Chapter 3

Principal Component Analysis

Introduction: Goal and Method

The market presents itself to the observer through a surface of incommensurably many data and movements. The links between those data and movements, for example the tendency of two-year (2Y) and 10Y interest rates to rise and fall together, point toward a more or less systematic mechanism hidden in the core of the market. Our goal is to see through the surface of the market and into its structural core.

To do so, we face the problem of finding the right (degree of) formal assumptions about the structure of a market. As an extreme position, if we made no formal assumptions at all, we would remain stuck to the surface of incommensurable market data, unable to understand anything about their structural core. So, in order to reach our goal we will need to use structural terms (i.e. mathematical formalism) and hence to impose some assumptions on the market. However, we shall use those assumptions which have turned out to fit the structure of the real market well and otherwise keep them as minimal as possible, thereby leaving enough space for the market to express its own mechanisms in our form.

Principal component analysis (PCA) has only one (main) assumption: that the market is driven by a set of uncorrelated linear factors. This is not only a relatively weak assumption (allowing the market to fill in the remaining structural information, in particular about the shape and strength of each factor) but also a very useful one both for relative value (RV) analysis and hedging. It satisfies the condition of asset pricing theory (see Chapter 1) and allows us to construct portfolios that are exposed to or hedged against any factor, just as specified by the investor. For the purpose of RV analysis,[1] PCA therefore appears to be a useful tool.

A main goal of this chapter is the empirical illustration of the way PCA leads us through the surface of the market into its core, where we can see its inner driving forces, gaining meaningful and deep insights in market mechanisms. After developing the mathematics, we shall spend a large part of this chapter exploring the application of PCA to actual market mechanisms, thereby illustrating the way the structure of real markets reveals itself through PCA.

On one hand, the relationship between market and mathematical form can involve problematic assumptions. On the other hand, however, it also connects the real world to its mathematical representation. PCA is therefore a link between the economy and mathematics, generating economically relevant statistics. In the example of 2Y and 10Y rates from above, a PCA could identify one factor behind moves of both rates in the same direction, and another, uncorrelated factor behind moves of both rates in opposite directions. Interpreting the identities of those factors (perhaps linked to economic data, like inflation) gives a deep understanding of the driving forces of the yield curve. In the next step, an investor who has a view on those factors (e.g. inflation) could construct through PCA the best trading position and hedge it against factors on which he has no view.

With PCA linking statistics to fundamentals, it also links the first to the second part of the book. While the focus in the preceding section was on statistics, it now shifts to statistics linked to economic insights. Correspondingly, the present goal is not an optimization of Sharpe ratios but to gain insights into the fundamental mechanisms. This will complement statistical optimality with the confidence of understanding the real economic driving forces behind a trade.

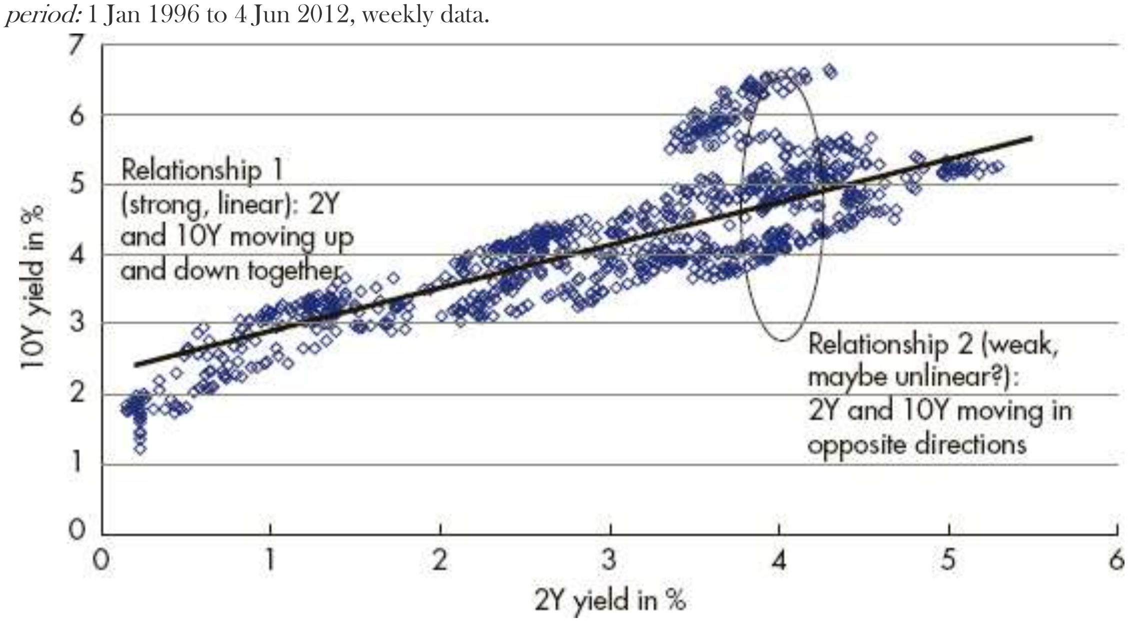
An Intuitive Approach toward PCA

An assumption-free start to market analysis could be to simply plot data observations into a scatter chart. Figure 3.1 shows an example for 2Y versus 10Y Bund yields (from 1996 to mid-2012). Then, one can try to distinguish structural relationships. In the current example, there seems to be a strong relationship driving both yields up and down together, which has been in force more or less over the whole time period. Moreover, that relationship seems to be rather linear and can thus be approximated quite well by a straight line. In addition, there seems to be a second mechanism driving both yields in opposite directions. Compared to the first relationship, however, it appears to be weaker than the first one over the whole time period and with an impact that varies significantly over time, being quite strong when 2Y yields were around 4% and less so in other sectors. Furthermore, it is not clear whether this relationship can be reasonably assumed to be linear as well.[2]

FIGURE 3.1 Structure of point cloud of 2Y and 10Y Bund yields.

Sources: data – Bloomberg; chart – Authors.

Data period: 1 Jan 1996 to 4 Jun 2012, weekly data.



Broadly speaking, PCA assumes the relationships to be linear and uncorrelated[3] and provides a *quantification* of the intuitive approach from above, telling us in mathematical form:

- *which* (linear) relationships exist in a given set of market data (i.e. what the *shape* of market mechanisms looks like). In the current example, this would correspond to quantifying the impact of the relationship 1 and 2 (short and long Bund yields moving up and down together or in opposite directions) on 2Y and 10Y yields.

- *how strong* the relationships are relative to each other, that is how much of market action (yield curve variation) is explained by a particular mechanism.

In mathematical terms, the shape of market mechanisms corresponds to the eigenvectors and their relative strength to the (scaled) eigenvalues of a PCA.

Before moving on, we would like to note that a chart like the one above can serve as a check on the suitability of the assumptions of a model for a particular market segment. In case of the assumptions of a PCA (such as linearity) seeming to violate actual market behavior, caution is advisable, and the diagnostic techniques discussed in the previous chapter could be used.

Furthermore, this intuitive approach to PCA may give the wrong impression that PCA works like a regression, with relationship 1 being the regression line and relationship 2 representing the residuals. However, the mathematics is different, and so (usually) are the results (i.e. the relationships calculated via a PCA will in most cases differ from the regression line). We shall discuss the reasons for the difference in the section about appropriate hedging, which will also show the superiority of using PCA-based relationships for calculating hedge ratios.

Factor Models: General Structure and Definitions

We shall now repeat the discussion from above in formal terms, thereby constructing PCA. The goal of understanding the few key mechanisms behind all market moves can be mathematically addressed by extracting the most relevant information from a given set of market data. Expressing this a bit more formally, it means reducing the dimensionality, with the remaining dimensions containing most of the information. Thus, the result of this exercise will reveal the number, strength, and shape of the market mechanisms.

We observe a number n of market data (e.g. yields) y_i ($i = 1, \ldots, n$) at time t. The general form of a k-factor linear model is given by:

$$\begin{pmatrix} y_1^t \\ \vdots \\ y_n^t \end{pmatrix} = \sum_{i=1}^{k} \alpha_i^t \cdot \begin{pmatrix} f_{i1} \\ \vdots \\ f_{in} \end{pmatrix} + \begin{pmatrix} \varepsilon_1^t \\ \vdots \\ \varepsilon_n^t \end{pmatrix}$$

where α_i (a number which changes over time) is called the i-th factor (at time t), $\begin{pmatrix} f_{i1} \\ \vdots \\ f_{in} \end{pmatrix}$ (a vector which does not change over time) is called the i-th factor loading and $\begin{pmatrix} \varepsilon_1^t \\ \vdots \\ \varepsilon_n^t \end{pmatrix}$ (a vector which changes over time) is called the k-factor-residual (at time t), that is the portion unexplained by the factors.

The factor loadings, which do not change over time, can be considered as containing the market mechanisms, while the factors show how much of a specific market mechanism is active at a certain point in time. As an intuitive comparison, consider a sound mixer: the (invariable) individual sounds correspond to the factor loadings, while their (variable) strengths at a certain point in time (as adjusted by the volume regulator on the mixer board) correspond to the factors. The other way around, we can decompose the overall sound we hear into its individual components by looking at the regulators and labels on the mixer. Likewise, we are able to decompose the overall market action we observe into its individual driving forces by looking at the factors and factor loadings of a PCA.

Example: Imagine we decide to model the yield curve from one to 10 years ($n = 10$) by

$$(n = 10) \text{ by } \begin{pmatrix} y_1^t \\ \vdots \\ y_{10}^t \end{pmatrix} = \alpha_1^t \cdot \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

(a one-factor model). This particular factor loading allows only parallel shifts. This is an example for a very strong assumption from our side, which limits our perceptive ability to observing parallel shifts only. The reality of the market would protest against our imposed form by exhibiting large residuals.

So far, we have only assumed that market mechanisms are linear factors. The question is how much more we should assume. The answer to that question classifies factor models into two categories.

In the first category, the analyst determines the factor loadings himself. The Nelson–Siegel (NS) model is a prominent example of this approach, assuming a priori that the first factor loading is the vector $\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ and the second a vector whose entries follow the discount factor curve. Thus, it expands the simple one-factor model from our example above (parallel shifts) with a second factor accounting for curve steepening and flattening.

The advantage of that approach is that by choosing the factor loadings appropriately one can ensure that the model exhibits the desired properties (e.g. is arbitrage-free). This is probably the main reason academic market analysis often prefers factor models falling into the first category.

The disadvantage is that the model results may reflect the assumptions of the analyst rather than the true market mechanisms. In the case of the NS model, for example, curve steepening and flattening moves that do not follow the shape of the assumed discount factor curve cannot be explained and will appear as residuals. Hence, the steepening mechanism of the real market will be ignored in the factor loadings (overwritten by an arbitrary a priori assumption) and pushed into the residuals unexplained by the factors of the model.

In the second category, the factor loadings are extracted from the market: rather than making his own a priori assumptions about the factor loadings (i.e. market mechanisms), the analyst lets the market reveal its own dynamics and *then* a posteriori *interprets* the factor loadings. For our goal of seeing into the core of the actual market (rather than imposing our hypotheses on it), the second category is thus the right place to look.

PCA is the main representative of the second category: it forces the market to reveal its mechanisms under the form of uncorrelated linear factors but usually leaves enough freedom for the market to reveal its real dynamics within that formal framework. For example, the steepening mechanism (net of direction) of the market will show up in the second factor loading just as it is, revealing the actual market dynamics rather than overwriting them with a priori assumptions of the analyst. Consequently, market mechanisms will be visible in the factor loadings rather than pushed into the residuals.

PCA: Mathematics

Since PCA is a tool from linear algebra, we first need to represent the market in the form of a matrix. The straightforward approach is therefore to express the structural information contained in the market under consideration in the form of a covariance matrix.[4]

Technical points

- Market data should be of sufficient length to allow the parameters of the model to be estimated with sufficient accuracy.
- Depending on the goal of the analysis, either level or change data can be used as inputs.
- It is important to use the co*variance* rather than co*rrelation* matrix since the difference in volatility (sensitivity) is a key element of the analysis and must not be netted out by using correlations.

Now that we have the market information in the form of a covariance *matrix*, we can extract information by applying the powerful tools of linear algebra, transforming the covariance matrix into the orthonormal basis of its eigenvectors.

Definition: If $Ax = \lambda x$ ($x \neq 0$), for a matrix A, then the vector x is called an eigenvector of A and the number λ is the associated eigenvalue of A.

Hence, an eigenvector does not change its direction, only its length, when the matrix is applied to it. For example, if A represents the rotation of a globe by 90 degrees, then the rotation axis through the two poles is an eigenvector with an associated eigenvalue of one. Another example: if one holds a ball with both hands and squeezes it, the line connecting both hands contains eigenvectors, whose associated eigenvalue is less than one.

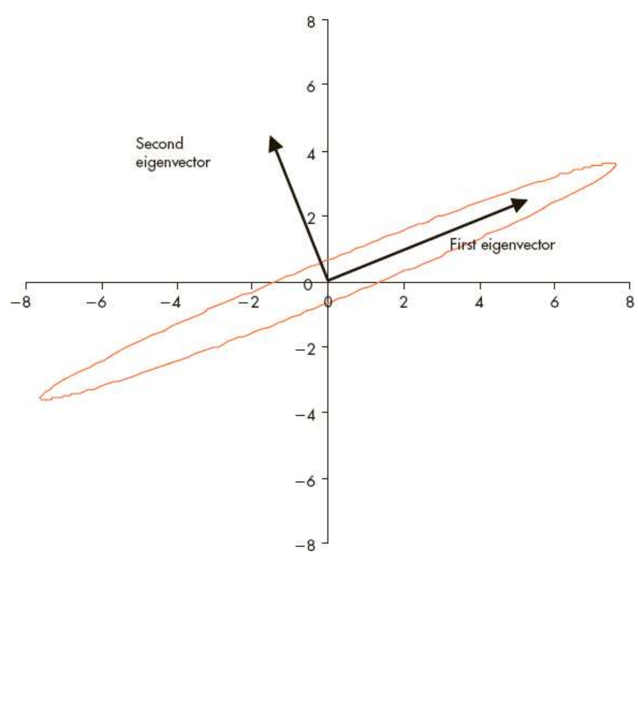
Theorem: For every covariance matrix[5] Cov, it is true that $Cov = B^{-1}\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{pmatrix} B$, where λ_i are the eigenvalues of Cov and the columns of B consist of the eigenvectors of Cov.

Intuitive interpretation: The matrix B acts like a transformation of the coordinate system and allows us to consider the covariance matrix from the perspective of the orthonormal basis given by its eigenvectors. Since these eigenvectors are orthogonal, they decompose the covariance matrix into uncorrelated relationships. Moreover, the eigenvector associated with the greatest absolute eigenvalue points into the direction of the highest variation, that is it represents the most important structural relationship in the market.

Example: $Cov = \begin{pmatrix} 7 & 3 \\ 3 & 2 \end{pmatrix}$ has the eigenvalues 8.4 and 0.6 with eigenvectors (e.g.) $\begin{pmatrix} 0.91 \\ 0.42 \end{pmatrix}$ and $\begin{pmatrix} -0.42 \\ 0.91 \end{pmatrix}$. Figure 3.2 shows the image of the unit circle under Cov and illustrates that most variation occurs along the direction of the first eigenvector.

FIGURE 3.2 *Cov* and examples for its two eigenvectors.

Source: Authors.



Technical point: Numerical issues involved in eigenvalue calculation

Eigenvalue calculation is famous for its numerical challenges. The main problem is to distinguish between two different eigenvalues and two slightly different numerical representations of the same eigenvalue. Since some of the eigenvalues tend to get very small when their number increases, this distinguishing becomes quite difficult for large covariance matrices and requires numerical representations with many digits and very high accuracy.

When an in-house PCA tool is required, the following hints could be useful:

- Scale the covariance matrix appropriately.
- Use a combination of numerical methods. In particular, since the largest eigenvalue is of crucial importance for the application of PCA to markets, confirm it through the Lanczos algorithm.
- Use as starting points for a Newton algorithm on the characteristic polynomial of the covariance matrix after a Householder transformation the following set of numbers: $a + 2 \cdot b \cdot \cos$ $a+2\cdot b\cdot\cos\left(i\cdot\frac{\pi}{n+1}\right)$ for $i = 1, \ldots, n$, with $a=\frac{1}{n}\sum_{i=1}^{n} a_i,\ b=\frac{1}{n-1}\sum_{i=2}^{n} b_i$, and with $\begin{pmatrix} a_1 & b_2 & 0 & 0 \\ b_2 & a_2 & \ddots & 0 \\ 0 & \ddots & \ddots & b_n \\ 0 & 0 & b_n & a_n \end{pmatrix}$ being the Householder transformation of the covariance matrix.
- Additionally, run a systematic search for eigenvalues, whereby the largest eigenvalue found via the Lanczos algorithm determines the width of steps and range of starting points for the Newton algorithm.
- Run checks, whether the computed eigenvectors are really eigenvectors (i.e. fulfill $Ax = \lambda x$) and orthogonal to each other.

The PCA sheet on the website accompanying this book implements these tricks. However, also due to numerical restrictions of Excel, it reaches its limitations rather soon and should not be expected to work for larger covariance matrices.

PCA as Factor Model

PCA becomes a factor model by using the eigenvectors e_i of the covariance matrix as factor loadings, that is by defining $f_i := e_i$, with the factor model being thus:

$$\begin{pmatrix} y_1^t \\ \vdots \\ y_n^t \end{pmatrix} = \sum_{i=1}^{n} \alpha_i^t \cdot \begin{pmatrix} e_{i1} \\ \vdots \\ e_{in} \end{pmatrix}$$

where $\begin{pmatrix} e_{i1} \\ \vdots \\ e_{in} \end{pmatrix}$ is the i-th eigenvector.

We sort the eigenvalues (and associated eigenvectors) by the percentage of total variation explained, that is

$$|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|$$

Hence, the first factor explains most of the market variation, the second explains most of the market variation not explained by the first factor, and so on. Correspondingly, the first factor loading, that is the first eigenvector, reveals the structure of the most important market mechanism, the second eigenvector the structure of the second most important, and so on. The importance of a market mechanism, that is the strength of its impact on the overall market variation, is quantified by the eigenvalues.

Technical point

If x is an eigenvector of Cov, then for every $a \neq 0$, ax is also an eigenvector of Cov. Thus, for every eigenvalue, there exist infinitely many eigenvectors, pointing in the same direction (or, if $a < 0$, in the opposite direction), but of different length. Any one of those can be chosen as a factor loading, and there is no reason to prefer one over the other.

In the end, one needs to arbitrarily decide the length of a particular eigenvector. For example, eigenvectors often are scaled to have unit length. However, there still exist two eigenvectors, x and y, with length one, with $x = -y$. Again, there is no criterion to decide for one or the other and one has to arbitrarily choose one. However, it is therefore important for the analyst to conduct the analysis and interpret the results with the particular choices of eigenvectors in mind. For example, if rates are an increasing function of factor 1 when x is chosen as the eigenvector, then rates will be a decreasing function of factor 1 when $y = -x$ is chosen as the eigenvector. Many analysts have fallen into this trap, which can be avoided only by constantly remembering the eigenvectors. As in this example, factors are only meaningful when analyzed in conjunction with the factor loadings.

Insight into Market Mechanisms through Interpretation of the Eigenvectors

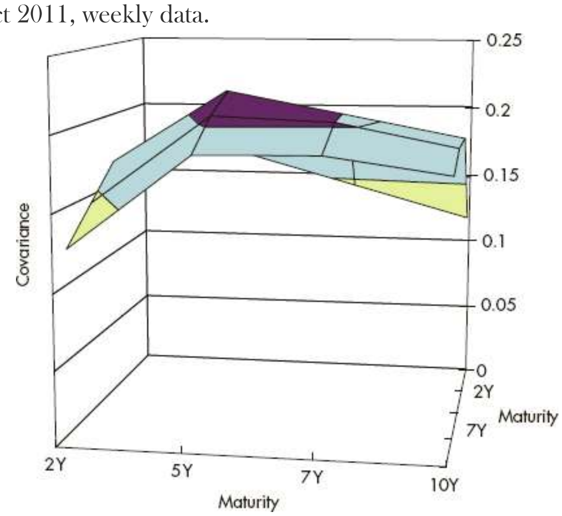
So far, we have developed the mathematical framework for PCA. For the rest of the chapter, we will see how PCA reveals the inner structural relationships of the market and how it can be applied to trading.

For the following example, we use weekly data from 4 Jan 2010 to 3 Oct 2011 for generic Bund yields (yield level, not change) for two, five, seven and 10 years. The input data and the PCA can be seen in the PCA sheet accompanying this book. In the first step, the covariance matrix is calculated and depicted in Figure 3.3. While this is not a necessary step in the analytical process based on PCA, displaying and examining the covariance matrix (i.e. the input into a PCA) can already give an intuition about the market mechanisms (i.e. the output of a PCA). In Figure 3.3, we can observe an area of maximal covariance in the medium part of the yield curve.

FIGURE 3.3 Covariance across the Bund yield curve.

Sources: data – Bloomberg; chart – Authors.

Data period: 4 Jan 2010 to 3 Oct 2011, weekly data.



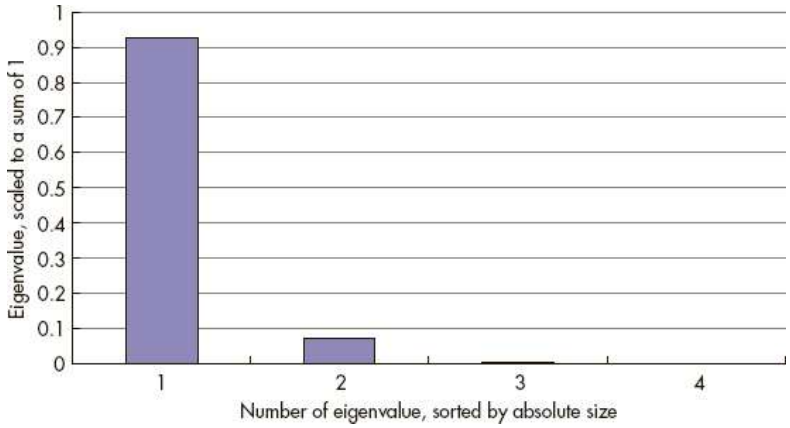
Then, the information about the market (as represented in the covariance matrix) is extracted through a PCA. As with any factor model, the eigenvectors of the covariance matrix represent the structural relations and mechanisms between the data series, while the eigenvalues show the relative significance of these factors in explaining the behavior of the data.

A typical situation for the relative strength of the factors, as quantified by the relative size of their eigenvalues, is depicted in Figure 3.4. As can be seen in that figure in the case of the Bund yield curve, the first factor explains more than 90% of the yield curve variation; and the first three factors together, almost everything. This means that basically the whole information the Bund market provides can be reduced to, captured, and expressed in three numbers (the first three factors), with the factor loadings translating the information between the full Bund market and the three factors (numbers) back and forth.

FIGURE 3.4 Scaled eigenvalues of a PCA on the Bund yield curve.

Sources: data – Bloomberg; chart – Authors.

Data period: 4 Jan 2010 to 3 Oct 2011, weekly data.



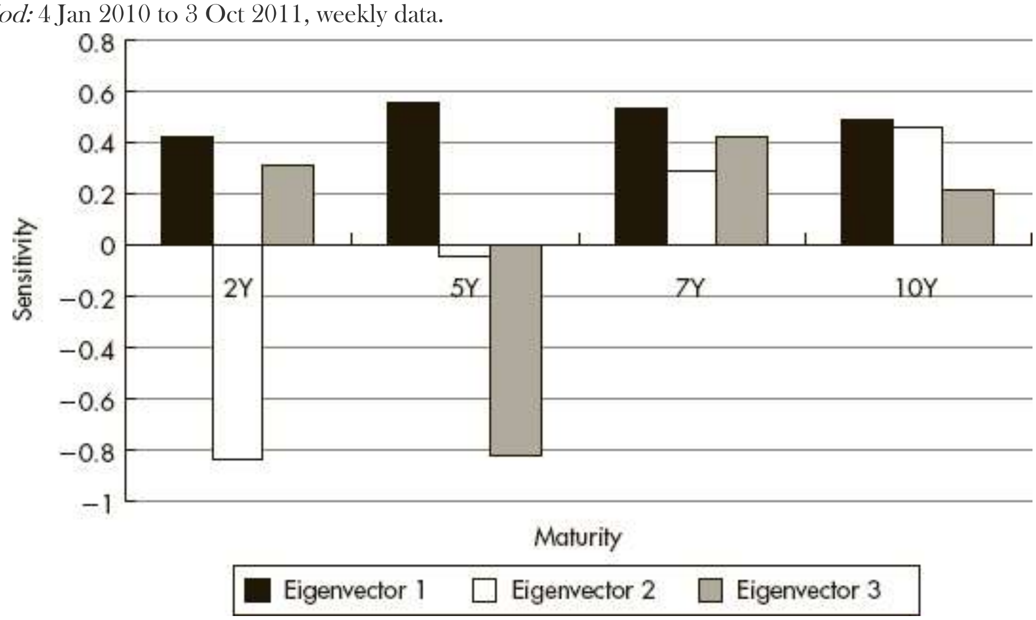
The next step to extract information about market mechanisms from a PCA is to examine the shape of the eigenvectors. The interpretation of the structural information contained in the eigenvectors can be done by applying the following scheme: if the *i*-th factor (α) increases by one, what happens to the rate curve? The answer to this question translates the market mechanisms from the mathematical language of eigenvectors of the covariance matrix into everyday terms.

In the following, we provide this interpretation for the first three eigenvectors of a PCA on the Bund curve, which are depicted in Figure 3.5. In this example, we find that a unit increase in the first factor corresponds to an increase in every point along the yield curve, since every entry in the first eigenvector has the same sign. Hence, we can interpret the first factor as representing the directional dynamics of the yield curve.

FIGURE 3.5 First three eigenvectors of a PCA on the Bund curve.

Sources: data – Bloomberg; chart – Authors.

Data period: 4 Jan 2010 to 3 Oct 2011, weekly data.



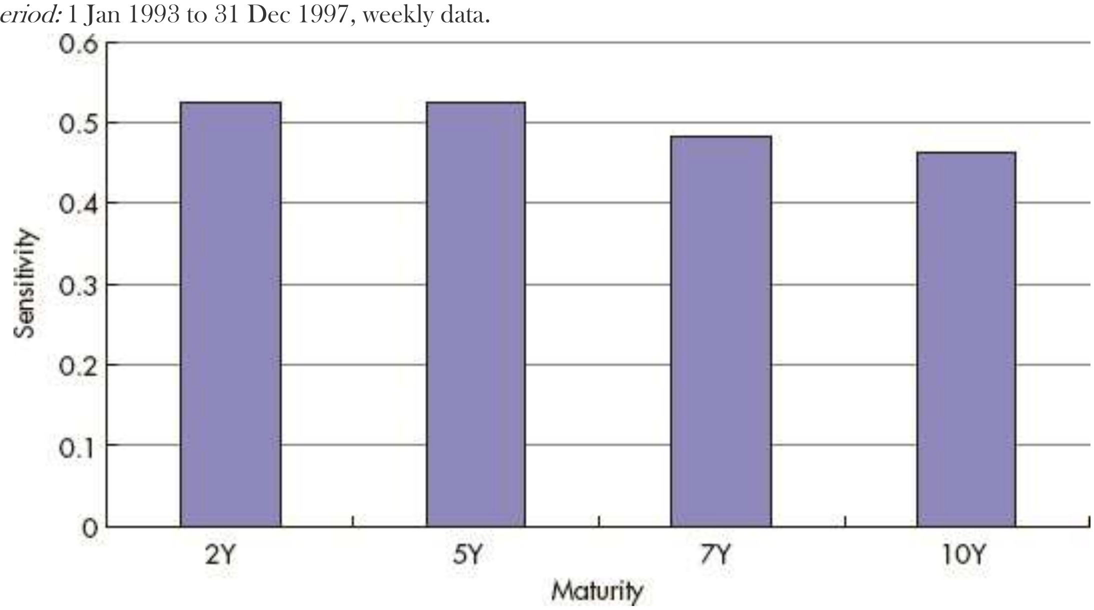
Moreover, the *shape* of the first eigenvector in Figure 3.5 represents the *shape* of directional moves: if α_1 increases by one, all yields increase, but medium yields (five-year, or 5Y) increase more than both short (2Y) and long (10Y) yields. This can be translated into everyday terms by saying that the pivotal point of directional moves of the yield curve is the 5Y area. We note that the output (eigenvector) of a PCA corresponds to the input (covariance matrix), with the shape of the first eigenvector reflecting the maximal covariance in the medium part of the yield curve we observed in Figure 3.3.

Now, let's compare this with the first eigenvector of a PCA on the Bund curve with data from 1993 to 1997, which is depicted in Figure 3.6. While for 5Y, 7Y, and 10Y the sensitivity is almost identical to the current situation, back in the 1990s, the sensitivity of 2Y yields was as high as that of 5Y yields (i.e. the market exhibited a bull-steepening/bear-flattening pattern). This is typical for a market with an active central bank driving the yield curve up and down, which corresponds to a higher sensitivity of shorter yields to directional moves, hence a higher entry in the first eigenvector. Thus, the decreasing activity of the central bank from the 1990s until now (partly a function of rates approaching zero) is reflected by the decreasing sensitivity of 2Y yields to the first factor. Note how the mathematics (entry in eigenvector) corresponds to real economic mechanisms. Even if we did not know anything about central bank history, one look at the PCA results would reveal both the structural shift in market mechanisms (pivotal point of directional moves shifting from the short end to the medium curve sector) and its likely source. Also note the stability displayed by the eigenvector over the past 20 years (apart from the short end), which we shall analyze in more detail later in this chapter.

FIGURE 3.6 First eigenvector of a PCA on the Bund curve with data from 1993 to 1997.

Sources: data – Bloomberg; chart – Authors.

Data period: 1 Jan 1993 to 31 Dec 1997, weekly data.



Furthermore, note the difference to the approach of a factor model in the first category, like the NS model: rather than *assuming* a priori that direction is the strongest market dynamic, the market reveals that this is the case for this particular market segment. And contrary to the a priori assumption of parallel shifts in the NS model, it shows that the real directional market mechanism follows a 2Y-5Y bear-steepening/5Y-10Y bear-flattening shape.

Also, together with the information about the eigenvalues shown in Figure 3.4, we have found (rather than assumed) that directional market mechanisms have indeed the strongest impact on the yield curve and can quantify their strength relative to other factors. Moreover, we could compare the strength of directional mechanisms in different markets and how it evolves over time by following the evolution of the scaled first eigenvalue. Interestingly, we found that the development of markets is characterized by a decreasing scaled first eigenvalue. In undeveloped markets, almost 100% of the action is explained by one single variable (direction), while increasing sophistication results in other mechanisms gaining strength. For example, while 10 years ago the Indian domestic bond market had a scaled first eigenvalue of virtually one, it has now decreased to levels more in line with Western government bond markets, reflecting its increasing development. The Indian MIFOR–MIBOR basis swap market, on the other hand, has always been similar to the Bund market depicted in Figure 3.4. This could be interpreted as evidence that the basis swap market has induced the development of the domestic Indian bond market. In a sense, the scaled first eigenvalue is a universal (reverse) indicator of the sophistication of a market, allowing comparisons over space and time.

Moving on with the interpretation to the second eigenvector, if α_2 increases by one, short yields decrease and long yields increase. Thus, the second factor represents the slope element of the curve that is not explained by the first factor (i.e. by directional impacts on the slope). This is due to the entries in the second eigenvector crossing the x-axis once. Again, the shape of the second eigenvector reveals and quantifies the shape of steepening moves.

And if α_3 increases by one, short yields increase, medium yields decrease, and long yields increase, which we interpret as the curvature dynamics of the curve (not explained by the first and second factors). This corresponds to the entries in the third eigenvector crossing the x-axis twice.

These results are typical, with the i-th eigenvector crossing the x-axis $i - 1$ times.[6] However, it is not always the case, as in a particular market steepening moves could explain more of the overall yield curve variation than directional moves. Sometimes, no reasonable interpretation is possible at all, which could indicate that modeling this particular market through PCA is not useful. An advantage of PCA versus factor models falling into the first category is that it reveals these issues as they are, including cautioning against its own use when appropriate.

Applying Eigenvector Interpretation in Different Markets

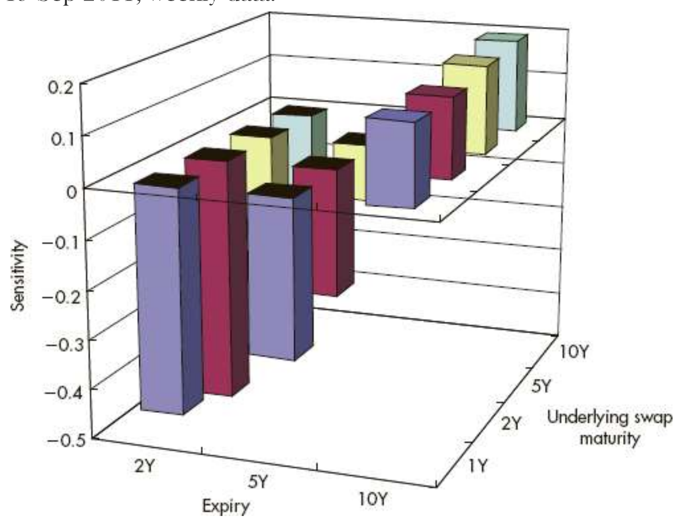
Statistical models like PCA require no specific knowledge about the instrument that is being modeled and are hence universally applicable. PCA only needs to know the time series, not whether the time series represents yields, swap spreads, or volatilities, or what drives that time series. Thus, the range of its possible applications is far larger than the yield curve example we have been using so far. In fact, interpretation of eigenvectors may become less predictable and straightforward, hence more interesting and revealing, when PCA is applied to other markets. Here, we illustrate the use of PCA in a number of different contexts, starting with volatility data as input variables.

The two-dimensional surface of at-the-money-forward (ATMF) volatilities (or, if skew is considered as well, the three-dimensional volatility cube) must first be transformed into a one-dimensional vector. After the PCA is conducted, the outcome in vector form can be displayed again in a two-dimensional format. The complete results are discussed in detail in Chapter 17, from which we pick the chart of the second eigenvector as a typical result in Figure 3.7.

FIGURE 3.7 Second eigenvector of a PCA of the vega sector of the JPY implied volatility surface.

Sources: data – Bloomberg; chart – Authors.

Data period: 5 Jan 2009 to 19 Sep 2011, weekly data.



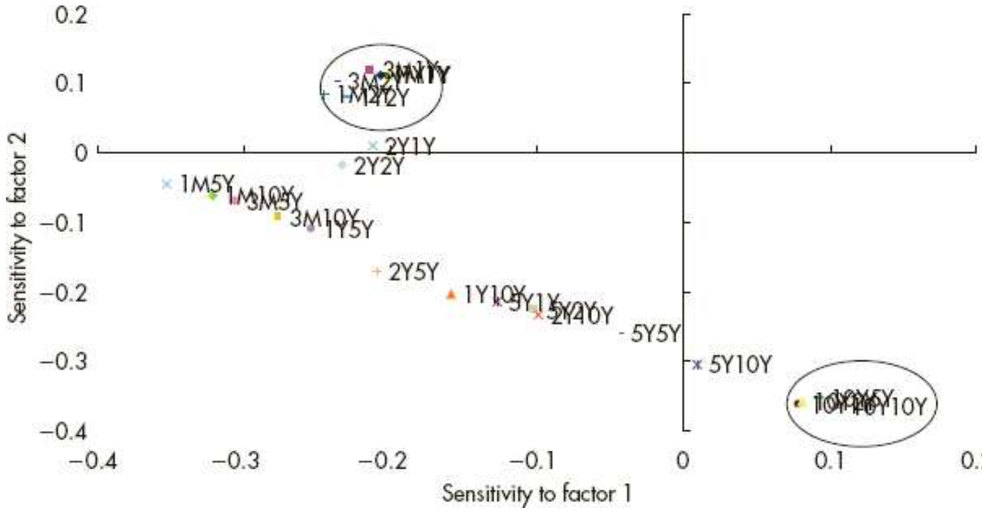
Usually, the first factor represents the overall level of volatility, the second factor differentiates across expiry, and the third one across underlying swap maturity of the options. However, the second and third factors sometimes change place (i.e. the differentiation across underlying maturity explains more of the overall variation of the volatility surface than the differentiation across expiry). Thus, PCA decomposes the volatility surface into its two dimensions (expiry and underlying swap maturity), with the first factor affecting both dimensions and the second and third factor reflecting the dimension-specific information.

When the number of instruments in the data set increases, as in the case of volatility analysis, the ability of PCA to reduce the dimensionality of the data becomes increasingly important. One can use that ability to detect relationships in large data sets. One tool supporting this process is cluster analysis, which depicts each instrument as a function of its sensitivity to various factors (typically the first and second, as in Figure 3.8). If the first and second eigenvalues are large relative to the others, then the behavior of an instrument will be largely determined by its sensitivity to the first and second factors. Thus, if two instruments have similar sensitivities to the first and second factors (i.e. they are close to each other in Figure 3.8), they can be expected to behave similarly (i.e. to form a "cluster"). In the example in Figure 3.8, we can observe two clusters: one containing all options with both short expiries and short underlying swap maturities and the other containing all options with both long expiries and long underlying swap maturities. We therefore conclude that volatilities at both ends of the diagonal of the volatility surface usually move closely together, while further away from those two corners of the volatility surface, options behave more individually. The same sort of analysis can be applied to other combinations and numbers of factors as well, for example by forming three-dimensional clusters of the first three sensitivities.

FIGURE 3.8 Example for a cluster analysis of the whole JPY volatility surface.[7]

Sources: data – Bloomberg; chart – Authors.

Data period: 5 Jan 2009 to 19 Sep 2011, weekly data.

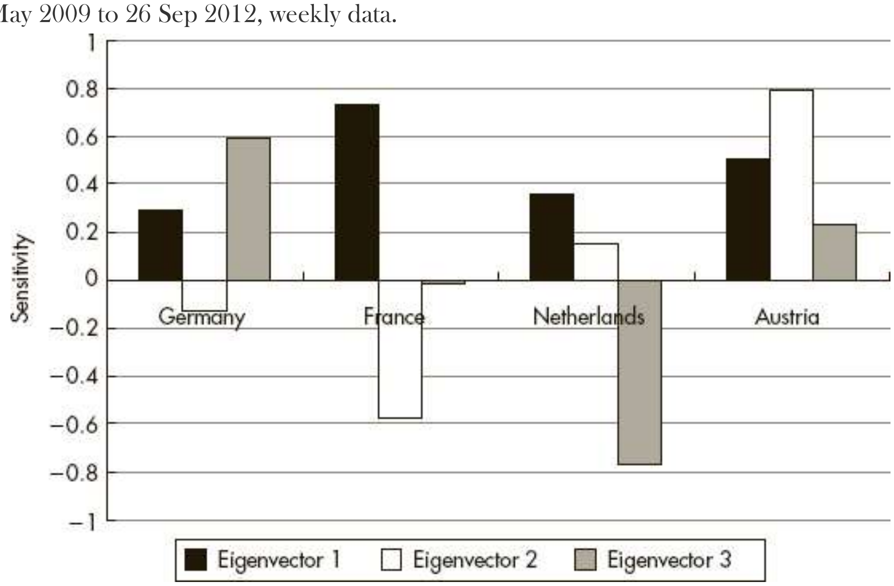


In Chapter 15, we shall use PCA to gain insights into the structure of various credit default swap (CDS) markets. Picking one result as an example, Figure 3.9 shows the first three eigenvectors of a PCA with the 5Y CDS quotes of sovereign issuers in the core Eurozone as input. The first eigenvector (in which all entries are positive) represents the overall level of Eurozone CDS quotes, with the sensitivities measuring the impact a general widening of Eurozone CDS quotes has on individual countries: it affects France more than Austria, Austria more than the Netherlands, and the Netherlands more than Germany. The second eigenvector groups together Germany and France (negative sensitivities) versus the Netherlands and Austria (positive sensitivities). Thus, if factor 2 increases, the CDS of the small core countries widen relative to the CDS of the big core countries. Therefore, factor 2 can be interpreted as differentiating between the big and small countries in the core Eurozone. This means that the size of the bond market is the second-most-important determining factor of core Eurozone countries' CDS levels (after the overall CDS level as measured by factor 1).

FIGURE 3.9 First three eigenvectors of a PCA on 5Y CDS for core Eurozone sovereign issuers.

Sources: data – Bloomberg; chart – Authors.

Data period: 6 May 2009 to 26 Sep 2012, weekly data.



For an example from the commodity market we have run a PCA on weekly data for the front month contract on the three soy-related series (soybeans, soybean meal, and soybean oil) from 2000 onward. The scaled eigenvalues shown in Figure 3.10 indicate that almost everything is explained by factor 1, with factor 2 having only 0.2% of explanatory power and factor 3 virtually nothing. This indicates that differentiation across the three soy products is limited relative to the variability of the overall price changes in the three commodities.

FIGURE 3.10 Scaled eigenvalues of a PCA on the soy market.

Sources: data – Bloomberg; chart – Authors.

Data period: 3 Jan 2000 to 6 Aug 2012, weekly data.

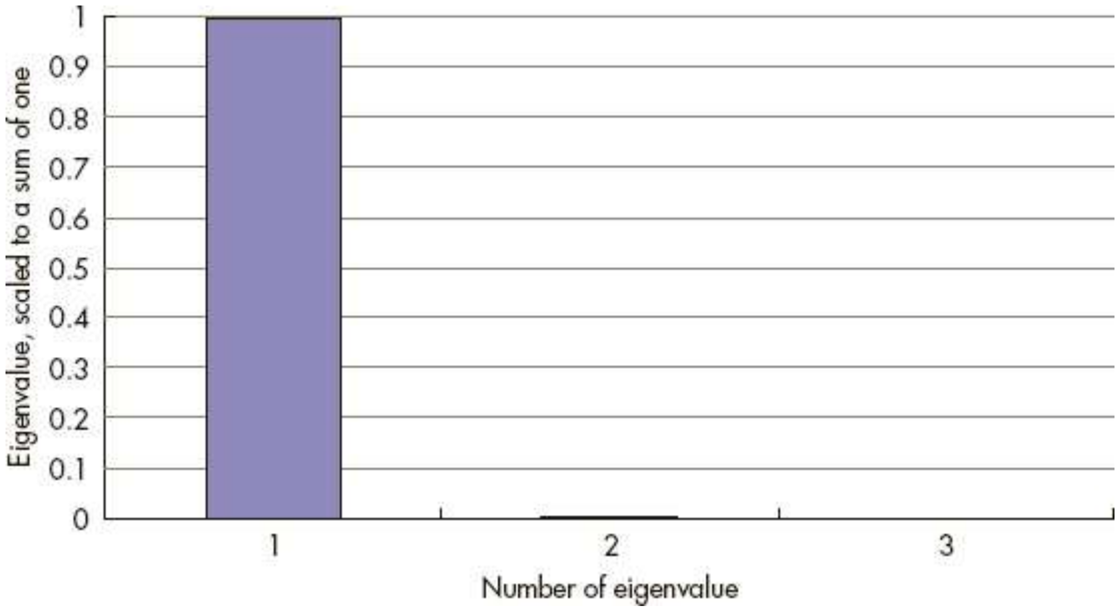
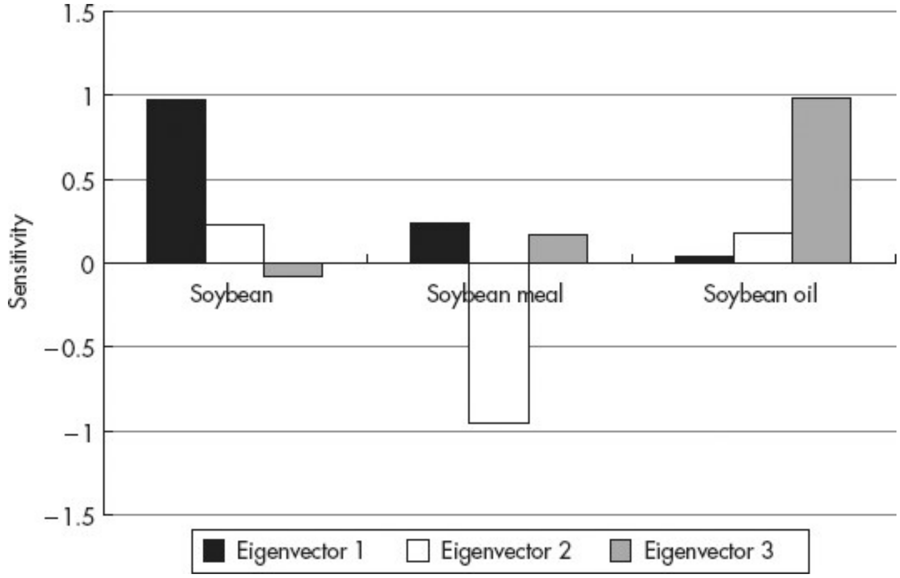


Figure 3.11 displays the eigenvectors. Note that the difference in sensitivities to the first factor is a function of the difference in the size of the input numbers (e.g. 1,600 for soybeans versus 51 for soybean meal). This could be avoided by creating synthetic time series with similar numbers, for example all starting with a value of one. It turns out that the overwhelming factor 1 affects all soy products in the same way (i.e. all rise and fall together). The (little) differentiation between different soy products which is measured by factor 2 shows that soybeans and soybean oil move together versus soybean meal.

FIGURE 3.11 Eigenvectors of a PCA on the soy market.

Sources: data – Bloomberg; chart – Authors.

Data period: 3 Jan 2000 to 6 Aug 2012, weekly data.

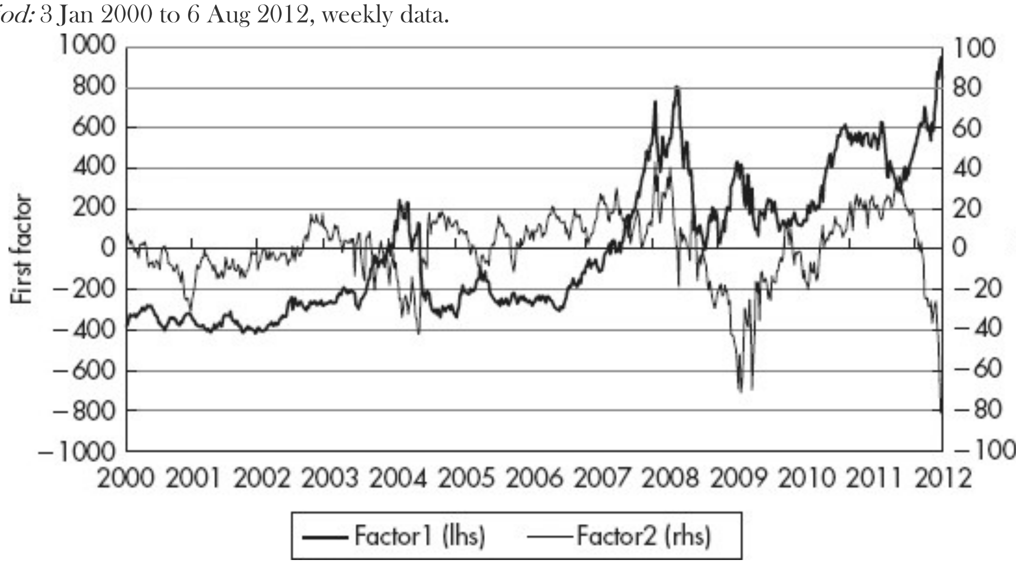


Finally, a look at the evolution of the PCA factors over time (Figure 3.12) reflects the increasing demand for soy products over the last decade in a rise of factor 1, which has led to all three products richening. Factor 2, on the other hand, has not exhibited a clear trend, which means that the price differentiation between soybeans and soybean oil on the one hand and soybean meal on the other tends to be temporary (i.e. factor 2 tends to be mean reverting). From this historical perspective, the current unprecedented deviation of factor 2 from its mean could be seen as a good trading opportunity. Factor 2 being too low (historically) translates through the sensitivities to eigenvector 2 in soybean meal being too expensive versus soybeans and soybean oil. Hence, a low factor 2 reflects the relative richness of soybean meal versus other soy products.

FIGURE 3.12 Historical evolution of the first and second factor of a PCA on the soy market.

Sources: data – Bloomberg; chart – Authors.

Data period: 3 Jan 2000 to 6 Aug 2012, weekly data.



Up until this point, the analysis has been purely statistical, revealing the driving forces of historical and current pricing. This provides a good basis for a trading decision, which needs to incorporate elements besides statistical properties as well. For example, is the current deviation of factor 2 an outlier that can be expected to revert to mean as quickly as it did in the past? Or is there a potential for the current drought to cause a permanent regime shift in the soy markets? While this assessment is beyond the reach of statistics, PCA both enables us to detect and formulate these trading decisions and indicates that any cause for a permanent regime shift would need to be extraordinarily strong – stronger than anything that has happened over the past decade. Thus, PCA shows that arguing for a regime change (i.e. for no reversion of factor 2 to the longstanding mean) would require very good reasons.

An investor believing in the current drought not to be such an extraordinary regime-changing event and thus in factor 2 continuing its decade-long behavior and returning to its mean may therefore want to consider selling soybean meal versus either soybeans or soybean oil. Which of the two is better? And precisely how much should be bought for one short soybean meal futures contract? The answer to these questions is again within the realm of statistics. Given the much stronger impact of factor 1 on the soy market, there is actually a high risk that a position that intends to exploit the current mismatch in factor 2 could end up being driven mainly by factor 1 and not 2 (i.e. overall direction of soy, not the spreads between different soy products), if hedge ratios are not calculated properly. Subsequently, we shall develop PCA into tools that answer these questions.

These examples could serve to give an impression of the analytical strength of PCA, thereby empirically proving our claim that PCA is a very useful tool for the relative value craftsman to dig out and clarify the key market mechanisms hidden below a muddy surface of incommensurable market actions.

So far, we have used PCA to gain insights into the structural core of markets. From now on, we shall focus on applying these insights to find, analyze, and construct trading ideas. This will take place in the framework of PCA and thus show the way PCA supports the translation of its insights into practical trading positions.

Decomposing Markets into Uncorrelated Factors

As the eigenvectors are orthogonal, the factors are uncorrelated by construction. Hence, one can decompose a complex market into individual, uncorrelated, and simple relationships. This ability is the key to a variety of analyses.

Technical points

- Factors are uncorrelated, not necessarily independent. This issue has led to the development of independent component analysis recently, though we have not seen a convincing application to market analysis yet.
- While factors are uncorrelated over the whole sample period, there can occur significant measured correlation within subperiods. Thus, the rolling correlation of factors should be investigated before relying on the model. This is a key point, which we shall discuss in detail toward the end of this chapter.

The decomposition into uncorrelated factors isolates a particular relationship from others. This makes it possible to analyze and trade individual market mechanisms, such as yield curve steepness, without being influenced by other effects, such as direction.[8] The most important application of this ability is the exclusion of directional effects (typically associated with the factor with the greatest explanatory power). If relative value trades are defined as offering return opportunities that are uncorrelated to market direction, then the ability of PCA to produce and analyze time series uncorrelated with market direction is a key aspect of the analysis.[9] In other words, if the first factor represents the market direction (beta), then alpha (P&L uncorrelated to the market direction) can be found in the other factors.

Using the example of the Bund yield curve from Figure 3.5, the second factor has been interpreted as explaining the steepness not explained by the first, directional factor (i.e. representing the *non-directional steepness*). Note that also the first factor has a steepness component (i.e. direction impacts steepness). If yields increase, the 5Y-10Y yield curve tends to flatten. The second factor is uncorrelated to direction *and to directional impacts on the steepness* and therefore shows the steepness of the curve *net* of directional impacts.

To repeat this key point: given the current yield level, the curve should have a certain steepness, given by the first eigenvector. The second factor shows the steepness of the curve that remains after taking into account the steepness already explained by the direction. There were many instances in which the 5Y-10Y yield curve has been steep but in which the steepness was due entirely to the low yield level, while on a non-directional basis the curve was actually too flat. In this case, one could argue for a non-directional steepening trade hedged against directional impacts, despite the steepness of the 5Y-10Y curve.

This example illustrates the importance of factor decomposition for RV analysis:

- Given the all-pervasive directional effects (e.g. on curve steepness, curvature, swap spreads, volatility) the indispensable prerequisite for any RV analysis is a measure for steepness or curvature *unaffected* by directional effects. The second and third factors provide that measure and thus the *starting point* for any RV analysis.
- As the first eigenvector shows the way yield levels affect slope, we can hedge against these effects, thereby formulating RV trades with no directional exposure.
- PCA is thus a direct way to gain access to true RV trades and thereby to much more trading, sales, and research opportunities than straight directional positions. Simply put, by using a three-factor model, one obtains three uncorrelated time series and thus three times as many possibilities for trades.

Embedding PCA in Trade Ideas

So far, we have discovered two important features of PCA: its ability to decompose a market into uncorrelated factors and the possibility to interpret these factors economically (via examining the shape of the eigenvectors).[10] Together, PCA decomposes a market into uncorrelated factors with an economic meaning.

The process of taking a view on the market therefore can be achieved by taking a view on each of the factors. Hence, for each factor, the analyst can decide independently whether to take a view on that factor. His decision should reflect:

- statistical criteria like mean reversion
- fundamental and structural criteria
- flow and other concerns.

A key benefit of PCA is that it links all those criteria (separately for each of the separate factors). PCA factors are not only mean reverting but often also have a meaningful economic interpretation. Hence, they do not only fulfill desired statistical properties but also reveal (in fundamental terms) *why* they have those properties. For example, knowing that the first factor is linked to GDP growth explains its mean reversion by the business cycles and its slow speed of mean reversion by the length of those cycles.

Hence, PCA decomposes the market into uncorrelated mean-reverting factors, which often carry an economic meaning. In other words, PCA allows us to combine statistical analysis with economic analysis so as to associate statistical features of a trade with particular fundamental considerations. For example, the second factor may be negatively correlated to the EUR exchange rate. In this case, a statistically attractive steepening trade looks even more appealing to an investor if he expects the EUR to weaken for fundamental reasons.

While statistical analysis is by construction backward looking, linking it to fundamental variables enables traders to incorporate forward-looking (economic) expectations. This ability to incorporate potential future risks in the analysis is a key benefit of linking statistics to external driving forces. In the example above, a steepening trade may well look statistically attractive, but this could all be due to EUR strength. Thus, a further EUR strengthening, which may be caused by political decisions independent of any statistical properties, is a risk to the trade. The link of statistics through a PCA to these external driving forces can identify these risks. An analyst knowing that the statistically attractive steepness is due to EUR strength will refrain from the trade if he sees a significant risk for further EUR appreciation. As in this example, outside information about macroeconomic events can help with both explaining some observed statistical properties and incorporating forward-looking information into the analysis.

The general form of this analysis is to investigate the following link, (e.g. via a regression):

Factor ~ External explaining variable(s)

While by construction this type of analysis is outside of the statistical reach of PCA, PCA both *enables* this analysis by generating the dependent variable and *facilitates* the search for relevant external explaining variables by revealing the meaning of its factors. One could investigate the link to external variables by heuristic methods, for example by trying all available financial time series in a regression table against all factors. Note that some financial time series could be trending and therefore not suitable for a regression.[11] The interpretation of the eigenvectors can facilitate the search for the "right" explanatory variables. As these relationships evolve, we recommend monitoring them over time, for example via rolling correlations. For the example of a PCA on the Bund yield curve, the correlations versus some candidates for external driving forces[12] are summarized in Table 3.1.

TABLE 3.1 Correlations of the First Three Factors of a PCA on the Bund Yield Curve versus Candidates for External Explaining Variables

	Factor 1	Factor 2	Factor 3
Factor 1 of a PCA on USD swaps	0.73	0.62	−0.10
Factor 2 of a PCA on USD swaps	0.89	0.40	−0.02
Factor 3 of a PCA on USD swaps	−0.21	−0.59	0.31
5Y Bund vol (6M rolling)	−0.52	−0.65	0.03
S&P500	0.60	−0.55	0.11
VIX	−0.72	0.01	−0.16
EUR FX rate	0.44	−0.63	−0.23
Oil	0.58	−0.63	0.08

It can be seen that both factor 1 and 2 are significantly influenced by a number of macroeconomic variables. We note in particular the strong link to the US yield curve (as represented by the factors of a PCA) with the surprising fact that factor 1 of the Bund PCA is most correlated to factor 2 of the USD PCA and factor 2 of the Bund PCA to factor 1 of the USD PCA. This could be the starting point for a further investigation, which may reveal interesting differences in the driving forces of global bond markets and their interconnections. Furthermore, the link between factor 2 and currencies as well as commodities jumps out at you, indicating that further analysis may well yield valuable results. Factor 3, on the other hand, seems to be rather uncorrelated to external driving forces. This could indicate that factor 3 is a relatively "pure" relative value factor (i.e. with little correlation to macroeconomic events). This is a typical result, that is while the lower factors often exhibit a high correlation to macroeconomic variables (reflecting the high impact of economic events on markets), higher factors are usually less correlated. As a rule of thumb, the higher the factor, the more weight statistical analysis carries, while the examination of potential external economic risks and the incorporation of forward-looking analysis described above become less important.

As an alternative to the regression of PCA factors versus explaining variables, which is done outside of the PCA itself, one could also include the candidates for explaining variables in the PCA, for example by running a PCA on input data consisting of time series for Bund yields and USD swap rates simultaneously.

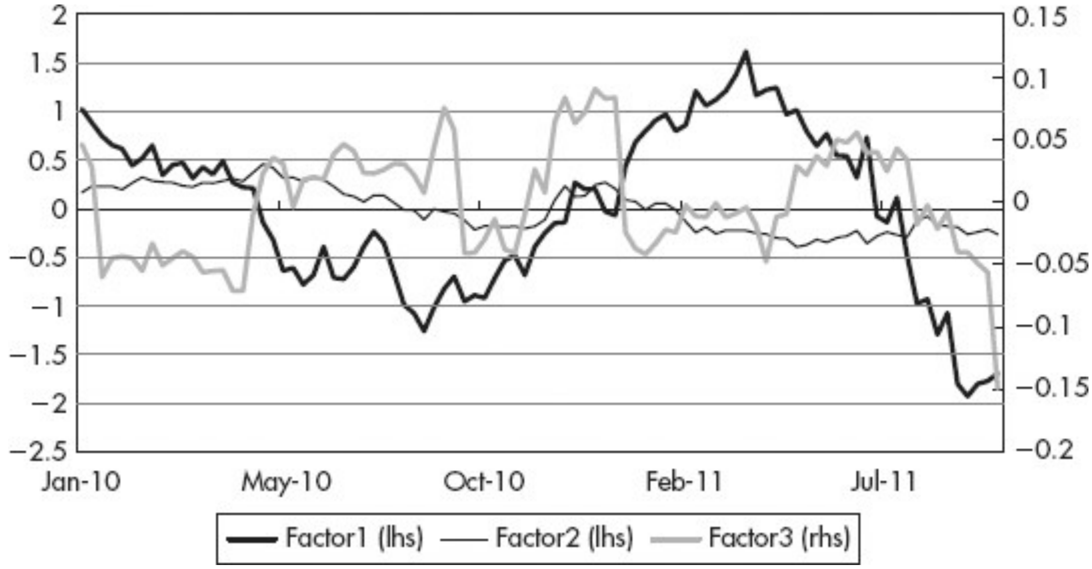
Furthermore, PCA may reveal how flows affect the pricing. For example, how much does the k-factor residual for 5Y rates move when a new 5Y bond is issued? Is this a stable pattern? Is the spike linearly dependent on the issuance size? This allows traders to incorporate more technical issues into the framework of a PCA.

Let's see how the ability of PCA to decompose the Bund yield curve into economically meaningful uncorrelated and statistically mean-reverting factors could support finding trade ideas in practice. Figure 3.13 shows the evolution of the first three factors of the Bund yield curve over time.

FIGURE 3.13 History of the first three factors of a PCA on the Bund yield curve.

Sources: data – Bloomberg; chart – Authors.

Data period: 4 Jan 2010 to 3 Oct 2011, weekly data.



Using mean reversion models, we can assess for all factors the distance from their long-run means and the speed with which they are likely to return to these means. In our current example, we may conclude that factor 1 has insufficient speed of mean reversion and thus we take no view. Similarly, factor 2 is relatively close to its mean, so we take no view. Factor 3, however, seems to be considerably away from its mean and to have a high speed of mean reversion. Hence, we decide to investigate further for trade ideas based on factor 3 (i.e. butterflies hedged against factors 1 and 2).

The statistical side of that investigation is done by a formal mean reversion model, which can calculate the expected return and risk attributes of various trades for specific time horizons.

Factor 3 has little exposure to economic variables and can thus be treated as a "pure" relative value play on mean reversion, with little risk of its statistical properties being thrown off track by macroeconomic events. In this case, the regression Table 3.1 above can serve as justification for restricting the analysis to statistics.

If we took a view on factor 1, on the other hand, we would need to consider the link of the time series of factor 1 and its statistical properties to its external driving forces. Since we found factor 1 to be positively correlated to USD rates and negatively to volatility, among others, the spike in the VIX on 3 Oct 2011 could help us understand the current statistical deviation from the mean. And if we thought that volatility was going to decrease again soon (e.g. because we expected the ECB to calm fears about the euro crisis), we would have both a statistical (mean reversion) and a fundamental reason to bet on an increase in factor 1. Conversely, the mean reversion of factor 1 (as given by statistics) certainly supports our fundamental expectation for volatility to decrease again (i.e. for the current point to be an outlier, i.e. a good entry opportunity). Additionally, one could run a regression between the VIX and factor 1 and decide via the current residual whether the fundamental expectation of decreasing volatility would be better expressed by a short option or a short Bund position.

As in this example, PCA guides the analyst through the process of taking informed views (or the informed absence of a view) on each of the uncorrelated factors: it reveals the statistical properties and economic meaning of each factor behind a position and gives thereby a good basis for a reasonable and substantiated decision, separately for each factor, whether or not to take a view on it.

Once the analyst has decided on which factors he wants to take a view, PCA does both jobs that are needed to execute that view via the most suitable position:

- PCA provides the hedge ratios to immunize against exposure to factors to which he does not want to be exposed.
- The $(n-1)$-factor residuals indicate those individual instruments that provide the best execution for his view on factor n. This can also be used to apply PCA for general asset selection purposes.

In the following, we will develop these features of PCA, before we integrate them back into the process of finding and analyzing trades.

Appropriate Hedging

PCA not only breaks down the driving forces of a trade (like a 5Y-10Y steepening position) into uncorrelated factors but also quantifies their impact and thereby allows us to hedge against specific factors. In order to create a 5Y-10Y steepening position which is not affected by the first factor (i.e. by direction and by directional impacts on the slope), we can simply see how changes in the first factor impact 5Y and 10Y yields and choose the hedge ratios in such a way that both net out. This leaves a position, which is hedged against changes in the first factor, that is a steepening trade which is only affected by the non-directional steepness (and higher factors like curvature), not by directional impacts on the curve and thus a source of P&L uncorrelated to market direction (alpha).

In the PCA framework, hedge ratios are calculated in order to immunize a portfolio against changes in factors. In order to execute the non-directional 5Y-10Y steepening position of the example above, we need to hedge against changes in the directional factor α_1, and the ratio of notionals for 5Y and 10Y is thus: $\frac{n_5}{n_{10}} = \frac{BPV_{10}}{BPV_5} \cdot \frac{e_{110}}{e_{15}}$.

In the formula above, the ratio $\frac{e_{110}}{e_{15}}$ represents the quotient of 10Y and 5Y sensitivities to changes in the first factor. This may sound similar to the "beta-adjustment" of basis point value (BPV) hedge ratios by the slope of the regression line of a regression between the two instruments involved. However, as a regression minimizes the *conditional* expected value of deviations, it is conceptually quite different from a PCA. This is reflected in the practical problem that, unless the correlation is 1 or −1, the hedge ratio determined by beta adjustment changes when the dependent and independent variable in the regression are exchanged. PCA works without conditional expected values and is hence free from this problem.

Hedge ratios against more factors are best calculated via matrix inversion. For example, the hedge ratio for a 2Y-5Y-10Y butterfly which is neutral to changes in the first and second factor can be calculated for a given notional n_5 for 5Y by:

$$\begin{pmatrix} n_2 \\ n_{10} \end{pmatrix} = \begin{pmatrix} BPV_2 \cdot e_{12} & BPV_{10} \cdot e_{110} \\ BPV_2 \cdot e_{22} & BPV_{10} \cdot e_{210} \end{pmatrix}^{-1} \cdot \begin{pmatrix} -n_5 \cdot BPV_5 \cdot e_{15} \\ -n_5 \cdot BPV_5 \cdot e_{25} \end{pmatrix}$$

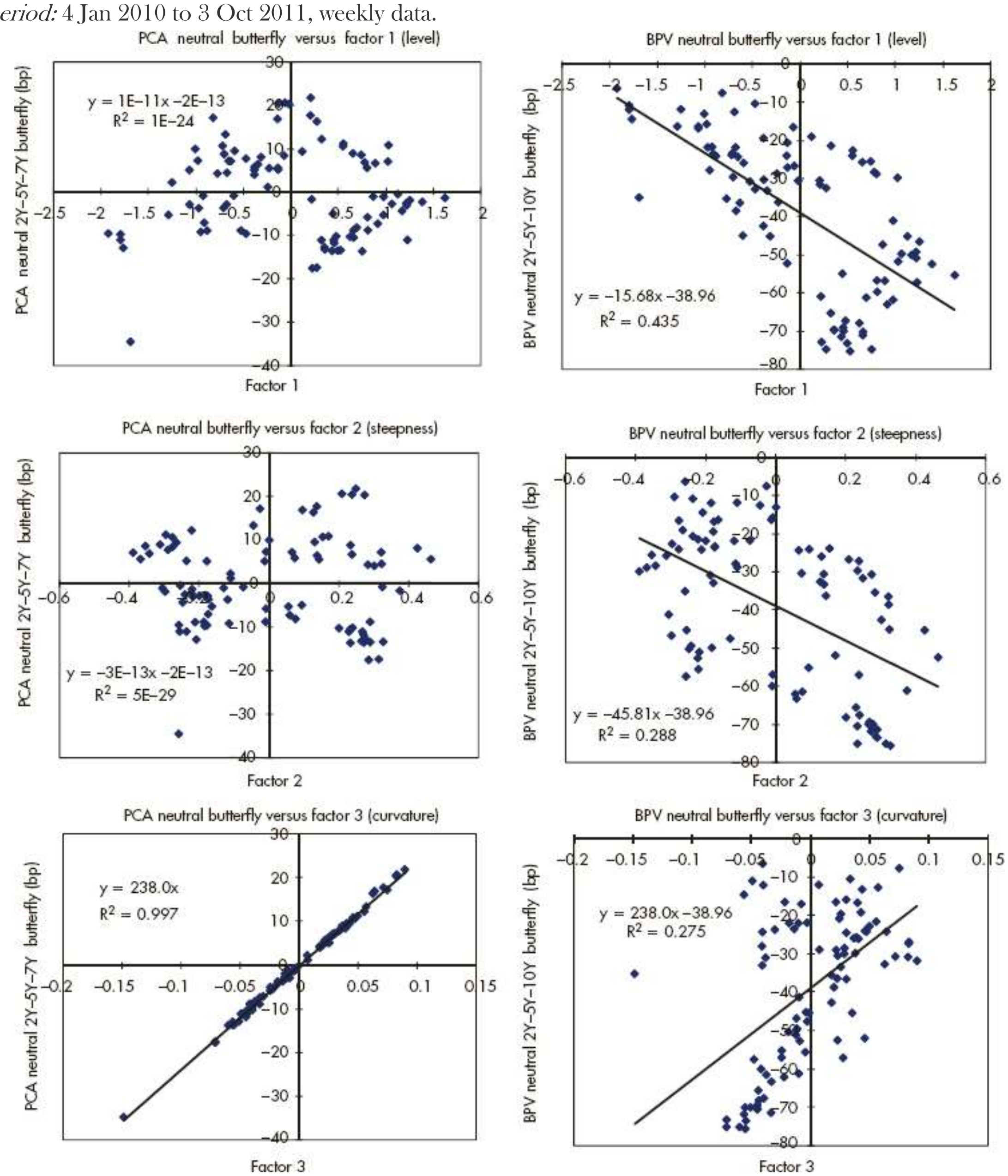
We argue that hedge ratios should be based on PCA rather than on BPV neutrality. BPV neutrality corresponds to assuming arbitrarily that all entries in the first eigenvector are the same (which is in fact our very crude example for a factor model of the first category at the beginning of this chapter). The market tells us that they are not. In other words, directional shifts do affect different points on the yield curve in different ways. If the 5Y Bund yield increases by one basis point (bp), the 10Y yield is expected to increase by 0.87 bp. True neutrality with respect to the level of yields needs to take these impacts of the direction on the curve shape into account by using hedge ratios based on PCA eigenvectors. By contrast, BPV neutrality results in positions with directional exposure.

To illustrate our point, we have regressed a BPV-neutral weighted 2Y-5Y-7Y butterfly and a PCA-neutral weighted 2Y-5Y-7Y butterfly on the Bund curve against the first factor (representing directionality), the second factor (representing non-directional steepness), and the third factor (representing curvature net of the impacts of the first and second factor on curvature, i.e. net of the impacts of direction and of non-directional steepness on curvature[13]). The results are depicted in Figure 3.14. It turns out that the BPV-neutral butterfly is not only directional but also provides little exposure to the third factor. Hence, by choosing BPV-neutral weights for a 2Y-5Y-7Y butterfly, the investor ends up with a fuzzy exposure to direction, non-directional steepness, and just a bit of net curvature. Given the strength of directional impacts, virtually all of his performance will be driven by the market direction, thus transacting Bund futures would have resulted in almost the same exposure (with fewer costs and a cleaner understanding of the risk profile). Presumably, he wanted to be exposed to factor 3 rather than a fuzzy combination of risks; however, by BPV-neutral weighting, he lost almost all net curvature exposure and ended up with a risk profile he probably did not even know of (because PCA is needed for the decomposition into factors) and is now heavily dependent on the market direction against which he thinks he is "hedged". Note that this example is still benign in the sense that at least the correlation of the BPV-neutral butterfly to the third factor is positive. However, this does not need to be the case and there are instances where BPV-neutral steepeners are actually non-directional *flatteners*. Thus, BPV-neutral "hedging" may give any exposure, by chance the right one, a completely different one (like in the butterfly example above, which is mainly dependent on factor 1 rather than 3), or even an exposure which is opposite to the intended one. Only PCA allows the decomposition of the exposure into its components – and hence also the calculation of the "right" hedge ratio (i.e. the one leaving only the desired exposure).

FIGURE 3.14 Driving forces of a BPV-neutral and a PCA-neutral 2Y-5Y-7Y butterfly on the Bund curve.

Sources: data – Bloomberg; chart – Authors.

Data period: 4 Jan 2010 to 3 Oct 2011, weekly data.



This is depicted in Figure 3.14 in the regression charts of the PCA-neutral butterfly, which is fully exposed to the net curvature (factor 3). Further, this is the only factor to which the PCA-neutral butterfly is exposed.[11] Hence, only the PCA hedge produces a position with a clean and clear exposure to precisely those factors, to which an exposure is intended.

Note the difference of goals in PCA-neutral hedging and hedge ratios, which are calculated in order to maximize Sharpe ratios. While hedge ratios determined by maximizing Sharpe ratios (or another desired feature) obviously have the best statistical properties, PCA-neutral positions give a clear picture, to which factors a trade is exposed. And since these factors often have a meaningful economic interpretation, the investor can assess and adjust his risk profile in economic terms. For example, he knows that the PCA-neutral butterfly is hedged against directional impacts on the yield curve. One cannot say that statistically optimized hedge ratios or PCA-neutral hedge ratios are "better", they just achieve different goals by construction: statistical optimality of a position or a position with a clear exposure to economically interpretable variables. In a manner of speaking, the investor choosing PCA-neutral hedge ratios rather than those optimizing Sharpe ratios gives up some statistical benefits in order to gain the confidence of knowing the exposure of his trade in economic terms, for example the confidence of knowing that it is immune from directional impacts.

Analyzing the Exposure of Trading Positions and Investment Portfolios

Just as in the example of Figure 3.14, any position can be regressed against the uncorrelated factors of a PCA. Thereby, its driving forces reveal themselves. This gives a clear picture of what really impacts the P&L of complex trades. For those investors using BPV-neutral hedging, the result can be quite surprising, as in the example above. Fortunately, if the actual exposure turns out to be different from the desired one, applying the appropriate PCA hedge ratios solves the problem.

For example, if a market maker has just transacted a BPV-neutral 2Y-5Y-7Y butterfly with an investor, he could run the regressions from above, find that he is now exposed 44% to factor 1, 29% to factor 2, and 28% to factor 3, see how this nets out with other positions in his portfolio and initiate the appropriate hedging (against those factors to which he does not want to be exposed).

Moreover, basing analysis on orthogonal eigenvectors "orthogonalizes" the mind of analysts. It helps us provide a clean and clear analysis. Which trade is exposed to which factors? Do I like that exposure or not? If not, how can I hedge against it? Do I recommend a butterfly due to its directionality or due to its curvature exposure? This breakdown of any trade into its uncorrelated driving forces prevents the all-too-common mix-up of arguments, for example an analyst arguing via curvature for a butterfly while it has, in fact, a purely directional exposure.

Market Reconstruction and Forecasting

Reversing the decomposition of a complex market into its uncorrelated components, factor models in general allow an approximate reconstruction of the larger market from a limited number of factors.

And among all factor models, PCA does the best job in that it provides the closest possible reconstruction to reality given limited information. This is the basis for a variety of applications, a couple of which we briefly highlight now.

Often, market participants are interested in knowing how a certain event would impact the overall market. For example, how would the yield curve be expected to react if CPI were to increase by 1%? Or what would it mean for the yield curve, if 2Y Bund yields increased by 25 bp? In mathematical terms, this corresponds to reconstructing the whole set of information (yield curve) from one single piece of information (short rates go up by 25 bp). The first eigenvector contains just this transformation, as it translates the movement of one point on the yield curve into the move of the whole curve. Using the sensitivities of Figure 3.5, an increase of 2Y yields by 25 bp would be associated with 5Y yields rising by 33 bp, 7Y by 24 bp, and 10Y by 23 bp on average. These numbers are given by multiplying 25 bp with the ratio of sensitivities to the first factor, just as in the calculation of PCA hedge ratios. Since by construction, the first factor contains maximum information, the PCA reconstruction minimizes the residuals (i.e. provides the best possible picture of the overall market given the informational constraints).

If the reaction of the yield curve to external driving forces should be estimated, the change of an external variable can in a first step be translated via a regression into changes of factors and then the yield curve can be reconstructed from those factors. For example, a change in the EUR FX rate affects all three factors of a PCA on the Bund curve (see regression Table 3.1). Thus, the impact of a one-point change in the EUR FX rate on the whole Bund yield curve can be assessed by calculating its impact on the first three factors (via a regression) and then reconstructing the whole yield curve from those factors.

A similar method can be used to hedge and price trading books. Imagine that since the last close, Schatz, Bobl, and Bund futures have moved by 2, 3, and 2 bp respectively. Now, a client asks a trader to buy an illiquid 7Y Bund, in which no market action has occurred since the last close, from him. The best hedge the trader can achieve in that situation is to reconstruct the yield curve (or only the 7Y point) from the information about the 2Y, 5Y, and 10Y points via the first three factors of a PCA and hedge his position in an illiquid 7Y Bund via a combination of Schatz, Bobl, and Bund futures. The hedge ratios are calculated as described above in order to achieve neutrality against the first three factors. Again, given the constraints (just three liquid instruments), this is the best hedge possible (i.e. it has smallest residual to the actual moves of the 7Y Bund).

A Yield Curve Model Based on PCA

PCA is a deterministic linear algebra tool that simply transforms the basis of the covariance matrix without introducing any stochastic process. Correspondingly, if there are n input variables, PCA will return n factors ($k = n$), and the PCA model outlined above has no residual.

One can now *artificially* introduce a residual by redefining factors as residuals. With $\begin{pmatrix} \varepsilon_1^t \\ \vdots \\ \varepsilon_n^t \end{pmatrix} := \sum_{i=k+1}^{n} \alpha_i^t \cdot \begin{pmatrix} e_{i1} \\ \vdots \\ e_{in} \end{pmatrix}$, the PCA model $\begin{pmatrix} y_1^t \\ \vdots \\ y_n^t \end{pmatrix} = \sum_{i=1}^{n} \alpha_i^t \cdot \begin{pmatrix} e_{i1} \\ \vdots \\ e_{in} \end{pmatrix}$ becomes $\begin{pmatrix} y_1^t \\ \vdots \\ y_n^t \end{pmatrix} = \sum_{i=1}^{k} \alpha_i^t \cdot \begin{pmatrix} e_{i1} \\ \vdots \\ e_{in} \end{pmatrix} + \begin{pmatrix} \varepsilon_1^t \\ \vdots \\ \varepsilon_n^t \end{pmatrix}$ and now *looks like* a k-factor stochastic yield curve model.

This specific approach has two important consequences. First, the analyst can decide the number of factors to use for a PCA-based model. There is no argument that forces him to limit the factor decomposition to two, three, or seven factors. Hence, he can choose the number of factors freely in his model. He will typically base his choice on the goal of the analysis and the structure of the eigenvalues. In particular, external knowledge about market mechanisms can be introduced via the selection of the number of factors. If, for example, in the process of eigenvector interpretation it turns out that the yield curve is driven by the first three factors, then the analyst can conduct a three-factor decomposition and justifiably treat the remaining factors as stochastic residuals.

Second, and key for the following, the residuals will continue to exhibit a factor structure when the number of factors used in the model is lower than the number of factors exhibited by the data. In the example above, the two-factor residuals will consist of both the third factor and the three-factor residuals (which, in this example, are assumed to have no clear factor structure and are therefore indiscriminately treated as noise). In general, if the eigenvalues are decreasing quickly, it will hold true that $\begin{pmatrix} \varepsilon_1^t \\ \vdots \\ \varepsilon_n^t \end{pmatrix} \approx \alpha_{k+1}^t \cdot \begin{pmatrix} e_{k+11} \\ \vdots \\ e_{k+1n} \end{pmatrix}$, that is:

k-factor residual ≈ (k + 1)-th factor × (k + 1)-th factor loading.

In particular, the shape of k-factor residuals will correspond to the shape of the (k+1)-th factor loading and the size of k-factor residuals to the size of the (k+1)-th factor. Thus, *k*-factor residuals are high if and only if factor (k+1) is high (in absolute terms). And as high residuals indicate candidates for trading opportunities, the analyst only needs to follow a few factors in order to keep track of all potential trades. This obviously greatly simplifies the task of screening the market for trade ideas.

PCA as a Tool for Screening the Market for Trade Ideas

This factor structure means in practice that by following a few uncorrelated factors the analyst can monitor all relevant market developments in an easy and orderly manner. Outliers in the factors can be directly translated into candidates for trade ideas. By contrast, the set of all BPV-neutral butterflies, which may at first glance appear to offer more independent trading possibilities (through different combinations, like 1-2-3, 1-2-5, 2-5-7), do not contain *more* information than the three factors (and their residuals) but just represent the same information in different and meaningless *mixtures*. For example, a 1-3-5 butterfly may mix 70% of factor 1, 25% of factor 2, and 5% of factor 3. A 2-7-10 butterfly may mix 80% of factor 1, 10% of factor 2, and 10% of factor 3. All these mixtures are endless repetitions of the same information, which can be identified and expressed cleanly, clearly, and easily through a few uncorrelated PCA factors.

Therefore, the yield curve model based on PCA can be used to screen the market for statistically attractive trading opportunities.

Applying this theory to the example of the Bund yield curve, we see with one glance on Figure 3.13 that factor 2 is close to its mean and hence it makes no sense to look for non-directional steepness positions.[15] Factor 3, on the other hand, seems to be significantly away from its mean (and, unlike factor 1, exhibits a sufficient speed of mean reversion). Thus, we would focus our attention on butterfly positions (neutral against factor 1 and 2).

Note again that almost all the relevant information needed to assess the statistical features and attractiveness of yield curve trades is contained in the single Figure 3.13. A screenshot like this can therefore greatly simplify the efforts of an analyst and direct them toward the most promising targets. This is a direct exploitation of the mathematics of a PCA, which reduce almost all the information contained in the Bund curve to three numbers (factors). Correspondingly, the single Figure 3.13 contains all information about the historical evolution of the Bund market, in an orderly and clear way. This is the basis that enables in practical terms an easy and systematical screening of the whole Bund market for trades, missing none and counting no one twice. After getting used to it, one does not like to follow a market in a different form than a factor decomposition like Figure 3.13.

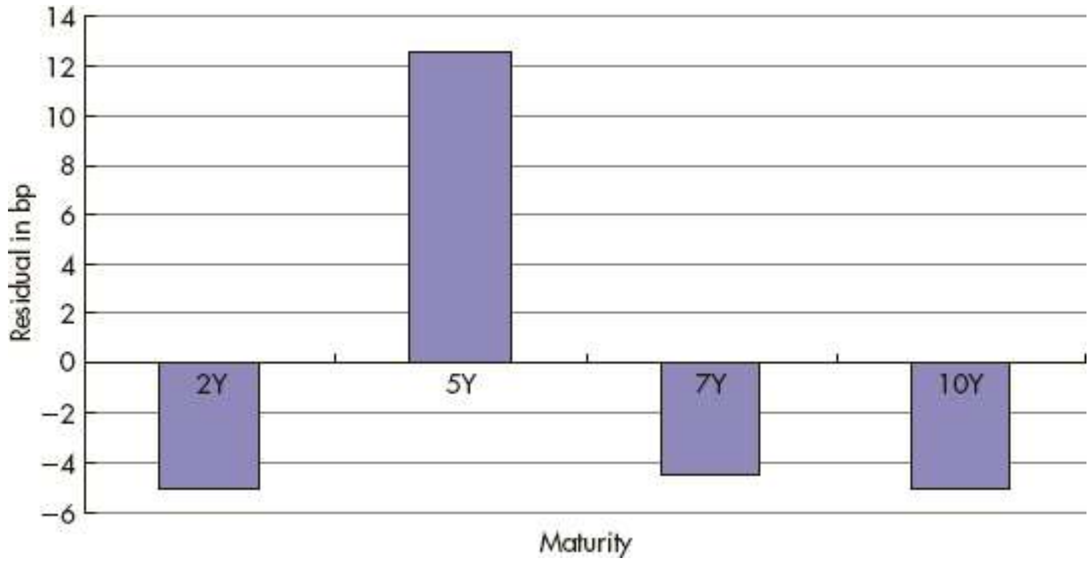
PCA as a Tool for Asset Selection

Having identified butterflies as candidates for curve trades, we can select the best maturities (from a statistical point of view) by looking at their two-factor residuals (i.e. the sum of the third factor times the third factor loadings and the three-factor residuals). In case of the eigenvalues decreasing quickly, this will almost equal the *shape* of the third factor loadings, and indeed the shape seen in Figure 3.15 is very similar to that of the third-factor loadings depicted in Figure 3.5 (mirrored by the x-axis since the third factor is currently negative). Consequently, a 2Y-5Y-10Y butterfly maximizes the potential profit from the current deviation of factor 3 from its mean.

FIGURE 3.15 Two-factor residuals.

Sources: data – Bloomberg; chart – Authors.

Data period: 4 Jan 2010 to 3 Oct 2011, weekly data; current residuals as of 3 Oct 2011.

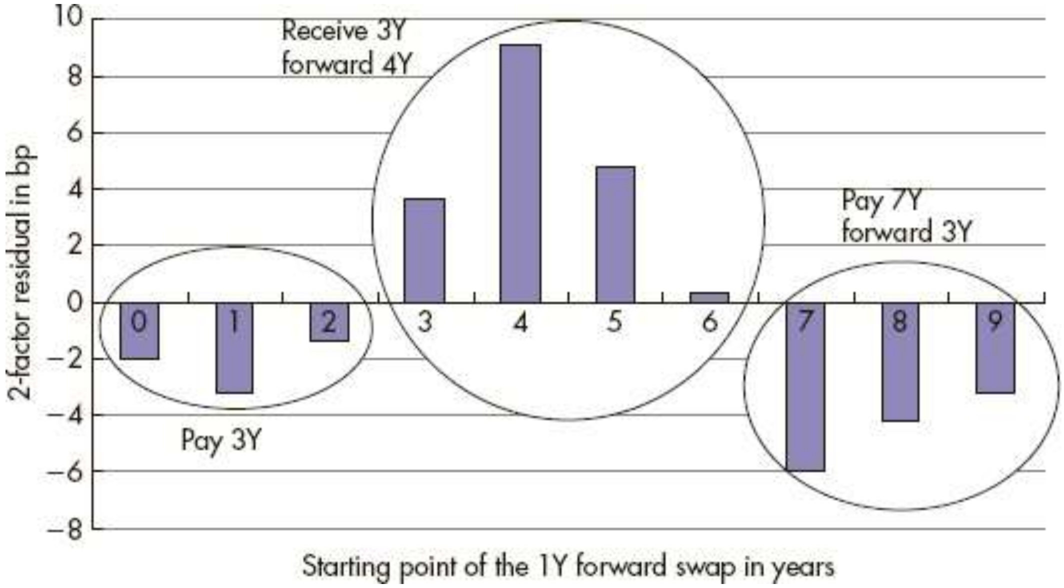


Moreover, the third factor mainly determines the *size* of the two-factor residuals (i.e. the potential profit from mean reversion). Given that factor 3 is currently quite exceptionally far away from its mean (Figure 3.13), the residuals (Figure 3.15) are large by historical standards and hence the 2Y-5Y-10Y butterfly looks attractive also from a historical point of view. Note how favorably such a PCA-based top-down analysis compares to the bottom-up approach of screening all possible butterfly combinations.[16]

When analyzing swap curves, the selection of the best maturities can of course also be done by displaying a residual chart like Figure 3.15 for the swap rates of each maturity. Additionally or alternatively, one can use the consecutive one-year forward swap rates (i.e. the 1Y, 1Y forward 1Y, 2Y forward 1Y, 3Y forward 1Y etc. swap rate) as input in a PCA. This is like running the analysis on the consecutive building blocks of the yield curve rather than on the combinations of those building blocks into usual swap rates. Since there is less information overlap between the consecutive one-year forward swap rates than between usual swap rates (the 29Y and 30Y swap rates contain the same consecutive 29 one-year forward swap rates and differ just in the last one), usually the correlation between the consecutive one-year forward swap rates is lower, which may result in better statistical properties of the PCA. Furthermore, the output of that PCA will be a residual chart showing richness and cheapness *of the individual building blocks* of the yield curve. For the example of a factor 3 trade on the consecutive one-year forward swap rate curve, the picture of two-factor residuals could look like the one shown in Figure 3.16.

FIGURE 3.16 Two-factor residuals of a PCA on consecutive one-year forward EUR swap rates.

Sources: data – Bloomberg, authors; chart – Authors.



Then, the areas of richness and cheapness can be combined. In the example of Figure 3.16, the resulting butterfly trade would be pay 3Y (the combination of 1Y, 1Y forward 1Y, and 2Y forward 1Y), receive 3Y forward 4Y (the combination of 3Y forward 1Y, 4Y forward 1Y, 5Y forward 1Y, and 6Y forward 1Y), pay 7Y forward 3Y (the combination of 7Y forward 1Y, 8Y forward 1Y, and 9Y forward 1Y). Note that this approach allows a much sharper selection of rich and cheap areas than the usual swap rates. For example, the cheap 5Y swap rate contains the rich 2Y swap rate, while the 3Y forward 4Y rate does not. Thus, combining the building blocks of the rich and cheap rate curve areas together offers the maximal exploitation of the residuals. On the other hand, transaction costs are usually higher for awkward forward rate combinations than for plain vanilla swap rates. There is hence a tradeoff between the higher residuals and the higher transaction costs for trades on the consecutive forward rate curve. Altogether, using consecutive forward rates is particularly advisable in case the statistical properties of the PCA need to be improved and in case the profit from asset selection (residuals) is high.

In general terms, once an investor has defined the desired factor exposure of his portfolio, the PCA residuals to these factors will show the best way to get this exposure. And since the residuals are uncorrelated to the factors, this asset selection method offers a source of profit which is uncorrelated to the factors (i.e. to whether the view on factors turns out to be right or wrong).

For example, imagine an investor is bullish on Bunds in general and thus decides to buy a Bund[17] (i.e. to get exposure to factor 1). He can then look at the chart shown in Figure 3.17 of one-factor residuals (i.e. the difference between actual Bund yields and where they typically should be, given the overall yield level). (Note again how closely the shape of one-factor residuals follows the shape of the second eigenvector.) A residual of 17 bp for 2Y means that a 2Y Bund trades 17 bp too cheaply relative to the overall yield level (factor 1). These 17 bp are a source of profit uncorrelated to the direction (i.e. alpha), enhance the return in case of the bet on direction turning out to be right, and serve as a cushion in case of the bullish view turning out to be wrong.

Of course, it is also possible to trade only the residuals, without exposure to the factors. In the case of a three-factor model, this corresponds in our framework to taking no view on any of the three factors, hence hedging against all of them (by using a combination of four instruments, what is sometimes called a *condor*), leaving exposure only to the residuals. As we have seen already, the influence of economic variables is typically concentrated on the lower factors; hedging against all factors and just exploiting the residuals (i.e. the combination of all higher "factors") usually leads to "pure" relative value trading with a focus on statistical properties.[18]

Example of a PCA-based Trade Idea

Let us now put together the elements developed above to illustrate the flow of analysis that could lead to a PCA-based trade recommendation. For the sake of simplicity, we present it in the form of a step-by-step guide (Box 3.1), but note that this cooking recipe is only a rough template that may need to be adjusted to different situations. Also, a reader wishing to get started with PCA would do well to set up an IT environment that was able to reproduce the charts used here. To facilitate this task, we have put the Excel sheet that produced these screenshots on the website accompanying the book, while referring to the disclaimer and warnings contained in it.

Box 3.1 Constructing trade ideas with PCA: A step-by-step guide

Step 1: Decide on the relevant input data for the PCA, depending on the goals of the analysis: time horizon, type and number of variables, change or level data. Usually, at least one year of data is advisable. In our example, we use weekly level data for 2Y, 5Y, 7Y, and 10Y Bund yields from 4 Jan 2010 to 3 Oct 2011.

For heuristic purposes (i.e. to find the best instruments), use a large number of variables (e.g. all yields from 1 to 30 years); if the instruments of the trade are already known, restrict the input to these. Actually, one can run the PCA twice, first with a large set of variables (like from 1 to 30 years), and after the best instruments have been found for a second time (e.g. just with 2Y, 5Y, and 7Y for a butterfly trade), in particular to calculate the hedge ratios.

Step 2: Run the PCA. Check numerical stability and the results: are the calculated eigenvectors really eigenvectors and orthogonal to each other?

Step 3: Display the eigenvalues (Figure 3.4) and assess the factor structure of the market analyzed. In particular, does factorization makes sense at all (i.e. do the eigenvalues exhibit a clear factor structure)? A strong decrease of eigenvalues (i.e. $|\lambda_1| >> |\lambda_2| >> \ldots >> |\lambda n|$) together with a high correlation within the data, corresponds to a clear factor structure and allows a meaningful reduction in the dimensionality of the data, of information into factors. On the other hand, if correlation within the data is small, meaningful information reduction will be impossible. For example, if $Cov = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & 1 \end{pmatrix}$, then every point on the yield curve will have its own factor.

Step 4: Display the (relevant) eigenvectors (Figure 3.5) and interpret them.

Step 5: Display the time series of the (relevant) factors (Figure 3.13). This is a crucial result that serves as the basis for a number of subsequent actions, for example:

- Assess the statistical qualities of each factor; in particular, is it close to its mean or far away?
- Check the correlation between factors over subperiods. We discuss this issue in detail later.
- Use the time series of a factor as input into an Ornstein–Uhlenbeck (OU) process, for example to assess its speed of mean reversion.
- Use that time series as explaining variable in regressions, as in Figure 3.14, in order to check the exposure of a certain position. In particular, after a trade idea has been formulated, run those regressions in order to confirm that it really has the desired exposure (especially that an RV trade is really non-directional).
- Use that time series as dependent variable in regressions versus external (candidates for) explaining variables. This is the link of the statistical analysis to the fundamental and structural analysis, which allows complementing

backward-looking statistics with forward-looking expectations about macroeconomic events and potential risks. For the current example, we have provided that analysis in Table 3.1.

Step 6: Based on statistical, fundamental/structural and flow/other considerations, decide on which factor you want to take a view. In the example above and restricting ourselves to statistical reasons only, we might conclude that factor 1 has too little speed of mean reversion, thus take no view. Factor 2 is close to its mean, thus we take no view. Factor 3, however, seems to be significantly away from its mean and to have a high speed of mean reversion. Hence, we decide to investigate further for trade ideas based on factor 3 (i.e. butterflies hedged against factors 1 and 2).
Step 7: In order to select the best points on the yield curve to express that view, display the relevant residual chart, in our example the two-factor residual chart (Figure 3.15). Decide on the instruments.[19] In our example, we may want to choose sell 2Y, buy 5Y, sell 7Y (10Y has a little more negative residual than 7Y, but we might decide that the additional 0.5 bp residual is not worth the risk of going out three years further on the yield curve).
Step 8: Calculate and display the time series of that specific residual butterfly (Figure 3.18). This represents the actual performance of the trade in the past and the future performance will depend on that series. Now, run the regressions from Figure 3.14 in order to check that the individual selection of instruments offers the desired factor exposure.[20]

FIGURE 3.17 One-factor residuals.

Sources: data – Bloomberg; chart – Authors.

Data period: 4 Jan 2010 to 3 Oct 2011, weekly data; current residuals as of 3 Oct 2011.

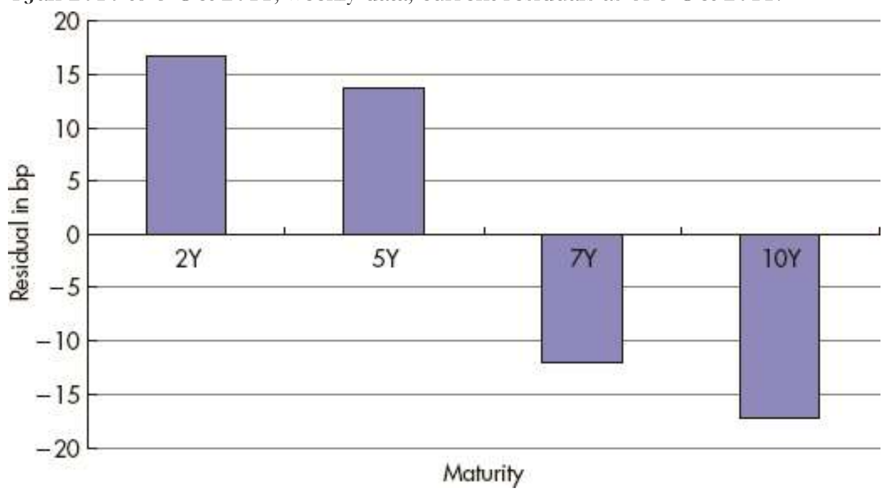
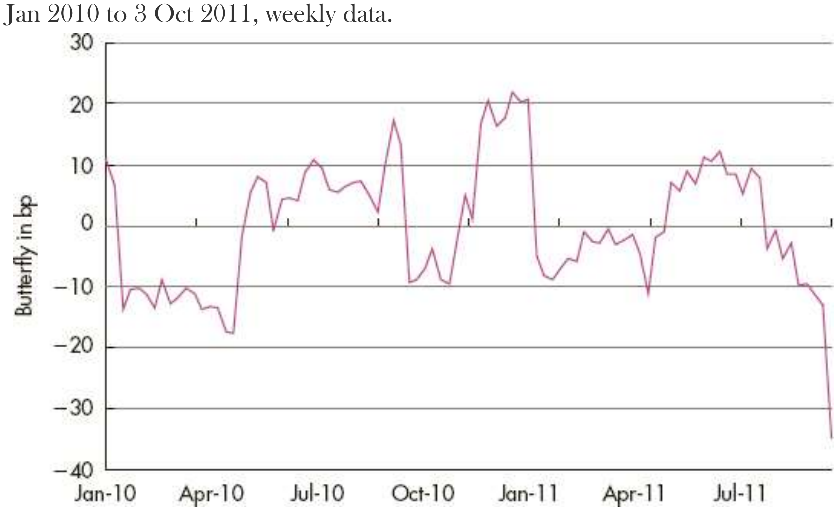


FIGURE 3.18 2Y-5Y-7Y butterfly of two-factor residuals.

Sources: data – Bloomberg; chart – Authors.

Data period: 4 Jan 2010 to 3 Oct 2011, weekly data.



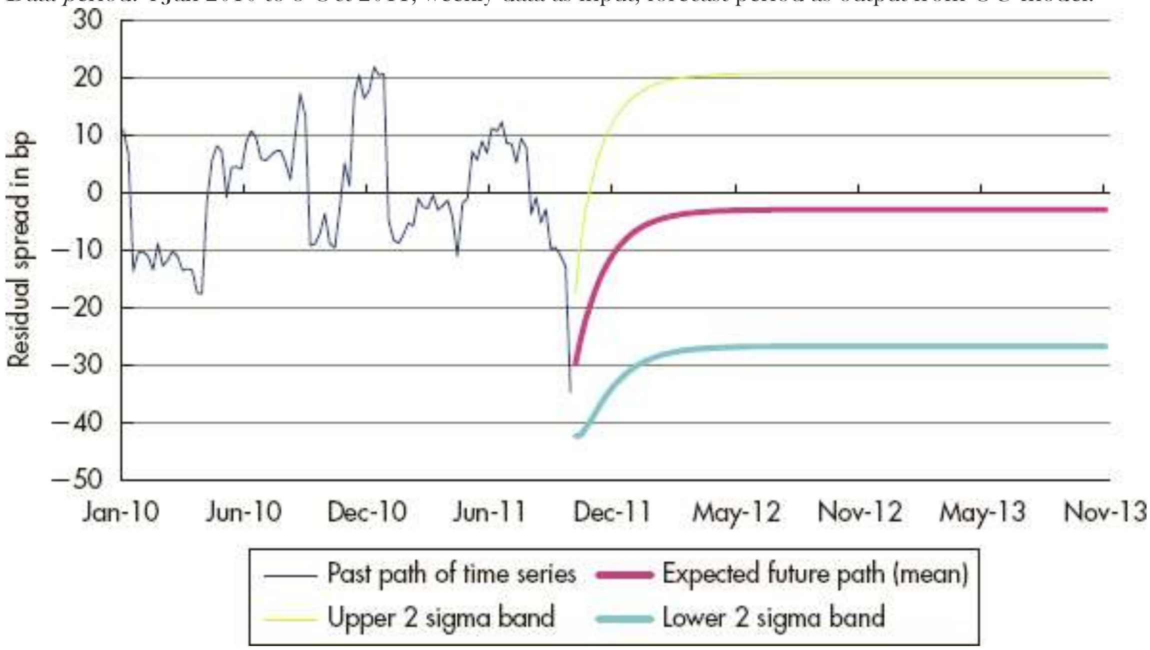
Step 9: Run the OU model for the time series of Figure 3.18 in order to assess the expected performance of the trade. As outlined in Chapter 2, calculate:

- the expected profit. In our example: 32 bp;[21]
- the expected downside risk. In the OU framework, the stop loss could be set at the two-sigma level (Figure 3.19), resulting in a stop loss level that moves over time. In our example this approach would result in a stop loss level at −42 bp (loss of 7 bp) after one week.

FIGURE 3.19 PCA-neutral 2Y-5Y-7Y Bund butterfly and its future path as modeled by an OU process.

Sources: data – Bloomberg; chart – Authors.

Data period: 4 Jan 2010 to 3 Oct 2011, weekly data as input; forecast period as output from OU model.

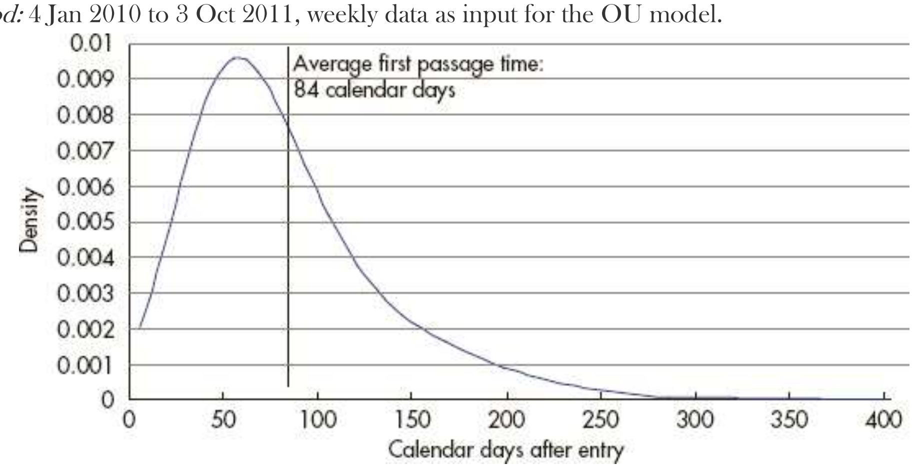


- The first passage time density and the expected time until mean reversion (Figure 3.20). In our example, the trade returns to its (estimated) mean (yields 32 bp profit) on average over 84 calendar days.

FIGURE 3.20 First passage time density for the PCA-neutral 2Y-5Y-7Y butterfly as modeled by an OU process.

Sources: data – Bloomberg; chart – Authors.

Data period: 4 Jan 2010 to 3 Oct 2011, weekly data as input for the OU model.



Step 10: Calculate the hedge ratios. In our example, we achieve factor 1 and 2 neutrality (i.e. the exposure to the time series of Figure 3.18) by selling 80m 2Y, buying 100m 5Y, and selling 58m 7Y Bunds. This has been calculated by a matrix inversion as discussed above. The numbers worked out as follows: with a BPV for 2Y of 1.98, for 5Y of 4.86, and for 7Y of 6.62 and sensitivities of 0.42, 0.55, and 0.53 to the first and of −0.84, −0.04, and 0.29 to the second factor of 2Y, 5Y, and 7Y (these numbers are the entries of the eigenvectors, which can be taken from the sheet on the website), the matrix $\begin{pmatrix} BPV_2 \cdot e_{12} & BPV_7 \cdot e_{17} \\ BPV_2 \cdot e_{22} & BPV_7 \cdot e_{27} \end{pmatrix}$ is $\begin{pmatrix} 0.84 & 3.51 \\ -1.66 & 1.91 \end{pmatrix}$ and its inverse thus $\begin{pmatrix} 0.26 & -0.47 \\ 0.22 & 0.11 \end{pmatrix}$. Multiplying to that matrix the vector $\begin{pmatrix} -n_5 \cdot BPV_5 \cdot e_{15} \\ -n_5 \cdot BPV_5 \cdot e_{25} \end{pmatrix}$ (with n_5 to be assumed to be one), that is $\begin{pmatrix} -2.69 \\ 0.22 \end{pmatrix}$, gives the vector of weights for 2Y and 7Y: $\begin{pmatrix} -0.80 \\ -0.58 \end{pmatrix}$. Then, the relative weights can be scaled to the desired trade size, in the example above by multiplying with 100m.

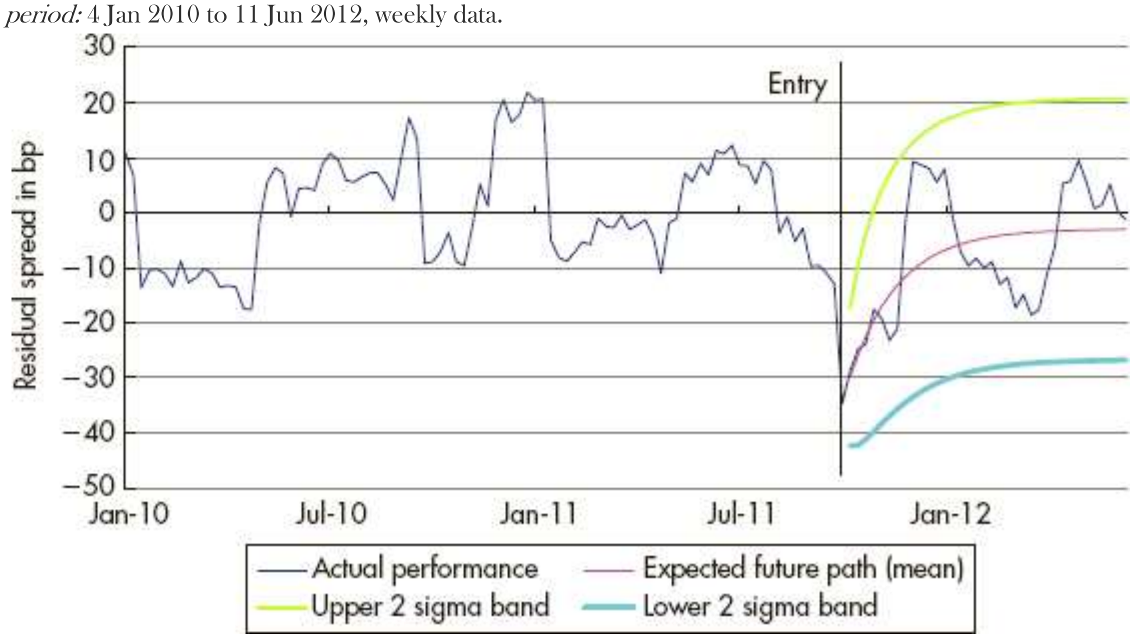
Step 11: Based on the expected holding horizon (84 calendar days) of the OU process and the hedge ratios, calculate the carry. Note that in the absence of BPV neutrality carry cannot naturally be expressed in bp terms anymore. Therefore, calculate the carry in money terms (i.e. euros and cents), based on the PCA-neutral hedge ratios. In a PCA framework, it is natural to express carry in money terms; should expression in bp terms be required, it must be stated, to *which* maturity the bps refer (as PCA does away with the false assumption of parallel curve shifts). In our example, the 84 calendar day carry is EUR 70,000 positive,[22] thus rather negligible compared to the profit potential from mean reversion. If the carry had been significantly negative, calculating additionally the carry until the 90% quartile of Figure 3.20 could have helped assess whether the trade was still attractive in case it took unusually long to revert back to its mean.[23]

And how did the trade in Box 3.1 perform in reality? Figure 3.21 compares the actual evolution with the forecast of the OU model. It may be no surprise to the reader that we have picked an example that worked well. In the following we will discuss instances where PCA might not work that smoothly.

FIGURE 3.21 Performance of Bund butterfly after entry compared to OU model forecast.

Sources: data – Bloomberg; chart – Authors.

Data period: 4 Jan 2010 to 11 Jun 2012, weekly data.



Problems and Pitfalls of PCA 1: Correlation between Factors during Subperiods

By construction, factors are uncorrelated over the whole sample period used as input data into a PCA, which allows one to construct a position hedged against certain factors. However, there could occur correlation between factors during subperiods. In such a case, the hedge could break down during that subperiod. For example, a trade on factor 2 hedged against factor 1 would be exposed to factor 1 during a time period, in which ephemeral correlation between factor 1 and 2 occurred. In this case, the performance of what was intended to be a non-directional steepening position would be driven (ephemerally) by direction.

To see how the theoretical problem of correlation between factors during subperiods can affect trades, let's consider the example of a 2Y-10Y PCA-neutral steepening trade on the Bund curve. In October 2010 and October 2011, the time series shown in Figure 3.22 may well have looked too good to resist from entering the trade.

FIGURE 3.22 PCA-neutral 2Y-10Y Bund steepening position.

Sources: data – Bloomberg; chart – Authors.

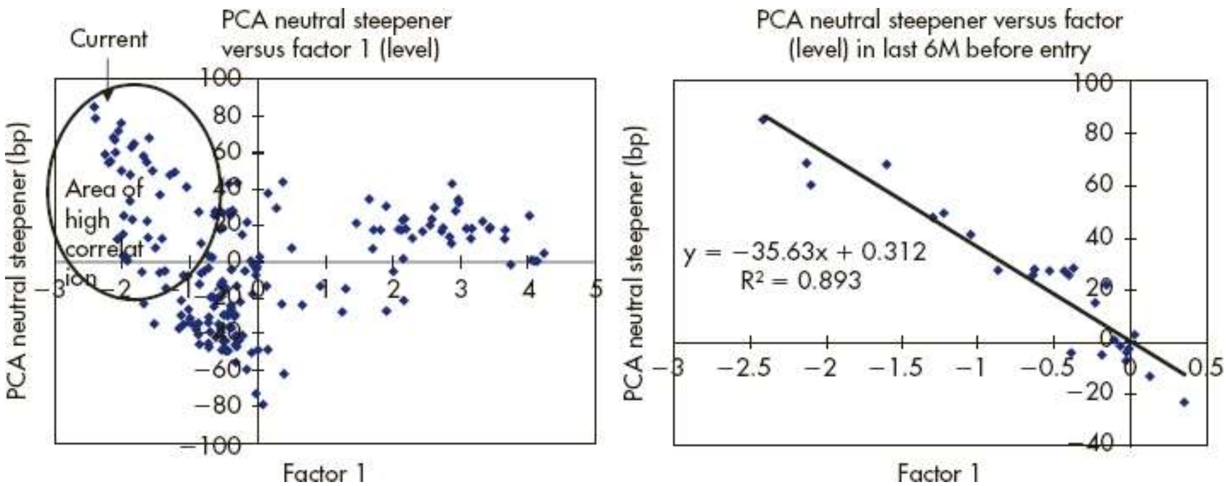
Data period: 7 Jan 2008 to 11 Jun 2012, weekly data.

However, before hitting the target (reverting back to mean), a steepening position entered in October 2011 broke through the OU stop loss and subsequently underperformed further. The reason for this misbehavior can be seen in the graph in Figure 3.23, which regresses the residual spread (PCA-neutral steepening trade) against the first factor, just like in Figure 3.14.

FIGURE 3.23 PCA-neutral steepener versus factor 1.

Sources: data – Bloomberg; chart – Authors.

Data period: 7 Jan 2008 to 22 Aug 2011, weekly data (left chart), 28 Feb 2011 to 22 Aug 2011, weekly data (right chart).



It turns out that, while overall the correlation is zero, there has been a high correlation to the first factor on the left-hand side of the chart in Figure 3.23. Further investigation reveals that in the six months before entry the trade (a factor 2 position) followed factor 1 very closely. Thus, the reason for the underperformance of the PCA-neutral Bund steepening trade was that it was in fact a directional trade, and, since the direction happened to work against it, its less-strong factor 2 exposure could not save it from losing.[24] Compare this situation with the regressions depicted in Figure 3.14, where we have argued against BPV-neutral weighting due to its factor 1 exposure: in times of ephemeral correlation between the factors, a PCA-neutral trade shares the fate and criticism of BPV-neutral trades.

The good news is that in many cases correlation between factors during subperiods can be spotted before entering into a trade. To do so, it is crucial to run the regressions (step 5 of the flow above). Often, a potential problem with factor correlation announces itself, as in the current example. Thus, if there has been a high correlation between the factors in the time period before (as in Figure 3.23), caution is advisable. By contrast, the 2Y-5Y-7Y butterfly entered on 3 Oct 2011 did not exhibit a significant correlation of factors before entry: the entry level is the isolated point in the lower left corner of Figure 3.14. Unlike the cloud of points along a regression line in Figure 3.23, in the case of the butterfly from Figure 3.14, there is no correlation problem visible – and this was one reason why one trade worked (Figure 3.21) and the other did not (Figure 3.22).

In recent times, the reason for correlation of factors between subperiods is often the shift in the credit assessment of sovereign issuers. Imagine that a PCA is calculated on data before credit risk became an issue in Western government bond markets. From the point in time onwards when it did become an issue, it will impact factor 1 (the yield level tends to increase when credit risk increases) *and* factor 2 (since the longer the maturity, the more it will be affected by increasing credit risk,[25] thus an increase in credit risk tends to lead to a higher non-directional steepness of the yield curve). Consequently, the emergence of credit concerns has often resulted in significant correlation between factors 1 and 2. We shall address this issue in detail in Chapter 15 and show how running a PCA on CDS-adjusted yield curves (bond yield minus CDS) can solve the problem.

In our experience, ephemeral correlation between factors is the main pitfall of PCA. Conversely, consistently checking for factor correlation, for example through the method described above, reduces the risk of PCA leading to unsatisfactory trades significantly. When we started using PCA systematically as a tool to find and construct trades in global bond markets about 15 years ago, we did not pay enough attention to the correlation problem and produced a ratio of profitable to overall PCA-based trade ideas of 82%.[26] After figuring out that many of the 18% of losing trades were due to the correlation issue, we avoided this trap better and could increase the success ratio to just over 90%, with the number of trades obviously going down. In general, experience is required to strike the right balance between having too many trades and being overly cautious. In the current case, striking this balance requires experience to judge the level of correlation between factors in the subperiod that is acceptable before entering a trade. As a rough guide, Figure 3.14 (isolated current point out of cloud of points with no correlation) and Figure 3.23 (correlated cloud of points leading to current point) provide an illustration of unsuspicious and suspicious situations.

Problems and Pitfalls of PCA 2: Instability of Eigenvectors over Time

If eigenvectors change after entering into a trade with PCA hedge ratios, the trade will become exposed to unintended factor risk. Imagine again that we have entered into a trade on factor 2 and hedged against factor 1. If the first eigenvector for the time period during which we hold the trade turns out to be different from the first eigenvector calculated on the sample period before the trade (and thus used for determining the hedge ratios for factor 1 neutrality), then we will be hedged against the "wrong" first eigenvector and exposed to the first factor. Thus, a change of the first eigenvector results in the hedge breaking down and in directional exposure. While the cause can be different, the problem is the same as in the case of correlation between factors occurring after entering into a PCA-neutral position: it loses its neutrality and becomes exposed to factors it was not intended to.

First of all, it is important to distinguish changes in factors from changes in factor loadings: changes in factors occur all the time and are no problem for the PCA model and hedges based on it. For example, yields could fluctuate between 1% and 10% (i.e. α_1 exhibits a high volatility), while the eigenvectors remain stable. If the central bank is driving yields both up and down, then the first eigenvector should maintain its downward sloping shape (depicted in Figure 3.6). However, if the central bank were to cut rates to zero and then announce that it would maintain a zero policy rate for a number of years, the sensitivity of short rates to directional moves would be expected to decrease versus the sensitivity of long rates. This corresponds to a change in the first factor loading.

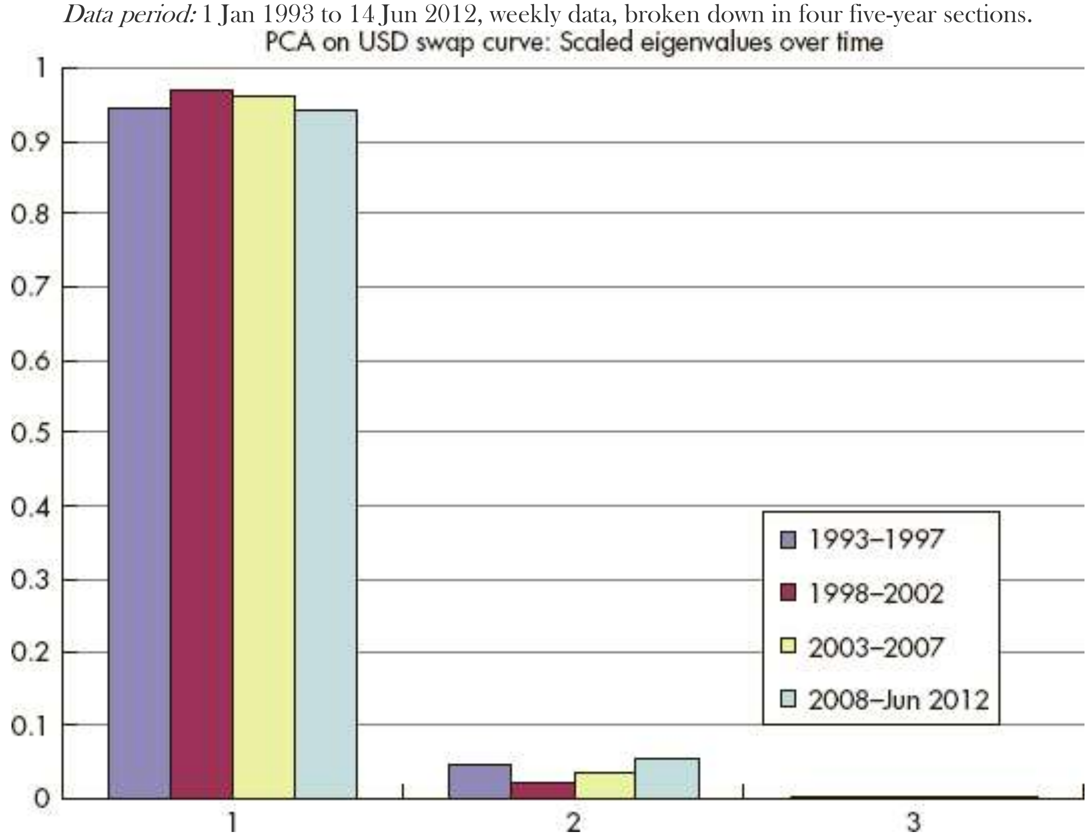
Moreover, sometimes a perceived instability of eigenvectors is the result of a too-short sample period for the PCA calculation. This can be easily avoided by choosing a longer time period (step 1) (in case of yield curve analysis usually at least one year).

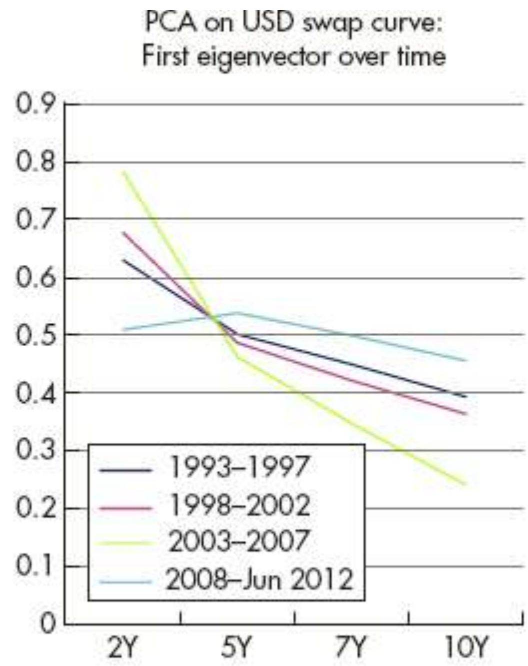
In order to get a sense of the empirical stability of eigenvectors, we have run a PCA on yield data for consecutive five-year periods and calculated the eigenvalues and eigenvectors. The charts in Figure 3.24 depict the evolution of eigenvalues and eigenvectors over time for the Bund, USD swap, and JPY swap markets.

FIGURE 3.24 Evolution of eigenvalues and eigenvectors from 1993 to 2012 in several rate markets.

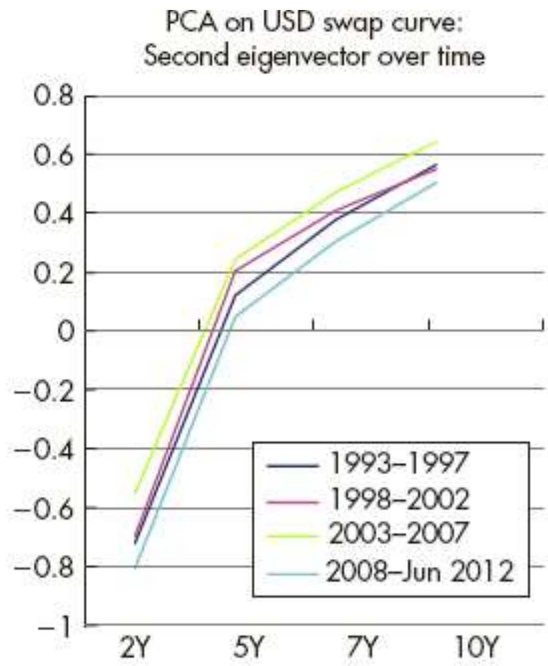
Sources: data – Bloomberg; chart – Authors.

Data period: 1 Jan 1993 to 14 Jun 2012, weekly data, broken down in four five-year sections.

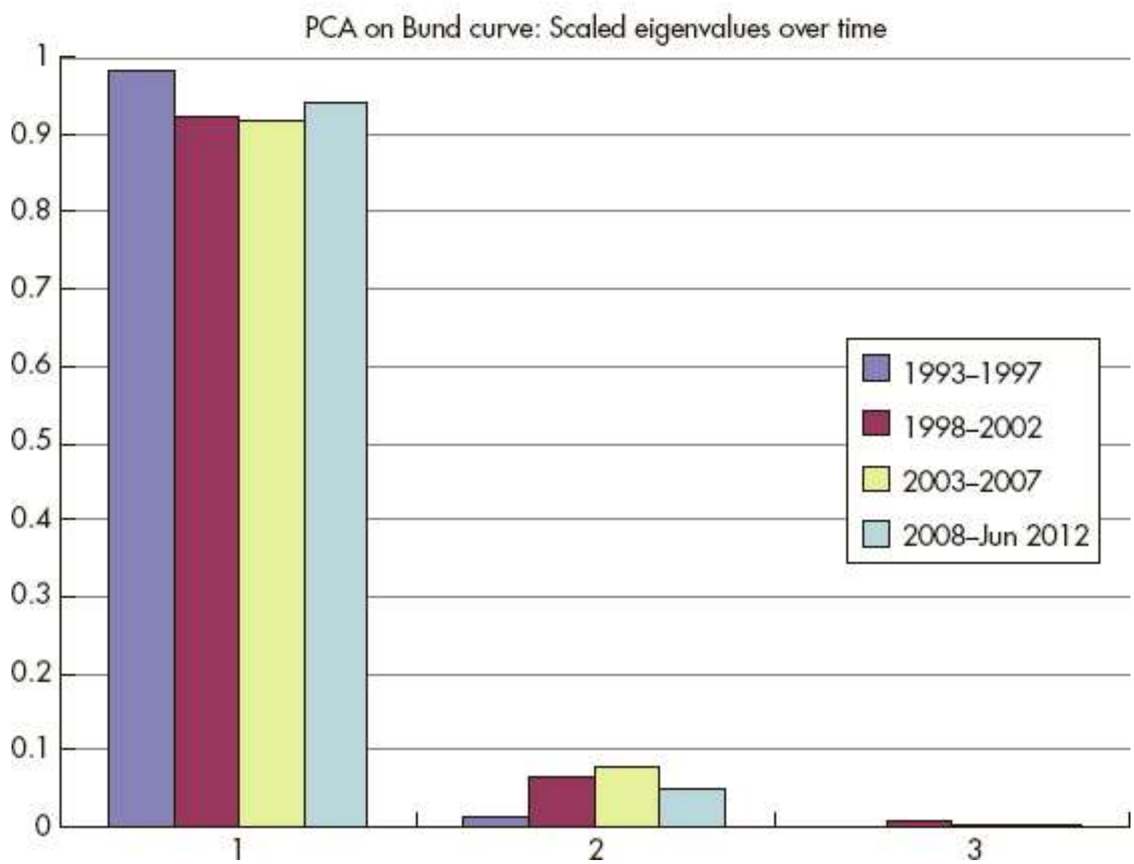




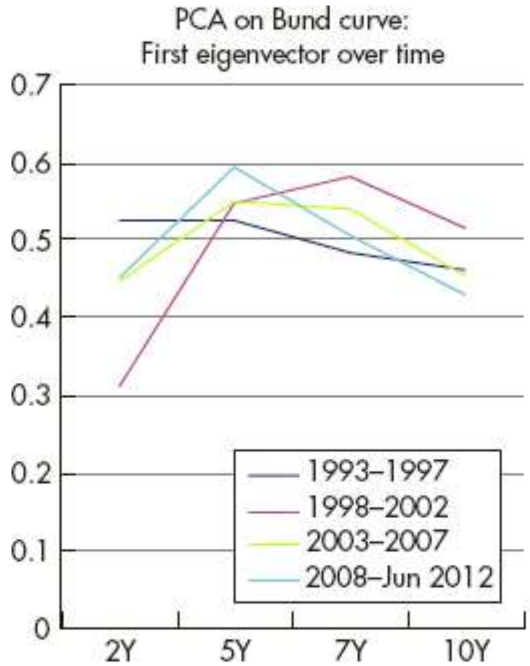
PCA on USD swap curve:
First eigenvector over time
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
2Y
5Y
7Y
10Y
1993–1997
1998–2002
2003–2007
2008–Jun 2012

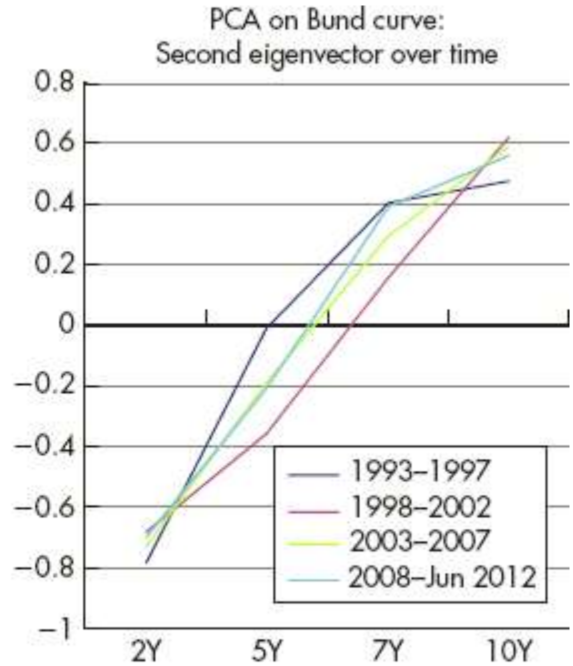


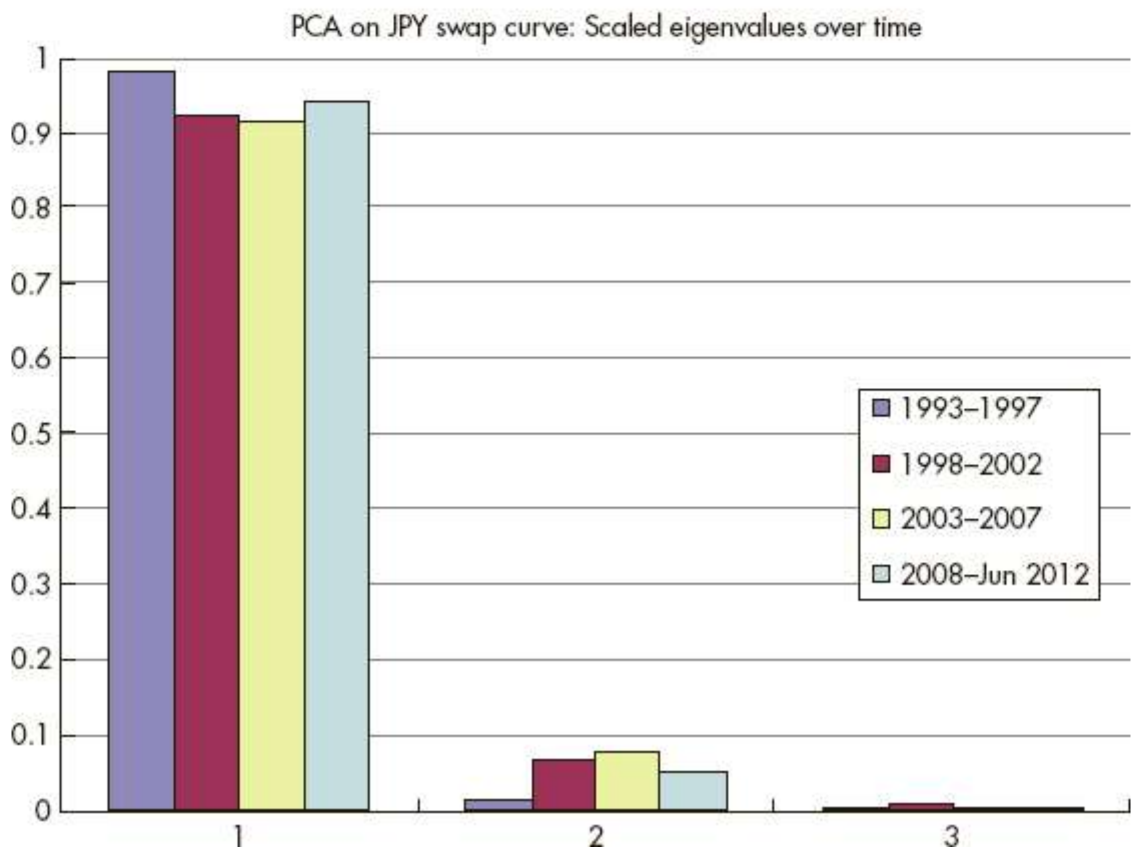
PCA on USD swap curve:
Second eigenvector over time
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−1
2Y
5Y
7Y
10Y
1993–1997
1998–2002
2003–2007
2008–Jun 2012

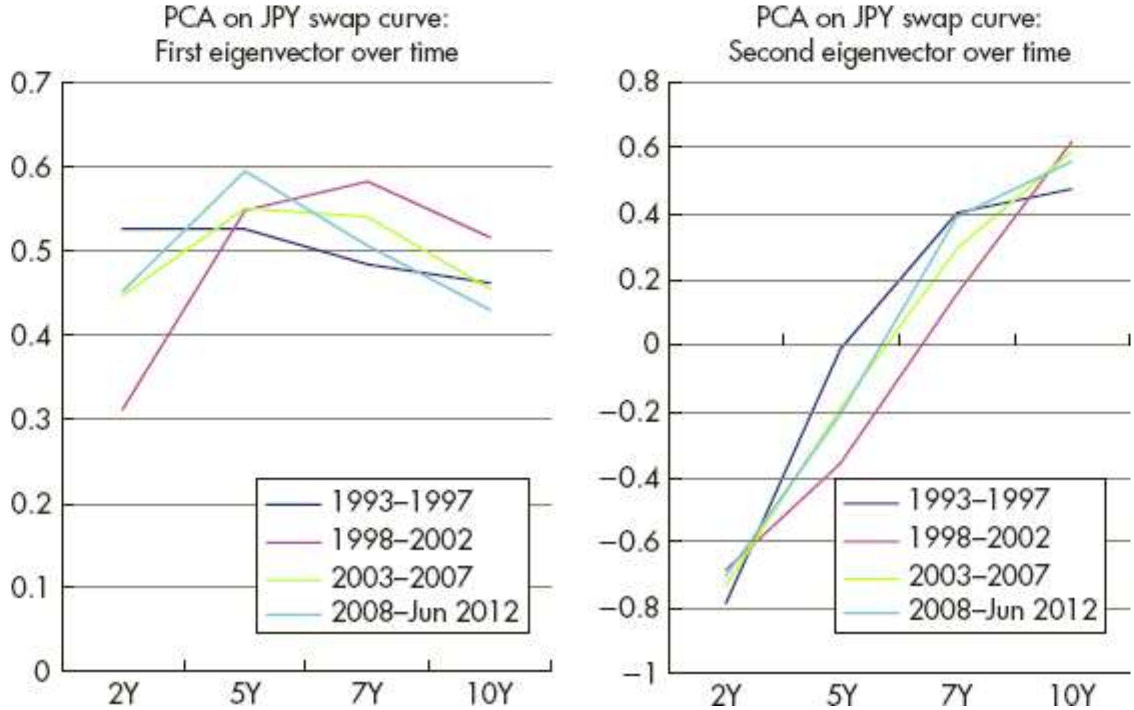


PCA on Bund curve: Scaled eigenvalues over time
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1
2
3
1993–1997
1998–2002
2003–2007
2008–Jun 2012


PCA on Bund curve:
First eigenvector over time
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
2Y
5Y
7Y
10Y
1993–1997
1998–2002
2003–2007
2008–Jun 2012


PCA on Bund curve:
Second eigenvector over time
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−1
2Y
5Y
7Y
10Y
1993–1997
1998–2002
2003–2007
2008–Jun 2012


PCA on JPY swap curve: Scaled eigenvalues over time
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1
2
3
1993–1997
1998–2002
2003–2007
2008–Jun 2012



In general, the result of this empirical study comes as a relief: despite the sea-changing events that occurred during the last 20 years, eigenvalues and eigenvectors have been remarkably stable (i.e. the factors changed and not the factor loadings), just as a factor model like PCA assumes and requires. With the exception of the short end, this is even the case for Japan, which saw the transition from a very active central bank to a BoJ which has removed virtually all volatility from its policy rates.

In particular, the recent debt crisis did not leave significant traces in the eigenvectors – while it is an important cause for the problem of correlation between factors during subperiods (see above). Overall, correlation between factors, not eigenvector instability, is the main pitfall of PCA. Accordingly, the crisis starting in 2008 produced a break in the market that often manifested itself in a changing correlation pattern between factors and not so much in an instability of factor loadings. PCA-based trades which used a sample period of pre-crisis data to calculate hedge ratios sometimes faced trouble when the crisis erupted, but almost always due to the factors becoming correlated, seldom due to eigenvectors changing shape. As the problem of factor correlation has receded afterward, we recommend avoiding basing a PCA on sample data which include both pre- and post-crisis data (or using the CDS adjustment outlined in Chapter 15). For analyzing the current market situation, a starting point like January 2010 (as in the butterfly trade example) could work well.

The one exception from stability of eigenvectors is the short end of the curve. This is probably due to the high impact of an external driving force (i.e. central bank action) on that part of the yield curve. Correspondingly, with central banks keeping to a low rate policy and thereby intentionally removing uncertainty, equal volatility, and equal sensitivity from the short end of the yield curve, we currently observe in all major rates markets a relatively low sensitivity of 2Y rates to the first factor. One can react in a couple of ways to this finding:

- Refrain from PCA-based trades on the short end. For example, restrict the input variables to a PCA (step 1) and thus the universe of trades analyzed to maturities above 3Y or 5Y.
- If, nevertheless, a PCA-based trade involves the short end, ensure that its speed of mean reversion is high enough for it to be likely to perform before a (rather long-term) change in central bank

policy could affect the eigenvector. Our butterfly trade example from above could fall into that category.

- Intentionally position for the short end of the eigenvector to change. We know that a change in the eigenvector will result in the PCA hedge breaking down. But if we can forecast how the eigenvector is going to change, we can position ourselves in such a way that we will end up being overhedged in a falling and underhedged in a rising market (i.e. turn the problem of changing eigenvectors into a profitable trading strategy that has similarities with delta hedging a long option position). For example, if BoJ ends the zero interest rate policy, an increase in rates (factor 1) should be linked to the short end becoming more volatile (sensitivities in first eigenvector increasing at the short relative to the long end). A curve-flattening position dynamically hedged against (the changing) factor 1 should therefore in fact be overhedged in declining and underhedged in increasing markets.[27]

While these results can serve to increase the confidence of analysts in the stability of eigenvectors of yield curves, it is important to note that in other markets the stability could be less pronounced. In this case, one could compare the stability of eigenvectors with the expected holding horizon of a trade and focus on those positions, which are expected to perform before eigenvectors become unstable. For example, if a PCA-based trade shows an expected holding horizon of a few weeks (step 9 above), it requires the eigenvectors to be stable over at least that period of time (and ideally longer). In markets where stability of eigenvectors could be an issue, an investor may want to refrain from that trade if he sees a risk of eigenvectors changing during the next few weeks, but may still feel comfortable enough, if he perceives potential instability to be a longer-term issue only.

PCA as a Tool to Construct New Types of Trades

Finally, we provide an example of how PCA could be used in the currency market, thereby underlining its universal applicability. At the same time, we shall illustrate how creatively a PCA can be applied, using the step-by-step guide from above as stepping stones to a trade idea, which only PCA enables but in which PCA is just one of several parts.

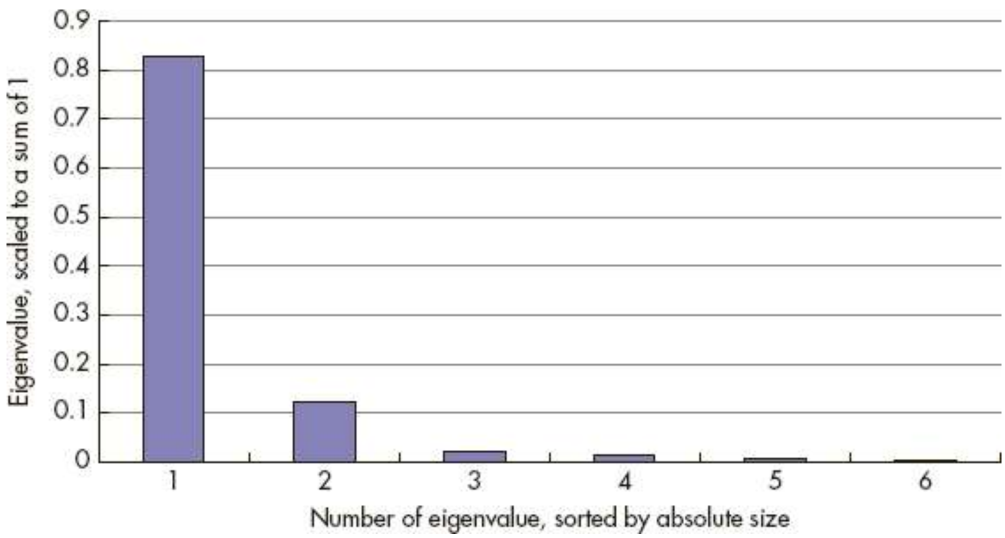
When running a PCA on JPY, GBP, SEK, CHF, AUD, and SGD (versus USD), it is useful to adjust the series for the difference in absolute values in order for the charts to be legible (otherwise, 100 JPY per USD versus 0.6 GBP per USD would show up as huge difference in sensitivities to the eigenvectors, for example). Thus, we have run the PCA on synthetic currencies, starting all at a value of 1 on 4 Jan 1999. The PCA then uses weekly data of these synthetic series from 4 Jan 1999 until 25 July 2011 and can be seen on the Excel sheet on the website accompanying this book.

Figure 3.25 depicts the scaled eigenvalues. It turns out that the relative explanatory strength of the first and second factor in the FX market is similar to the bond market, while the eigenvalues above 3 decrease slower than in the case of bonds.

FIGURE 3.25 Scaled eigenvalues of a PCA on currencies.

Sources: data – ECB; chart – Authors.

Data period: 4 Jan 1999 to 25 Jul 2011, weekly data.

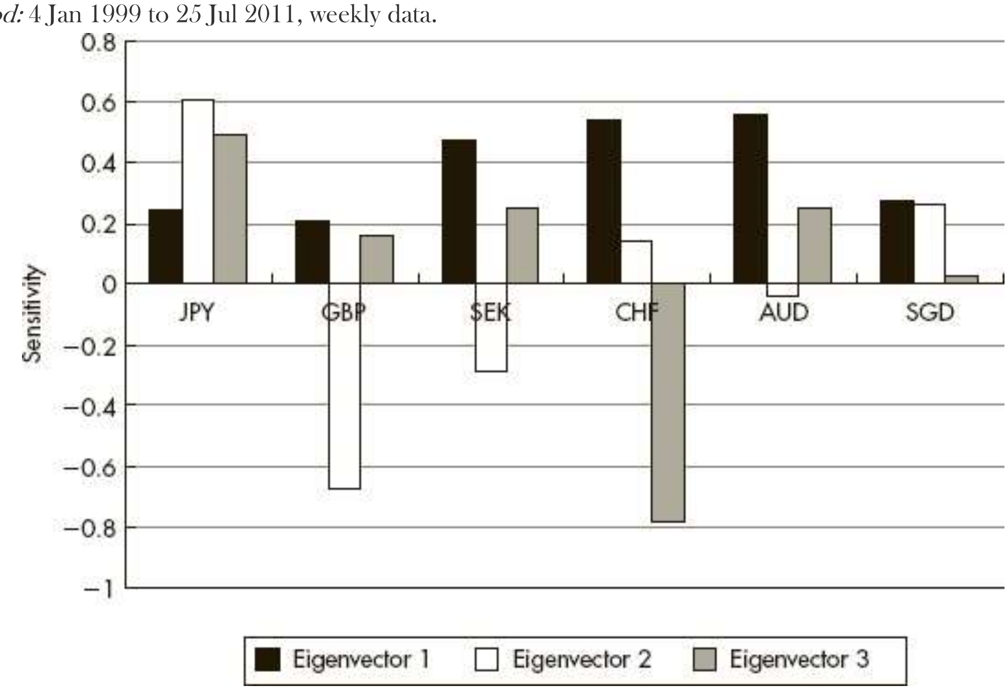


In Figure 3.26, the eigenvectors are displayed. As in the case of a PCA on bonds, the first eigenvector has only positive entries (i.e. a change in factor 1 affects all currencies similarly). However, the sensitivities may seem puzzling: if factor 1 increases, SEK, CHF, and AUD increase a lot, while JPY, GBP, and SGD less so. Since few FX traders think of grouping currencies in such a manner, this puzzle may represent an interesting, PCA-induced insight, but requires further examination, which will be provided below by looking simultaneously at the time series of factor 1.

FIGURE 3.26 Eigenvectors of a PCA on currencies.

Sources: data – ECB; chart – Authors.

Data period: 4 Jan 1999 to 25 Jul 2011, weekly data.



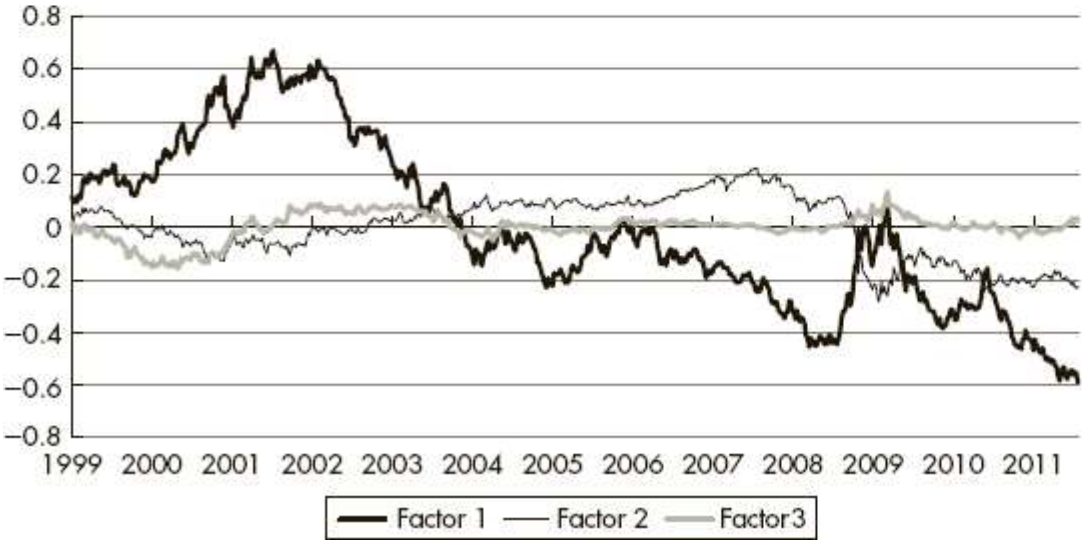
Likewise, factors 2 and 3 cause differentiation among currencies (e.g. if factor 2 rises, JPY, CHF, and SGD increase, while GBP, SEK and AUD decrease). This grouping may sound more familiar, as it puts "low risk" and "high risk" currencies together.

Since the interpretation of eigenvectors (i.e. market mechanisms) is less straightforward than in the case of Bunds, we need to consider additionally the time series of the factors and their link to external variables in order to be able to complete the interpretation of the PCA results. The evolution of the factors is shown in Figure 3.27.

FIGURE 3.27 Factors of a PCA on currencies.

Sources: data – ECB; chart – Authors.

Data period: 4 Jan 1999 to 25 Jul 2011, weekly data.



This graphical representation may well evoke the visual memory of the analyst. For example, he may find that the evolution of factor 1 is almost a mirror image of the USD–EUR FX rate. This visual discovery can be confirmed (e.g. via a regression, which returns a correlation of −0.94 between factor 1 and the EUR exchange rate). Thus, an increase of factor 1 is strongly linked to the EUR weakening versus the USD. An interesting consequence of that result is that the exchange rate of the USD versus the EUR (which we excluded from the PCA input series) explains 83% (scaled first eigenvalue) of the moves of the USD versus the other currencies. Hence, one may conclude that the USD–EUR rate is by far the most important driving force of all USD FX rates. Moreover, the sensitivities versus the first factor (first eigenvector) group the currencies with a strong (SEK, CHF, AUD) and not so strong (JPY, GBP, SGD) tendency to weaken versus the USD in case of factor 1 going up (i.e. the EUR weakening against the USD).

The time series of factor 2 corresponds well to a more psychological variable like "risk on/risk off". And as expected, factor 2 falling significantly in 2009 (and staying low) has resulted in those currencies with a positive sensitivity to factor 2 ("safe havens") (JPY, CHF, SGD) to outperform those with a negative sensitivity to factor 2 ("risky") (GBP, SEK, AUD). This is probably the picture most traders have in mind when they think about the FX market. Note, however, that a PCA relegates this market mechanism to number two, revealing that it only explains 12% of the FX market action – and reveals the more important structure given by eigenvector 1.

In cases of puzzling eigenvectors, it is often useful to calculate a table of heuristic regressions, with the factors as dependent variables and with various candidates serving as independent or explanatory variables. We have tried a couple of those regressions and present the results in Table 3.2.[28]

TABLE 3.2 Correlations of the First Three Factors of a PCA on Currencies versus Candidates for External Explaining Variables

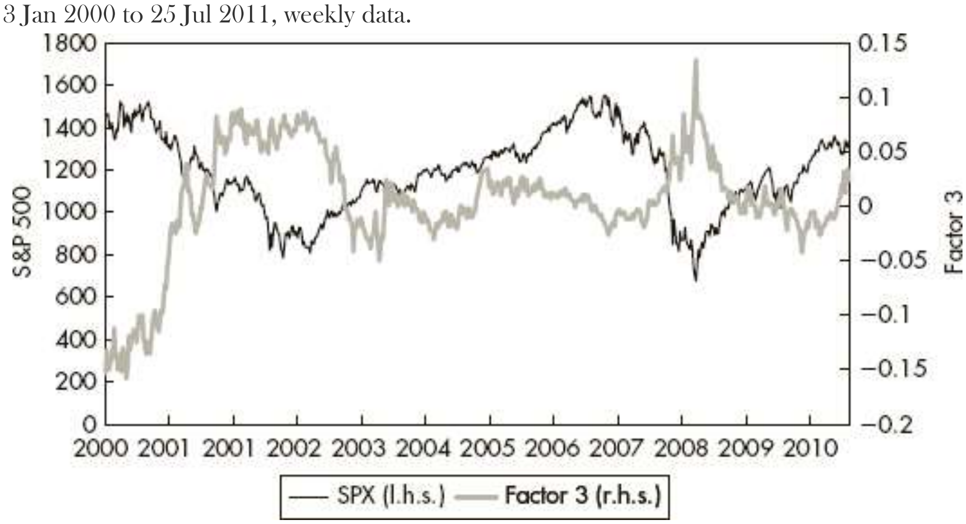
	Factor 1	Factor 2	Factor 3
EUR	−0.94	−0.01	0.13
S&P500	−0.20	0.27	−0.66
VIX	0.15	−0.46	0.28
US swap PCA factor 1	0.64	0.55	−0.49
US swap PCA factor 2	0.42	−0.10	0.16
US swap PCA factor 3	−0.13	−0.04	−0.41
Oil	−0.86	−0.12	−0.07

This table confirms our optical interpretation of factor 1 being closely linked to the EUR and factor 2 being a “risk on/risk off” factor, which has therefore some correlation to the direction of USD interest rates (as represented in the first factor of a PCA on USD swaps) and the VIX. Furthermore, this table reveals that the third factor is linked to the S&P500 index – and, indeed, taking another look at the time series of factor 3 in the chart above confirms that it is a close mirror image to the stock market. This relationship is depicted in Figure 3.28 in more detail.

FIGURE 3.28 Factor 3 of a PCA on currencies versus the S&P500 index.

Sources: data – ECB, Bloomberg; chart – Authors.

Data period: 3 Jan 2000 to 25 Jul 2011, weekly data.



These new insights, which a PCA provides into the FX market, could be used to model currencies by a three-factor model, with the factors linked to external variables like EUR, risk adversity (e.g. VIX), and stock prices.

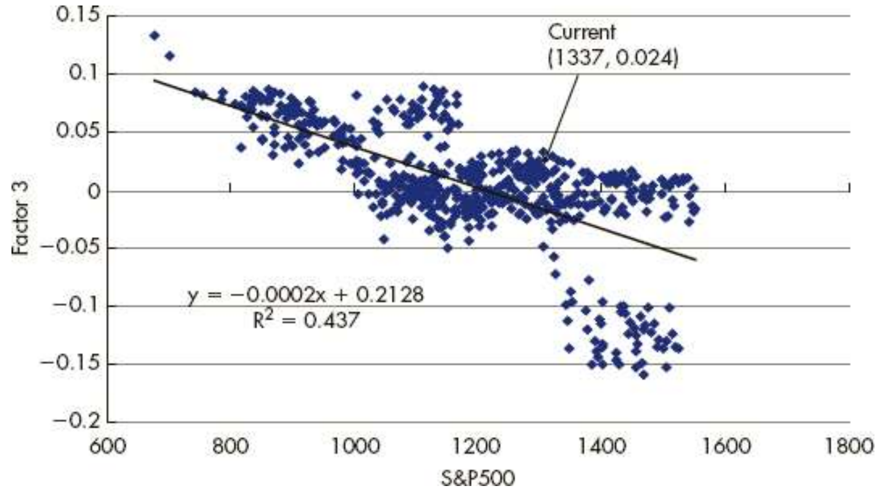
In the following, however, we would like to show how the insights of a PCA into market mechanisms can be used to construct new trading positions, which would not be possible or understandable outside of the PCA framework.

While there is a reasonable correlation between factor 3 and the S&P500, currently (25 July 2011), the residual of a regression is rather high. Figure 3.29 illustrates that the current point is quite far away from the regression line. Note that in case we use shorter time horizons for the regression we still witness a significant residual. This means that the relationship between factor 3 and stock prices, which has been relatively stable over the past 12 years, is currently disturbed. Thus, in case we believe that the long-term relationship will hold in future (and we see no reason not to do so), we may want to bet on the residual disappearing.[29] Hence, we investigate for a trade of factor 3 of a PCA on currencies versus the S&P500 index.

FIGURE 3.29 Regression of factor 3 of the PCA on currencies versus the S&P500 index.

Sources: data – ECB, Bloomberg; chart – Authors.

Data period: 3 Jan 2000 to 25 Jul 2011, weekly data.

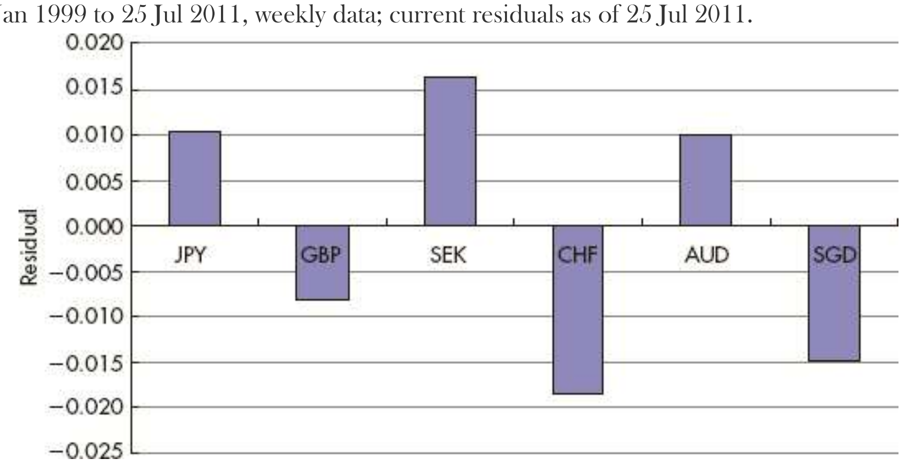


Since we cannot trade factor 3 directly, we need to find a portfolio of three currencies, hedged against factors 1 and 2 and highly correlated to factor 3. Given the relatively high strength of factor 4 in particular, not every two-factor neutral combination of currencies can be expected to work, as it could be mainly a function of factor 4 rather than factor 3. In addition, we would like to improve our return by choosing a combination of currencies with a high two-factor residual. These residuals are shown in Figure 3.30, and thus both a JPY–CHF–SEK and a JPY–SGD–SEK PCA-neutral combination of currencies (similar to a butterfly on the yield curve) seem attractive. However, only the first one has a strong correlation to the third factor, so we choose this one.

FIGURE 3.30 Two-factor residuals of a PCA on currencies.

Sources: data – ECB; chart – Authors.

Data period: 4 Jan 1999 to 25 Jul 2011, weekly data; current residuals as of 25 Jul 2011.



Now, we can formulate our trading strategy, which is to trade a two-factor PCA-neutral portfolio of JPY, CHF, and SEK versus the S&P500 index. The hedge ratio between the portfolio of currencies and the S&P500 index is given by the slope of the regression line, while the weightings of the currencies in the portfolio are determined by the conditions of neutrality versus factors 1 and 2 of the PCA on currencies. The result is the exposure to the residual of a regression between the portfolio of currencies and the S&P500 index, whose time series is shown in Figure 3.31.

FIGURE 3.31 Residual of a regression between a PCA-neutral portfolio of currencies and the S&P500 index.

Sources: data – ECB, Bloomberg; chart – Authors.

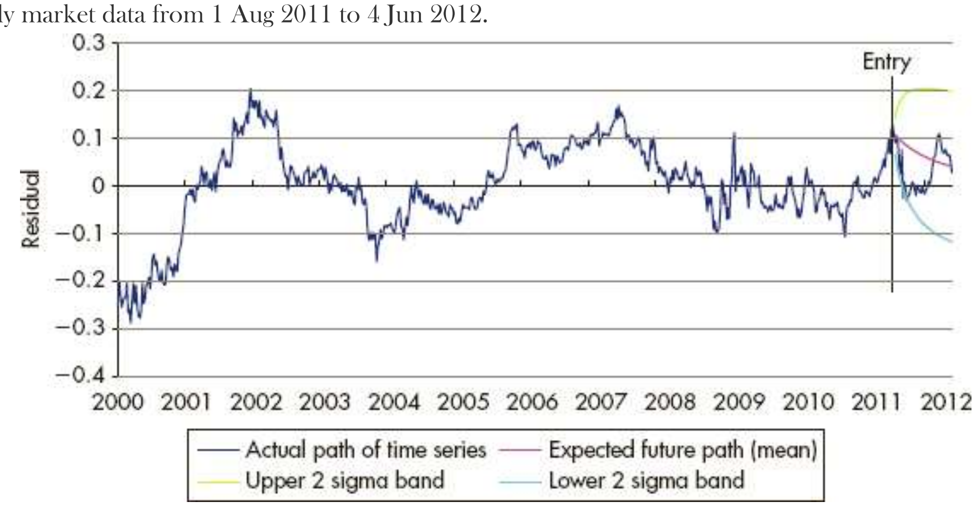
Data period: 3 Jan 2000 to 25 Jul 2011, weekly data.

Of course, we can now run a mean reversion model on that series and judge whether the statistical properties, like speed of mean reversion, seem sufficiently good to enter the trade. Figure 3.32 depicts the actual performance of the trade versus the forecast of the OU model. Since the actual performance followed the lower 2-sigma band of the OU simulation closely, one can say that an investor entering the trade was quite lucky to be able to realize the profit potential from the time series returning to its mean (i.e. the residual of the regression disappearing) soon.

FIGURE 3.32 Actual performance of the trade versus the OU model forecast.

Sources: data – ECB, Bloomberg; chart – Authors.

Data period: 3 Jan 2000 to 25 Jul 2011, weekly data as input in OU model; forecast period as output of OU model versus weekly market data from 1 Aug 2011 to 4 Jun 2012.



This example demonstrates how PCA both yields new insights into a variety of markets and is the basis for new trading opportunities. By producing time series (like factor 3) that are not directly observable in the market but can be traded nevertheless (in this example through a two-factor neutral portfolio), PCA opens the door to a multitude of new and creative analyses and trading opportunities. The advantages of these new positions, besides their analytical challenges and illuminating results, are that they offer:

- a good chance to find a source of profit which is uncorrelated to more common positions;
- trading ideas that are unlikely to be analyzed by many people and therefore likely to be profitable. While lots of analytical effort is invested in 2Y-5Y-10Y butterflies on the yield curve, probably no one ever thought of the trade developed above. Hence, PCA is also a tool to find value off the beaten track, as it allows constructing new trade ideas individually.

Our goal is not to promote trading stock indices versus portfolios of currencies but to encourage using PCA as a compass to venture into unknown territory, where golden fruits may be growing on the trees.

[1]Other purposes may require other tools.

[2]One could also remark a systemic divergence from the line for very low yield levels. The discussion following Figure 3.24 will reveal the background of that deviation.

[3]The absence of correlation is a result of the linear algebra behind a PCA. It is also a very convenient feature for the usual purposes of relative value analysis, as it allows decomposing a market into uncorrelated factors, which is the basis for analyzing them individually. However, it is also conceivable to construct a factor model with correlated factors. For example, in order to obtain a particular analytical goal it might be useful to model a market as a function of two macroeconomic variables, even if they are correlated.

[4]Obviously, we lose some statistical information in this step, in particular above the second order. However, in practice this is seldom of relevance for the goals of relative value analysis.

[5]In fact, this theorem is true for every matrix which is symmetric and positive semi-definite.

[6]One can back up this statement theoretically by linking PCA to a Fourier analysis.

[7]Note that the input data for Figure 3.8 consist of the whole volatility surface, while for Figure 3.7 only options with an expiry of at least 2Y are used. The reason for this is explained in Chapter 17. As there are two different PCAs behind the two charts, the sensitivities to factor 2 displayed in Figure 3.7 and Figure 3.8 are also different.

[8]This requires the assumption that the future curve dynamic will be the same as the current curve dynamic expressed in PCA eigenvectors. Toward the end of this chapter we shall investigate the validity of this assumption.

[9]Likewise, one could use the same techniques to analyze and execute trading positions, which are uncorrelated to, for example, non-directional steepness (represented here by the second factor). However, in practice most of the time the main concern will be to prevent the all-pervasive directional effects from influencing relative value positions, in other words to create alpha.

[10]While the shape of the eigenvector always allows an economic interpretation of the factors such as "yield curve steepness", it may not always be possible to link each of the factors with an obvious and specific macroeconomic variable like "inflation". In fact, the latter relationship will be the subject of a heuristic regression below.

[11]Over the time period used for Table 3.1, the independent variables – including the S&P500, EUR, and Oil – did not exhibit a significant trend and are therefore included. Of course, this could be different over other time periods.

[12]Additionally, economic variables like inflation or GDP (growth) could be included. However, given the low frequency of these data (e.g. quarterly), it is statistically meaningful only for time series spanning several years, not in the current example, which has less than two years of data.

[13]These verbal monstrosities are the translation of mathematical factor decomposition into everyday terms.

[14]In fact, it is also exposed to factors 4 and higher, but they have much less impact on the overall yield curve variation (and butterfly trade performance) than factor 3 does. The crucial point is to hedge against all factors of lower order than the one to which exposure is wanted.

[15]This discussion is obviously restricted to *statistical* reasons for a trade. Of course, there can be other reasons, like fundamental views or anticipated flows, for or against a trade.

[16]Actually, PCA was first applied in engineering with the aim to find the common behavior of a production series and to detect outliers, quite similar to our application in finance, where common behavior = market mechanisms and outliers = candidates for trading opportunities.

[17]Of course, he could also receive swaps instead, or buy futures, or calls on futures etc. We shall discuss asset selection among different classes of instruments in the following chapters.

[18]Any correlation a high factor (e.g. factor 7) might have to an external economic variable is likely to disappear in the combination of the high factors (e.g. factors 4–10) to a residual.

[19]Again, other considerations than statistics play a major role, too, in particular flow information and carry/roll-down concerns.

[20]It could happen, for example, that the sensitivities to factor 3 of all the specific instruments selected are zero and that therefore the performance of the specific butterfly is uncorrelated to factor 3 and rather a function of factor 4.

[21]This chart has been generated by applying the general tool on the website, which *estimates* the mean. In the current case, the estimated mean of −3 bp is slightly different from the actual mean, which is 0 for any factor or residual time series from a PCA. If the actual mean is known, as in the current case, the estimation of the mean from the general tool on the website could be overwritten. Then, the expected profit would be 3 bp higher.

[22]As this is for illustration purposes only, we have used general collateral (GC) rates for the carry calculation.

[23]Imagine a trade with a 10 bp profit potential and a three-month negative carry also of 10 bp. If the OU model shows an expected holding horizon of just two weeks (over which the negative carry is, say, 2 bp), the negative carry might be acceptable. The confidence in this position could increase further if the OU process suggests a 90% chance of the 10 bp profit to materialize over one month. Alternatively, one could calculate the probability of the mean reversion taking longer than three months (i.e. the likelihood that the negative carry would exceed the profit from mean reversion).

[24]Also in October 2010, there was a high correlation to factor 1, but this time the direction worked in favor of the steepening trade. Thus, in 2010, the trade made money, but for the wrong reason: not because it was a factor 2 position hedged against factor 1 as intended but because we were lucky that the direction worked in our favor.

[25]As can be seen from Moody's transition matrix, for example, this statement is true for good credits, such as Western government bonds at the start of the recent crisis. When the credit becomes bad, however, the dynamics can change and increasing credit risk can then affect shorter maturities more. See Chapter 15 for more details.

[26]We generated these trade ideas by screening mainly Western government bond markets through an approach similar to the step-by-step guide from above, including macroeconomic analysis (as in Table 3.1) when necessary. A trade was counted as profitable when it reverted back to its mean before hitting the stop loss. Since we pursued an analytical goal, we disregarded issues arising in a trading context, such as bid–ask spreads and capital charges.

[27]This strategy has first been published in the ABN AMRO Research note from 17 Nov 2006 "Exploiting the regime shift with PCA weighted flatteners" and is mentioned here with kind permission from RBS.

[28]For some series, the time period for the regression is slightly different.

[29]Ideally, we would like to see an even higher correlation, though.