Better Betas

Lisa R. Goldberg∗ Alex Papanicolaou†

Alex Shkolnik‡ Simge Ulucam§

August 24, 2020

Forthcoming in the Journal of Portfolio Management

Abstract

We introduce the data-driven Goldberg, Papanicolaou and Shkolnik (GPS) adjustment for estimated betas, which leads to material improvements in the accuracy of weights and risk forecasts of minimum variance portfolios. Like the widely used Blume 2/3-rule and Vasicek adjustment developed in the 1970s, the GPS adjustment for estimated betas shrinks raw beta estimates toward one. Unlike its antecedents, the GPS adjustment operates on the dominant factor of a sample covariance matrix, and this adjustment adapts dynamically to varying levels of beta dispersion. We illustrate the power of the GPS adjustment in a simulation that is calibrated to calm and stressed market regimes.
Finance practitioners have routinely used mean-variance optimization (the “E-V rule”) to construct portfolios since the publication of Markowitz (1952). In that seminal article, Markowitz comments on the subtlety inherent in determining the estimates of expected return and risk required to run mean-variance optimization.

To use the E-V rule in the selection of securities we must have procedures for finding reasonable $\mu_i$ and $\sigma_{ij}$. These procedures, I believe, should combine statistical techniques and the judgment of practical men. My feeling is that the statistical computations should be used to arrive at a tentative set of $\mu_i$ and $\sigma_{ij}$. Judgment should then be used in increasing or decreasing some of these $\mu_i$ and $\sigma_{ij}$ on the basis of factors or nuances not taken into account by the formal computations.

The problem of accurately estimating expected return and risk has proven to be a rich area of research, combining judgment and statistics. A vast library of statistical techniques, many of which had not yet been invented in 1952, has turned out to be relevant.

In this article, we focus on risk, specifically, on the propagation of errors in an estimated covariance matrix to weights and risk forecasts of minimum variance portfolios. As is intuitive and well known, risk-minimizing optimization tends to overweight those securities whose volatilities and correlations are underestimated, so that forecast risk of an optimized portfolio is biased downward. But how do we recognize those securities?

It may not be possible to identify the precise set of securities whose risk is underforecast. It is possible, however, to mitigate the impact of estimation error in a covariance matrix on a minimum variance portfolio by correcting the dispersion bias. This is the tendency of beta estimates for securities with high estimated betas to be too high, and of beta estimates for securities with low estimated betas to be too low.

Goldberg, Papanicolaou & Shkolnik (2020) show that the dispersion bias is a substantial source of error in the weights and risk forecasts of minimum portfolios, and develop a dynamic method for mitigating the excess dispersion. Here, we build on that article in two ways. First, we provide an easy-to-implement, data-driven algorithm that produces better betas by shrinking the dispersion of naively estimated betas by a prescribed amount. The optimal degree of shrinkage depends on the level of dispersion of the betas, which varies over time. As we illustrate with simulations calibrated to calm and stressed markets, optimization with covariance matrices built with better betas generates better minimum variance portfolios. Second, we extend the analysis in Goldberg, Papanicolaou & Shkolnik (2020), which focuses on a single-factor model with homogenous specific risk, to the more
realistic setting of a multi-factor model with heterogenous specific risk.

Empirical motivation and background

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 Exhibit 1, which shows scatter plots of minimum variance portfolio weights against market betas. Two portfolios are constructed from securities in the S&P 500 at the end of December, 2015. The first uses a multi-factor model, and the second is a trimmed version that retains only the market factor and the specific risks. The example shows that, absent interference by a portfolio manager, minimum variance weights are, to a large extent, determined by the market betas. The effect is dramatic in a single-index model, and it persists even when the model has many more factors.\(^2\)

Beta adjustments for time-series regression.

A collection 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 and Vasicek remedied this shortcoming by an adjusted beta

\[
\beta_{\text{adj}} = c \beta_{\text{raw}} + (1 - c),
\]

where \(\beta_{\text{raw}}\) 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

\[
\beta^{\text{Blume}} = \frac{2}{3} \beta_{\text{raw}} + \frac{1}{3}
\]

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

\(^1\)Articles explaining how market betas influence weights of optimized portfolios include Green & Hollifield (1992), Jagannathan & Ma (2003), Clarke, De Silva & Thorley (2006).

\(^2\)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.
Exhibit 1. Scatter plots of minimum variance 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 US-SLOW Model and a trimmed version that retains only the market factor and specific risk.

- Prepared by the authors from Aperio Group data.
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). In contrast, the proposal in Vasicek (1973), which is used in commercial models including 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.

Dispersion shrinkage.

Adjusted betas are examples of “shrinkage” estimators. The archetype is the James-Stein estimator, which vastly expanded statistical thinking in the 1950s and 1960s by showing how to improve the accuracy of a collection of estimates, each obtained by minimizing mean-squared error. The thesis underlying the James-Stein estimator is similar to one of the principles that motivated Blume and Vasicek in the 1980s: the largest estimates in a series tend to be biased upward while the smallest estimates tend to be biased downwards. The word shrinkage refers to the fact that $β_{adj}$ is the $β_{raw}$ shrunk towards its mean. Note, formula (1) matches

$$β_{adj} = μ_{β_{raw}} + c(β_{raw} - μ_{β_{raw}}),$$

when the mean entry $μ_{β_{raw}}$ of $β_{raw}$ is one. (3) shrinks $β_{raw}$ by subtracting the mean, scaling the entries by c and then redistributing them about $μ_{β}$.

In our treatment of beta adjustments, we focus on the concept of “beta dispersion”, or coefficient of variation. It is the standard deviation of the betas divided by their mean. Beta dispersion allows us to compare the variation in market betas across time, as the mean changes. The impact of formulas (1) and (3) on raw beta estimates is a reduction in their dispersion.

---

3Estimation error, a statistical artifact Blume referred to as “order bias”, has a secondary role in Blume (1975). Indeed, it has to be removed to reveal the more significant “regression to the mean” tendency of the betas (Klemkosky & Martin (1975), Blume (1979), Elgers, Haltiner & Hawthorne (1979), Eubank & Zumwalt (1979) and Mantripragada (1980)).

4See Elton, Gruber, Brown & Goetzmann (2009, Chap. 7) and Sparks (2010) for a survey of various beta adjustment methodologies for time-series regression.

5Primary sources for the James-Stein estimator are Stein (1955) and James & Stein (1961). An accessible application of the James-Stein estimator is Efron & Morris (1977).

6A James-Stein estimator for regression is in Efron & Hastie (2016, Formula 7.43).

7To our knowledge, the first application of the James-Stein estimator for beta adjustment is by Lavely, Wakefield & Barrett (1980).
PCA betas versus market betas.

Like the James-Stein estimator, principal component analysis (PCA) is an essential element of data science. A tool for exploring high-dimensional data sets, PCA was developed independently to fit lines and planes to points in space (Pearson 1901) and to analyze large sets of correlated variables (Hotelling 1933). A statistical analog to the market factor, the dominant PCA factor is a portfolio of securities that maximizes explained variance, normalized to have its weights squared sum to one.

We define PCA betas as the dominant PCA factor normalized to a unit mean. In public equity markets, PCA betas are generally all of the same sign, which suggests a parallel to market betas that is illustrated in Exhibit 2. This scatter plot of Barra versus PCA betas for the S&P 500 demonstrates a significant degree of correlation. This 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 most important source of risk in the US equity market.

An advantage of PCA over methods relying on observable factors (e.g., time-series regression) lies in the ability of PCA to adapt to structural changes in market conditions. This approach to factor identification can uncover emerging drivers of correlation, like the internet factor in the late 1990s, the housing factor that drove the financial crisis in 2008, or a climate factor that may soon be discovered. However, estimated PCA betas, like estimated market betas, tend to be overly dispersed. This tendency is identified in Goldberg, Papanicolaou & Shkolnik (2020). The authors characterize the amount of excess dispersion to be removed, to reveal the better betas.

Better betas and minimum variance.

Arguments that the Blume, Vasicek and similar beta 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. The comparisons test whether adjusted betas are more

---

8 A thorough exposition of PCA is Jolliffe (2002).
9 Principal component analysis is sometimes used to estimate factors in the Arbitrage Pricing Theory pioneered in Ross (1976); see also Chamberlain & Rothschild (1983).
Exhibit 2. Barra USSLOW predicted betas versus PCA betas on December 31, 2015. Both betas are renormalized to have mean one.

- Prepared by the authors from Aperio Group data.
accurate predictors than their raw counterparts.\textsuperscript{10}

An alternative approach to evaluating betas builds on their tight relationship with minimum variance portfolios, which is carefully considered in Clarke et al. (2006), Clarke, De Silva & Thorley (2011) and Goldberg, Papanicolaou & Shkolnik (2020), and illustrated in Exhibit 1. As pointed out in Michaud (1989), variance minimizers are “estimation-error maximizers.” Indeed, estimation error in betas in particular has a profound influence on optimized portfolios.\textsuperscript{11} We argue that better betas should neutralize this influence, and we propose metrics that test for such a capacity in simulation.


– Prepared by the authors from Aperio Group data.

\textbf{The challenges in betas adjustment.}

A challenge to the design of optimal beta adjustments is the dynamics of betas over time. Exhibit 3 shows the means, standard deviations and dispersions of predicted S&P 500 security betas from June, 1995 through


Exhibit 4. 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.

– Prepared by the authors from Aperio Group data.
Empirically, we see that betas vary over time. Exhibit 3 suggests that a static rule like Blume’s $2/3$-adjustment may not be optimal. But, the relationship between changes in true betas and the amount of estimation error is not apparent from empirics alone.

Another challenge relates to the impact that beta adjustments have on diversification. More adjustment, a lower $c$ in formula (1), reduces the level of beta dispersion, concentrating the betas. This tends to generate more extreme portfolio weights, and thus, more concentrated portfolios.\(^{12}\) In other words, beta shrinkage is the enemy of diversification, and a portfolio manager must choose between greater accuracy and greater concentration. The trade-off is illustrated in Exhibit 4, which shows scatter plots of minimum variance portfolio weights versus market betas in 2000, when betas were more dispersed, and in 2015, when betas were more concentrated.\(^{13}\) We conclude that there are dangers that beta adjustment will lead to minimum variance portfolios that are both inaccurate and needlessly concentrated.

**PCA betas**

Principal component analysis (PCA) is applied to a sample covariance matrix $S$ of security returns to extract factors that drive equity return and risk. As mentioned in the introduction, the dominant PCA factor for US equities is market-like, meaning that it has a relatively large variance and mostly positive exposures.\(^{14}\) The dominant PCA factor is the portfolio that explains as much variance as possible, subject to the constraint that the sum of the weights squared is one. Mathematically, it solves the optimization

$$\max_{b \in \mathbb{R}^N} b^T S b \quad \sum_n b_n^2 = 1.$$  

\(^{12}\) This relationship can be mathematically verified in a single-index model for “reasonable” calibrations, but it does not hold in all circumstances. In an atypical example of a single-index homogeneous specific variance model with all betas equal to one, any sufficiently small perturbation of the betas will concentrate a minimum variance portfolio.

\(^{13}\) The minimum variance portfolio was more diversified in 2000, with an effective security count (inverse Herfindahl index) of 39, than in 2015, with an effective security count of 16.

\(^{14}\) The market-like qualities of PCA betas from sample covariance matrices of US equity returns amount to an empirical argument for the existence of a large common factor. Any assertion that this factor is itself, the market, would require additional argument.
From the two solutions of (4), we pick the one with a positive mean, and thus more positive than negative exposures to this factor. The remaining PCA factors explain as much of the remaining variance as possible, subject to a mutual orthogonality constraint.

The variance of the dominant PCA factor is the maximum in (4), i.e.,

\[ \sigma_1^2 = b^\top S b. \]  

(5)

It is the dominant (that is, the largest) eigenvalue of \( S \). We denote by \( \sigma_2^2 \geq \sigma_3^2 \), etc., the variances of the second and third PCA factors, and so on.

Market betas are distributed about one, and so, we define PCA betas to share this property. This normalization of beta estimates also plays a role in the adjustment given by formula (1). Since its effect is to shrink the raw betas toward one, the raw estimates should have a unit mean. Consequently, with \( b \) denoting the dominant PCA factor, we define PCA betas by

\[ \beta_{\text{PCA}} = \frac{b}{\mu_b}; \quad \mu_b = \frac{\sum_n b_n}{N}. \]  

(6)

We assume the mean entry \( \mu_b \) of \( b \) 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. A scaling corresponding to (6) must be applied to the variance \( \sigma_1^2 \) of the dominant PCA factor in (5). Let

\[ \sigma^2 = \sigma_1^2 \mu_b^2. \]  

(7)

be the PCA estimate of the market variance. Both estimates ((6) and (7)) are plagued by finite sample error. But, while many methods exist to correct PCA factor variances such as \( \sigma^2 \), much less attention has been devoted to their exposures. We address this problem for PCA betas in particular.

**Beta dispersion**

Goldberg, Papanicolaou & Shkolnik (2020) show that estimated PCA betas tend to be more dispersed than their true counterparts. Here, we extend the results in that paper and adapt them for the beta adjustments in formula (1). We consider cross-sectional market betas \((\beta_n)\) of \( N \) securities, and define

\[ \tau_\beta^2 = \frac{\sum_n (\beta_n - \mu_\beta)^2}{\mu_\beta^2 N}. \]  

(8)
We call $\tau_\beta$ the dispersion of $\beta$. It is the standard deviation of the $(\beta_n)$ divided by their mean $\mu_\beta = \sum_n \beta_n / N$.\(^\text{15}\) The dispersion $\tau_\beta$ expresses the variation of the betas about one after renormalizing the entries to have unit mean. Since $\beta^{\text{PCA}}$ is already so normalized, $\tau^2_{\beta^{\text{PCA}}} = \sum_n (\beta^{\text{PCA}}_n - 1)^2 / N$.

It may be established under quite general conditions that, as the $N$ grows relative to the sample size $T$, it becomes exceedingly likely that

\begin{equation}
\tau^2_{\beta^{\text{PCA}}} > \tau^2_\beta,
\end{equation}

where $\tau_\beta$ in (8) is the dispersion of true betas in any multi-factor model. The excess dispersion of PCA betas becomes more pronounced when $T \leq N$. Such settings are frequently encountered in finance, where the histories of observed security returns can be dwarfed by the size of the market.

Motivated by (9), define an excess dispersion $\mathcal{D}_{\beta^{\text{PCA}}}$ of PCA betas by

\begin{equation}
\mathcal{D}^2_{\beta^{\text{PCA}}} = \tau^2_{\beta^{\text{PCA}}} - \tau^2_\beta,
\end{equation}

which is nearly always positive for $T \leq N$.\(^\text{16}\) The adjustment in formula (1) can shrink the excess dispersion in $\beta^{\text{raw}} = \beta^{\text{PCA}}$. The associated vector formula $\beta^{\text{adj}} = c \beta^{\text{raw}} + (1 - c)$ accomplishes this for $c \in (0, 1)$ because,

\begin{equation}
\tau^2_{\beta^{\text{adj}}} = c^2 \tau^2_{\beta^{\text{raw}}} < \tau^2_{\beta^{\text{raw}}}.
\end{equation}

Choosing an optimal $c$ for the PCA betas $\beta^{\text{PCA}}$ amounts to learning how much excess dispersion $\mathcal{D}_{\beta^{\text{PCA}}}$ there is, and how much should be removed.

Remarkably, the excess dispersion $\mathcal{D}_{\beta^{\text{PCA}}}$ can be accurately determined without knowing the true dispersion $\tau_\beta$. For $\sigma^2$, an estimate of the realized market variance, and $\delta^2_{\text{ave}}$, the estimate of the average specific variance, let

\begin{equation}
\rho^2 = \frac{T \sigma^2}{\delta^2_{\text{ave}}},
\end{equation}

which may be interpreted as a signal-to-noise ratio. This quantity can be used to estimate $\mathcal{D}^2_{\beta^{\text{PCA}}}$, and both $\sigma^2$ and $\delta^2_{\text{ave}}$ can be obtained from the variances of PCA factors, such as the dominant factor variance $s^2_1$ in (5).

\(^{15}\)This is commonly referred to as the coefficient of variation in statistics. The division by zero in (8) may be taken to mean an infinite dispersion (and $0/0 = 0$).

\(^{16}\)We slightly abuse the notation by allowing the (squared) variable $\mathcal{D}^2_{\beta^{\text{PCA}}}$ to be negative. But, as $N$ grows relative to a fixed $T$, the event $\mathcal{D}^2_{\beta^{\text{PCA}}} > 0$ becomes a certainty.
Given a single-index model, we derive the approximation

\[ \tau_2^2 \approx \tau_{\beta_{\text{PCA}}}^2 - (1/\rho^2), \]

which tends to equality as \( N \) grows relative to \( T \). A smaller signal-to-noise ratio \( \rho^2 \) leads to a greater dispersion in PCA betas. Conversely, a stronger signal, as measured by \( T\sigma^2 \), yields smaller levels of excess dispersion.

From (13) and (14), the following approximation is immediate.

\[ \mathcal{D}_{\beta_{\text{PCA}}}^2 \approx (1/\rho^2) = \delta_{\text{ave}}^2/(T\sigma^2) \]

This retains its validity for a multi-factor model, but the accuracy may be reduced when more factors are added. Extracting \( K \) largest PCA factor variances \( \delta_1^2, \ldots, \delta_K^2 \) from the returns sample covariance matrix, we have

\[ \delta_{\text{ave}}^2 = \frac{\delta^2 - (\delta_1^2 + \cdots + \delta_K^2)}{(N/T) \cdot (T - K)}, \]

where \( \delta^2 \) is the sum of the diagonal entries of the sample covariance matrix. Further, the realized market variance estimate (cf., footnote 17) is

\[ \sigma^2 = \frac{\delta_1^2/N}{1 + \tau_{\beta_{\text{PCA}}}^2}. \]

The expressions for \( \sigma^2 \) in (16) and (7) are equal, so either form suffices.

Combining these considerations with (11), it is not difficult to determine the optimal adjustment \( c \), for which formula (1) removes the excess dispersion \( \mathcal{D}_{\beta_{\text{PCA}}} \) from PCA betas. But, this is suboptimal. The optimality properties concern the accuracy of weights and risk forecasts for estimated minimum variance portfolios (Goldberg, Papanicolaou & Shkolnik 2020). Our proposed choice of adjustment parameter targets the minimization of

\[ \sum_n \left( \frac{\beta_n}{\mu_\beta} - \frac{\beta_{\text{adj}}}{\mu_{\beta_{\text{adj}}}} \right)^2 \]

over \( c \) in adjustments \( \beta_{\text{adj}} \) in formula (1). We mention that minimum variance portfolios are highly sensitive to the choice of parameter \( c \) used to adjust market beta estimates. These portfolio weights demand better betas.

---

17 This is a one-factor model where the return to the \( n \)th security is \( R_n = \beta_n \psi_M + \epsilon_n \) where \( \psi_M \) and \( \epsilon \) are the market and specific returns respectively. Approximation (13) is distribution-free subject to some moment conditions, and takes effect when \( \psi_M \) and \( \epsilon \) are uncorrelated for \( N \) large. In the context of (12), the \( \delta_{\text{ave}}^2 \) is an estimate of the average of the \( \text{VAR}(\epsilon_n) \) and \( \sigma^2 \) estimates the realized market variance. In a PCA setup, market returns \( (\psi_M^t) \) are not utilized, but the realized variance is \( \sum_t (\psi_M^t - \mu_M)^2/T \) for \( \mu_M = \sum_t \psi_M^t/T \).

18 Collinear (true) factor exposures reduce the validity of the approximation.
Better betas

We develop a PCA beta adjustment based on the theoretical results in Goldberg, Papanicolaou & Shkolnik (2020) that we term GPS. Like the adjustments of Blume and Vasicek for time-series beta estimates, the GPS adjustment is a special case of formula (1). It removes excess dispersion.

The Blume adjustment.

Blume’s rule is associated with a static choice of parameter $c$ in formula (1) to compute the adjusted betas. This choice is

$$c^{\text{Blume}} = \frac{2}{3}. \tag{18}$$

This oversimplification was not Blume’s intention, yet, the wide adoption of (18) and (2) is perhaps due to the rule’s simplicity and apparent universality.

In Blume’s actual procedure (see Blume (1971)), estimates of beta 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, their mean (cf., Exhibit 3). The value $2/3$ turns out to be approximately equal to this historical rate of regression for the period Blume studied. As noted in Blume (1971), this rate of regression changes over time. But, Blume (1971, pg. 9) also points out that, “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.” This provides some justification for the static choice in (18), but not for reasons involving estimation error. Blume’s rule corrects for changes in true betas.\(^{20}\)

The Vasicek adjustment.

Estimation error in betas is the primary focus of Vasicek (1973). Vasicek’s rule relies on a Bayesian prior for the market beta that is Normal with mean $\mu_\beta$ and variance $s_\beta^2$. Assuming the security returns are Gaussian, Vasicek derived a posterior distribution of the betas. For $\mu_\beta = 1$, we

\(^{19}\)Estimates for several periods in 1933–1961 are provided in Blume (1971, Table 4).

\(^{20}\)Blume (1975) explains an approach to disentangling the movement of the true market betas from the estimation error of beta estimates, concluding that estimation error is negligible. This finding led to a vigorous debate in the literature.
obtain an adjustment in formula (1) with $c$ given by
\[
(19) \quad c_{\text{Vasicek}} = \frac{s_{\beta}^2}{s_{\beta}^2 + s_{\text{raw}}^2},
\]
where $s_{\text{raw}}$ is the standard error of the (raw) regression estimate $\beta_{\text{raw}}$. In this time-series setting, the adjustment is applied to each security separately. Each $c_{\text{Vasicek}}$ is security specific. When $\mu_\beta$ is thought to deviate from one, shrinkage towards it (or its estimate) per formula (3) is recommended.

Central to Vasicek’s approach is prior knowledge of $\beta$, and in particular, the $s_\beta^2$ in (19) is unknown. One can leverage time-series for many securities to obtain (cross-sectional) estimates of the prior $s_\beta^2$. However, this estimate would itself be biased, reducing the effectiveness of the adjustment. Alternatively, reasonable priors may be hypothesized. Vasicek suggests taking $\mu_\beta = 1$ and $s_\beta^2 = 0.5$ for the NYSE betas (Vasicek (1973, Section III)).

The GPS adjustment.

It turns out that the parameter $c$ for the Vasicek and GPS adjustments can be expressed in the same general form. We set
\[
(20) \quad c = \frac{\tau_\beta^2}{\tau_\beta^2 + \theta_{\beta\text{PCA}}^2}
\]
in formula (1) with the raw estimate $\beta_{\text{raw}} = \beta_{\text{PCA}}$ in (6). This similarity to the Vasicek adjustment parameter in (19) is evident, but there are two noteworthy differences. First, the security specific, standard error $s_{\text{raw}}$ in (19) is replaced by the excess dispersion $\theta_{\beta\text{PCA}}$ in (14), which is the same for all securities. Second, Vasicek’s prior on the standard deviation of the market betas, $s_\beta$, is naturally replaced in (20) by the dispersion $\tau_\beta$ of the true market betas in (8).

Since $\theta_{\beta\text{PCA}}^2 = \tau_{\beta\text{PCA}}^2 - \tau_\beta^2$, we have $c = \tau_\beta^2/\theta_{\beta\text{PCA}}^2$, but this choice is not directly implementable as $\tau_\beta$ is not known. To carry out the GPS adjustment in practice, we apply the estimate for $\tau_\beta^2$ in (13) to approximate (20) by
\[
(21) \quad c_{\text{GPS}} = 1 - \frac{1}{\tau_{\beta\text{PCA}}^2 \cdot \rho^2}
\]
where $\rho^2 = T\sigma^2/\delta_{\text{ave}}^2$, is the signal-to-noise ratio in (12).\(^{21}\) The $\delta_{\text{ave}}^2$ and $\sigma^2$ are computed via (15) and (16) from the PCA factor variances. Note, the

\(^{21}\)It is plausible, but unlikely when $T \leq N$, that $c_{\text{GPS}}$ is negative. This indicates that $\tau_\beta$ is close to zero and taking $c_{\text{GPS}} = 0$ may be preferred, but not necessary.
true dispersion $\tau_\beta$ cannot be naively estimated by $\tau_{\beta^{\text{PCA}}}$ because $\beta^{\text{PCA}}$ is overly dispersed, and doing so leads to the choice $c = 1$, i.e., no adjustment.

It is important to mention that adjusting $\beta^{\text{PCA}}$ by using $c = c^{\text{GPS}}$ does not remove all the excess dispersion, $\mathcal{D}_{\beta^{\text{PCA}}}$. In fact, it removes more, as

$$
\tau^2_\beta - \tau^2_{\beta^{\text{GPS}}} \approx \frac{c^{\text{GPS}}}{\rho^2}
$$

which is nearly always positive for $N$ large. Choosing $c$ for the adjustment $\beta^{\text{adj}}$ of $\beta^{\text{raw}} = \beta^{\text{PCA}}$ as the square-root of $c^{\text{GPS}}$, produces $\tau^2_{\beta^{\text{adj}}} \approx \tau^2_\beta$. Instead, the parameter $c$ in (21) targets the minimization of (17) over all $c$ in $\beta^{\text{adj}}$.

**Minimum variance**

The composition of a minimum variance portfolio is governed predominantly by the distribution of market betas, and the weights of these portfolios are highly sensitive to estimation error. It is for these two reasons that we test the impact of better betas on minimum variance in simulation.

**Portfolio construction.**

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

$$
\min_{x \in \mathbb{R}^N} \omega^\top V \omega
$$

subject to:

$$
\sum_n \omega_n = 1,
$$

where $V$ is the true $N \times N$ covariance matrix of security returns. We use $\omega^{\text{opt}}$ to denote the optimal weights computed via (23) and write $\omega$ for the weights computed with an estimated covariance matrix $\Sigma$ in place of $V$. In practice, estimates $\Sigma$ are all we have available. In simulation experiments that we perform, we are afforded the knowledge of the true covariance $V$.

**Accuracy metrics.**

Estimation error distorts the minimum variance weights and more specifically, it biases the risk of these portfolios downward. The metrics we use to quantify estimation error can be used only in simulation, as they rely on the true covariance matrix $V$ of the security returns. We define,

$$
\mathcal{F}^2(\omega) = (\omega^{\text{opt}} - \omega)^\top V (\omega^{\text{opt}} - \omega),
$$
the squared tracking error of an estimated portfolio $\omega$. It measures the distance between the optimal and estimated portfolios, $\omega_{opt}$ and $\omega$, as the square of the width of the distribution of the return difference of $\omega_{opt} - \omega$.

Next, we define the variance forecast ratio $R(\omega)$ of an estimated portfolio $\omega$. Note that the estimated variance of portfolio $\omega$ is given by $\omega^T \Sigma \omega$, and its realized variance is $\omega^T V \omega$. Consequently, we consider their ratio

$$R(\omega) = \frac{\omega^T \Sigma \omega}{\omega^T V \omega}.$$  

(25)

The ratio is smaller than one when the risk of the portfolio $\omega$ is underforecast. This is the case when the sample covariance matrix is used for the estimate $\Sigma$. While a PCA-estimated covariance matrix improves over this naive choice, as we shall see, underforecasts persist.

Tracking error $T$ penalizes mistakes in the weights of the riskiest of securities, and the variance ratio $R$ quantifies mistakes in portfolio risk forecasts. One can show (24) stays above zero and (25) tends to zero as $N$ grows for most estimates $\Sigma$. To avoid this, very specific and precise adjustments are required. Somewhat surprisingly, beta adjustments have such a capacity.

Calibration of market conditions

To illustrate how beta adjustments affect the accuracy of estimated minimum variance portfolios, we specify and calibrate a security return generating model. Simulation allows for a direct comparison of estimates to ground truth, but conclusions drawn are useful only to the extent that the output of the model is realistic. We calibrate a factor model to two market conditions, calm and stressed. It is reasonable that the beta adjustment parameter in formula (1) should change with changing conditions. In the next section, we investigate if this holds for our calm and stressed scenarios in simulation.

In this section, we specify a multi-factor model for generating security returns and then calibrate it to calm and stressed market scenarios.

Return generating model and covariance matrix.

par We consider a multi-factor model with $K$ factors in which the excess return $R_n$ to the $n$th security is

$$R_n = \sum_k B_{nk} \psi_k + \epsilon_n,$$  

(26)

where $\psi_k$ is the return to the $k$th factor and $\epsilon_n$ is the specific return to the $n$th security. The constant $B_{nk}$ is the exposure of the $n$th security to the
The sum in (26) ranges over $1 \leq k \leq K$ factors. We have,

$$
(27) \quad \beta = (\beta_n) \quad \beta_n = B_{n1}, \quad n = 1, \ldots, N
$$

constituting the true market betas of the $N$ securities in our model.

Under our assumption that the $(\psi_k, \epsilon_n)$ are pairwise uncorrelated, the (true) $N \times N$ covariance matrix $V$ of the security returns takes the form

$$
(28) \quad V = BB^T + \Delta
$$

where $B = (B_{nk})$ is a $N \times K$ matrix of exposures, and the $(\Psi, \Delta)$ are covariance matrices of the factor and specific returns that we assume to be diagonal with entries $\text{VAR}(\psi_1), \ldots, \text{VAR}(\psi_K)$ and $\text{VAR}(\epsilon_1), \ldots, \text{VAR}(\epsilon_N)$, respectively.

<table>
<thead>
<tr>
<th></th>
<th>Calm market</th>
<th>Stressed market</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market volatility</td>
<td>15</td>
<td>30</td>
</tr>
<tr>
<td>Style volatility</td>
<td>${2, 4, 8}$</td>
<td>${2, 4, 8}$</td>
</tr>
<tr>
<td>Specific volatility</td>
<td>(20, 50)</td>
<td>(40, 70)</td>
</tr>
<tr>
<td>Beta dispersion</td>
<td>1/5</td>
<td>1/3</td>
</tr>
<tr>
<td>Average correlation</td>
<td>0.16</td>
<td>0.22</td>
</tr>
<tr>
<td>Portfolio volatility</td>
<td>6.48</td>
<td>7.25</td>
</tr>
</tbody>
</table>

**Exhibit 5.** Simulation parameters for the calm and stressed market regimes. All factor and specific volatilities are %-annualized. The average pairwise correlation and the volatility of the minimum variance portfolios are implied by the model with 500 securities.

**Specifying the calm and stressed regimes.**

Exhibit 5 summarizes the calibration of model (26) to our two market scenarios. It follows the empirical observations from Barra models (Bayraktar, Mashtaler, Meng & Radchenko 2014).

The market volatility (square-root of $\text{VAR}(\psi_1)$) increases from the calm to the stressed regime in accordance with Barra forecasts. We take
$K = 4$ factors total, with the three non-market factors modeled on equity styles such as volatility, earnings yield and size.\textsuperscript{22} Bayraktar et al. (2014, Table 4.1) provides guidance on the volatility of equity style factors, with estimates typically under 10% per year. While these factors undergo volatility regimes, they need not coincide with the market (van Dijk 2011, Figure 2). Thus, we set the style factor volatilities (square-roots of $\text{Var}(\psi_2), \text{Var}(\psi_3), \text{Var}(\psi_4)$) to the same values in both market regimes. The specific volatilities tend to increase materially with more turbulence (Bayraktar et al. 2014, Figure 4.5). We draw their values uniformly from the intervals specified by Exhibit 5.

All factor exposures ($B_{nk}$) are drawn independently from the normal distribution. Since cap-weighted market betas are necessarily distributed about one, we standardize the mean of the betas ($\beta_n$) in (27) to this value. We standardize their dispersion $\tau_\beta$ (see (8)) to the two values in Exhibit 5. Sefton, Jessop, De Rossi, Jones & Zhang (2011, Chart 2) indicates higher beta dispersion during stressed market times and we adopt this guideline.\textsuperscript{23} For the style factors (last three columns of $B$), the exposures ($B_{nk}$) are standardized to be mean zero and variance one (see Bayraktar et al. (2014)).

The last two rows of Exhibit 5 show some consequences of our calibration. The average pairwise security correlation and the volatility of the minimum variance portfolio both increase with stressed market conditions.\textsuperscript{24}

\section*{Better betas in any weather}

We examine the properties of the GPS adjustment for PCA betas under varying market conditions. Does adjusting PCA betas dynamically, based on a calm or stressed market scenario, yield benefits over a static adjustment, such as Blume’s rule? We take advantage of the sensitivity of minimum variance portfolio weights to estimation error to shed light on this question.

\textsuperscript{22}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).

\textsuperscript{23}However, Exhibit 3 points to a more complicated behavior of the dispersion of market betas. Notably, betas compressed amidst the market turbulence during the first half of 2020, as discussed in Goldberg, Ulucam, Shkedi, Branch, Hand, Leshem & Tymoczko (2020).

\textsuperscript{24}See Clarke et al. (2006), Clarke et al. (2011) and Goldberg, Leshem & Geddes (2014) for empirical properties of minimum variance portfolios.
Experimental design.

We simulate $T = 100$ i.i.d. Gaussian observations of returns to $N = 500$ securities from a $K = 4$ factor model based on (26). We write (26) as $R_{nt} = \sum_k B_{nk} \psi_k + \epsilon_{nt}$ indexed by the $n$th security and the $t$th observation. The calibration of this model to the calm and stressed market regimes is summarized in Exhibit 5. In each simulation path, we assemble three estimates of the (true) covariance matrix $V$ in (28). One utilizes PCA with no beta adjustment. The second is based on a time-series regression with Blume’s adjustment and the third, on the GPS beta adjustment for PCA.

1. PCA observes only the security returns ($R_{nt}$) (neither the factor nor the specific return are used). A sample covariance matrix $S$ is constructed from the returns. Four factors and their variances are extracted from $S$. The PCA betas are the exposures to the dominant factor normalized per (6). The specific variances are obtained by regression.

2. Blume’s rule utilizes the observations of the security return ($R_{nt}$) and the factor return ($\psi_{kt}$). The latter yield unbiased estimates of the factor variances. The factor exposures ($B_{nk}$) are estimated by a standard time-series regression. The ($B_{n1}$) form the raw beta estimates $\beta^{\text{raw}}$, to which formula (2) is applied (i.e., the 2/3’s rule). Given the adjusted factor exposures, the specific variances are estimated the same way as for PCA.

3. The GPS adjustment picks up where PCA leaves off. It does not have the benefit of observing the factor returns (as Blume’s adjustment does). It proceeds by adjusting the PCA betas $\beta^{\text{PCA}}$ by applying the formula

$\beta^{\text{GPS}} = c^{\text{GPS}} \beta^{\text{PCA}} + (1 - c^{\text{GPS}})$

where $c^{\text{GPS}} = 1 - 1/(\tau_3^{\beta^{\text{PCA}}} \cdot \rho)^2$ per (21) for a signal-to-noise ratio $\rho^2$ in (12). The rest of the estimates are routinely adjusted given this modification.

Given the estimates $\Sigma^{\text{PCA}}$, $\Sigma^{\text{Blume}}$ and $\Sigma^{\text{GPS}}$ of the true covariance $V$, the accuracy metrics (24) and (25) are readily computed. We run $10^5$ such simulation paths to obtain granular estimates of the distributions of all metrics and parameters. All parallelized simulations are implemented in Python and completed on a 8-core CPU (1.90GHz) in under one hour.

Simulation results.

We examine the dynamic aspects of the GPS adjustment. A key marker, that tracks the excess dispersion $\mathcal{D}_{\beta^{\text{PCA}}}^2 = \tau_{\beta^{\text{PCA}}}^2 - \tau_3^2$ in PCA betas, is the signal-to-noise ratio $\rho^2 = T\sigma^2/\delta_{\text{ave}}^2$. Then $\mathcal{D}_{\beta^{\text{PCA}}}^2 \approx 1/\rho^2$ by (13), and
from Exhibit 5, we calculate the idealized values involved in GPS. These values employ formula (20) for $c_{GPS}$ and are collected in Exhibit 6.\(^{25}\)

<table>
<thead>
<tr>
<th></th>
<th>Calm</th>
<th>Stressed</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho^2$</td>
<td>18.4</td>
<td>29.8</td>
</tr>
<tr>
<td>$\mathcal{D}<em>{\beta}^2</em>{PCA}$</td>
<td>0.054</td>
<td>0.034</td>
</tr>
<tr>
<td>$\tau_{\beta}^2$</td>
<td>0.040</td>
<td>0.111</td>
</tr>
<tr>
<td>$c_{GPS}$</td>
<td>0.424</td>
<td>0.768</td>
</tr>
</tbody>
</table>

**Exhibit 6.** The idealized parameters involved in the GPS adjustment.

The simulated distributions corresponding to the values in Exhibit 6 are pictured in Exhibit 7. We see the idealized values aligning relatively well with the centers of the simulated distributions. The agreement is better for the stressed regime, which exhibits less excess dispersion $\mathcal{D}_{\beta}^2_{PCA}$ in the PCA betas. This is due to a higher signal-to-noise ratio in the stressed market. The width of $\mathcal{D}_{\beta}^2_{PCA}$ is also smaller in the stressed regime, which suggests that smaller but more precise adjustments are needed in the stressed scenarios of our model. This is confirmed by the simulated distributions of $c_{GPS}$. Larger and more concentrated $c_{GPS}$ values try to correct smaller and more concentrated dispersion levels under stress. This finding aligns with the fact that $c_{GPS}$ must increase with $\rho \tau_{\beta}$ on average. Larger values of $\rho^2 \tau_{\beta}^2 = T \sigma^2 / \delta_{ave}^2$, and hence $c_{GPS}$, imply more trust in PCA betas by the GPS adjustment per (29). The static Blume 2/3-rule strikes a reasonable balance between the $c_{GPS}$ values in the two regimes.

We next examine the performance of these adjustments. Distributions of the tracking error and the variance ratio of estimated minimum variance portfolios are displayed in Exhibits 8 and 9. We see that unadjusted PCA has by far the worst performance on all metrics. It does not have the benefit of observing the factor returns like Blume’s method, and its betas are not adjusted. Its minimum variance weights are significantly affected by estimation error. We observe significant portfolio risk underforecasts by PCA in both sets of market conditions in Exhibit 8. In the calm market scenario, PCA forecasts less than 25% of the realized portfolio risk. Remarkably, the calm market calibration presents a

\(^{25}\)These are not the optimal values to use for the GPS beta adjustment. As an example, the idealized value of $\rho^2 = T \sigma^2 / \delta_{ave}^2$ is calculated as 100 · (15/35)\(^2\) and 100 · (30/55)\(^2\) in the two regimes. However, the adjustment should use an estimate $\bar{\sigma}^2$ of the realized market variance, not the true variances (cf., footnote (17) the optimal $\rho^2$ is truly data-driven).
Exhibit 7. Distributions of the signal-to-noise ratio $\rho^2 = T\sigma^2/\delta_{\text{ave}}^2$, excess dispersion squared $\mathcal{D}^2_{\text{PCA}} = \tau^2_{\text{PCA}} - \tau^2_\beta$ and parameter $c^{\text{GPS}} = 1 - 1/(\tau_{\beta_{\text{PCA}}} \cdot \rho)^2$. 

22
Exhibit 8. Variance ratio $\mathcal{R}(\omega)$ distributions for minimum variance portfolio weights estimated with PCA, Blume’s method and the GPS adjustment.

A more challenging scenario, not just for PCA, but for all methods. Blume’s method outperforms on tracking error in the calm regime, but forecasts only 50% of the true portfolio risk. In the stressed market, Blume’s rule provides accurate risk forecasts, but does not improve on PCA tracking error significantly. The GPS adjustment performs well on all metrics and scenarios. It is noteworthy that its (relatively) large, concentrated values of $c_{GPS}^{G}$ during market stress, manage to make the small but precise adjustments required to correct the PCA betas.

Exhibit 9. Tracking error $\mathcal{T}(\omega)$ distributions (% units) for minimum variance portfolio weights estimated with PCA, Blume’s method and the GPS adjustment. The optimal variance $\omega_{opt}^T V_{\omega_{opt}}$ lies on the horizontal line.
Implementation

For reproducibility, we provide a complete specification of our simulations.

Principal component analysis estimates.

Take the sample covariance matrix $S = (S_{ij})$ of the security returns as our starting point. With $R_{nt}$ denoting the return to the $n$th security that is observed at time $t$, using the centralized return $Y_{nt} = R_{nt} - \sum_t R_{nt}/T$, (30)

$$S_{ij} = \sum_t Y_{it}Y_{jt}/T$$

where the sum ranges over times $t = 1, \ldots, T$ and $1 \leq i, j \leq N$. In principal component analysis (PCA) for equities, the eigenvectors and eigenvalues of $S$ are identified with the factors that drive risk. An exposition of PCA for equity risk factor models is found in Litterman et al. (2003, Chapter 20).

Assuming $K$ factors in our model, we extract $K$ eigenvectors (called principal components) denoting them by $b^{(1)}, \ldots, b^{(K)}$, and ordering them by the associated eigenvalues $s_{1}^{2} \geq \cdots \geq s_{K}^{2}$ (called variances). We have,

$$s_{k}^{2} = \sum_{ij} b_{i}^{(k)} S_{ij} b_{j}^{(k)}$$

and the associated factor $b^{(k)}$ is unique up to sign. By our convention, we take the dominant PCA factor $b^{(1)}$ to be positively oriented, i.e., $\sum_n b_n \geq 0$.

The PCA estimate of the covariance for a $K$-factor model takes the form

$$\Sigma^{PCA} = B^{PCA} \Psi^{PCA} (B^{PCA})^\top + \Delta^{PCA}.$$ 

The $K \times K$ factor variance matrix $\Psi^{PCA}$ consists of variances $(s_{1}^{2}, \ldots, s_{K}^{2})$ on the diagonal and zeros elsewhere. The $N \times K$ exposure matrix $B^{PCA}$ has the columns $(b^{(1)}, \ldots, b^{(K)})$. Regressing the returns $(Y_{ij})$ onto the factors in $B^{PCA}$ yields the $N \times T$ matrix of specific returns with entries $(Z_{nt})$.²⁶

The specific variance $\delta_{n}^{2}$ of the $n$th security is computed as the row average

$$\delta_{n}^{2} = \sum_t Z_{nt}^{2}/T.$$ (32)

Lastly, $\Delta^{PCA}$ is a $N \times N$ matrix with $\delta_{1}^{2}, \ldots, \delta_{N}^{2}$ on the diagonal and zero elsewhere. Note, that PCA betas did not enter the covariance matrix. This is accomplished, without altering $\Sigma^{PCA}$, by replacing $b = b^{(1)}$ in $B^{PCA}$ by

$$\beta^{PCA} = \frac{b}{\mu_b}; \quad \mu_b = \sum_n b_n / N.$$ (33)

²⁶A regression of $Y$ onto $B$ has $Z = Y - B(B^\top B)^{-1}B^\top Y$ for $N \times T$ matrices $Z = (Z_{ij})$ and $Y = (Y_{ij})$. For PCA, conveniently, $(B^{PCA})^\top B^{PCA} = I$ is the $K \times K$ identity matrix.
and the variance $s^2_1$ in the first entry of $\Psi^{\text{PCA}}$ by the market variance

$$\sigma^2 = s^2_1 \mu_0^2 = \frac{s^2_1/N}{1 + \tau^2_{\beta^{\text{PCA}}}}. \tag{34}$$

Again, these substitutions do not change $\Sigma^{\text{PCA}}$ itself, only $B^{\text{PCA}}$ and $\Psi^{\text{PCA}}$.

**GPS adjustment for PCA estimates.**

Here we recapitulate the GPS adjustment from “Better Betas” and assemble the $K$-factor model covariance matrix. Taking $\beta^{\text{PCA}}$ from (33),

$$\beta^{\text{GPS}} = c^{\text{GPS}} \beta^{\text{PCA}} + (1 - c^{\text{GPS}}) e \tag{35}$$

where $e$ is the $N$-vector all ones and the parameter $c^{\text{GPS}}$ as in (21) reads

$$c^{\text{GPS}} = 1 - \frac{1}{\tau^2_{\beta^{\text{PCA}}} \cdot \rho^2}. \tag{36}$$

The $c^{\text{GPS}}$ may be taken to be zero in the rare event that the right side is negative. As before $\tau^2_{\beta^{\text{PCA}}} = \sum_n (\beta_n^{\text{PCA}} - 1)^2/N$ and the signal-to-noise ratio $\rho^2$ is

$$\rho^2 = \frac{T \sigma^2}{\delta^2_{\text{ave}}} \tag{37}$$

as in (12) with $\sigma^2$ originally in (16), reproduced above in (34), and $\delta^2_{\text{ave}}$, originally appearing in (15), is assembled from the eigenvalues of $S$ in (5) as

$$\delta^2_{\text{ave}} = \frac{\delta^2 - (\delta^2_1 + \cdots + \delta^2_K)}{(N/T) \cdot (T - K)}. \tag{38}$$

Here, the quantity $\delta^2$ is the sum of the diagonal entries of $S$.

It remains to assemble the $K$-factor covariance matrix of the form

$$\Sigma^{\text{GPS}} = B^{\text{GPS}} \Psi^{\text{GPS}} (B^{\text{GPS}})^\top + \Delta^{\text{GPS}}$$

The $K \times K$ factor variance matrix $\Psi^{\text{GPS}}$ has $(\sigma^2, \delta^2_2, \ldots, \delta^2_K)$ on its diagonal and is zero elsewhere, for $\sigma^2$ the market variance in (34). The corresponding $N \times K$ factor exposure matrix is $B = (\beta^{\text{GPS}}, b_2, \ldots, b_K)^{\top}$.

Note, we use

$^{27}$The PCA factors $(b_2, \ldots, b_K)$ may be adjusted analogously to $\beta^{\text{PCA}} = b_1/\mu_b_1$, but with diminishing returns. These factors are orthogonal and so, the second, third and remaining PCA factors are not market-like. Adjusting them will yield little additional benefit.
the unit-mean normalization on the dominant factor. Regressing returns $(Y_{ij})$ onto the factors $B^{GPS}$ yields the $N \times T$ matrix of specific returns with entries $(Z_{nt})$ (cf., footnote 26).

The specific variance $\delta^2_n$ of the $n$th security is computed as the $n$th row average of $(Z_{nt})$ adjusted so that $\sum_n \delta^2_n / N = \delta^2_{\text{ave}}$ in (38). That is,

$$\delta^2_n = \left( \frac{\sum_t Z^2_{nt}}{T} \right) - \left( \frac{\sum_n \sum_t Z^2_{nt}}{NT} \right) + \delta^2_{\text{ave}}.$$ 

Taking $\Delta^{GPS}$ to be the $N \times N$ matrix with $\delta^2_1, \ldots, \delta^2_N$ on the diagonal and zero elsewhere, we obtain the adjusted covariance matrix estimate $\Sigma^{GPS}$.

Some limitations and open problems

We endow the estimates in our simulation with the knowledge of the true number of factors, and estimating it would be more realistic. It is inevitable that the use of an algorithm to judge the number of factors would add to the overall level of estimation error. Relying again on Exhibit 1 for guidance, suggests that, given the dominance of the market-like factor, our adjustment may retain its effectiveness in this more challenging setting.

The Gaussian distribution used to generate factor returns, and especially specific returns, is unrealistic. The theory developed in Goldberg, Papanicolaou & Shkolnik (2020) and experiments with simulations that rely on heavy-tailed returns indicate that the dispersion bias is still present and, to a great extent, correctable. The theory requires finite fourth moments of the returns however, and larger portfolios are needed in practice. The accuracy of the GPS adjustment increases with the number of securities.

Measures beyond a beta dispersion bias correction need to be taken in order to generate even more accurate risk forecasts. Dispersion involves only the first two moments of the beta distribution, which has tended to be right skewed and heavy-tailed. This suggests that there may be a non-linear adjustment that improves upon formula (1). Future research could examine these issues, both through simulation and empirical study. The analysis would benefit from the initial benchmarking exercise carried out here, and hence merits treatment in a separate forum. Coincidentally, there is literature on non-linear adjustments for factor variances (i.e., eigenvalues; see Ledoit & Wolf (2017b) and Ledoit & Wolf (2017a), for example).

Counter-intuitively, the estimated variance $\sigma^2$ of the dominant PCA factor plays little to no role in the weights of minimum variance portfolios.\(^{29}\)

\(^{28}\)For example, Bera & Kannan (1986) suggest adjusting square-roots of betas.

\(^{29}\)This observation, communicated to us by Stephen Bianchi, was one of the primary
Their dependence on $\sigma^2$ vanishes as the number of securities grows. However, this estimate is needed to implement the GPS adjustment. Moreover, it does influence the risk forecasts of the equally-weighted portfolio. Further study is required to understand the impact of model parameters, such as $\sigma^2$, on the different aspects of portfolio construction and risk forecasting.

Finally, there are interesting connections between the GPS beta adjustment and the popular covariance matrix shrinkage approach of Ledoit & Wolf (2004) and a series of related, more technical articles. Covariance matrix shrinkage implicitly enacts a beta adjustment. This observation deserves a thorough comparison to the adjustment 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 believe that betas are better if they guard against the impact of estimation error on optimized portfolios. The weights of these portfolios are highly sensitive to incorrect estimates of market betas. This perspective facilitates the development of the GPS adjustment, which lowers excess dispersion in PCA betas. It is designed specifically for minimum variance portfolios, but has other optimality properties. The procedure can be viewed as an analog of the Vasicek’s adjustment of time-series beta estimates. The efficacy of the GPS adjustment is a step toward making PCA-based risk models useful to financial practitioners. It also sheds light on the meaning of “better betas.”

We extend the GPS adjustment in Goldberg, Papanicolaou & Shkolnik (2020) to more general models, and develop simple formulas that are easily implemented. We illustrate the dynamic nature and the effectiveness of the GPS adjustment in a simulation. For this, we calibrate the simulation to two different market scenarios, and demonstrate how errors in betas affect the accuracy of minimum variance portfolio weights and risk forecasts.

Four features distinguish the GPS adjustment from its analogs. The first is that it addresses PCA, a powerful data-driven approach for factor identification. The second is the precise notion of optimality. The GPS adjustment mitigates the impact of estimation error on minimum variance portfolios. The third is the purely data-driven formulation of the GPS adjustment. No knowledge of the true betas is required. The fourth is clues that led us to our formulation of excess dispersion for PCA betas.
flexibility. The GPS adjustment adapts to volatility regimes. This point is essential, given the empirical observation that distributions of betas vary over time.

Acknowledgement

We are grateful to Alec Kercheval for substantial contributions to the development of the ideas in this article. We thank Torben Andersen, Hubeyb Gurdogan, Ravi Jagannathan, David He, Bob Korajczyk, Olivier Ledoit, Peter Shepard, Viktor Todorov and Matti von Turk for insightful dialogs. 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. We are grateful to MSCI for the use of the Barra USSLOW model and the Aegis Optimizer.

References


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


