSLOW *historical betas vs* PCA *betas on December* 31, 2015. *Both are normalized to have mean one. Source: Aperio Group.*

mula (1). Since the effect of formula (1) is to shrink the raw betas toward one, the shrinkage target and the estimate must have the same means. However, by convention, each PCA factor is normalized so that the sum of the squares of its entries is equal to one.[17] Therefore, we rescale the *raw* betas β^{raw} to be normalized to have mean one. So, with β denoting the PCA betas extracted from a sample covariance matrix of security returns, we take[18]

$$(4) \qquad \beta^{\text{raw}} = \frac{\beta}{\mu_\beta}; \quad \mu_\beta = \frac{1}{N}\sum_{n=1}^{N} \beta_n .$$

We have assumed the mean entry μ_β of β is not equal to zero. While it is not a mathematical certainty, this assumption is virtually always satisfied by market betas and their PCA counterparts.

Beta dispersion

As discussed in the introduction, Goldberg et al. (2018) show that PCA betas tend to be overly dispersed. This calls for a precise definition of dispersion and it also presumes there are some true market betas that have a dispersion of their own. We denote these true market betas by β^* and define their dispersion as[19]

$$(5) \qquad \tau_{\beta^*} = \sqrt{\frac{1}{N}\sum_{n=1}^{N}\left(\frac{\beta_n^*}{\mu_{\beta^*}} - 1\right)^2}.$$

The mean μ_{β^*} of β^* (defined analogously to μ_β in (4)) is assumed not equal to zero. The dispersion τ_{β^*} is similar to the standard of deviation of the enties of the true betas, but it first normalized them to mean one. This facilitates a proper comparison to the dispersion of the estimate β^{raw}. Note, since the estimate β^{raw} is already normalized in this manner (per (4)), we have $\tau_\beta = \tau_{\beta^{\text{raw}}}$, both sides defined analogously to equation (5).

As shown in Goldberg et al. (2018), raw betas β^{raw} coming from PCA tend to have a larger dispersion than their true counterparts and this becomes more likely as N grows. In particular,

$$\tau^2_{\beta^{\text{raw}}} = \tau^2_{\beta^*} + \mathscr{D}^2_{\text{bias}},$$

where the dispersion bias $\mathscr{D}^2_{\text{bias}}$ is typically positive.[20] This is the PCA analog of the observation by Blume, Vasicek and others that time series estimates of market betas are overly dispersed. Accordingly, we define the dispersion bias to be

$$(6) \qquad \mathscr{D}^2_{\text{bias}} \triangleq \tau^2_{\beta^{\text{raw}}} - \tau^2_{\beta^*} .$$

Since $\mathscr{D}^2_{\text{bias}}$ is typically greater than zero, the adjusted beta β^{adj} must shrink the excess dispersion in β^{raw}. Formula (1) accomplishes

[17] Furthermore, the standard PCA factor normalization determines its entries only up to sign (multiplication by -1). In other words, the multiplicative inverse of a PCA factor is a PCA factor.

[18] The equally weighted mean in (4) can be replaced with a cap-weighted mean, or it can be weighted according to another heuristic; see Appendix B for a weighting we use for our general method to correct PCA betas.

[19] Our definition for dispersion coincides with the coefficient of variation (standard deviation divided by the mean).

[20] This is theoretically shown in the case when the security returns are Gaussian. We slightly abuse the notation in the variable $\mathscr{D}^2_{\text{bias}}$ by allowing this quantity to be negative in (6).

this, since

$$\tau^2_{\beta^{\text{adj}}} = c^2 \tau^2_{\beta^{\text{raw}}} \leq \tau^2_{\beta^{\text{raw}}}$$

and $0 \leq c \leq 1$.

In the case of PCA, this bias can be mitigated without knowledge of the true dispersion τ_{β^*}. Surprisingly, the sample covariance matrix from which PCA extracts factors also contains all the information required to estimate this unknown.

Goldberg et al. (2018) find the optimal value of c that removes all excess dispersion in $\mathscr{D}^2_{\text{bias}}$. The optimality properties concern accuracy of weights and risk forecasts for estimated minimum variance portfolios are discussed in that article along with necessary assumptions and theoretical guarantees. Here, we mention only that minimum variance portfolios are highly sensitive to the choice of the parameter c used to construct β^{adj}. These portfolio weights demand better betas.

Better betas

We introduce the GPS adjustment to raw PCA betas. Like Blume's and Vasicek's adjustments for time-series estimates, the GPS adjustment is a special case of formula (1). It is a form of dispersion shrinkage.

Blume's and Vasicek's adjustments. Recall that Blume's rule is associated with a static choice of

$$c^{\text{Blume}} = \frac{2}{3} \tag{7}$$

for the parameter c in (1) to compute the adjusted betas β^{adj}. This oversimplification was clearly not Blume's intention, yet, the wide adoption of (7) (and so (2)) is perhaps due to the rule's simplicity and apparent universality. In Blume's actual procedure, estimates of betas from later periods are regressed on those from an earlier period to identify the rate of regression (or reversion) of the true market betas to one. The value 2/3 turns out to be approximately equal to this historical rate of regression for the period Blume studied.[21] Blume (1975) explains an approach to disentangling the movement of the true betas from the estimation error that plagues beta estimates.

Estimation error in betas is the main focus of Vasicek (1973). Vasicek's rule relies on a Bayesian prior for the market beta that is normal with mean μ_{β^*} and standard deviation $s^2_{\beta^*}$. Assuming the security returns are Gaussian, Vasicek derives a posterior distribution for betas. Further assuming that $\mu_{\beta^*} = 1$ leads to an adjustment given by formula (1) with c defined

$$c^{\text{Vasicek}} = \frac{s^2_{\beta^*}}{s^2_{\beta^*} + s^2_{\beta^{\text{raw}}}}, \tag{8}$$

where $s_{\beta^{\text{raw}}}$ denotes the standard deviation of β^{raw} (its ordinary least squares standard error).

Central to Vasicek's approach is some prior knowledge of β^*, and for example, Vasicek suggests a reasonable prior for the NYSE betas has $\mu_{\beta^*} = 1$ and $s^2_{\beta^*} = 0.5$. The Bayesian perspective in Vasicek's rule becomes apparent when the mean μ_{β^*} is thought to deviate from one and then formula (1) must be modified. For further discussion see Vasicek (1973, Sections II and III).

The GPS adjustment. Goldberg et al. (2018) devise an adjustment in the spirit of Vasicek, but adapted for PCA betas. To this end, here, we assume β^{raw} is the vector of PCA betas as in (4).

The shrinkage parameter c for the GPS adjustment takes the same form as (8), but with two important alterations. First, we must replace the standard error of the true betas $s^2_{\beta^*}$ with their dispersion $\tau^2_{\beta^*}$; these are closely related but not equal. Second, the squared standard error of observed betas $s^2_{\beta^{\text{raw}}}$ is replaced with an estimate of the dispersion bias $\mathscr{D}^2_{\text{bias}}$ as defined in (6).

The next formula makes precise the analogy between the Vasicek adjustment in for-

[21] As noted in Blume (1971), this rate of regression is not stationary (i.e., static). It changes over time, and some estimates for various periods from 1933 to 1961 are provided (see, Blume (1971, Table 4)). Blume also points out: *"an improvement in the accuracy of one's assessments of risk can be obtained by adjusting for the historical rate of regression even though the rate of regression over time is not strictly stationary."*

mula (8) and the GPS adjustment for PCA betas.

$$c^{\mathrm{GPS}} = \frac{\tau^2_{\beta^*}}{\tau^2_{\beta^*} + \mathscr{D}^2_{\mathrm{bias}}} \tag{9}$$

This formula draws a parallel between $\mathscr{D}^2_{\mathrm{bias}}$ and the squared standard error $s^2_{\beta \mathrm{raw}}$ in Vasicek's adjustment.[22] While both c^{Vasicek} and c^{GPS} depend on unknowns s_{β^*} and τ_{β^*}, the latter may be accurately estimated from data in a PCA setting.

Substituting $\mathscr{D}^2_{\mathrm{bias}}$ into (9) gives

$$c^{\mathrm{GPS}} = \frac{\tau^2_{\beta^*}}{\tau^2_{\beta \mathrm{raw}}}. \tag{10}$$

To provide an expression for τ_{β^*}, let σ^2 and δ^2 denote the estimates of the variance of the market and the specific variance (the variance of security return that is residual to the factor(s)). Then,

$$\tau^2_{\beta^*} \approx \tau^2_{\beta \mathrm{raw}} - \left(\frac{\delta^2}{\sigma^2 + \delta^2}\right)\left(1 + \tau^2_{\beta \mathrm{raw}}\right). \tag{11}$$

Substituting (11) into (10) yields the following data-driven approximation for (10),

$$c^{\mathrm{GPS}} \approx 1 - \left(\frac{\delta^2}{\sigma^2 + \delta^2}\right)\left(1 + \frac{1}{\tau^2_{\beta \mathrm{raw}}}\right). \tag{12}$$

We emphasize the fact that no knowledge of the true betas nor τ_{β^*} is required for the estimate.

Formulas (11) and (12) are predicated on the unrealistic assumption that all securities have the same specific variance. To account for heterogeneous specific variances, one needs a more generalized definition of the mean and dispersion. We refer the reader to Appendix B for this extension.

[22] A difference between the dispersion bias $\mathscr{D}^2_{\mathrm{bias}}$ in the GPS adjustment and the squared standared error $s^2_{\beta \mathrm{raw}}$ in the Vasicek adjusment is that the former is applied to all securities while the latter is estimated separately for each security.

Minimum variance

As argued in the introduction, the composition of minimum variance portfolios is governed predominantly by the market betas. It is for this reason that we chose this portfolio as the means of testing for better betas in simulation.

Portfolio construction. A fully invested but otherwise unconstrained minimum variance portfolio is the solution to the minimization problem

$$\begin{aligned} &\min_{w \in \mathbb{R}^N} w^\top \mathbf{\Sigma} w \\ &\text{subject to:} \\ &\textstyle\sum_n w_n = 1, \end{aligned} \tag{13}$$

where $\mathbf{\Sigma}$ is our estimated security returns covariance matrix. We denote by $\mathbf{\Sigma}_*$ the true returns covariance matrix, which is not observed. Throughout, we denote by w the estimated minimum variance portfolio, the solution of (13). We write w_* for the true minimum variance portfolio, the one computed by replacing $\mathbf{\Sigma}$ with $\mathbf{\Sigma}_*$.

Accuracy metrics. Estimation error causes two types of difficulties in optimized portfolios. It distorts portfolio weights and it biases the risk of optimized portfolios downward. We use two metrics to capture these distortions.

The following two metrics can only be used in simulation as they rely on the true covariance matrix $\mathbf{\Sigma}_*$ of the security returns. We define,

$$\mathscr{T}^2_w = (w_* - w)^\top \mathbf{\Sigma}_* (w_* - w), \tag{14}$$

the (squared) tracking error of w. It measures the distance between the optimal and estimated portfolios, w_* and w, as the square of the width of the distribution of the return difference $w_* - w$.

Next, we define the variance forecast ratio $\mathscr{R}_w$ of w. Note that the estimated variance of portfolio w is given by $w^\top \mathbf{\Sigma} w$, and its true variance by $w^\top \mathbf{\Sigma}_* w$. Consequently, we define

$$\mathscr{R}_w = \frac{w^\top \mathbf{\Sigma} w}{w^\top \mathbf{\Sigma}_* w}. \tag{15}$$

The ratio $\mathscr{R}_w$ is less than one when the risk of the portfolio w is under-forecast. This is typically the case when the sample covariance matrix is used as the estimate $\mathbf{\Sigma}$. While a PCA-estimated covariance matrix (see Appendix A) improves over this naive choice, as we shall see, under-forecasts persist.

Metrics (14) and (15) quantify both types of errors that we see in minimum variance portfolio weights. We use both in our simulations.

Model calibration

We specify and calibrate a return generating model that allows us to evaluate the impact of beta shrinkage on minimum variance portfolios. Simulation allows for a direct comparison of estimates to ground truth, but the value of the insights from a simulation depends on the realism of the underlying model.

To illustrate the challenges inherent in developing a realistic simulation, consider the interactions of greater beta dispersion, higher market variance and higher specific variance, all of which are characteristic of stressed markets. All else equal, greater beta dispersion leads to lower volatility of a minimum variance portfolio in "reasonable" calibrations. At the same time, higher specific variance leads to higher volatility of minimum variance and lower security correlation. A delicate balance of these three effects is essential to any realistic simulation, and that the efficacy of a beta shrinkage operator must be considered in both calm and stressed markets.

Covariance matrix. We assume a generating process R of (excess) returns to N securities that follows a K-factor model. More precisely, we let

$$R = \mathbf{B}_* \psi + \epsilon, \tag{16}$$

where $\psi = (\psi_1, \ldots, \psi_K)$ is a vector of returns to the factors, $\mathbf{B}_*$ is a $N \times K$ matrix of exposures to the factors, and $\epsilon = (\epsilon_1, \ldots, \epsilon_N)$ is the vector of heterogeneous specific returns. While the returns ψ and ϵ are random, we treat each entry of $\mathbf{B}_*$ as a constant to be estimated. Assuming all the ψ_k and ϵ_n are mean zero and pairwise uncorrelated, the $N \times N$ covariance matrix of R take the form[23]

$$\mathbf{\Sigma}_* = \mathbf{B}_* \mathbf{\Omega} \mathbf{B}_*^\top + \mathbf{\Delta}. \tag{17}$$

Here, $\mathbf{\Omega}$ and $\mathbf{\Delta}$ denote the covariance matrices of ψ and ϵ respectively. Under our assumptions, both matrices are diagonal. The covariance matrix $\mathbf{\Sigma}_*$ is not known, and we assume only the returns R are observed.

Calm and stressed regimes. Our experiments are based on $N = 500$ securities and $K = 4$ factors. We assume the returns R are Gaussian and we calibrate our model to calm and stressed market regimes.

The dominant factor is market-like, so the first column of $\mathbf{B}_*$ can be thought of as the vector of market betas, we denote by β^*. Since market betas are necessarily distributed around one, we set the mean of the entries of β^* to that value, and we use the empirical data from the US market as a guide to calibrating the dispersion of the betas around their mean.

The three remaining factors are modeled on equity styles such as volatility, earnings yield and size.[24] We draw the exposure of each security from a normal distribution and standarize to a z-score with mean 0, variance 1. Bayraktar et al. (2014, Table 4.1) provides guidance on the volatility of equity style factors, with estimates typically less than 10% per year. While style factors undergo volatility regimes, they need not coincide with market regimes, as shown in van Dijk (2011, Figure 2). Therefore, we set the style factor volatilities to the same values in our calm and stressed regimes, which are governed by the market factor.

Both security correlations and portfolio volatilities have tended to be higher in stressed markets than in calm markets.[25] As illustrated in Figure 2 and Sefton et al. (2011, Chart 2), the same has been true for beta dispersion. For realistic parameter values, higher

[23] $\mathbf{B}_*^\top$ denotes the transpose of $\mathbf{B}_*$.

[24] References on the importance of volatility, earnings yield and size in explaining cross-sectional correlations are Black, Jensen & Scholes (1972), Basu (1983) and Banz (1981).

[25] See, for example, Clarke et al. (2006), Clarke et al. (2011) and Goldberg, Leshem & Geddes (2014).

beta dispersion leads to lower security correlations and lower minimum variance portfolio volatility. The effect of the higher beta dispersion on the portfolio volatility is more than offset by higher factor and specific volatilities. In the return generating process (16), we calibrate specific volatilities in accordance with Bayraktar et al. (2014, Figure 4.5). The regime-specific parameters we use in our simulation are in Table 1.

As discussed in the introduction, betas have tended to be more dispersed in more volatile periods. Simple experiments reveal the extreme sensitivity of minimum variance portfolio weights and risk to beta dispersion, so we used the empirical results shown in Figure 2 to guide the calibration.

The implied parameters in the last two rows of Table 1 show some of the more reasonable consequences of our calibration.[26] Average pairwise security correlations are higher in the stressed regime than in the calm regime, as are the volatilities of minimum variance portfolios.

Simulation results

We examine the impact of correcting the dispersion bias in PCA betas on simulated minimum variance portfolios.

Experimental design. We evaluate the impact of beta shrinkage methods in a numerical simulation based on the 4-factor model specified in (16). We emphasize how errors in betas affect optimized portfolios. Specifically, we use tracking error and the variance forecast ratio to evaluate the accuracy of the weights and risk forecasts of minimum variance portfolios.

In each experiment, we simulate a year's worth of independent daily returns, $T = 250$, to $N = 500$ securities. From this data set, we construct a sample covariance matrix $\mathbf{S}$ from which we construct estimators $\mathbf{\Sigma}$ with three methods. For the raw beta method, we use principal component analysis to extract 4 factors from a sample covariance matrix $\mathbf{S}$.

Parameter	Calm	Stressed
market volatility	16%	32%
style volatility	$\{4\%, 4\%, 8\%\}$	$\{4\%, 4\%, 8\%\}$
specific volatility	$[32\%, 64\%]$	$[48\%, 96\%]$
beta dispersion	0.33	0.40
average correlation	0.10	0.16
minvol volatility	6.2%	8.1%

Table 1. *Simulation parameters for calm and stressed regimes. Market, style and specific volatilities, as well as beta dispersion are set. Average correlation and volatility of minimum variance portfolios are implied by the model.*

[26] Less reasonable consequences are discussed below in our discussion of limitations and open problems.

The extracted factors are used to estimate the first term on the right hand side of (17), and the residuals are used to construct the second term. In the Blume method, we adjust the dominant factor with (7). In the Blind GPS Adjustment, we adjust the dominant factor with (12). We repeat this procedure 400 times to generate the results presented below.

Results. We begin by looking at the shrinkage weights that the methods apply to the raw PCA betas via formula (1). Figure 5 shows the histograms of the weights placed on raw betas by the Blume 2/3 Rule (7) as well as the Blind GPS Adjustment (12) in the calm and stressed regimes specified in Table 1. Relative to the GPS adjustments, the Blume 2/3 Rule overshrinks. To see that the error is in the Blume 2/3 rule and not in the Blind GPS Adjustment, we add the optimal shrinkage weights generated by the Exact GPS Adjustment (9) to Figure (5). Errors made by Blume 2/3 rule are uniformly greater than the errors made by the Blind GPS Adjustment, and the difference is exacerbated by market stress.

We turn next to portfolio accuracy metrics. Figure 6 reports results for minimum variance portfolios. The top two panels focus on tracking error, which measures the accuracy of weights. Tracking error was lowest for the Blind GPS Adjustment in both the calm and stressed regimes. The Blume 2/3 Rule improved on raw betas in the calm regime but it raised tracking error in the stressed regime. The bottom two panels focus on the variance forecast ratio. The Blind GPS Adjustment had the most accurate risk forecasts in both regimes, but it was biased slightly downward. This stems from a downward bias in specific risk estimates, and it must be addressed independently of the dispersion bias in betas.

Limitations and open problems

We highlight several of the many areas in which our study might be extended.

There is no consensus on the number of factors required to explain equity returns. Industry standard models of the US equity market have dozens of factors, suggesting that the 4 we use in our simulations may not be adequate for some purposes. Still, as indicated by Figure 1, the market-like factor has a dominant influence on minimum variance portfolio weights, and a 4-factor model (or even a 1-factor model) may well be enough to illustrate our point. Further, we endow the estimates in our simulation with knowledge of the true number of factors. It is inevitable that the use of an algorithm to estimate the number of factors from data would add to the overall level of estimation error. Relying yet again on Figure 1 for guidance suggests that, given the dominance of the market-like factor, our adjustment may well retain its effectiveness in this more challenging setting. The Gaussian distribution used to generate factor returns and especially specific returns is unrealistic. Experiments with simulations that rely on heavy-tailed returns indicate that the dispersion bias is still present and, to a great extent, correctable. However, heavy-tailed distributions tend to exacerbate risk underforecasts in simulation, and measures beyond a beta dispersion bias correction need to be taken in order to generate accurate risk forecasts. Dispersion involves only the first two moments of the beta distribution, which has tended to be right skewed and heavy-tailed.[27] This suggests that there may be a non-linear shrinkage formula that improves on (1). Future research could examine these issues both through simulation and empirical study. That analysis would benefit from the initial benchmarking exercise carried out here, and hence merits treatment in a separate forum.

Counterintuitively, the estimated volatility σ^2 of the dominant PCA factor plays no role in the weights or forecast risk of minimum variance portfolios.[28] However, σ^2 does influence the risk of the equally weighted portfolio, as discussed in Goldberg et al. (2018). Further study is required to understand the impact

[27] This is noted in Bera & Kannan (1986), which suggests shrinking square roots of betas rather than betas themselves.

[28] This observation, communicated to us by Stephen Bianchi, was one of the primary clues that led us to our formulation of the dispersion bias. A rigorous proof that σ^2 plays no role in the weights or forecast risk of minimum variance portfolios is in Goldberg et al. (2018).

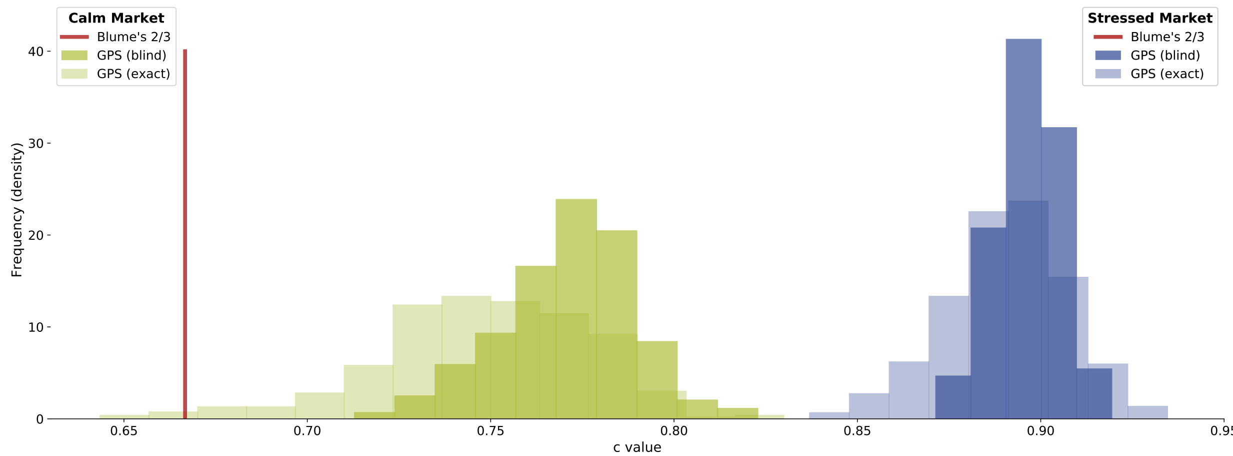

Figure 5. *Shrinkage weights for the Blume and GPS methods. Each histogram is based on* $N = 500$ *securities,* $T = 250$ *observations and* 400 *simulation paths.*

of model parameters, such as σ^2, on different aspects of portfolio construction risk forecasting.

Finally, there is a connection between our beta shrinkage and the covariance matrix shrinkage discussed in Ledoit & Wolf (2004) and a series of related, more technical articles. Covariance matrix shrinkage implicitly determines beta shrinkage, which deserves a thorough comparison to the shrinkage methods studied in this article.

Summary

Betas are fundamental to investing but there is no standard for deciding which methods, among the vast range of possibilities, produce better betas. We argue for "bottom line" criteria that take account of the impact of betas on quantitatively constructed portfolios. Specifically, we think that betas are better if they reduce errors in weights of optimized portfolios, or risk forecasts of any portfolio. This perspective facilitates the development of the blind GPS adjustment, which lowers dispersion in PCA betas by shrinking them toward 1. This adjustment can be viewed as an analog of the Blume and Vasicek rules for shrinking time series estimates of betas.

Three features distinguish the blind GPS adjustment from its analogs. The first is the precise notion of optimality: the blind GPS adjustment mitigates the impact of estimation error on minimum variance portfolios, absent knowledge of true betas. The second is the purely data-driven formulation of the blind GPS adjustment. The third is flexibility: the blind GPS adjustment adapts to volatility regimes. This last point is essential, given the empirical observation that betas tend to be more dispersed in stressed markets than in calm markets. In other words, absent interference by a portfolio manager, optimized portfolios tend to exhibit more concentrated positions in calm markets than in stressed markets. This counter-intuitive finding forces us to examine the rule of thumb that diversification disappears in a crisis. To explain the higher security correlations observed in stressed markets, factor volatility must increase by enough to overcome both the increase in specific risk and the increase in beta dispersion that generally accompany turbulence. The finding also implies that beta shrinkage, widely practiced throughout financial services, is the enemy of diversification. Portfolio managers must choose between accuracy and diversification.

We illustrate the effectiveness of the GPS adjustment in simulations tuned to calm and stressed markets, focusing on the way errors in betas affect the accuracy of portfolio weights and risk forecasts for minimum variance portfolios. The efficacy of our adjustment is a step toward making PCA-based risk

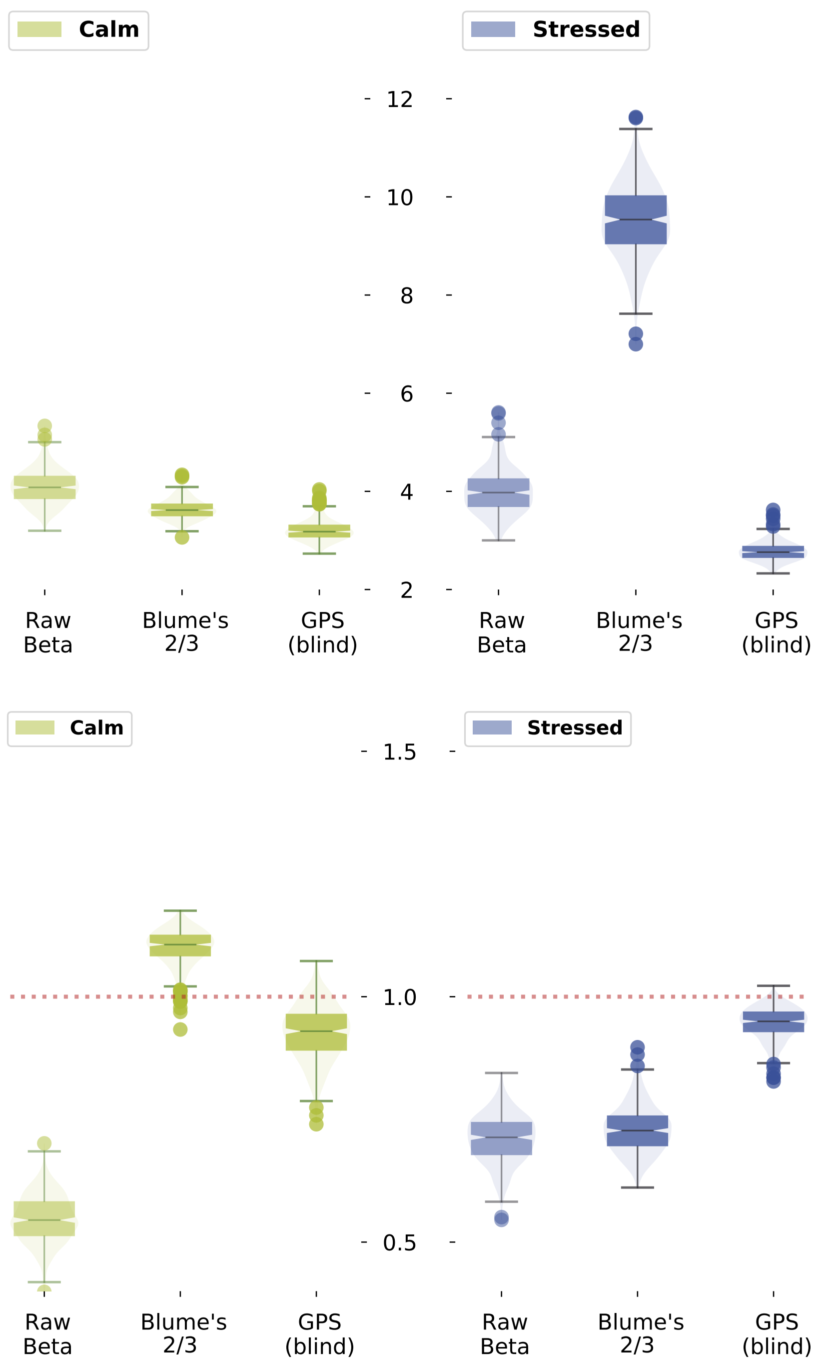

Figure 6. *Tracking error (top row) and variance forecast ratio (bottom row) for the minimum variance portfolio with* $N = 500$ *securities estimated from* $T = 250$ *observations. Each boxplot is based on* 400 *simulation paths.*

models useful to financial practitioners, and it sheds light on the meaning of "better betas."

A. Principal Components

With a distinguished history dating back, at least, to Pearson (1901) and Hotelling (1933), principal component analysis (PCA) is a statistical method used to identify factors that maximize explained in-sample variance. A more recent exposition of PCA for equity risk factor modeling may be found in Litterman et al. (2003, Chapter 20). We explain briefly how this technique works in our setting.

Take the sample covariance matrix $\mathbf{S}$ of the security returns as our starting point. The first principal component is the portfolio, or weighted combination of securities, that explains as much variance as possible, subject to the constraint that the sum of the squared portfolio weights is one. Mathmatically, this portfolio is the solution to the following problem.

$$\max_{v \in \mathbb{R}^N} v^\top \mathbf{S} v \quad \sum_n v_n^2 = 1. \tag{18}$$

The estimated PCA betas are the entries of the first principal component β normalized to have mean one, and hence are given by

$$\beta^{\text{raw}} = \frac{\beta}{\mu_\beta}; \quad \mu_\beta = \frac{1}{N}\sum_{n=1}^{N} \beta_n. \tag{19}$$

as in (4), where the vector β is the maximizer of (18). In the generalized GPS adustment provided in Appendix B, we require the following quantity.

$$\ell^2 = \beta^\top \mathbf{S} \beta \tag{20}$$

This is the largest eigenvalue of the matrix $\mathbf{S}$.

The second principal component explains as much of the remaining variance as possible, subject to the same constraint and the additional requirement that it be orthogonal to the first principal component. And so on.

The estimated specific returns $\mathbf{Z}$ are the residuals to the regression of the security returns onto the factors. Consequently, the specific variance of security n is estimated to be

$$\delta_n^2 = \frac{1}{T}\sum_{t=1}^{T} \mathbf{Z}_{nt}^2. \tag{21}$$

B. Generalized GPS

Here we generalize the calculations of Section 3.3 to a model with heterogeneous specific risk. The estimated adjustment parameter c^{GPS} given in (12) is correct as is with slight generalizations for the dispersion $\tau_{\beta^{\text{raw}}}$. We also provide the estimate σ^2 for the market variance, and δ^2 for the average specific variance.

For δ_n^2 as in (21), define

$$\delta^2 = \frac{1}{N}\sum_{n=1}^{N} \delta_n^2. \tag{22}$$

The estimated market variance is given by

$$\sigma^2 = \ell^2 - \delta^2, \tag{23}$$

where ℓ^2 is the largest eigenvalue as in (20).

To estimate the dispersion for a model with heterogeneous specific risk, we generalize (4) (repeated in (19)) and (5) to utilize the inverse specific variances as weights. The weighted mean is

$$\mu_\beta = \frac{1}{N_\delta}\sum_{n=1}^{N} \beta_n / \delta_n^2, \tag{24}$$

where $N_\delta = \sum_n 1/\delta_n^2$. The PCA betas β^{raw} are then assembled as in (19) with μ_β as above. The dispersion of the raw betas is then computed via

$$\tau_{\beta^{\text{raw}}}^2 = \frac{1}{N_\delta}\sum_{n=1}^{N} \frac{1}{\delta_n^2}\left(\frac{\beta_n}{\mu_\beta} - 1\right)^2, \tag{25}$$

When the returns are believed (or assumed) to follow the Gaussian distribution, we can impove the estimates δ^2 and σ^2 in (22) and (23) by using a Marchenko-Pastur correction (we do so in our numerical experiments). In this case,

$$\delta^2 = \frac{S - (\ell_1^2 + \cdots + \ell_K^2)}{N - K(1 - N/T)} \tag{26}$$

where S denotes the sum of the diagonal entries of $\mathbf{S}$ and ℓ_k^2, the kth largest eigenvalue of $\mathbf{S}$. Here, K is the total number of factors and note that ℓ_1^2 is equal to ℓ^2 in (20). Finally,

$$\sigma^2 = \ell_1^2 - \delta^2 (1 + N/T)\,, \tag{27}$$

using the δ^2 in (26). We recommend using these adjusted values for σ^2 and δ^2 in practice even if the returns are not believed to be Gaussian.

References

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

Basu, S. (1983), 'The relationship between earnings' yield, market value and return for nyse common stocks', *Journal of Financial Economics* **12**, 129–156.

Bayraktar, M. K., Mashtaler, I., Meng, N. & Radchenko, S. (2014), Barra us total market equity model for long-term investors: Empirical notes.

Bender, J., Lee, J.-H., Stefek, D. & Yao, J. (2009), 'Forecast risk bias in optimized portfolios'.

Bera, A. K. & Kannan, S. (1986), 'An adjustment procedure for predicting systematic risk', *Journal of Applied Econometrics* **1**, 317–332.

Best, M. J. & Grauer, R. R. (1991), 'On the sensitivity of mean-variance-efficient portfolios to changes in asset means: some analytical and computational results', *The review of financial studies* **4**(2), 315–342.

Bianchi, S. W., Goldberg, L. R. & Rosenberg, A. (2017), 'The impact of estimation error on latent factor models of portfolio risk', *The Journal of Portfolio Management* **43**(5), 145–156.

Black, F., Jensen, M. C. & Scholes, M. (1972), The capital asset pricing model: some empirical tests, *in* M. C. Jensen, ed., 'Studies in the Theory of Capital Markets', Praeger.

Blume, M. E. (1971), 'On the assessment of risk', *The Journal of Finance* **26**(1), 1–10.

Blume, M. E. (1975), 'Betas and their regression tendencies', *The Journal of Finance* **30**(3), 785–795.

Blume, M. E. (1979), 'Betas and their regression tendencies: Some further evidence', *Journal of Financial and Quantitative Analysis* **34**(1), 265–267.

Blume, M. E. & Friend, I. (1973), 'A new look at the capital asset pricing model', *The Journal of Finance* **28**(1), 19–33.

Britten-Jones, M. (1999), 'The sampling error in estimates of mean-variance efficient portfolio weights', *The Journal of Finance* **54**(2), 655–671.

Brown, S. (1979), 'The effect of estimation risk on capital market equilibrium', *Journal of Financial and Quantitative Analysis* **14**(2), 215–220.

Chamberlain, G. & Rothschild, M. (1983), 'Arbitrage, factor structure, and mean-variance analysis on large asset markets', *Econometrica* **51**(5), 1281–1304.

Clarke, R., De Silva, H. & Thorley, S. (2006), 'Minimum-variance portfolios in the U.S. equity market', *Journal of Portfolio Management* **33**(1).

Clarke, R., De Silva, H. & Thorley, S. (2011), 'Minimum-variance portfolio composition', *Journal of Portfolio Management* **2**(37), 31–45.

Connor, G. (1995), 'The three types of factor models: A comparison of their explanatory power', *Financial Analysts Journal* **51**(3), 42–46.

Connor, G. & Korajczyk, R. A. (1986), 'Performance measurement with the arbitrage pricing theory: A new framework for analysis', *Journal of financial economics* **15**, 373–394.

Connor, G. & Korajczyk, R. A. (1988), 'Risk and return in equilibrium apt: Application of a new test methodology', *Journal of Financial Economics* **21**, 255–289.

Connor, G. & Korajczyk, R. A. (2010), 'Factor models of asset returns', *Encyclopedia of Quantitative Finance, Rama Cont, ed., Chicester: Wiley* .

Efron, B. & Hastie, T. (2016), *Computer Age Statistical Inference*, Cambridge University Press.

Elgers, P. T., Haltiner, J. R. & Hawthorne, W. H. (1979), 'Beta regression tendences: statistical and real causes', *The Journal of Finance* **3**(1), 261–263.

Elton, E. J., Gruber, M. J., Brown, S. J. & Goetzmann, W. N. (2009), *Modern portfolio theory and investment analysis*, John Wiley & Sons.

Elton, E. J., Gruber, M. J. & Urich, T. J. (1978), 'Are betas best?', *The Journal of Finance* **33**(5), 1375–1384.

Eubank, A. A. & Zumwalt, J. K. (1979), 'An analysis of the forecast error impact of alternative beta adjustment techniques and risk classes', *The Journal of Finance* **34**(3), 761–776.

Fama, E. F. & French, K. R. (2004), 'The capital asset pricing model: theory and evidence', *Journal of Economic Perspectives* **18**(3), 25–46.

Goldberg, L., Leshem, R. & Geddes, P. (2014), 'Restoring value to minimum variance', *Journal of Investment Management* **12**(2), 32–39.

Goldberg, L., Papanicolaou, A. & Shkolnik, A. (2018), The dispersion bias. CDAR working paper.

Gray, S., Hall, J., Diamond, N. & Brooks, R. (2013), The vasicek adjustment to beta estimates in the capital asset pricing model. SFG Consulting and Monash University.

Green, R. C. & Hollifield, B. (1992), 'When will mean-variance efficient portfolios be well diversified?', *The Journal of Finance* **47**(5), 1785–1809.

Hawawini, G. A. & Vora, A. (1983), 'Is adjusting beta estimates an illusion?', *The Journal of Portfolio Management* **10**(1), 23–26.

Hotelling, H. (1933), 'Analysis of a complex of statistical variables into principal components.', *Journal of Educational Psychology* **24**, 417–441.

Jagannathan, R. & Ma, T. (2003), 'Risk reduction in large portfolios: Why imposing the wrong constraints helps', *The Journal of Finance* **58**(4), 1651–1683.

Jobson, J. D. & Korkie, B. (1980), 'Estimation for markowitz efficient portfolios', *Journal of the American Statistical Association* **75**(371), 544–554.

Klein, R. W. & Bawa, V. S. (1976), 'The effect of estimation risk on optimal portfolio choice', *Journal of Financial Economics* **3**, 215–231.

Klemkosky, R. C. & Martin, J. D. (1975), 'The adjustment of beta forecasts', *The Journal of Finance* **30**(4), 1123–1128.

Lally, M. (1998), 'An examination of Blume and Vasicek betas', *The Financial Review* **33**, 183–199.

Lavely, J. A., Wakefield, G. & Barrett, B. (1980), 'Toward enhancing beta estimates', *The Journal of Portfolio Management* **6**(4), 43–46.

Ledoit, O. & Wolf, M. (2004), 'Honey, i shrunk the sample covariance matrix', *The Journal of Portfolio Management* **30**, 110–119.

Levy, R. A. (1971), 'On the short-term stationarity of beta coefficients', *Financial Analysts Journal* **27**(6), 55–62.

Lintner, J. (1965*a*), 'Security prices, risk and maximal gains from diversification', *Journal of Finance* **4**(587–615).

Lintner, J. (1965*b*), 'The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets', *Review of Economics and Statistics* **73**, 13–37.

Litterman, B. et al. (2003), *Modern investment management: an equilibrium approach*, Vol. 246, John Wiley & Sons.

Mantripragada, K. G. (1980), 'Beta adjustment methods', *Journal of Business Research* **8**(3), 329–339.

Markowitz, H. M. (2005), 'Market efficiency: a theoretical distinction and so what?', *Financial Analysts Journal* **61**(5), 17–30.

Michaud, R. O. (1989), 'The markowitz optimization enigma: Is 'optimized' optimal?', *Financial Analysts Journal* **45**(1), 31–42.
URL: *https://doi.org/10.2469/faj.v45.n1.31*

Michaud, R. O. & Michaud, R. O. (2008), *Efficient asset management: a practical guide to stock portfolio optimization and asset allocation*, Oxford University Press.

Mossin, J. (1966), 'Equilibrium in a capital asset market', *Econometrica* **34**(4), 768–783.

Pearson, K. (1901), 'On lines and planes of closest fit to systems of points in space', *Philisophical Magazine* **2**(11), 559–572.

Ross, S. A. (1973), Return, risk and arbitrage, *in* I. Friend & J. Bicksler, eds, 'Return and Risk in Finance', Cambridge: Ballinger, pp. 189–217.

Ross, S. A. (1976), 'The arbitrage theory of capital asset pricing', *Journal of economic theory* **13**(3), 341–360.

Sefton, J., Jessop, D., De Rossi, G., Jones, C. & Zhang, H. (2011), Low-risk investing, Technical report, UBS.

Sharpe, W. F. (1964), 'Capital asset prices: A theory of market equilibrium under conditions of risk', *The Journal of Finance* **19**(3), 425–442.

Sparks, N. (2010), An analysis of the primary beta adjustment techniques: method reconciliations and extension to industries. Senior Thesis, Claremont McKenna College.

Treynor, J. L. (1962), Toward a theory of market value of risky assets. Presented to the MIT Finance Faculty Seminar.

van Dijk, M. A. (2011), 'Is size dead? A review of the size effect in equity returns', *Journal of Banking and Finance* **35**, 3263–3274.

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