Chapter 12
Credit Rating Score Analysis

Wolfgang Karl Härdle, K.F. Phoon and D.K.C. Lee

Abstract We analyse a sample of funds and other securities each assigned a total rating score by an unknown *expert* entity. The scores are based on a number of risk and complexity factors, each assigned a category (factor score) of Low, Medium, or High by the expert entity. A principal component analysis of the data reveals that based on the chosen risk factors alone we cannot identify a single underlying latent source of risk in the data. Conversely, the chosen complexity factors are clearly related to one or two underlying sources of complexity. For the sample we find a clear positive relation between the first principal component and the total expert score. An attempt to match the securities' expert score by linear projection of their individual factor scores yields a best case correlation between expert score and projection of 0.9952. However, the sum of squared differences is, at 46.5552, still notable.

W.K Härdle (✉)
C.A.S.E.-Center for Applied Statistics and Economics, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany
e-mail: haerdle@hu-berlin.de

W.K. Härdle
School of Economics, 6th Level, Research fellow in Sim Kee Boon Institute for Financial Economics, Singapore Management University, 90 Stamford Road, Singapore 178903, Singapore

K.F. Phoon
School of Business, Singapore University of Social Sciences, Singapore, Singapore
e-mail: kfphoon@suss.edu.sg

D.K.C. Lee
Singapore University of Social Sciences, Singapore, Singapore
e-mail: Davidleekc@suss.edu.sg

W.K. Härdle et al. (eds.), *Applied Quantitative Finance*, Statistics and Computing,
DOI 10.1007/978-3-662-54486-0_12

12.1 Introduction

We are provided with a sample of $n = 100$ funds and other securities that have been assigned a rating score by an unknown *expert entity* – the expert (rating) score in the following. We assume the rating score to depend on a set of six risk factors and five complexity factors, each modelled as random variables on an ordinal scale of Low, Medium, High. The risk factors are volatility, liquidity, credit rating, duration/cash flow, leverage, and diversification degree. The complexity factors comprise of the number of structural layers, expansiveness of derivatives, availability and known pricing models, number of return outcome scenarios, and transparency/ease of understanding. In addition to the rating score, we know the category (i.e. Low, Medium, High) assigned to each factor for any given security included in the sample. Figures 12.1 and 12.2 show histograms for each of the risk and complexity factors, respectively.

To get a better impression regarding the relation between individual securities in the sample, we perform cluster analyses based on (i) only the risk factors, (ii) only the complexity factors, and (iii) both risk and complexity factors in the sample. In particular, we apply the Ward clustering algorithm using an Euclidean distance matrix. This algorithm is chosen to ensure that individual clusters are as homogenous as possible. However, other algorithms such as the single linkage or complete linkage

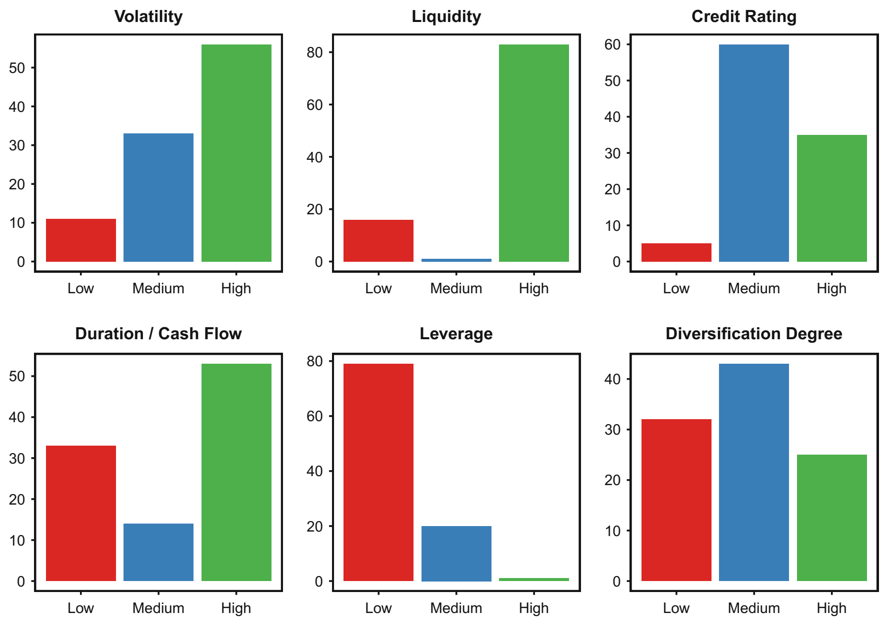

Fig. 12.1 Histograms of risk factor scores

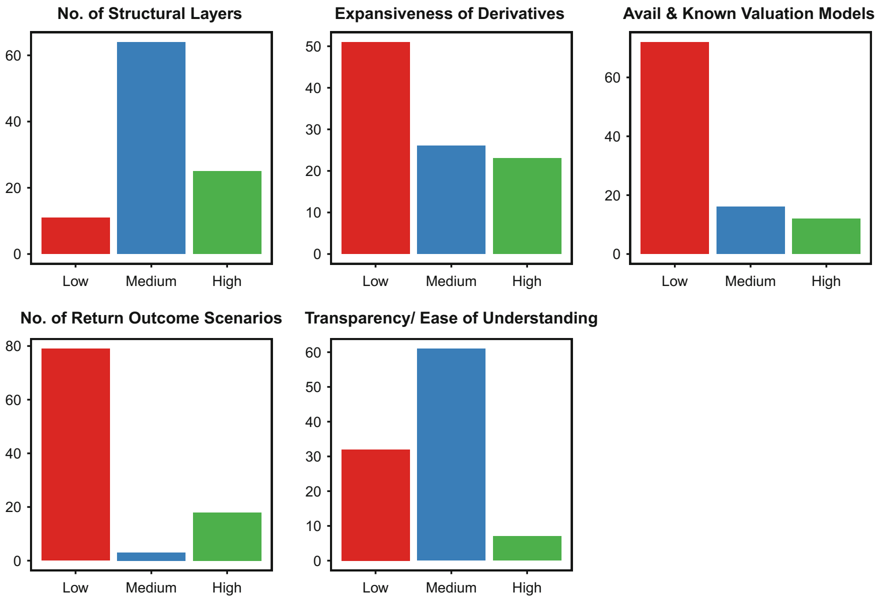

Fig. 12.2 Histograms of complexity factor scores

algorithms can be applied as well Härdle and Simar (2015). The results are depicted in Fig. 12.3.

12.2 Principal Components Analysis of Factor Scores

Principal components analysis (PCA) allows for the identification of uncorrelated latent factors that drive the variation in a sample of multivariate random variables. We consider a random variable $Y = (Y_1, \dots, Y_j, \dots, Y_k)^\intercal$ with $Y_j \in \{Low, Medium, High\}, 1 \le j \le k$. Y represents a vector of the risk and complexity categories assigned to a security i by the expert entity. To later be able to perform PCA on our sample we assign a discrete scale $\{1, 2, 3\}$ to each Y_j yielding a random variable $X = (X_1, \dots, X_j, \dots, X_k)^\intercal$ with $X_j \in \{1, 2, 3\}$, $1 \le j \le k$ (i.e. $Y_j = High$ is equivalent to $X_j = 3$). For easier reference let us refer to each of the X_j as a factor score.

Our sample is now represented by a discrete matrix $X \in \{1, 2, 3\}^{n \times k}$, with each row i representing a security and each column j representing a factor. The element $x_{i,j}$ is therefore security i's score for the j-th factor. We still cannot apply PCA to X directly, however, without violating the basic assumption of normally distributed continuous random variables made in PCA. To circumvent this issue, we apply a discrete PCA using the polychoric correlation matrix of the factor scores.

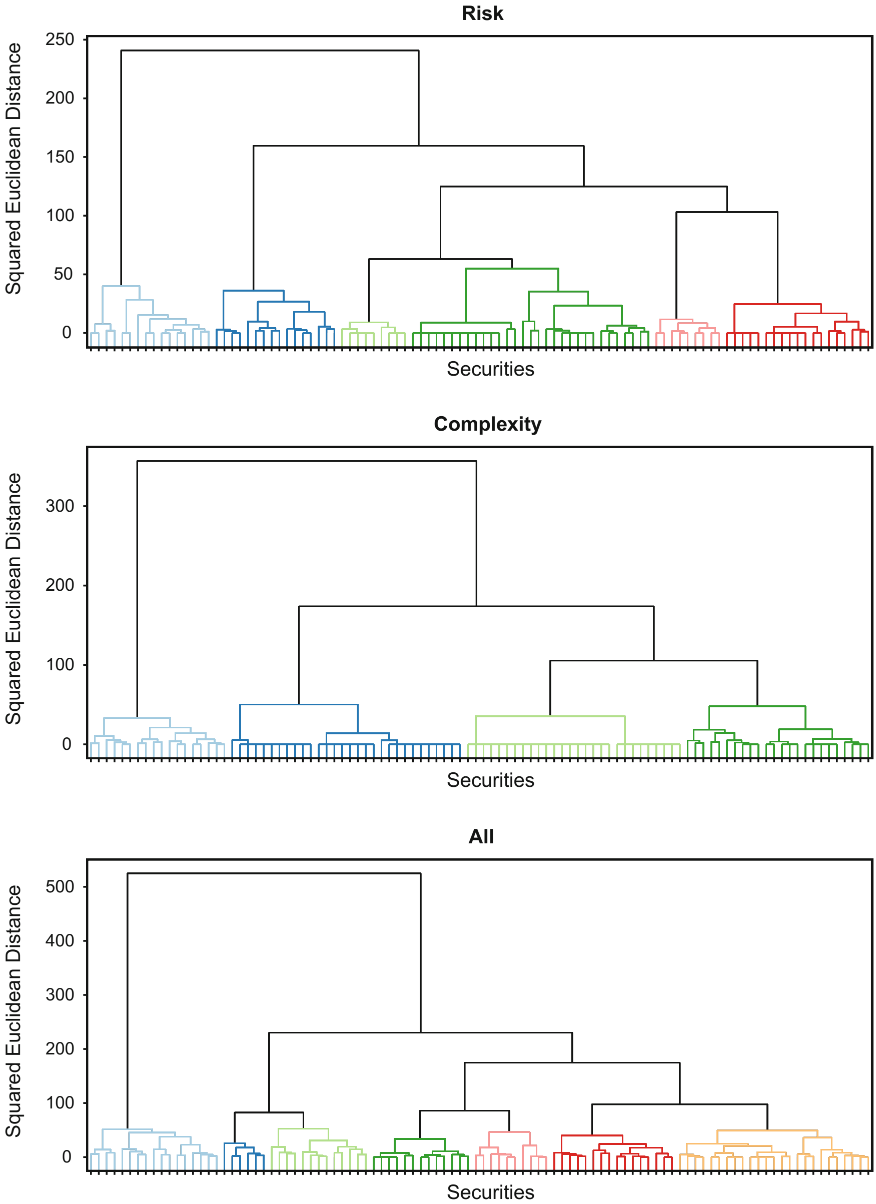

Fig. 12.3 Dendrograms of cluster analysis. Ward algorithm using Euclidean distances. Clusters formed below a threshold of 60 are coloured

Table 12.1 Projection Vector of PC_1 Projection vectors for PC_1 obtained from the eigendecompositions of the polychoric correlation matrices of X^{Risk}, X^{Comp}, and X^{All}

	X^{Risk}	X^{Comp}	X^{All}
w_1	−0.2141	0.3279	−0.1594
w_2	0.6013	0.4030	0.4275
w_3	0.0905	0.5185	0.1237
w_4	0.5106	0.4896	0.2687
w_5	0.1308	0.4707	−0.1087
w_6	0.5537		0.1166
w_7			−0.1929
w_8			−0.3142
w_9			−0.4440
w_{10}			−0.4553
w_{11}			−0.3722

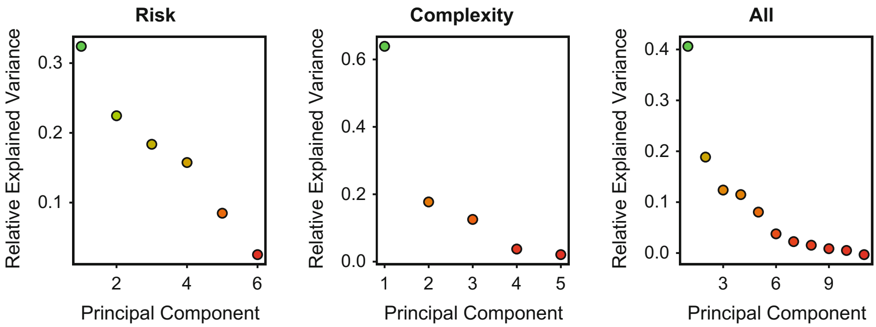

Fig. 12.4 Fraction of variance explained by each of the principal components

Just as the cluster analysis, PCA is performed on three sub-samples of X; X^{Risk}, X^{Comp}, and X^{All}. The number of columns of X therefore depends on the sub-sample (i.e. X^{Risk} is 100×6, X^{Comp} is 100×5, and X^{All} is 100×11). Table 12.1 shows the resulting projection vectors for the first principal component (PC), PC_1.

One method of analysing the relation between PCs and the underlying sample is to look at fractions of sample variance explained by each PC. This is possible, because the sum of PC variances matches the sum of variances of the underlying random variables in a sample (i.e. $\sum_{j=1}^{k} Var[PC_j] = \sum_{j=1}^{k} s_{x_j,x_j}$). The fraction of variance explained by each PC can therefore be measured as $\frac{Var[PC_j]}{\sum_{j=1}^{k} Var[PC_j]}$. If the fraction of explained variance for the first one or two PCs is very high, we know that the underlying random variables are in fact mainly driven by some latent factors represented by those two PCs. Figure 12.4 depicts the fractions of sample variance explained by each of the principal components (PCs).

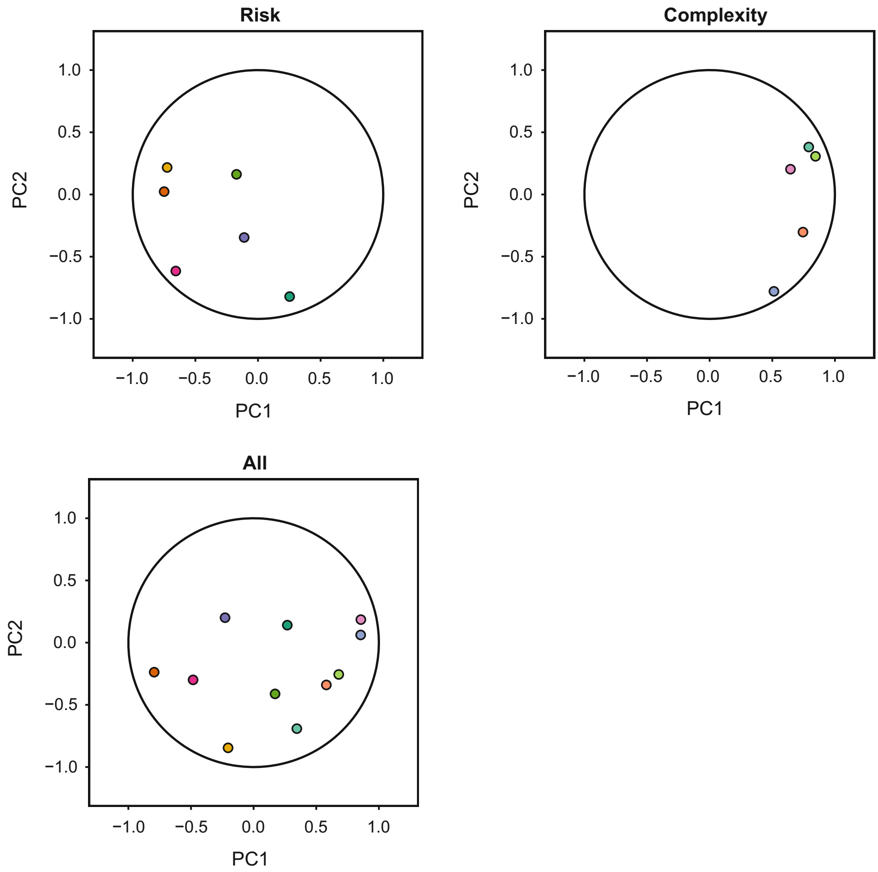

Fig. 12.5 Correlations of the factors with the first two PCs, based on the PCA of only risk, only complexity, and both risk and complexity factors. The risk factors are volatility, liquidity, credit rating, duration/cash flow, leverage, and diversification degree. The complexity factors comprise of the number of structural layers, expansiveness of derivatives, availability & known pricing models, number of return outcome scenarios, and transparency/ease of understanding

When only considering risk factors, the sample variance appears to be distributed fairly evenly among PCs. If we assume risk to be some latent variable that we expect the risk factors to be proxies of, the finding contradicts this assumption. Instead, the chosen risk factors appear to proxy for various independent latent factors. The opposite is true for the group of complexity factors, where the first PC explains more than 60 percent of the sample variation. All remaining PCs each explain less than 20 percent at the most. This reveals that the chosen complexity factors – at least in large parts – track the same underlying latent complexity factor. When including both risk and complexity factors in the PCA, the first PC explains around 40 percent of the sample variation and the next three or four PCs add another 10 to 20 percent each.

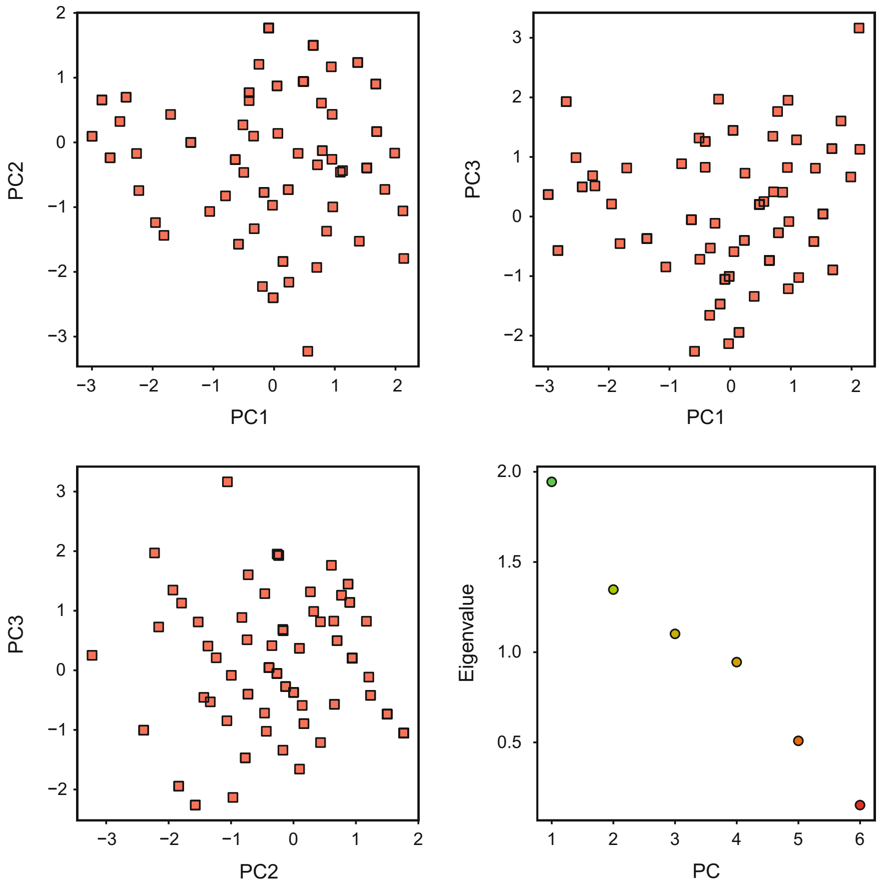

Fig. 12.6 The first three PCs derived from the PCA of the risk factors plotted against each other (*top left*, *top right*, and *bottom left*) and the eigenvalues of the polychoric correlation matrix of risk factors (*bottom right*)

In Fig. 12.5 we plot the correlation of each of the risk and complexity factors with the first two PCs for each of the factor sample subsets. Note that only the absolute correlation value is relevant when interpreting these correlations because PCs are not determined in their sign. Our results support the previous discussion regarding the explained sample variance. While the absolute correlation for risk factors with both PC_1 and PC_2 range from zero to 1.0 (top left panel), absolute correlations for complexity factors lie clearly within a range from 0.5 to 1.0 with a strong tendency towards higher values (top right panel). In the bottom left panel we note the absence of a clear correlation pattern between factors and the first two PCs. With the exception of the "number of structural" layers factor all complexity factors maintain a strong correlation with PC_1. Risk factors deviate very clearly from their correlations with both PCs in the top left panel. Figures 12.6, 12.7, and 12.8 plot the first three PCs

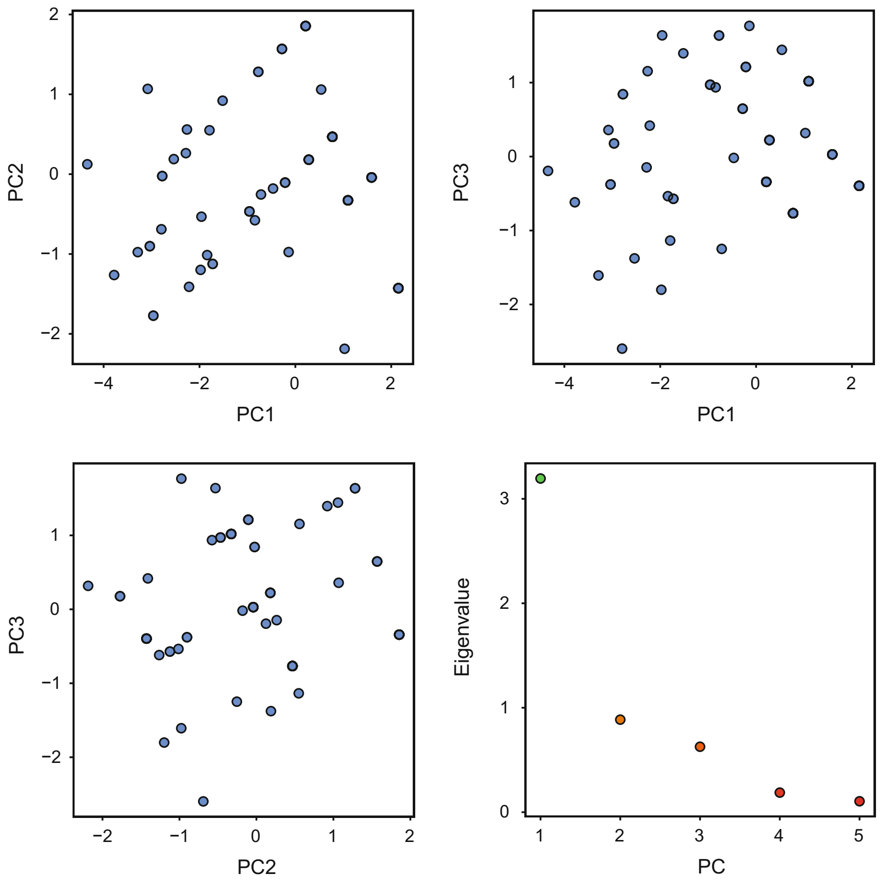

Fig. 12.7 The first three PCs derived from the PCA of the complexity factors plotted against each other (*top left*, *top right*, and *bottom left*) and the eigenvalues of the polychoric correlation matrix of complexity factors (*bottom right*)

against each other and show the correlation matrix eigenvalues associated with each principle component.

Finally, we plot the expert score of each security in the sample against its first PC in Fig. 12.9. As can be seen there is a clear relation between the total score and the first PC for risk, complexity, and both risk and complexity factors. This relation is most evident for the latter two groups.

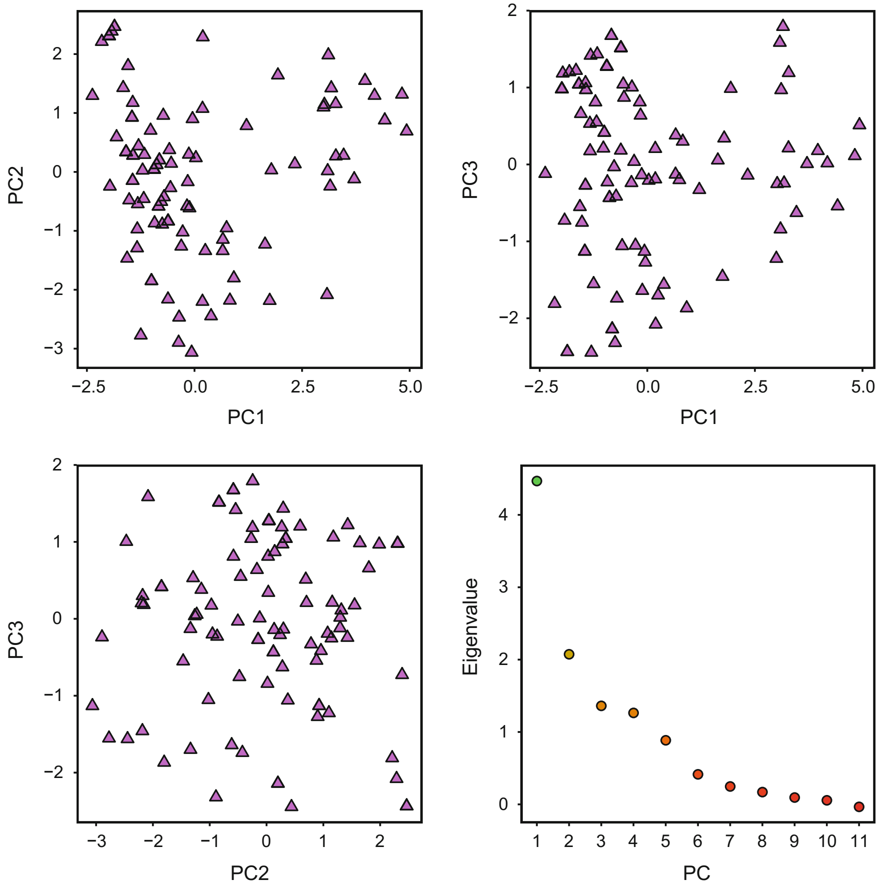

Fig. 12.8 The first three PCs derived from the PCA of the risk and complexity factors plotted against each other (*top left*, *top right*, and *bottom left*) and the eigenvalues of the polychoric correlation matrix of risk and complexity factors (*bottom right*)

12.2.1 Cross Validation via Leave-One-Out

The PCA results are cross validated by employing a leave-one-out (LOO) procedure. We compute the first PC for a security i based on weights obtained from a PCA of the sample excluding security i. In Fig. 12.10 we plot the LOO PCs against their regular counterparts. Additionally, we define a function

$$R_1 = \sum_{i=1}^{n} \left\{ f_1(x_i) - \hat{f}_1(x_i) \right\}^2, \tag{12.1}$$

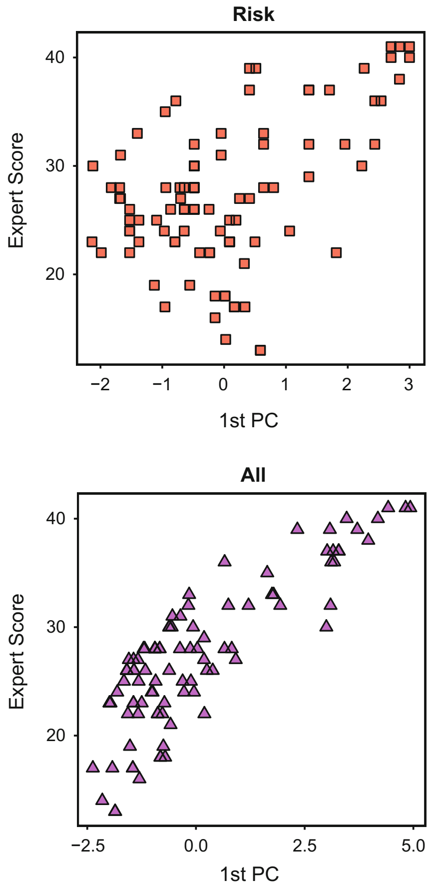

Fig. 12.9 The first PC from the PCA of risk (*top left*), complexity (*top right*), and risk and complexity (*bottom left*) factor scores plotted against the expert score of the corresponding securities

where $f_1(x_i)$ is the first PC for security i resulting from a PCA of the whole sample and $\hat{f}_1(x_i)$ is the first PC for security i computed from the weights of a PCA of the sample of $n-1$ securities (i.e. excluding security i). The values of R_1 for the three samples X^{Risk}, X^{Comp}, and X^{All} are 10.0655, 0.0219, and 0.2899, respectively. From these results we take that the PCA has some stability issues when only considering risk factors. Otherwise results are stable.

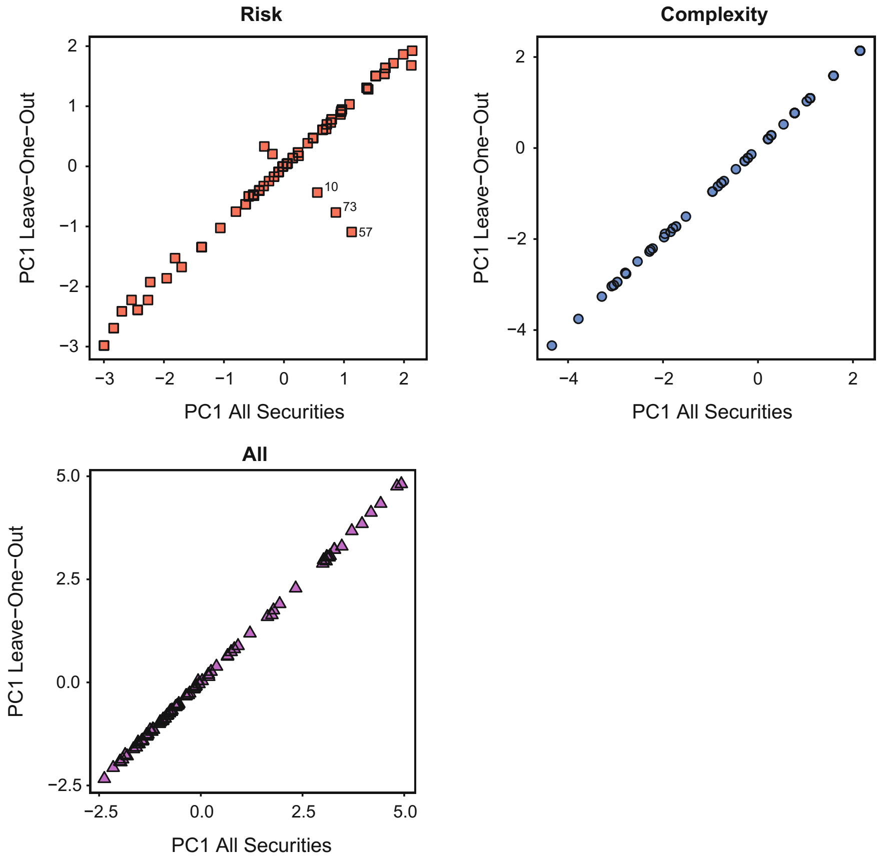

Fig. 12.10 $\hat{f}_1(x_i)$ plotted against $f_1(x_i)$ for risk factors (*top left*), for complexity factors (*top right*), for risk and complexity factors (*bottom*). 10.0655, 0.0219, and 0.2899 in each setup respectively Outliers are labeled with their security index in the sample

12.3 Adjusted Weighting of Factor Scores

In the following we consider two different applications of adjusting the weights applied to X. First, we try to find a weighting vector $w \in \mathbb{R}^k$ such that the projection $x_i\, w$ for each security i is as close as possible to its known expert score. Second, we evaluate the maximum distance between the projections of X through randomly chosen random vectors w.

12.3.1 Match Expert Score

Given a matrix $X_1 \in \{1, 2, 3\}^{n \times k}$, $X_2 \in \{1, 3, 5\}^{n \times k}$, or $X_3 \in \{1, 4, 9\}^{n \times k}$ and again considering the sub-samples X^{Risk}, X^{Comp}, or X^{All}, we can compute a function

$$R_2(X, w) = X\,w - f, \tag{12.2}$$

where w is an $k \times 1$ vector of weights and f is an $n \times 1$ vector of expert scores. From this we derive two optimisation problems (OPs) OP_1 and OP_2,

$$\widehat{w}_{OP_1} = \arg\min_{w_{OP_1}} \parallel X\,w_{OP_1} - f \parallel_1, \tag{12.3}$$

and

$$\widehat{w}_{OP_2} = \arg\min_{w_{OP_2}} \parallel X\,w_{OP_2} - f \parallel_2^2, \tag{12.4}$$

respectively. Table 12.2 shows the optimal weights for both OPs using one of X_1, X_2, or X_3 and either risk factors, complexity factors, or both risk and complexity factors. Figures 12.11, 12.12, 12.13, 12.14, 12.15 and 12.16 show the resulting weighted scores $X\widehat{w}$ plotted against the known expert scores.

As can be seen in our results, the linear approximation of expert scores is hard, even when using all 11 factors. The sum of squared approximation errors, R_2^*, in Table 12.2 is lowest for X_1 and the use of all factors. A discrete scale of $\{1, 2, 3\}$ thus appears better suited than the alternatives $\{1, 3, 5\}$ and $\{1, 4, 9\}$.

12.3.2 Cross Validation via Leave-One-Out

As with the PCA, we perform a LOO analysis to see how strongly the optimisation results for (12.4) depend on individual securities (Table 12.3).

We only consider OP_2 for X_1 because the overall results are best in this specification. The results, depicted in Fig. 12.17, are fairly robust against sample modifications. This is particularly true for X_1^{All}.

12.3.3 Widest Projection Spread

Given some random $k \times 1$ weighting vector we can compute the maximum spread between each projection in $X\,w$ and its nearest neighbour. We define $z = X\,w$ and then consider the order statistics of the elements z_i of z (i.e. $\forall i = 1, \ldots, n-1$: $z_{(i)} \leq z_{(i+1)}$). The maximum spread between all $z_{(i)}$ and their respective nearest

Table 12.2 Match Expert Score Weights. Optimal (normalised) weights $\widehat{w} \in \mathbb{R}^k$, correlations between $A\,\widehat{w}$ and f, as well as the optimal target function value R_2^* (this is the actual target function and not R_2 itself) for OP_1 and OP_2 using matrices $X_1 \in \{1, 2, 3\}^{k \times n}$, $X_2 \in \{1, 3, 5\}^{k \times n}$, and $X_3 \in \{1, 4, 9\}^{k \times n}$. The weights have been normalised to unit vectors to facilitate a comparison with PCA weights and simulation weights

Panel A: Risk factors

	X_1		X_2		X_3	
	OP_1	OP_2	OP_1	OP_2	OP_1	OP_2
$\widehat{w}_1$	0.7773	0.7882	0.8020	0.8123	0.6574	0.7511
$\widehat{w}_2$	−0.2064	−0.1632	−0.1433	−0.0702	0.0750	0.0529
$\widehat{w}_3$	0.0346	0.0289	0.1299	0.1182	0.1069	0.0946
$\widehat{w}_4$	−0.1650	−0.1170	−0.1433	−0.1328	−0.1737	−0.1412
$\widehat{w}_5$	0.4952	0.5499	0.4596	0.5077	0.6352	0.5829
$\widehat{w}_6$	0.2821	0.1876	0.2962	0.2142	0.3424	0.2536
$\rho_{A\,\widehat{w},f}$	0.7729	0.8296	0.7562	0.7920	0.6380	0.7664
R_2^*	334.8355	1845.7312	408.8966	2698.5647	487.7317	3924.8506

Panel B: Complexity factors

	X_1		X_2		X_3	
	OP_1	OP_2	OP_1	OP_2	OP_1	OP_2
$\widehat{w}_1$	0.5636	0.5290	0.6732	0.6777	0.5420	0.6216
$\widehat{w}_2$	0.2873	0.3255	0.2891	0.2960	0.3391	0.2848
$\widehat{w}_3$	−0.0136	0.1554	−0.0008	0.0916	0.0947	0.0498
$\widehat{w}_4$	0.5849	0.5154	0.4811	0.3873	0.1521	0.2456
$\widehat{w}_5$	0.5075	0.5695	0.4814	0.5429	0.7477	0.6854
$\rho_{A\,\widehat{w},f}$	0.9825	0.9924	0.9745	0.9755	0.9535	0.9515
R_2^*	339.0912	1755.0434	453.0134	3363.2401	677.2657	6798.8500

Panel C: Risk and complexity factors

	X_1		X_2		X_3	
	OP_1	OP_2	OP_1	OP_2	OP_1	OP_2
$\widehat{w}_1$	0.6622	0.6279	0.5914	0.6130	0.3821	0.5261
$\widehat{w}_2$	−0.0003	−0.0284	0.2187	0.1362	0.4382	0.2817
$\widehat{w}_3$	−0.2645	−0.2019	−0.1590	−0.0705	−0.1454	−0.0126
$\widehat{w}_4$	0.1326	0.1535	0.1182	0.1511	0.0754	0.1186
$\widehat{w}_5$	0.2651	0.3006	0.2368	0.2868	0.2879	0.3930
$\widehat{w}_6$	−0.1326	−0.1486	−0.1183	−0.1187	−0.0669	−0.0568
$\widehat{w}_7$	0.2652	0.2960	0.2368	0.2921	0.1633	0.2432
$\widehat{w}_8$	0.2650	0.2776	0.2365	0.2566	0.1853	0.2278
$\widehat{w}_9$	0.1329	0.1695	0.1950	0.2079	0.1391	0.2617
$\widehat{w}_{10}$	0.3969	0.3551	0.4945	0.3880	0.6406	0.3581
$\widehat{w}_{11}$	0.2653	0.3296	0.3139	0.3695	0.2390	0.4052
$\rho_{A\,\widehat{w},f}$	0.9942	0.9952	0.9768	0.9792	0.9117	0.9513
R_2^*	38.0562	46.5552	91.9700	225.6082	147.5615	589.0314

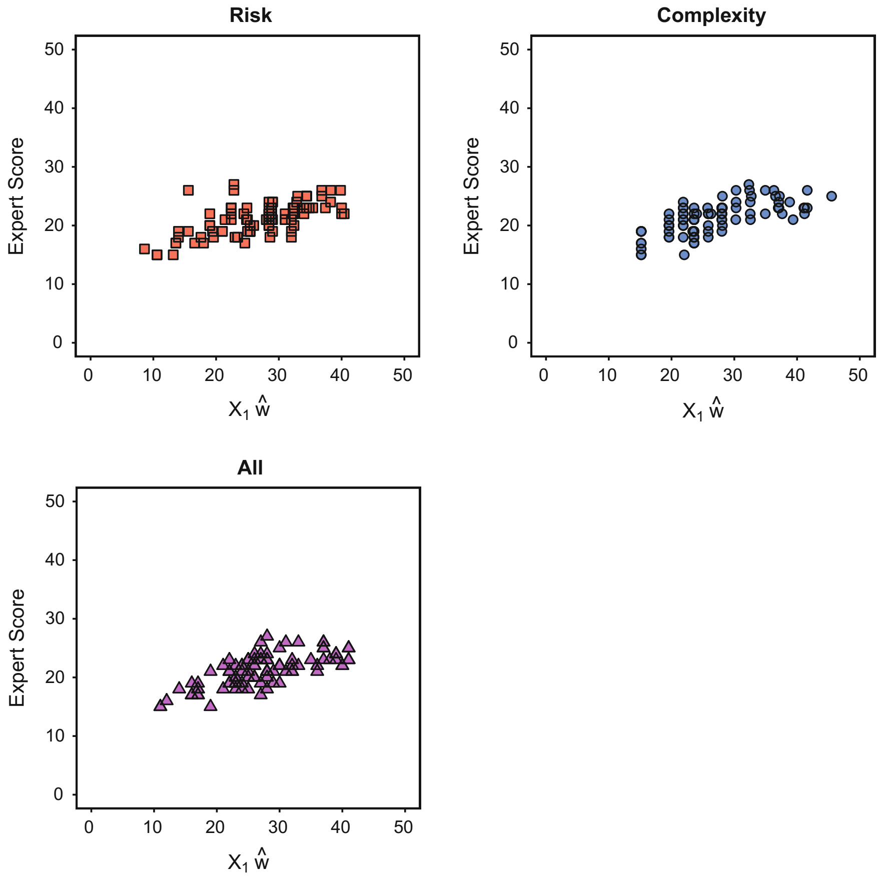

Fig. 12.11 The expert score (f) plotted against $X_1\,\widehat{w}$ for OP_1. We distinguish between results for risk factors (*top left*), complexity factors (*top right*), and risk and complexity factors (*bottom left*)

neighbour is then given by

$$R_3(z) = \max_{i}^{n-1} \left(z_{(i+1)} - z_{(i)} \right). \tag{12.5}$$

To examine the influence of the weighting vector w on the maximum projection spread we generate 1000 $k \times 1$ uniform random vectors ($w \sim \mathcal{U}(-1, 1)^k$). These vectors are then scaled to unit vectors.

Figure 12.18 shows the resulting 1000 simulated maximum spreads. The mean maximum spreads for the risk, complexity, and both risk and complexity cases are $\bar{s}^{Risk} = 0.6807$, $\bar{s}^{Compl} = 0.74725$, $\bar{s}^{All} = 0.6904$. A box plot of the results is shown in Fig. 12.19.

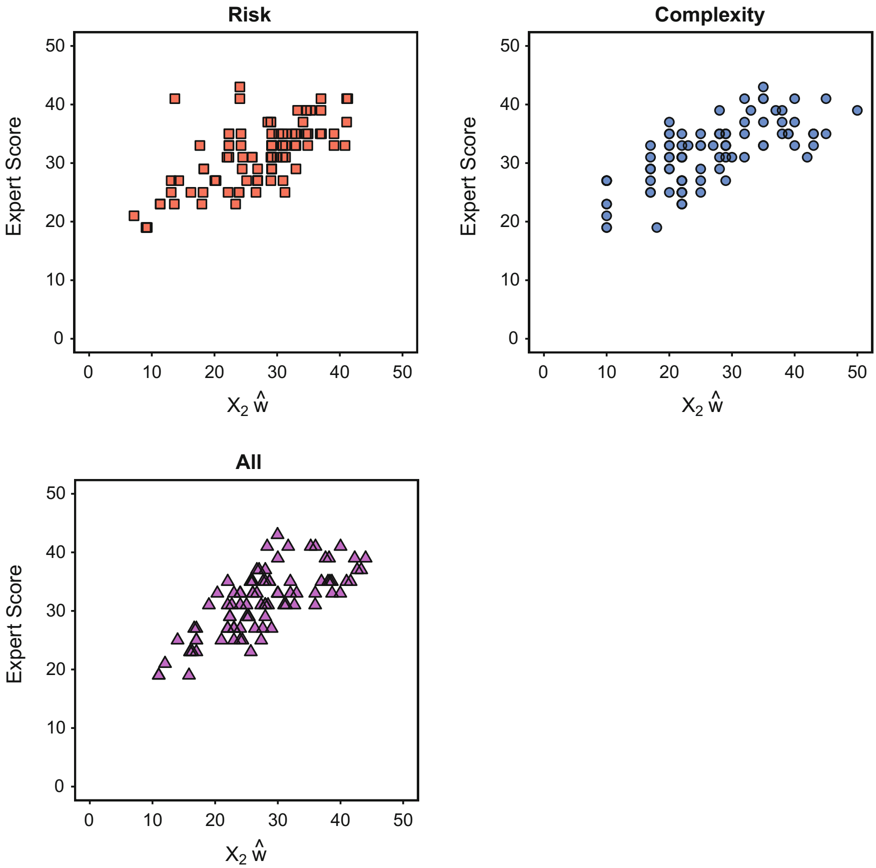

Fig. 12.12 The expert score (f) plotted against $X_2\,\widehat{w}$ for OP_1. We distinguish between results for risk factors (*top left*), complexity factors (*top right*), and risk and complexity factors (*bottom left*)

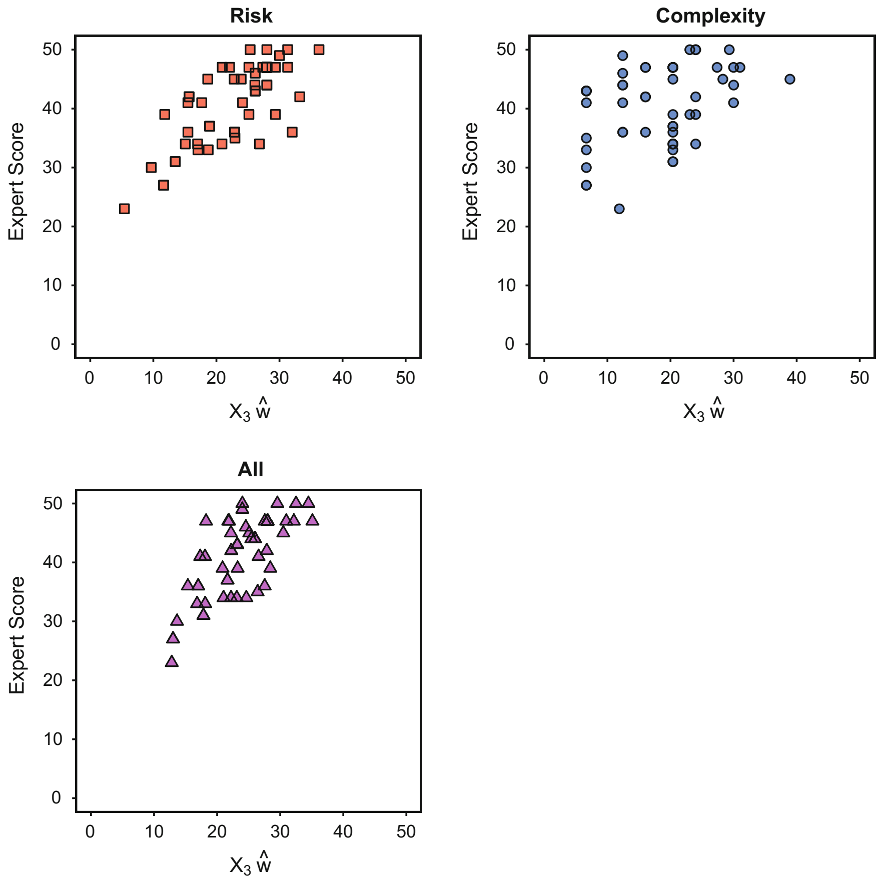

Fig. 12.13 The expert score (f) plotted against $X_3\,\widehat{w}$ for OP_1. We distinguish between results for risk factors (*top left*), complexity factors (*top right*), and risk and complexity factors (*bottom left*)

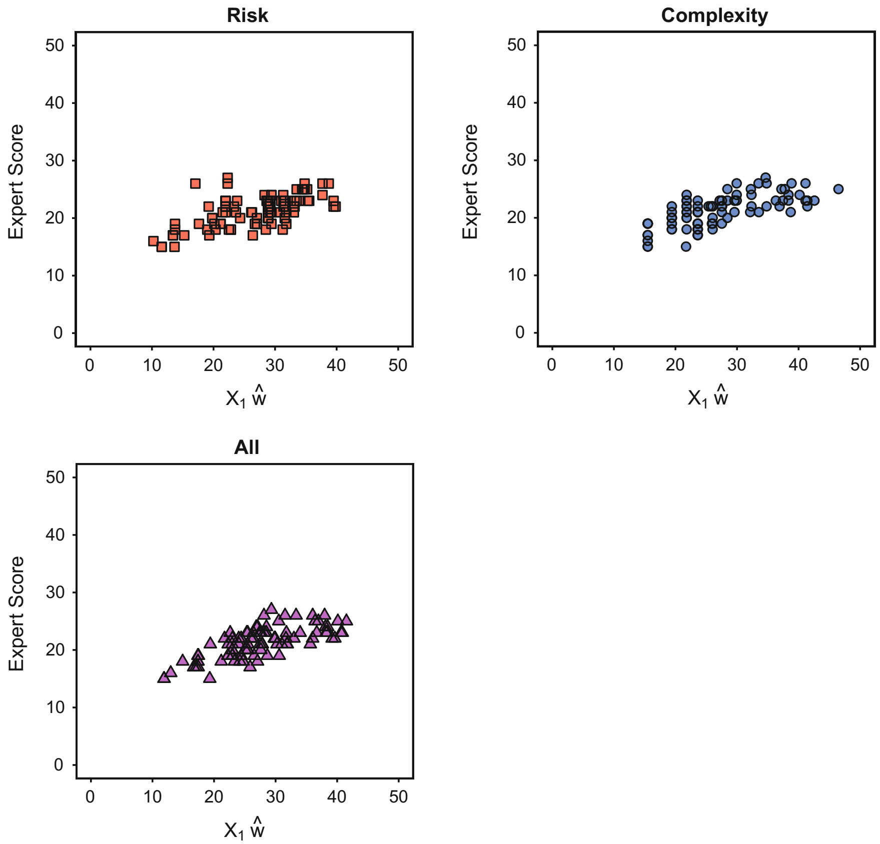

Fig. 12.14 The expert score (f) plotted against $X_1\,\widehat{w}$ for OP_2. We distinguish between results for risk factors (*top left*), complexity factors (*top right*), and risk and complexity factors (*bottom left*)

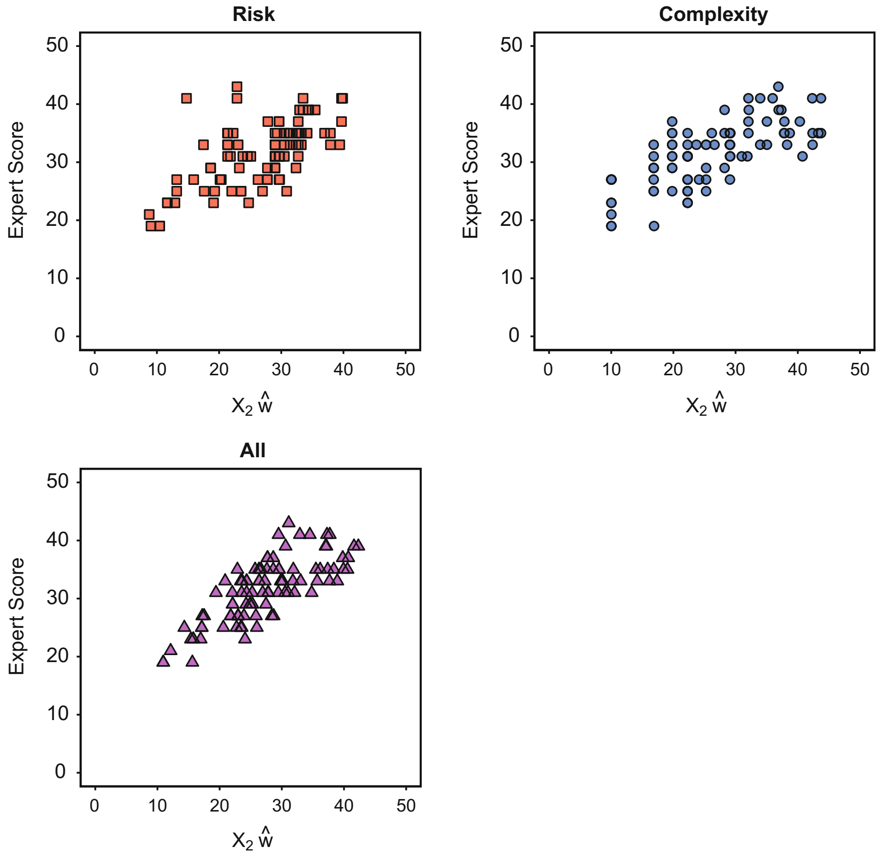

Fig. 12.15 The expert score (f) plotted against $X_2\,\widehat{w}$ for OP_2. We distinguish between results for risk factors (*top left*), complexity factors (*top right*), and risk and complexity factors (*bottom left*)

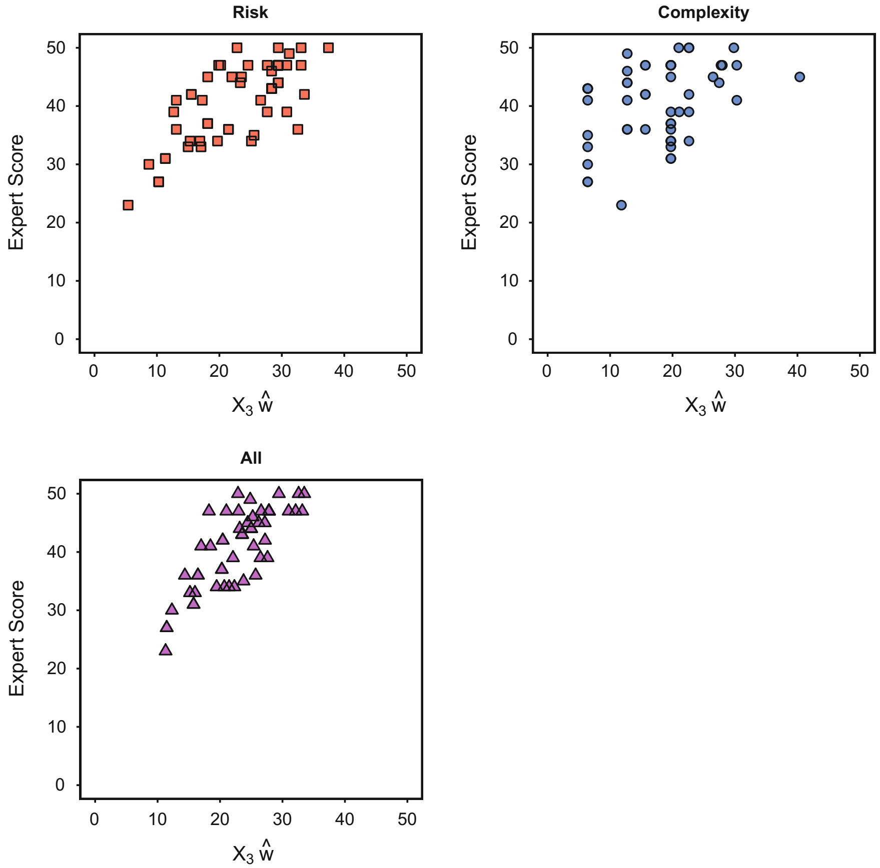

Fig. 12.16 The expert score (f) plotted against $X_3\,\widehat{w}$ for OP_2. We distinguish between results for risk factors (*top left*), complexity factors (*top right*), and risk and complexity factors (*bottom left*)

Table 12.3 Top Ten Mean Maximum Spread Simulation Weights The mean of the weighting vectors projecting the ten largest spreads from the original score matrix. The mean vectors for X^{Risk}, X^{Comp}, and X^{All} are normalised to unit vectors

	X^{Risk}	X^{Comp}	X^{All}
w_1	−0.9275	0.5161	0.4319
w_2	0.1824	0.4141	0.0833
w_3	−0.2483	0.3300	0.0257
w_4	0.1164	−0.3030	−0.2453
w_5	0.0130	0.6013	−0.3244
w_6	0.1764		−0.2807
w_7			−0.1371
w_8			−0.3083
w_9			0.1572
w_{10}			−0.5835
w_{11}			−0.2874

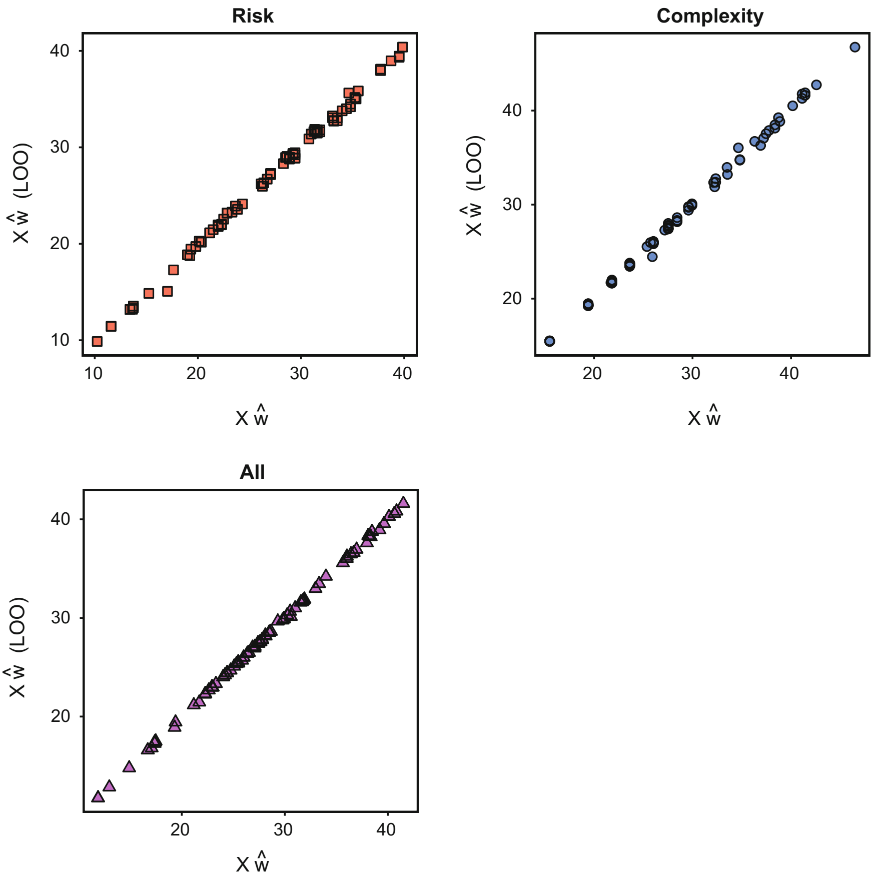

Fig. 12.17 $X_1\,\widehat{w}_{LOO}$ plotted against $X_1\,\widehat{w}$ for OP_2. We distinguish between results for risk factors (*top left*), complexity factors (*top right*), and risk and complexity factors (*bottom left*)

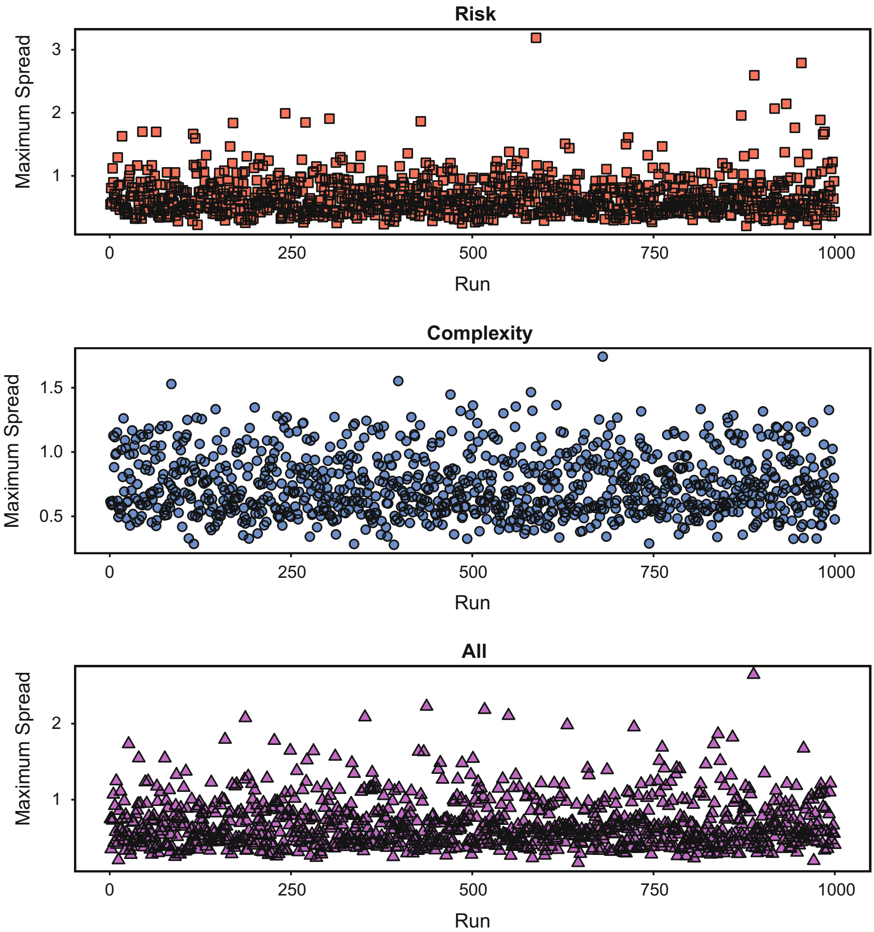

Fig. 12.18 Maximum spread among projections $\mathbf{X_1}\,\mathbf{w}$ for 1000 randomly chosen $\mathbf{w}$

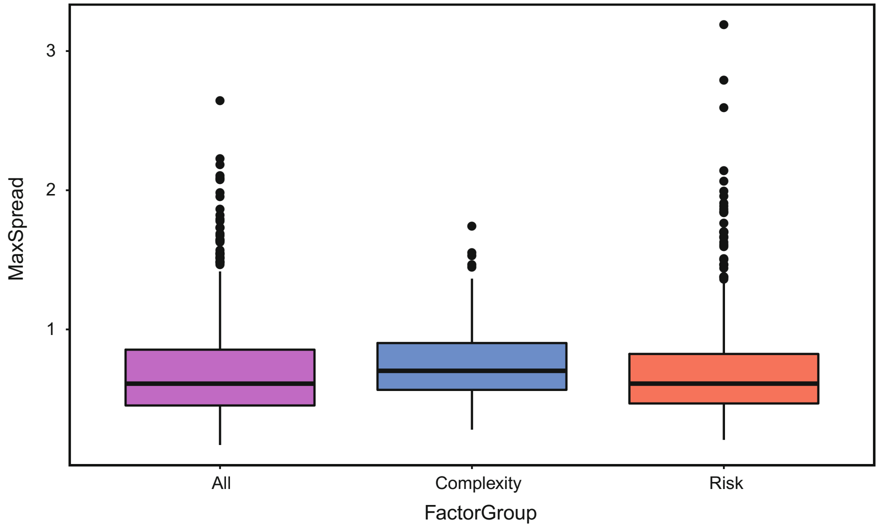

Fig. 12.19 Box plot of maximum spreads among projections $\mathbf{X_1}\,\mathbf{w}$ for 1000 randomly chosen $\mathbf{w}$

12.4 Conclusion

We can summarise our results in a few key points:

1. The choice of risk factors, as the PCA has revealed, does not seem to proxy for a single latent source of risk. The opposite is true for the choice of complexity factors.
2. Overall there is a clear positive relation between the first PC of the full PCA, involving all factors, and the expert score of a security as shown in Fig. 12.9.
3. Approximation of the total expert scores through linear projection of the score matrix is possible, but not perfect. We obtain best results by using a score scale of $\{1, 2, 3\}$ and applying the L^2 norm during optimisation.

Reference

Härdle, W. K., & Simar, L. (2015). *Applied multivariate statistical analysis* (4th ed.). Berlin: Springer.