Determining the Practical Dimension of an Interest Rate Model

Joel R. Barber[1]
Florida International University

Mark L. Copper
Titan Open Source Interface Company

March 15, 2005

[1]Department of Finance
Florida International University
College of Business Administration
Miami, FL 33176
Tel: 305-348-2027
FAX: 305-348-4245
barberj@fiu.edu

Determining the Practical Dimension of an Interest Rate Model

Abstract

An important issue in interest rate modeling is the number and nature of the random factors driving the evolution of the yield curve. This paper uses principal component analysis to examine the dimension of the time series of historical yield curve changes determined by a variance threshold. Unlike other studies investigating the number of factors using principal component analysis, we statistically test the explanatory power. While there is no clear demarcation between operative factors and noise, the first two principal components pick up between 92 and 94 percent of total interest rate variation in a 90 percent confidence interval, and the first six pick up between 99.2 to 99.5 percent. In contrast, a single component model explains between 67 to 78 percent of the total variation within a 90 percent confidence interval.

1 Introduction

An important issue in interest rate modeling is the number and nature of the random factors driving the evolution of yield curve. A number of studies employing factor analysis and principal component analysis have found that two to three factors explains most of the variation in returns. See for example, (Litterman and Scheinkman 1991), (Barber and Copper 1996), (Falkenstein and Hanweck 1997), (Geyer and Pichler 1999), (Golub and Tillman 2000), (Dungey, Martin, and Pagan 2000), (Lekkos 2001), (Brummelhuis, Cordoba, Quintanilla, and Seco 2002), and (Soto 2003).

Many researchers have interpreted the first factor as a parallel shift in the yield curve. Litterman and Scheinkman's (1991) comment is representative of received opinion:[1]

> [T]he yield changes caused by the first factor are basically constant across maturities. That is, the first factor represents essentially a parallel change in yields.... Thus, hedging against Factor 1 is close to duration hedging.... The impact of Factor 1

Likewise, the second factor is interpreted as a change in shape or a "twist" in the yield curve, and perhaps is indicative of so called "twist risk," to which a duration matched portfolio with high convexity is exposed (Fong and Vasicek 1984).

This paper seeks to fill a gap in the literature by applying statistical tests to the results of principal component analysis of historical yield curve changes. In particular, we examine the dimension of the time series of historical yield curve changes determined by a variance threshold. Based upon techniques developed by Lawley, James, and Anderson,[2] we are able to establish confidence intervals for the proportion of variation explained by models with $K = 1, 2, ..., m$ components. The lower bound on the confidence level provides a more reliable and conservative assessment of a K-component model's explanatory power than a point estimate. Further, our approach can be used to evaluate the trade-off among the number of components, the explanatory power, and the confidence interval. Finally, we can determine a practical dimension on an interest rate model by specifying the minimum acceptable explanatory power for a given level of significance. The practical dimension K equals the smallest number of components under which the null hypothesis that a K-component model explains at least the threshold level of variance is not rejected for a given level of significance.

2 Principal Component Analysis

It is our everyday experience that interest rate changes for different maturities are positively correlated. If the correlation is perfect, then the vector representing the time t change in the yield curve X_t at defined set of m maturities can be expressed in terms of a single component:

$$X_t = b_t U$$

where U is an $m \times 1$ vector independent of time and b_t is a scalar that changes over time. The vector U can be thought of as the direction of the shift and b_t as the component of the shift in the direction U (Reitano 1996). Indeed, we will interpret b_t as a random variable that determines the magnitude and sign of the shift, and U as vector indicating the direction. For example, if U is a vector of ones, then we have the so-called parallel shift model. If the sign of the all the elements of U are positive and decreasing, then short-term rates are more volatile than longer-term rates.

The number of factors and their respective shape is an empirical question. If a single component models explains most but not all the variation in X, then we can include an error or noise term:

$$X_t = b_{1t}U_1 + E_t$$

In general, for m maturities, we can exactly represent the yield curve change as linear combination of m basis vectors:

$$X_t = b_{1t}U_1 + ... + b_{mt}U_m$$

where b_{kt} is the projection of X_t onto U_k.

As a practical matter, we like to develop a parsimonious model in which the number of components is small $K \ll m$ but the explanatory power is high:

$$\begin{aligned} X_t &= b_{1t}U_1 + ... + b_{Kt}U_K + E_t \\ &= \widehat{X}_t + E_t \end{aligned}$$

where $\widehat{X}_t$ is the approximated (or explained) value of X_t by the K component model and E_t is the deviation of the approximation from the actual value. Notice that E_t is simply the sum of the terms we have deemed insignificant. Without loss of generality, assume the set of direction vectors $U_1, ..., U_m$ is orthonormal.[3] At time t the components by definition are the projection of X_t onto the vectors $U_1, ..., U_K$:

$$b_{kt} = U_k' X_t \text{ for } k = 1, 2, ..., K$$

The sum of squared errors at time t equals the difference between the total and explained sum of squared errors:

$$
\begin{aligned}
E_t' E_t &= X_t' X_t - \widehat{X}_t' \widehat{X}_t \\
&= X_t' X_t - \left(U_1' X_t U_1 + ... + U_K' X_t U_K\right)' \left(U_1' X_t U_1 + ... + U_K' X_t U_K\right) \\
&= X_t' X_t - \left(U_1' X_t' X_t U_1 + ... + U_K' X_t' X_t U_K\right)
\end{aligned}
$$

The objective of principal component analysis is to choose the set of direction vectors such that the average of the squared error over the historical sample $t = 1, 2, ..., N$ is minimum. Given that the sample X is fixed, this is equivalent to maximizing the sum of the explained variance of X:

$$
\frac{1}{T} \sum_{t=1}^{T} \widehat{X}_t' \widehat{X} = \frac{1}{T} \sum_{t=1}^{T} \left(U_1' X_t' X_t U_1 + ... + U_K' X_t U_K\right) = \left[U_1 \cdots U_K\right]' V \left[U_1 \cdots U_K\right] \tag{1}
$$

where V is the sample covariance matrix.

Based upon a well known result from linear algebra (see (Strang 1980)), the positive definite matrix V has positive distinct eigenvalues $\lambda_1, \ldots, \lambda_m$ and the eigenvectors $Q = [Q_1, ..., Q_m]$ are orthonormal. Further, the covariance matrix V can be factored as follows:

$$
V = Q \Lambda Q' \tag{2}
$$

where Λ is a positive diagonal matrix: $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_m)$. Substituting (2) in for V in equation (1), gives

$$
\frac{1}{T} \sum_{t=1}^{T} \widehat{X}_t' \widehat{X} = \left[U_1 \cdots U_K\right]' Q \Lambda Q' \left[U_1 \cdots U_K\right] \tag{3}
$$

Assume the eigenvectors and the elements of Λ are both arranged in decreasing order of the eigenvalues. Then for a single component model ($K = 1$), expression (3) is maximized if $U_1 = Q_1$. Further, then the sum of the explained variances equals the first eigenvalue. If $K = 2$, then clearly $U_1 = Q_1$ and $U_2 = Q_2$ and the sum of the explained variances equals the sum of the first two eigenvalues. In general, the first K direction vectors equal the first K eigenvectors of the covariance matrix. The sum of the explained variance equals the sum of the first K eigenvalues, and the sum of the error variances equals the sum of the remaining eigenvalues squared.

3 Estimating Principal components from Historical Data

The use of historical data to estimate principal components in the evolution of the yield curve requires two levels of abstraction. First, a zero-coupon yield curve must be inferred from a universe of treasury securities. Second, to define and compare yield curve changes, every constructed zero-coupon yield curve must be interpolated at a common set of nodes. For the estimates presented below, we use Bliss's unfiltered implementation of McCulloch's cubic spline regression for the zero-coupon yield curve for with monthly data for the ten year period from 1992 to 2001.[4] For the common set of nodes we use McCulloch and Kwon's selection (McCulloch and Kwon 1993). These nodes are given in years in Table 1.

Table 2 lists the twelve largest eigenvalues. The remaining eigenvalues are small, but as our tests show, still significantly different from both zero and each other. Graphs of the first four principal components are given in Figures 1 and 2. By construction all the curves are orthogonal. The first component can be interpreted as an approximate translation or parallel shift.[5] The second component contains a twist at roughly eight years, and perhaps is indicative of so called "twist risk."

0.083	0.167	0.250	0.333	0.417	0.500	0.583	0.667
0.750	0.833	0.917	1.000	1.083	1.167	1.250	1.333
1.417	1.500	1.750	2.000	2.500	3.000	4.000	5.000
6.000	7.000	8.000	9.000	10.000	11.000	12.000	13.000
14.000	15.000	16.000	17.000	18.000	19.000	20.000	21.000
22.000	23.000	24.000	25.000	26.000	27.000	28.000	29.000

Table 1: Common yield curve nodes, in years, used for principal component analysis. The discount function is estimated each month with a cubic spline; the spline is then interpolated at these nodes.

0.0247224	0.0067709	0.0012182	0.0005729	0.0002075	0.0001106
0.0000804	0.0000599	0.0000294	0.0000160	0.0000121	0.0000063

Table 2: The twelve largest eigenvalues of the sample covariance matrix for Treasury securities 1992–2001 at nodes listed in Table 1.

The statistical tests described in the followings sections were developed by the statisticians, Lawley, James, and Anderson.[6] Although principal component analysis does not assume the data are normal, like most parametric statistics, the tests depend on the assumption or normality.

4 Practical Dimension: Proportion of Variation

Because the covariance matrix is nonsingular, to fully explain the time series of yield curve shifts over the sample requires all the principal components. As a practical matter, we might expect that the components corresponding to the smallest eigenvalues are spurious. Consequently, we can limit the number of components required. However, to establish a limit, we need to specify a noise threshold. Roughly speaking, we start by fixing the fraction h of unexplained to total variance that can be tolerated. Then for a given confidence level, we determine the minimum number of components under which we fail to reject the null hypothesis that the fraction of unexplained variation equals h. Suppose we are satisfied with a K component model so long as the fraction of variance explained by the remaining principal components is less than or equal to h. Formally, we consider the set hypothesis,

$$H_K^* : \frac{\sum_{i=K+1}^{m} \lambda_i}{\sum_{i=1}^{m} \lambda_i} = h.$$

for $K = 1, 2, ..., m-1$. These hypotheses can be tested with the statistics

$$M_K = -h \sum_{i=1}^{K} l_i + (1-h) \sum_{i=K+1}^{m} l_i.$$

For if H_K^* is true, then (see (Muirhead 1982), page 416) $\sqrt{n} M_K$ is asymptotically $N(0, \tau^2)$ as $n \to \infty$ where

$$\tau^2 = 2h^2 \sum_{i=1}^{K} \lambda_i^2 + 2(1-h)^2 \sum_{i=K+1}^{m} \lambda_i^2.$$

Thus, replacing λ_i by l_i, $i = 1, \ldots, m$, in this expression for τ^2, we obtain a test of H_K^* for our sample data. By computing the choices of h and K that make H_K^* acceptable, we derive the confidence intervals.

These computations provide a practical measure of the dimension of the interest rate fluctuation space. Table 3 shows the total fraction of variation explained $(1-h)$ by the first six principal components, and the 90 percent confidence interval based upon our null hypotheses. Figure 3 displays the fraction of total variation explained by models with one, two, three, and four components versus the probability (p) of rejecting the null hypothesis.

The results indicate that the first principal component explains 73 percent of the observed interest rate. The 90 percent confidence interval ranges from 66 to 78 percent of the total variation. For a two component model, the observed variation increases to 93 percent with a confidence interval ranging from 92 to

	1 p.c.	2 p.c.'s	3 p.c.'s	4 p.c.'s	5 p.c.'s	6 p.c.'s
Lower 5%	67%	92%	96 %	98.1%	98.8%	99.2%
Observed	73%	93%	96.7%	98.4%	99.0%	99.4%
Upper 5%	78%	94%	97 %	98.7%	99.2%	99.5%

Table 3: Confidence intervals and actual observed values of proportions of variation captured by the first 6 principal components.

94 percent. A six component models gives a nearly perfect fit with a confidence interval ranging from 99.2 to 99.5 percent.

5 Conclusion

Principal components analysis indicates strong support for a multifactor model. While there is no clear demarcation between operative factors and noise, the first two principal components pick up between 92 and 94 percent of total interest rate variation in a 90 percent confidence interval, and the first six pick up between 99.2 to 99.5 percent. In contrast, a single component model explains between 67 to 78 percent of the total variation within a 90 percent confidence interval.

The confidence interval provides a basis for comparison with the measurements of explained variation observed by other authors. For example, our 90% confidence interval for the first component does not include the 80% proportion found in Barber and Copper (1996), or any of the proportions found by Litterman and Scheinkman (1991).[7]

The lower bound on the confidence interval provides a more reliable and conservative assessment of a model's explanatory power. In addition, Figure 3 can be used to evaluate the trade-off among the number of components, the explanatory power, and the confidence level. A practical dimension for an interest rate model can be determined by specifying the minimum acceptable explanatory power for a given level of significance. The practical dimension K equals the smallest number of components under which the null hypothesis that a K-component models explains at least the threshold level of variance is not rejected for a given level of significance. For our sample interest rate data, with a variance threshold of 91 percent at a 90 percent confidence level the practical dimension is two, because the lower bound on the 90 percent confidence interval for a two component model is greater than 91 percent. A higher variance threshold or higher confidence level could require more than two components.

References

Barber, J. R. and M. L. Copper (1996, Fall). Immunization using principal component analysis. *Journal of Portfolio Management 23*(1), 99–105.

Bliss, R. R. (1997). Testing term structure estimation methods. *Advances in Futures and Options Research 9*, 197–231.

Brummelhuis, R., A. Cordoba, M. Quintanilla, and L. Seco (2002, January). Principal component value at risk. *Mathematical Finance 12*(1).

Dungey, M., V. L. Martin, and A. R. Pagan (2000, Nov/Dec). A multivariate latent factor decomposition of international bond yield spreads. *Journal of Applied Econometrics 15*(6).

Falkenstein, E. and J. Hanweck (1997). Minimizing basis risk from non-parallel shifts in the yield curve. part ii: Principal components. *Journal of Fixed Income 7*, 85–90.

Fong, H. G. and O. A. Vasicek (1984, Dec). A risk minimizing strategy for portfolio immunization. *Journal of Finance*.

Geyer, A. L. J. and S. Pichler (1999, Spring). A state-space approach to estimate and test multifactor cox-ingersoll-ross models of the term structure. *The Journal of Financial Research 22*(1).

Golub, B. W. and L. M. Tillman (2000). *Risk Management: Approaches for Fixed Income Markets*. New York: John Wiley & Sons.

Lekkos, I. (2001, August). Factor models and the correlation structure of interest rates: Some evidence for usd, gbp, dem and jpy. *Journal of Banking & Finance 25*(8).

Litterman, R. and J. Scheinkman (1991). Common factors affecting bond returns. *The Journal of Fixed Income 1*.

McCulloch, J. H. and H.-C. Kwon (1993, March). U.S. term structure data, 1947-1991. Working Paper 93-6, Ohio State University.

Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory.* New York: John Wiley.

Reitano, R. R. (1996, Winter). Non-parallel yield curve shifts and stochastic immunization. *The Journal of Portfolio Management*.

Soto, G. M. (2003). Duration models and IRR management: A question of dimensions? *Journal of Banking & Finance*.

Strang, G. (1980). *Linear Algebra and its Applications.* New York, New York: Academic Press.

Willner, R. (1997, June). A new tool for portfolio managers: Level, slope, and curvature durations. *Journal of Fixed Income 6*(1), 48–59.

Notes

[1]Along the same lines, see (Barber and Copper 1996), Willner (Willner 1997) and Soto, Note 17 (Soto 2003).

[2] We rely on the exposition by Muirhead (Muirhead 1982).

[3]In other words

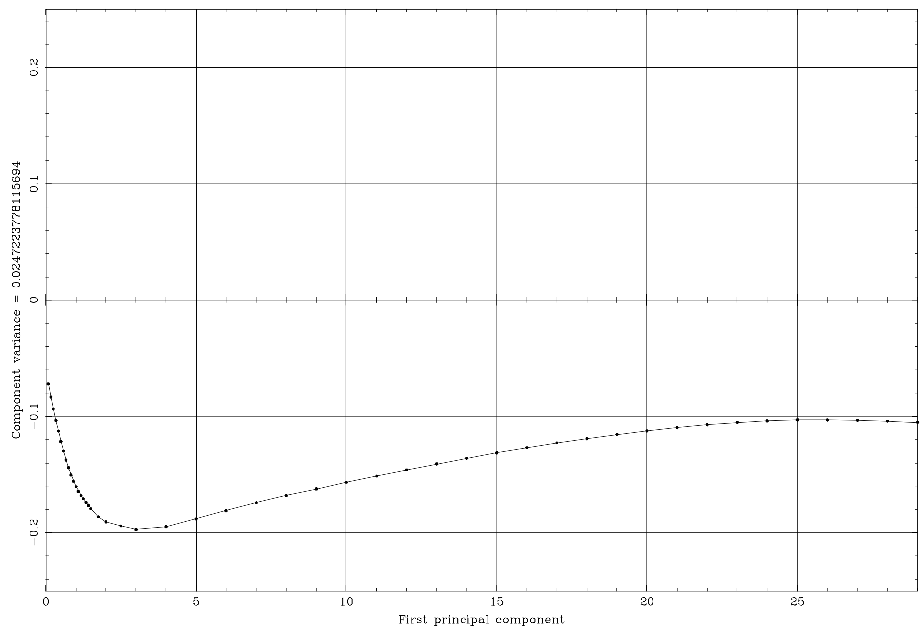
$$U_i'U_j = \begin{cases} 0 & \text{if} \quad i \neq j \\ 1 & \text{if} \quad i = j \end{cases}$$

[4]Details are given in Bliss's paper (Bliss 1997); Mr. Bliss kindly provide updated data.

[5]The sign of the principal component is arbitrary. Depending upon the sign of the coefficient at time t the shift could be up or down.

[6] We rely on the exposition by Muirhead (Muirhead 1982).

[7]Also compare to Soto (Soto 2003), note 16, for Spanish bonds.



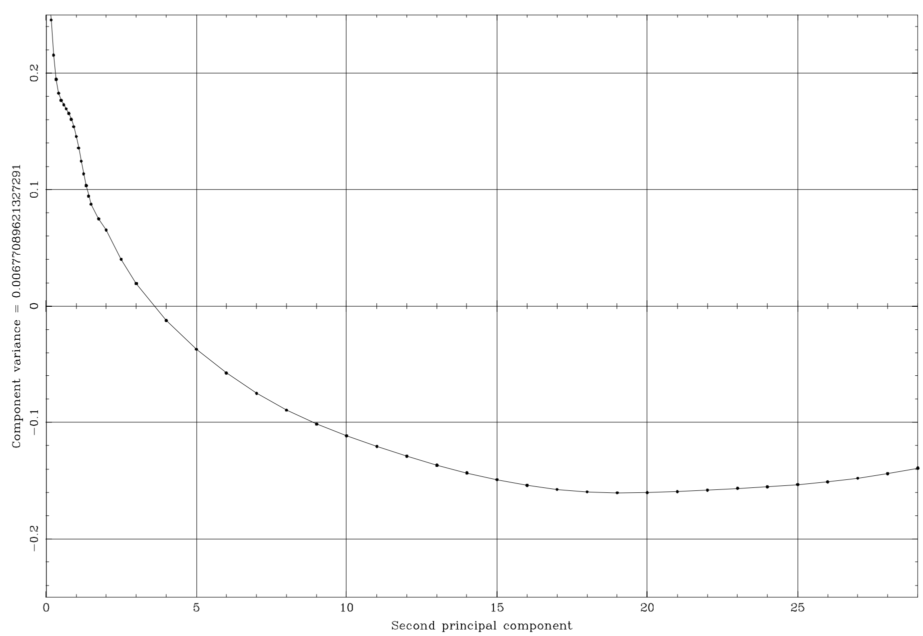


Figure 1: Principal components derived from monthly McCulloch yield curve estimate changes. Parameters derived by R. Bliss from 1992–2001 CRSP Treasury data using McCulloch's programs.

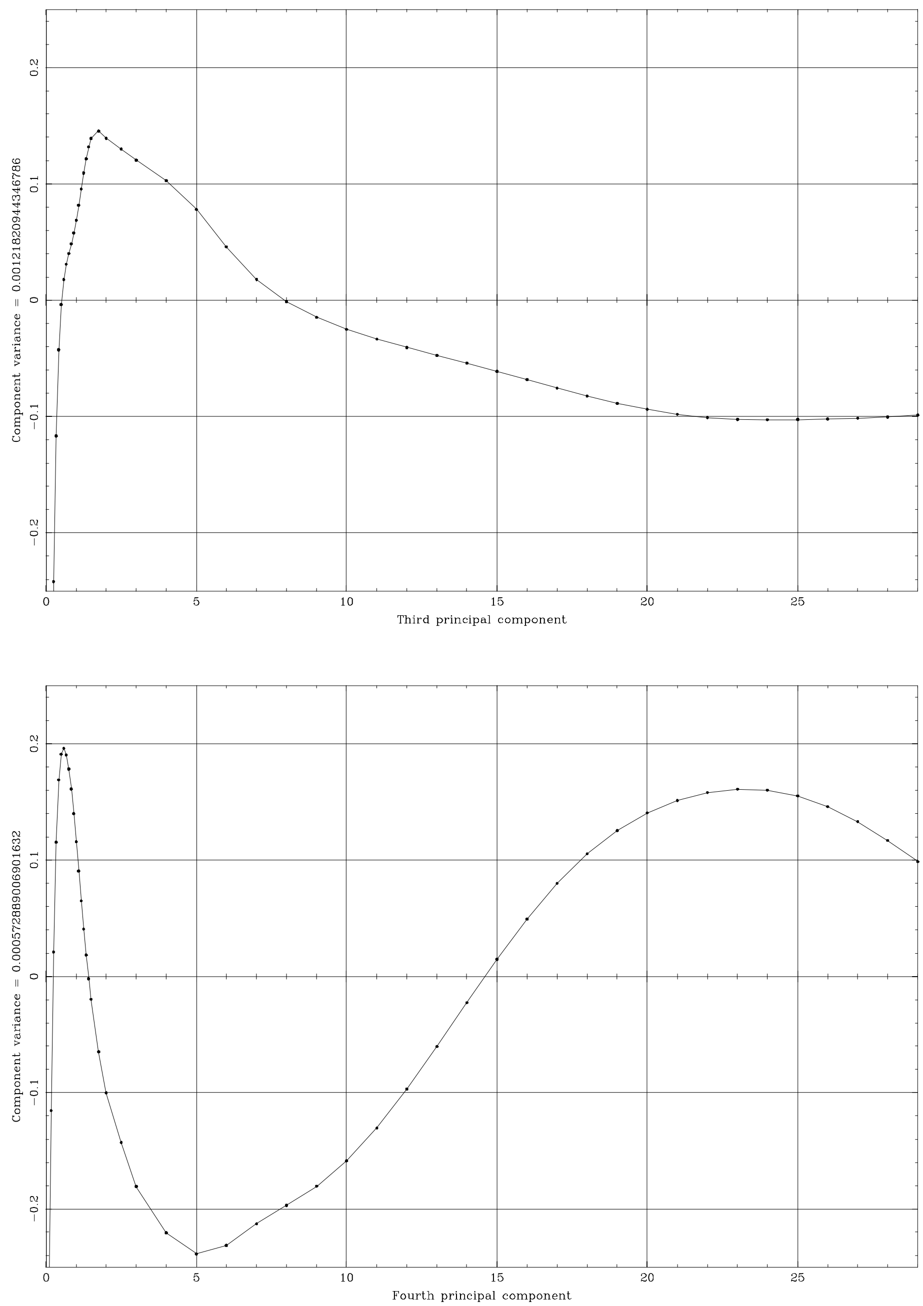


Figure 2: Principal components derived from monthly McCulloch yield curve estimate changes. Parameters derived by R. Bliss from 1992–2001 CRSP Treasury data using McCulloch's programs.

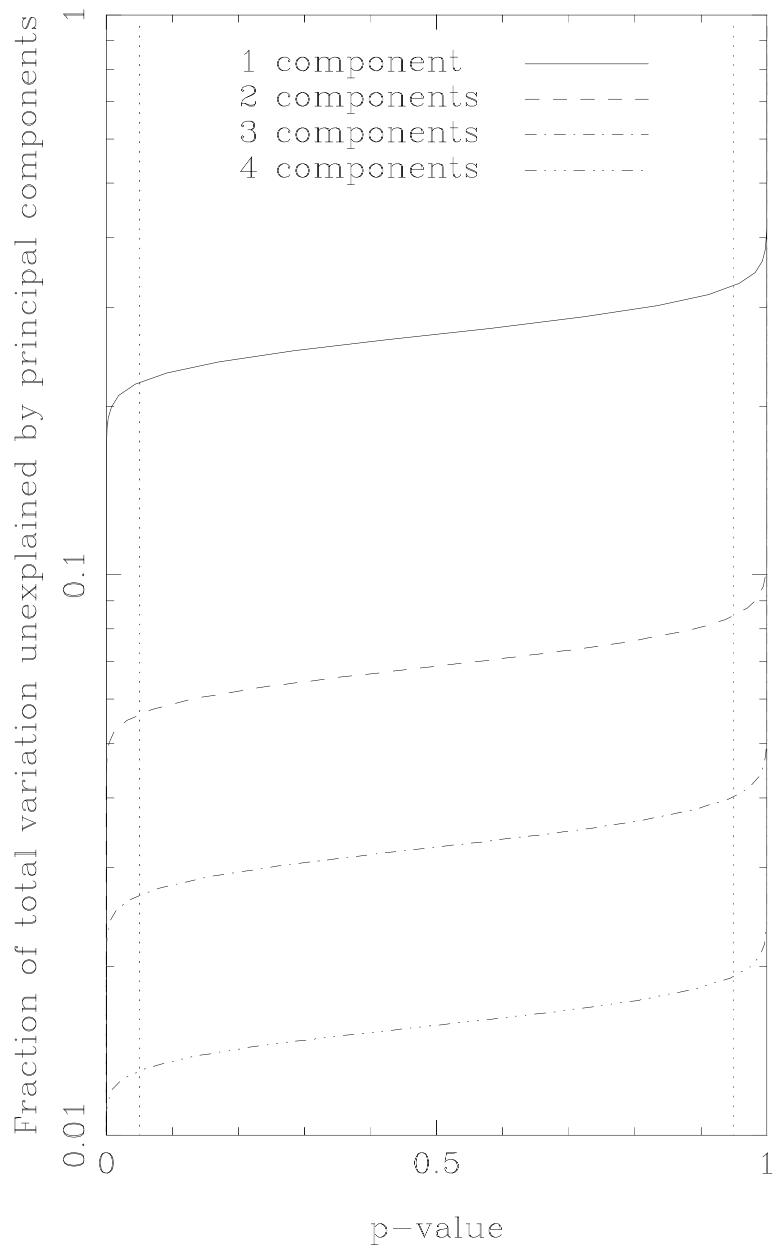


Figure 3: p-values for the explanatory power of the 4 largest principal components. 90% confidence intervals are indicated by the vertical dotted lines. The vertical axis is scaled logarithmically for better visualization.