DEPARTMENT OF ECONOMICS

DISCUSSION PAPER SERIES

Changes in the Risk Structure of Stock Returns: Consumer Confidence and the Dotcom Bubble

Lawrence A. Leger and Vitor Leone

WP 2007 - 15

Dept Economics
Loughborough University
Loughborough
LE11 3TU United Kingdom
Tel: + 44 (0) 1509 222701
Fax: + 44 (0) 1509 223910
http://www.lboro.ac.uk/departments/ec

Loughborough
University

Changes in the Risk Structure of Stock Returns: Consumer Confidence and the Dotcom Bubble

Lawrence Leger* and Vitor Leone**

Department of Economics
Loughborough University

June 2007

Abstract

Changes in the risk structure of stock returns may sometimes be very revealing. We examine economic variables that help explain principal components in UK stock returns, 01/1985 to 12/2001. The loading pattern on explanatory variables for the first component in a 'bubble' period is distinctive and consistent with a bubble/crash market. The second component shows a loading pattern on a Consumer Confidence variable in a pre-bubble period only. We observe apparently systematic changes in the structure of risk, and conjecture that Consumer Confidence captures a change in market sentiment that could be a signal for the evolution of stock prices.

*Corresponding author: L.A.Leger@lboro.ac.uk. Department of Economics, Loughborough University, LE11 3TU, UK. Telephone: (+44) (0)1509 222716, Fax: (+44) (0)1509 223910.
** Vitor.Leone@dtz.com. Economics & Real Estate Forecasting Team, DTZ, London, UK.

Changes in the Risk Structure of Stock Returns: Consumer Confidence and the Dotcom Bubble

1. Introduction

In this paper we explore the idea that analyzing changes in the risk structure of stock returns may sometimes be more fruitful than trying to identify specific economic underpinnings of risk. As an example we exploit the natural experiment created by the dotcom bubble in the late 1990s and ask whether this was characterised by changes in the risk structure of UK stock returns. We carry out principal components analyses (PCA) of monthly returns for three sub-periods between 01/1986 and 12/2000 and examine the loadings of the first and second principal components on innovations in explanatory economic variables in time series regressions. The loading pattern for the first component in the 'bubble/crash' period is both distinct from the preceding periods and plausibly characteristic of a market showing departure from fundamental pricing. While the second component shows no strong loading pattern on the explanatory variables in either the first or third sub-periods, in the 'pre-bubble' period it is characterized by significant loadings on a Consumer Confidence variable, leading us to conjecture that the changes in risk structure provided signals for the evolution of stock prices.

The relationship between risk and return in equity markets and the status of asset pricing models remains unresolved after four decades of debate. Multi-factor approaches (unlike the CAPM) have the apparent advantage of being based on observable sources of risk but there is as yet no satisfactory theory to say what these risks are, so they have to be identified empirically. For example, in their three-factor model, Fama and French (FF) (1993) use size (SMB) and book-to-market (HML) risk factors to explain the cross-section of average returns, but can only conjecture that loadings on these factors reflect firm profitability. Others (Mei, 1993; Brennan *et al.,* 1998) have focused on size and book-to-market ratios as 'non-risk' firm-level characteristics. Brennan *et al.* suggest that these attract premia incremental[1] to both the FF factors and five components of a PCA model. The appellation 'non-risk' distinguishes firm

characteristics from risk factors but the former are still subject to economic shocks and may not be wholly diversifiable. In one sense, therefore, a shift of focus towards firm characteristics simply moves the old question to a new domain – rather than identify the underlying sources of variation in stock returns we must now identify them for production, earnings and profitability.

Since they reflect the business cycle and the international competitive environment, macroeconomic variables with potential impact on the discounted dividend valuation model of stock prices seem to be attractive candidates as sources of risk. Two favoured methods for identifying risk factors in arbitrage pricing models have been either to carry out factor analyses of stock returns and relate the resulting factors to a selection of variables or to regress stock returns directly against pre-selected variables (Chan, *et al.*, 1985; Chen *et al.*, 1986; Chen and Jordan, 1993; for the US – Beenstock and Chan, 1988; Poon and Taylor, 1991; Clare and Thomas, 1994; Priestley, 1996; Antoniou *et al.*, 1998; for the UK). There are problems of interpretation here since a common premise, implicit or explicit, is that the economic variables represent specific sources of risk. This may not be reasonable, since it is hard to specify *a priori* the news impact of a shock of given size, to know what the risk implications might be or even to know whether the shock constitutes good or bad news for investors, because these things depend on the wider social, political and economic context. Indeed, Cheng (1996) argues that macroeconomic variables can themselves be clustered into common risk factors (again shifting the problem of interpretation to a new domain). More generally, even if there are strong short-term regularities and patterns, it is very difficult to believe that the risk structure of stock returns has real permanence or even that the sources of risk themselves remain the same over time[2]. Are the sources of common risk really invariant over periods of international tension and conflict, climate change, international terrorism, post-war reconstruction, financial liberalization, technological revolution, investment fads, oil shocks, inflation, electioneering[3], programmed trading, economic transition, and so on?

It is apparent in the literature that explanatory variables that work in one context very often fail to work in another. While it is possible that this can be ascribed to particular features of methodology or to sampling differences, we

argue that if there are continually evolving complex interactions among unpredictably shifting sources of risk then any search for their 'true' identity may yield only short-term results at best. On the other hand, the dynamic nature of risk may itself lead to other interesting opportunities. If we can observe systematic changes in the pattern of risk (however measured) we may be able to enhance the explanatory power of pricing models and reveal signals that suggest further patterns of evolution in stock prices. What is more, some of the burden of methodology may be relieved, since an instrument sufficient to detect meaningful changes in the pattern of risk is unlikely to require the precision of one that must identify its 'true' underpinnings.

The stock market bubble of the late 1990s forms a natural experiment in which to examine changes in the risk structure of stock returns. In particular we are drawn to the idea that a price bubble is unlikely to arise unless there is a pre-existing 'psychological climate' that allows it. Almost by definition, market participants are likely to over-estimate the significance of positive price signals in an optimistic climate and under-estimate their significance in a pessimistic climate (creating, *inter alia*, the psychological underpinnings for momentum in asset prices). But this begs the question of what it is that creates the market sentiment in the first place. Since a market bubble is highly unlikely to be a sudden random event without observable pre-cursors it is the evolution of the relevant market sentiment that is of particular interest. Indeed, compared to the progress of the bubble itself, the evolution of conditions favorable to a bubble may be relatively slow. It also seems to us that looking for bubble precursors in the time-path of stock prices is problematic, since the bubble and prior stock prices are clearly not independent and there is a strong danger of creating a circular explanation. A variable is needed that is both measured independently of stock prices and likely to capture the evolution of market sentiment. Since markets and investors are part of the wider economy, we argue that consumer confidence indicators make suitable candidates.

Our approach differs from other UK studies (Beenstock and Chan, 1988; Poon and Taylor, 1991; Clare and Thomas, 1994; Priestley, 1996), all of which sought to identify macroeconomic variables attracting risk-premia in cross-sectional pricing equations. We use a time-series approach similar to that of

Chen and Jordan (1993) to examine the macroeconomic underpinnings of risk factors but we differ by (i) using UK data and (ii) using PCA to extract components from individual stock returns instead of factor analysis on portfolios of stocks. More important, we differ from all the studies cited above by focusing on changes in the pattern of explanatory variables over time rather than seeking to identify the 'true' economic underpinnings of risk. Finally we differ from all other studies of which we are aware by the novel inclusion of a Consumer Confidence explanatory variable.

In the next section we briefly review standard PCA as a method of extracting risk factors and in section 3 we describe the data and the design of the analysis. We present PCA results in section 4 followed by a discussion in section 5 of the results of regressing component scores against the selected economic variables.

2. Risk Factors in Stock Markets and Principal Components Analysis

Arbitrage pricing models have been examined using factor-analytic techniques by a number of authors for the US (Roll and Ross, 1980; Chen, 1983; Dhrymes *et al.*, 1985a, 1985b; Shukla and Trzcinka, 1990) and the UK (Beenstock and Chan, 1986; Abeysekera and Mahajan, 1987; Cheng, 1996).

Although much of the literature has been focused on factor analysis, work by Chamberlain and Rothschild (1983) suggests that PCA may be the methodology of choice when stock returns are assumed to follow an approximate factor structure. In the APM (Ross, 1977) it is assumed that the return to asset *i*, R_i, is generated by a multi-factor process with a strict factor structure:

$$R_i = E(R_i) + \sum_{k=1}^{K} b_{ik} F_k + e_i \qquad i = 1 \ldots N \qquad (1)$$

Here $E(R_i)$ is the mean return to asset *i* while the F_k are independently distributed mean-zero, unit-variance factors with loadings b_{ik}. The F_k variables are uncorrelated with the residual disturbances, e_i, which are themselves uncorrelated with each other. The covariance matrix of returns to the *N* assets, Σ_N, can therefore be partitioned into

$$\Sigma_N = \mathbf{B}_N\mathbf{B}'_N + \mathbf{D}_N \tag{2}$$

where $\mathbf{B}_N$ is an $N \times K$ matrix of factor loadings, of rank K, and $\mathbf{D}_N$ is a diagonal matrix of finite residual variances. Assuming that no-arbitrage conditions hold and that there exists a riskless asset with return ρ, a strict factor structure implies that the mean return to asset i will be an approximate linear function of the factor loadings:

$$E(R_i) \approx \rho + \sum_{k=1}^{K} \gamma_k b_{ik} \tag{3}$$

If a sufficiently well-diversified portfolio can be formed that the residual risk of (1) vanishes then (3) will hold as an equality.

Chamberlain and Rothschild (1983) argue that the assumption of a strict factor structure with few factors is too strong, and show that an approximate factor structure will also lead to the APM relationship given in (3). Using an approximate factor structure the partition given by (2) is re-written as

$$\Sigma_N = \mathbf{B}_N\mathbf{B}'_N + \mathbf{R}_N \tag{4}$$

where for any number of assets, the residual matrix $\mathbf{R}_N$ is a positive semi-definite matrix with non-zero off-diagonal elements. Chamberlain and Rothschild show that the first K eigenvalues of Σ_N will increase without bound as N increases, with all other eigenvalues being bounded. The partition of Σ_N given in (4) is unique, with $\mathbf{B}_N\mathbf{B}'_N$ being obtained from the eigenvectors corresponding to the largest K eigenvalues of Σ_N. They then argue that it is both appropriate and convenient to identify these eigenvalues by using principal components analysis.

In PCA the principal components are written as N linear combinations of variables x_i ($i = 1 \ldots N$):

$$
\begin{aligned}
c_1 &= b_{11}x_1 + \cdots + b_{1N}x_N \\
&\vdots \\
c_N &= b_{N1}x_1 + \cdots + b_{NN}x_N.
\end{aligned}
\tag{5}
$$

The components have values or 'component scores' derived from their loadings b_{ji} on the observations x_{it} for each variable i. In what follows, the matrix $\mathbf{S}$ can be either the covariance or the correlation matrix of the x variables but it must be noted that there is no one-to-one correspondence between the components extracted in the two cases. In this study we use the correlation matrix because using the covariance matrix may (i) cause individual stocks with larger than average variances to dominate the first few principal components (Joliffe, 1986) and (ii) emphasise sample-specific noise in identifying common sources of risk.

The first component, $c_1 = \mathbf{b}_1'\mathbf{x}$, is chosen to maximise $Var(c_1) = \mathbf{b}_1'\mathbf{S}\mathbf{b}_1$ subject to the normalisation constraint $\mathbf{b}_1'\mathbf{b}_1 = 1$ (the eigenvectors have unit length). The first-order condition with respect to $\mathbf{b}_1$ of the Langrangean $L = \mathbf{b}_1'\mathbf{S}\mathbf{b}_1 - \lambda_1(\mathbf{b}_1'\mathbf{b}_1 - 1)$ is given by $(\mathbf{S} - \lambda_1\mathbf{I})\mathbf{b}_1 = 0$. Since a non-trivial solution requires λ_1 to satisfy $|\mathbf{S} - \lambda_1\mathbf{I}| = 0$, λ_1 is an eigenvalue of $\mathbf{S}$ with associated eigenvector $\mathbf{b}_1$. Imposing the normalisation constraint gives $\lambda_1 = \mathbf{b}_1'\mathbf{S}\mathbf{b}_1$, showing that λ_1 is also the variance of principal component c_1. Since $\mathbf{b}_1$ is chosen to maximise this variance it follows that λ_1 is the largest eigenvalue of $\mathbf{S}$. The second principal component is chosen similarly but subject to the additional constraint that it is orthogonal to the first. This gives λ_2 as the second-largest eigenvalue of $\mathbf{S}$, with eigenvector $\mathbf{b}_2$. The remaining $N - 2$ components are chosen analogously. The total variation of the sample is

$$\sum_{j=1}^{N} \lambda_j = \text{tr } \mathbf{S}. \tag{6}$$

If the covariation in stock returns can be summarised by a small number, K, of components, there will be K distinct eigenvalues (and principal components) that explain the majority of the variation in $\mathbf{S}$. Chamberlain and Rothschild (1983) show (i) that if $\mathbf{S}$ has K unbounded eigenvalues there is an approximate factor structure that is unique and (ii) that estimates of the K eigenvectors will converge.

The number of significant components, K, is likely to be substantially smaller than the number of variables (US cross-sectional studies suggest that about three or four risk factors are priced) but identifying this number is likely to

be complicated by sample-specific random variation, particularly when the PCA is carried out on small samples. While tests are available for the number of significant components (Connor and Korajczyk, 1993) this is not of concern here. We argue instead that if the multi-factor structure of common systematic risk has any strong underlying economic meaning then we should be able to detect this in the relationship between economic variables and scores on the first two or three components, even when using fairly small stock samples.

3. Data and Methodology

In this study, PCA is carried out on groups of UK stocks selected in alphabetical order. Patterns in the structure of risk are revealed by regressing component scores in time-series against economic variables. Since stock markets are presumed to respond to economic news, white noise 'innovations' in the variables are derived by using the Kalman Filter from STAMP (Koopman *et al.* 2000). The relationship between the innovations and component scores is established by using the general-to-specific methodology (Hendry, 1995; Hendry and Krolzig, 2001) of PcGets (OxMetrics^{TM}).

While we have argued that changes in socio-economic conditions are likely to cause the underlying risk structure of stock markets to change, it is implicit in standard PCA that the true component structure is invariant over the sample period for which the components are extracted. It is also assumed that sufficient stocks are sampled to allow valid inferences to be drawn (Stevens (1996) recommends that standard PCA should use at least twice the number of observations as variables) leading to an uneasy trade-off between a sufficient number of observations and a short enough time period to avoid within-sample instabilities in the true component structure. Shorter time periods reduce the possible number of sampled stocks, making it necessary to repeat the procedures, but replication across many different stock samples should allow reasonable inferences to be drawn with respect to the market as a whole.

The data used here came from a wider sample constructed from stocks with histories from 1975 to 2001. Monthly stock returns, adjusted for dividends and capital changes, were drawn from the London Business School Share Price Database (LSPD). 516 UK stocks showed returns over the entire period.

Financial companies, investment trusts and stocks with more than 20% missing or zero observations were excluded[4], resulting in a final sample of 240 companies and a compromise between reducing a thin-trading problem and retaining sufficient stocks for the PCA.

Inspection of the market index over the full sample period shows that the choice of 01/1997 as the starting point for our third sample period represents a compromise between the demands of PCA discussed above (which seem to us to suggest periods of 5 or 6 years) and other possible 'bubble' periods of around 7 or 3 years (starting either in early 1995 or late 1998). The final choice (01/1985 to 12/1990, 01/1991 to 12/1996, 01/1997 to 12/2001) allows the final 'bubble' period to include the major volatility of the later years of the sample (and possible departures from fundamental pricing) without extending too far the length of any sub-period. The 240 selected stocks were allocated in alphabetical order to groups of 24, allowing both sufficient observations for each PCA and replication across 10 non-overlapping stock samples. IT stocks are not identified separately in our analysis, for three reasons: first, only firms trading continuously from 1975 were selected, so IT stocks form a very small minority; second, the dotcom bubble affected all stocks; third, given the small number of IT stocks, the random allocation of stocks to samples and the high degree of replication, there is no reason to suppose that the presence of a few IT stocks in the sample creates any systematic effect. Descriptive statistics for the samples are given in Table 1.

Table 1. Descriptive Statistics for the Stock Samples

Period	1985-1990			1991-1996			1997-2001		
Group	$\overline{R}$	MV (£bn)	Beta	$\overline{R}$	MV (£bn)	Beta	$\overline{R}$	MV (£bn)	Beta
1	0.0101	9.15	0.92	0.0151	16.88	0.96	0.0002	19.46	0.91
2	0.0105	9.13	0.97	0.0119	14.54	0.97	-0.0018	47.44	0.85
3	0.0082	2.62	1.01	0.0101	5.13	0.99	0.0031	6.65	0.87
4	0.0127	11.68	0.92	0.0126	27.10	0.94	0.0025	57.85	0.88
5	0.0107	6.06	0.98	0.0094	8.00	1.04	-0.0008	8.77	0.90
6	0.0103	3.98	0.88	0.0094	7.60	0.91	0.0054	8.73	0.85
7	0.0110	5.02	0.92	0.0099	9.52	1.00	0.0027	16.47	0.93
8	0.0105	9.43	0.97	0.0121	18.91	0.97	0.0007	33.66	0.91
9	0.0126	4.28	0.93	0.0119	10.25	0.99	0.0012	16.72	0.93
10	0.0131	2.18	0.85	0.0102	4.97	0.85	0.0037	8.41	0.81
Mean	0.0110	6.35	0.94	0.0113	12.29	0.96	0.0017	22.41	0.88

It is evident that average returns are lower for the 'bubble' period (as a result of the eventual crash) with betas fairly close to unity in most cases (with a narrow spread). There is some heterogeneity in market capitalisation across groups and sub-periods but this does not produce any systematic effects in the results.

The PCA was repeated for each group of stocks and each period, producing a large number of principal components, eigenvalues and loadings. Adapting from expression (5) we can derive scores for the K principal components from the estimated component loadings, b_{gki}, and the monthly stock returns for each stock sample ($g = 1...10$) and each month, t:

$$c_{gkt} = \sum_{i=1}^{n} b_{gki} R_{it} . \qquad (7)$$

The PcGets General Unrestricted Model (GUM) for each group is given by

$$c_{gkt} = \gamma_{g0} + \sum_{h=1}^{H} \gamma_{gkh} z_{ht} + \varepsilon_{it} \qquad (8)$$

where c_{gkt} is the score for the kth component of stock sample g in month t, and z_{ht} is the observation for macroeconomic innovation h ($h = 1...H$).

The Economic Variables

Here we follow both traditional practice and published empirical findings (Chen *et al.*, 1986; Chen and Jordan, 1993; Poon and Taylor, 1991; Clare and Thomas, 1994; Priestley, 1996) to select macroeconomic variables thought to influence either expected dividends $E(D_t)$ or the discount rate R of the stock valuation model:

$$P_0 = \sum_{t=1}^{\infty} \frac{E(D_t)}{(1+R)^t} . \qquad (9)$$

The selected variables are shown in Table 2. We include the UK FTA All Share index and its Dividend Yield since the market index return is a generally powerful explanatory variable in time-series factor models and its exclusion

could lead to omitted variables bias (we are also interested in the incremental impact on component scores from other sources of economic news).

Table 2. Economic Variables

Variable	Innovation
Industrial Production	I
Unemployment Rate	U
Retail Price Index	P
Money Supply M0	S
UK Retail Sales	R
FT All Share Index	M
FT All Share Index Dividend Yield	Y
US$ Exchange Rate	X
Reuters Commodity Price Index	Co
Oil Price (£)	O
Gold Price (£)	G
20-Year Gilts Redemption Yield (Gilts)	
3 Month UK TBill Yield (TBill)	
UK Corporate Debenture and Loan Redemption Yield (Corp)	
Default Risk (Gilts − TBill)	D
Term Spread (Gilts − Corp)	T
UK Consumer Confidence Indicator	C

In addition to variables commonly selected in other studies, and following the motivation outlined in the introduction, we include a Consumer Confidence indicator. Chen *et al.* (1986) note that explanations of systematic risk should include any variables "necessary to complete the description of the state of nature" and that an example "would be one that has no direct influence on current cash flows but that does describe the changing investment opportunity set" (p384). Consumer confidence indicators that are intended to measure expectations seem to be a case in point, with various studies finding evidence of a relationship between these indicators and stock returns (Otoo, 1999; Fisher and Statman, 2002; Jansen and Nahuis, 2003). They also have the advantage of being constructed entirely independently of the macroeconomic series and of stock prices. Thus, the UK consumer confidence indicator is a composite of survey responses concerning expectations of the general economic situation and household financial position, perceptions of how these have changed over the last twelve months and views on major household purchases. If the risk

structure of stock returns is sensitive to changes in the wider socio-economic context it seems to us highly plausible that this should be indicated by effects arising through indicators of this sort.

While our other variables are expected to have an impact on the valuation of cash flows they are also intended to capture the business cycle and required risk premia (Industrial production, unemployment, retail prices, money supply, retail sales, term spread and default risk) as well as portfolio balance (market capital gain and dividend yield, commodity prices, oil and gold prices, exchange rate). Data for all explanatory variables are drawn from Datastream.

Since only unexpected changes in the variables are of importance ('news'), innovations were derived in every case by applying the Kalman Filter to the raw series – a technique that seen increasing use in recent years (Priestley, 1996; Cauchie *et al.*, 2004). Using STAMP, white noise residuals (the innovations) were generated by either an unobserved components model or (if this produced autocorrelated residuals) an autoregressive model with time-varying parameters. A detailed presentation of the Kalman filter is to be found in Harvey (1989), with brief details in Priestley (1996). The time series of the raw variables and the Kalman filter residuals are shown in Figures 1 and 2 respectively.

Figure 1. Raw Macroeconomic Variables 01/1985 to 12/2001

Market Index

Market Dividend Yield

Industrial Production

Retail Price Index

Money Supply

Unemployment Rate

Retail Sales

Commodity Price Index

Oil Price

Gold Price

Exchange Rate

Consumer Confidence

Term Spread

Default Risk
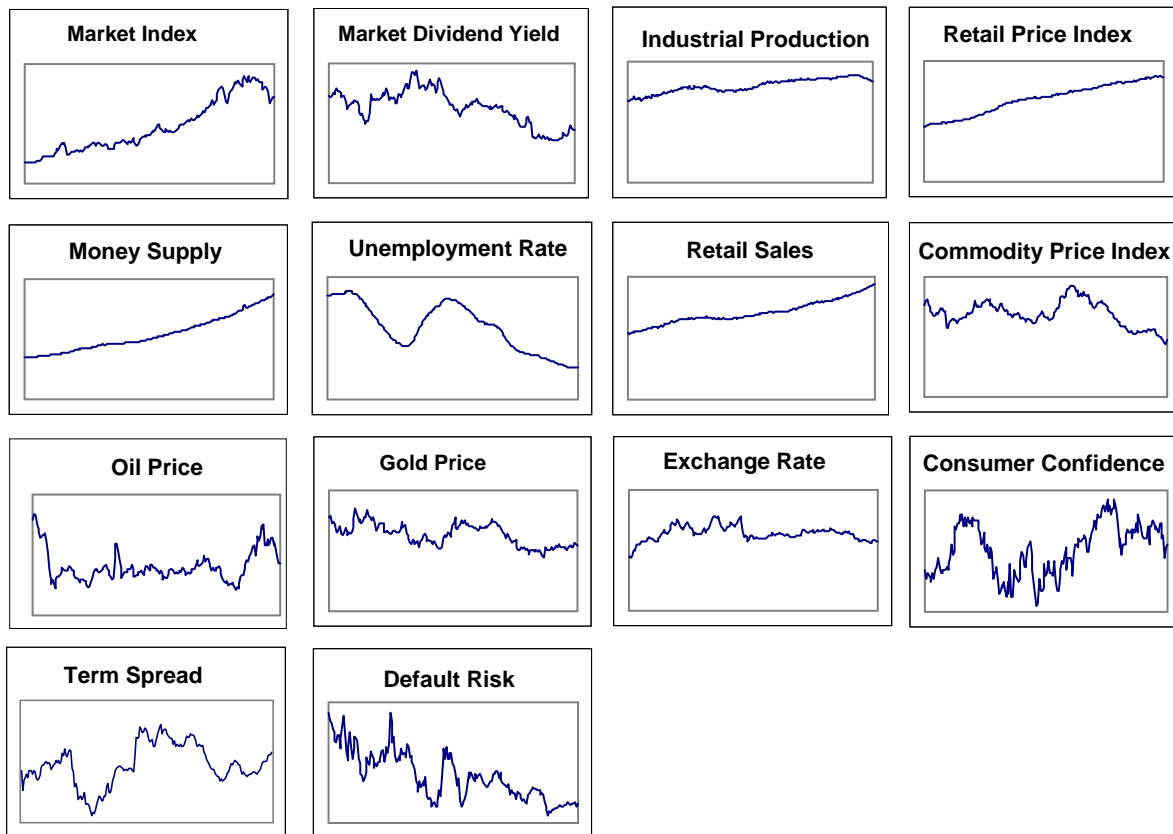
Figure 2. Kalman Filter Innovations 01/1985 to 12/2001

Market Index

Market Dividend Yield

Industrial Production

Retail Price Index

Money Supply

Unemployment

Retail Sales

Commodity Price Index

Oil Price

Gold Price

Exchange Rate

Consumer Confidence

Term Spread

Default Risk
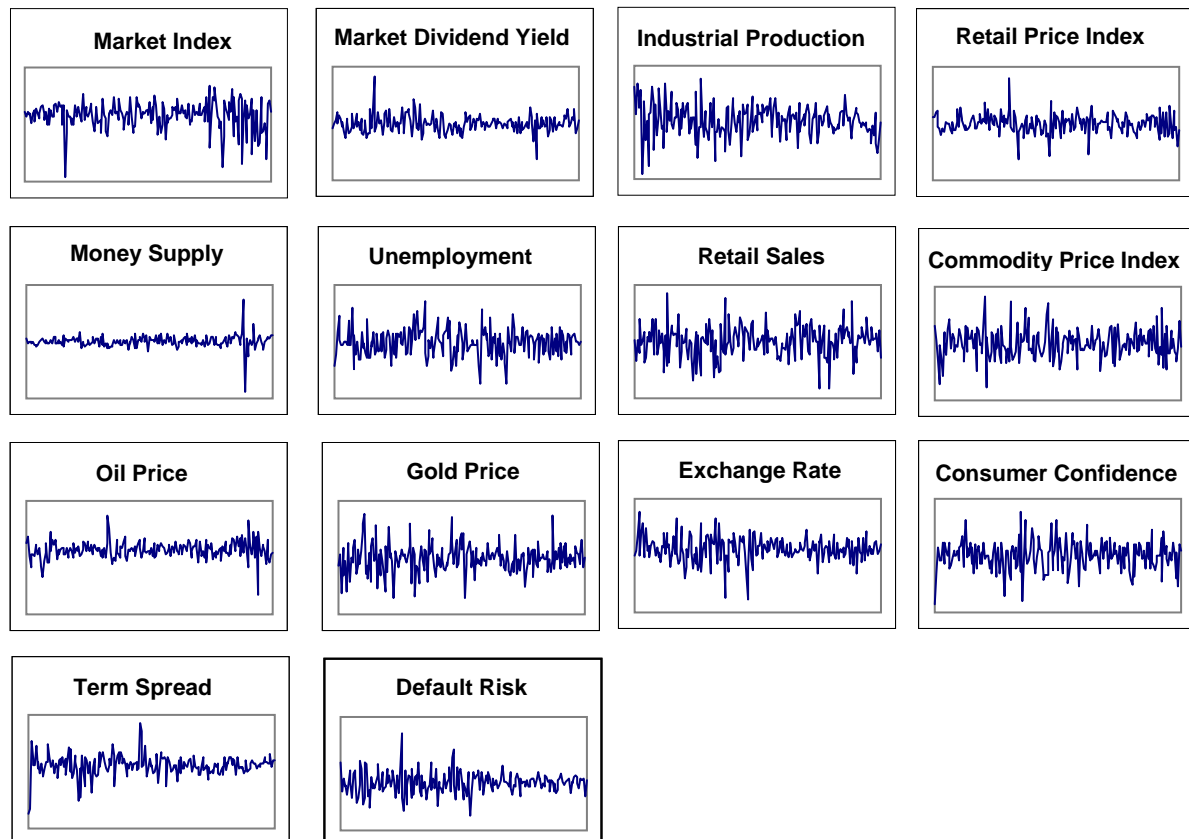
4. PCA Results

The full matrix of PCA results is quite large, so only the main results are given. Table 3 shows the incremental percent of total variation explained by the first three principal components. Table 4 shows the numbers of principal components with eigenvalues greater than unity and the percentage of total variation they explain.

Table 3. Percent of Variation Explained by Three Principal Components

Component	Period	Stock Sample Groups									
		1	2	3	4	5	6	7	8	9	10
PC1	85-90	45	43	47	42	45	39	41	48	44	33
PC2		6	6	6	8	6	7	6	6	7	7
PC3		5	6	5	5	6	5	6	5	6	6
PC1	91-96	32	28	30	31	30	29	31	33	36	26
PC2		12	7	8	8	7	9	8	8	9	6
PC3		7	8	6	7	6	6	7	6	6	6
PC1	97-01	23	22	24	24	25	24	24	24	26	21
PC2		11	9	9	9	9	9	9	9	8	10
PC3		7	8	8	7	8	8	8	8	7	8

(Rounded to whole numbers)

Table 4. Numbers of Components with Eigenvalues Greater than Unity, and Percent of Total Variation Explained

Period	Stock Sample Groups									
	1	2	3	4	5	6	7	8	9	10
85-90	5 (64)	5 (64)	5 (67)	5 (65)	5 (67)	6 (65)	7 (71)	4 (64)	5 (66)	6 (61)
91-96	6 (66)	8 (69)	7 (65)	7 (66)	8 (64)	7 (65)	6 (62)	7 (67)	6 (66)	8 (64)
97-01	8 (69)	9 (71)	7 65)	8 (68)	9 (74)	8 (67)	8 (68)	9 (69)	8 (69)	8 (67)

(% values in brackets, rounded to whole numbers)

From Tables 3 and 4 it is evident that the first component captures as much of the total variation as the second and third components combined and sometimes very much more (four times as much in the extreme case). However, the percent explained by the second and third components is not negligible (averaging 7 or 8% in each case).

The individual stock loadings on a principal component convey very little information and there is a very strong chance that they will capture sample-specific variation – (i.e. individual stock loadings have little economic significance). We therefore investigate the economic underpinnings of the overall component scores to see if any systematic explanatory patterns can be observed, with replication over different stock samples.

5. Regression of Economic Variables on Component Scores

The 10 sample groups provide 10 non-overlapping replications on stocks that should only be related via common sources of variation in returns. Similarity in the results across samples therefore suggests significant robustness of the conclusions with respect to both the methodology and any identified sources of risk. Our PCA results suggest that maybe three principal components are required to capture a substantial proportion of market risk, while cross-sectional studies suggest that about three to five risk factors may earn premia in the UK market. If these risk factors are indeed related to macroeconomic variables (as the cited studies have suggested) then this relationship should be revealed here. In particular, we might expect specific patterns of macroeconomic variables to be associated with particular principal components.

Since the market model generally describes any time series of stock returns reasonably well, we might expect Market Capital Gain (M) and Dividend Yield (Y) to be significant explanatory variables for the first principal component and for this to capture most of the systematic variation in the returns. Since a price rise will generate a positive return but a decrease in dividend yield (unless dividends are increased in proportion) we also expect M and Y to have positive and negative influences respectively on the scores for the first component.

Our null hypotheses are that: (i) regression analysis yields the same results for all 30 sample groups (the risk structure is invariant over time and across stock samples); (ii) the first component is explained by Market Capital Gain and Dividend Yield; (iii) the second and subsequent components are explained by other patterns of macroeconomic variables. Our alternative hypotheses are based on two considerations. First, Fama and French (1993) note that a much wider spread of betas is observed when the excess market return is used alone in time series regressions than when used in conjunction with their SMB and HML factors (when the market betas all move closer to unity). This suggests that the market return used alone is also likely to capture sources of risk otherwise represented by the SMB and HML factors. If the latter are driven by variation in earnings and profitability, as FF conjecture, we might expect to find additional explanatory contributions to the first principal component from the remaining variables (insofar as these affect cash flows).

Second, since we conjecture that the structure of risk varies with changes in the wider socio-economic context, we expect distinct patterns of significant variables in the various principal components for the bubble and pre-bubble periods, involving Consumer Confidence in particular.

Consumer Confidence should capture the aggregation of information contained in other variables (the wider economic context) into the economic expectations of individuals. The 'sentiment' of the market should in turn be the systematic influence on stock prices represented the aggregation of these individual expectations. It is not clear *a priori* whether Consumer Confidence should reflect the immediate or the future mood of the market, but both are possible. Thus, in the absence of a formal model of the aggregation of individual economic optimism into market sentiment, Consumer Confidence could emerge as an additional explanatory variable in any principal component (including the first component). Since it is measured independently of stock prices (and all other economic variables) we might also expect it to be the dominant explanatory variable in its 'own' principal component. Further, if investor sentiment (as measured by Consumer Confidence) evolves gradually over time, and is at least partly independent of stock prices, we should expect changes in the representation of Consumer Confidence over time.

Our alternative hypotheses are necessarily rather vague because, as far as we are aware, there is as yet neither a suitable formal model for aggregating Consumer Confidence into market sentiment nor any suitable formal structural model for clustering economic variables into sources of systematic risk. However, we propose as alternative hypotheses: (i) in addition to Market Capital Gain and Dividend Yield, other economic variables and Consumer Confidence may underpin the first principal component; (ii) Consumer Confidence is an independent systematic influence on stock prices; (iii) patterns of explanatory economic variables change over time in any principal component.

In the general-to-specific approach to economic modelling, a general unrestricted model (GUM) is formulated from the theoretical and empirical framework under consideration (in our case the relations between economic innovations and principal components). The GUM is automatically simplified by PcGets to a parsimonious congruent model containing individually significant

regressors, with each simplification stage being checked automatically by the diagnostic testing procedures of the programme. The 'black box' features of PcGets might seem questionable in some circumstances but, because they remove any possibility of subjectivity in the final selection of variables, they seem to us to be highly desirable for any research design involving replication. In our view, this substantially increases the robustness of the results.

The final models derived from the general-to-specific modeling procedure are shown in Table 5. This has been made more readable by identifying significant coefficients by letter and removing all non-significant coefficients (significance level set at 5%). Very few significant variables and no systematic patterns of any kind were observed for the third principal component, so results are reported for the first two components only. The first two columns of the table are reserved for Market Capital Gain (M) and Dividend Yield (Y) while other variables are positioned so as to maximize the visual impact of any pattern.

F-tests confirm the joint significance of the coefficients in every case for which either the initial GUM of equation (8) or an encompassing model is supported. Cases in which irrelevant variables are eliminated but no superior encompassing model can be found are denoted in Table 5 by the entry 'Yes'. The entry 'YES' indicates that a superior encompassing model can be found while 'Yes*' indicates 'YES' plus a significant intercept. The remaining category (NO) shows that the initial GUM was not significant.

Negative coefficients are flagged by (-). Inspecting for overall systematic patterns of significance among the individual variables reveals that Market Capital Gain (M), Dividend Yield (Y), Consumer Confidence (C), Retail Sales (R) and the Gold Price (G) may have some impact on the time series of returns. Of the remaining variables none occur with sufficient regularity across the 30 replications for any conclusions to be drawn.

Table 5. Significant Variables in the Regression of Component Scores on Economic Innovations

Group	First PC						R²	Gets	Second PC						R²	Gets
PERIOD 1985-1990																
1	M	Y-	G-				0.874	YES								NO
2	M	Y-					0.829	YES					G-		0.081	YES
3	M	Y-				X	0.866	YES						U-	0.055	YES
4	M	Y-				O-	0.895	YES	M-	Y-				I	0.245	YES
5	M	Y-					0.870	Yes					G-		0.067	YES
6	M	Y-	G-				0.843	YES						P-	0.079	YES
7	M	Y-					0.835	YES	M-	Y-					0.119	YES
8	M	Y-					0.877	YES	M-	Y-					0.192	YES
9	M	Y-	G-				0.855	YES								NO
10	M	Y-	G-	U-	C	Co	0.846	YES	M-	Y-		R-			0.191	YES
PERIOD 1991-1996																
1	M						0.682	YES	M		C-	R-			0.295	YES
2	M	Y-					0.623	YES						U-	0.062	YES
3	M	Y-		U			0.682	YES	M-		C				0.164	YES
4	M	Y-	G-				0.751	YES			C-				0.178	YES
5	M						0.565	YES	M-		C				0.217	YES
6	M	Y-					0.686	YES	M	Y				T	0.225	YES
7	M	Y-					0.685	YES	M-		C				0.236	YES
8	M	Y-		U		T	0.746	YES	M		C-				0.242	Yes
9	M	Y-					0.821	Yes*	M		C-				0.294	YES
10	M				C		0.414	YES								NO
PERIOD 1997-2001																
1	M					X-	0.486	YES								NO
2	M		R				0.409	YES	M					D	0.135	YES
3	M		R				0.382	YES		Y					0.096	YES
4	M						0.451	YES				R	G-		0.229	YES
5	M						0.315	Yes						S-	0.091	YES
6	M						0.366	YES								NO
7	M		R			P	0.539	YES			C				0.151	YES
8	M						0.465	YES			C				0.064	YES
9	M						0.580	YES				R			0.069	Yes
10	M		R				0.440	YES		Y				O	0.159	YES

Key

All listed variables are significant at the 5% level or better.
See Table 2 for the key to variable letter codes.
Gets: YES – final encompassing model found; Yes – final encompassing model found with significant constant term; No – irrelevant variables eliminated but no superior encompassing model found. NO – the GUM is insignificant.

For the first principal component a satisfactory model with generally high R^2 is found[5] in all 20 cases for the first two periods. Here the time series of component scores is strongly explained by M and Y(-) with weak support in the first period from the Gold Price G(-). While these results in isolation seem to support the null hypothesis, the results for the final period are quite different and are consistent with a market showing departures from fundamental valuation. Here dividend yield drops out completely and the component score is driven in

half of the groups by Market Capital Gain alone. The emergence of a weak Retail Sales (R) effect in some groups can perhaps be explained by traders overweighting sales signals when looking for confirmation of exaggerated expectations (optimistic or pessimistic) during a bubble/crash period, but the influence of this variable is restricted to only 4 of the 10 groups. The values of R^2 are generally lower in this period (suggesting that the model is less explanatory) and the significance of M is reduced (though it remains very strong). There are no obvious patterns of significance among the remaining variables and their presence probably reflects sample-specific noise. There is evidently a strongly replicated pattern of explanatory variables for the first component that distinguishes the final 'bubble' period from the prior periods. In addition, it appears that the first component is not explained solely by the combination of M and Y in any period. The null hypotheses are therefore not supported.

With respect to the second component, while the values of R^2 are much lower and the Gets algorithm fails to support an explanatory model for 5 of the 30 regressions, the results nonetheless show some interesting features. For the 'bubble' period the second component shows no systematic pattern whatsoever, suggesting that systematic risk in the market was dominated by the first component alone (this seems consistent with, for example, a market driven by psychological forces such as momentum). For the first period, while M(-) and Y(-) emerge as explanatory variables in 4 out of 10 samples this is weak replication to which it is difficult to ascribe strong economic significance. In marked contrast, for the pre-bubble period the second principal component results are quite striking: (i) although they are still rather low, the values of R^2 are generally higher than the other periods (ii) there is a strong joint contribution of M and Consumer Confidence (C) (significant in 7 out of 10 replications) and (iii) M and C always have opposite signs (so the fact that neither variable has constant sign across groups is probably not of any great concern). We therefore see sharp systematic changes over time in the pattern of explanatory variables for the second component, apparently related to Consumer Confidence.

In view of our general conjecture that the risk structure of stock returns varies over time to reflect changes in the wider socio-economic context, it is not

particularly surprising to find that the dotcom bubble is reflected in the underpinnings of principal components. The view that periods of high volatility (bubbles and crashes) are characterised by departures from fundamental asset pricing seems to be supported by our results – in the bubble period (with subsequent crash) the first component is dominated by a single market force while the second component is not explained by any of our variables.

The general non-significance of the macroeconomic variables is somewhat surprising, considering the significant risk premia that have been found elsewhere in UK cross-sectional studies. While this may arise from using small stock samples to generate the principal components, it seems to us that an economic variable that cannot easily be detected as a source of risk is probably not very important. Much more important, in our view, is that our procedures seem to have identified interesting changes in the pattern of risk during the 'natural experiment' of the dotcom bubble and crash. Of these results, the most interesting is the strong emergence of Consumer Confidence as an explanatory variable. While we expected this variable to reflect investor sentiment in a changing economic environment, we did not really expect the impact to be quite so sharply demarcated (pre-bubble period and second component only). Given that the principal components are orthogonal, and that the first component is, as expected, dominated by the actual market return, it is perfectly reasonable to suppose that an evolution of market sentiment would be detected by the second component – that is, during the pre-bubble period the gradual evolution of market sentiment and the behavior of the market index itself were measured independently. During the bubble period, however, we speculate that consumer confidence was directly translated into stock market sentiment, making it impossible to disentangle consumer confidence, market sentiment and the behavior of the market index, leading to the predominant importance of Market Capital Gain as the only clear explanatory variable and to the apparent departure from fundamental pricing.

Summary and Conclusion

In this paper we have investigated the time pattern of risk in the UK stock market, as revealed by changes in the relationship between the principal

components of stock returns and macro-economic 'news' (innovations in variables related to portfolio balance, fundamental pricing and the business cycle). Under the assumption that bubbles in asset prices require the prior evolution of a favorable climate of expectations, we examined progression in the structure of risk over three sub-periods, including a dotcom bubble period, using a UK Consumer Confidence indicator to capture the development of market sentiment. The principal components analysis and the subsequent regressions were also subjected to replication in 10 non-overlapping stock samples (making 30 replications in all) using stocks with continuous data from 1975. While three components appeared to have possible significance in the PCA only the first two showed any systematic relationship to the chosen economic variables. Of these, only Market Capital Gain, Dividend Yield and Consumer Confidence revealed any real explanatory power. The first principal component was explained by Market Capital Gain with general support from Dividend Yield in the first two sub-periods, as expected. Unsurprisingly, in the bubble period there appeared to be a departure from fundamental pricing with innovations in Market Capital Gain emerging as the only strong explanatory variable. Analysis of the second principal component also revealed strong systematic effects. In particular, Consumer Confidence emerged as strongly explanatory in the pre-bubble period, leading us to speculate that this period saw the evolution of a sentiment that pre-disposed markets to bubble. All these results were strongly robust with respect to replication across non-overlapping samples.

We have argued that the investigation of changes in risk structure is likely to be more fruitful than any search for the 'true' underpinnings of risk, speculating that these changes contain information signaling the future pattern of stock prices. Such speculation is admittedly strong and ultimately demands a suitable model of behavioral finance in which investor sentiment can both accumulate over time and aggregate into a dynamic macroeconomic influence on stock prices.

References

Abeysekera, S.P. and A. Mahajan (1987). "A Test of the APT in Pricing UK Stocks", *Journal of Business Finance and Accounting*, 14, 377-391.

Antoniou, A., I. Garrett and R. Priestley (1998). "Macroeconomic Variables as Common Pervasive Risk Factors and the Empirical Content of the Arbitrage Pricing Theory", *Journal of Empirical Finance*, 5, 221-240.

Avramov, D. and T. Chordia (2005). *Asset Pricing Models and Financial Market Anomalies*, working paper available at SSRN: http://ssrn.com/abstract=476781

Beenstock, M. and K-F. Chan (1986). "Testing the Arbitrage Pricing Theory in the United Kingdom", *Oxford Bulletin of Economics and Statistics*, 48, 121-141.

Beenstock, M. and K-F. Chan (1988). "Economic Forces in the London Stock Market", *Oxford Bulletin of Economics and Statistics*, 50, 27-39.

Bollerslev, T., R.Y. Chou and K.F. Kroner (1992). "Arch Modelling in Finance", *Journal of Econometrics*, 52, 5-59.

Booth, J.R. and L.C. Booth (2003). "Is the Presidential Cycle in Security Returns Merely a Reflection of Business Conditions?", *Review of Financial Economics*, 12, 131-159.

Brennan, M.J., T. Chordia and A. Subrahmanyam, (1998). Alternative Factor Specifications, Security Characteristics, and the Cross-Section of Expected Stock Returns", *Journal of Financial Economics*, 49, 231-262.

Cauchie, S., M. Hoesli and D. Isakov (2004). "The Determinants of Stock Returns in a Small Open Economy", *International Review of Economics and Finance*, 13, 167-185.

Chamberlain, G. and M. Rothschild (1983). "Arbitrage, Factor Structure and Mean-Variance Analysis on Large Asset Markets", *Econometrica*, 51, 1281-1304.

Chan, K.C., N-F. Chen and D.A. Hsieh (1985). An Exploratory Investigation of the Firm Size Effect, *Journal of Financial Economics*, 14, 451-471.

Chen, N-F. (1983). "Some Empirical Tests of the Theory of Arbitrage Pricing", *Journal of Finance*, 38, 1392-1414.

Chen, N-F., R. Roll and S.A. Ross (1986). "Economic Forces and the Stock Market", *Journal of Business*, 59, 383-403.

Chen, S.J. and B.D. Jordan (1993). "Some Empirical Tests in the Arbitrage Pricing Theory: Macrovariables vs Derived Factors", *Journal of Banking and Finance*, 17, 65-89.

Cheng, A.C.S. (1996). "Economic Factors and Stock Markets: Empirical Evidence from the UK and the US", *International Journal of Finance and Economics*, 1, 287-302.

Clare, A.D. and S.H. Thomas (1994). "Macroeconomic Factors, the APT, and the UK Stock Market", *Journal of Business Finance and Accounting*, 21, 309-330.

Connor, G. and R.A. Korajczyk (1993). "A Test for the Number of Factors in an Approximate Factor Model", *Journal of Finance*, 48, 1263-1291.

Dhrymes, P., I. Friend and B. Gultekin (1985a). "An Empirical Examination of the Implications of Arbitrage Pricing Theory", *Journal of Banking and Finance*, 9, 72-99.

Dhrymes, P., I. Friend and B. Gultekin (1985b). "New Tests of the APT and their Implications", *Journal of Finance*, 40, 659-675.

Fama, E.F. and K.R. French (1993). "Common Risk Factors in the Returns on Stocks and Bonds", *Journal of Financial Economics*, 33, 3-56.

Fisher, K.L. and M. Statman (2002). "Blowing Bubbles", *The Journal of Psychology and Financial Markets*, 3, 53-65.

Fraser, P. (1996). "UK Excess Share Returns: Firm Size and Volatility", *Scottish Journal of Political Economy*, 43, 71-84.

Harvey, A.C. (1989). *Forecasting Structural Time-Series Models and the Kalman Filter*, Cambridge University Press.

Hendry, D.F. (1995). *Dynamic Econometrics*, Oxford University Press.

Hendry, D.F. and H.M. Krolzig (2001). *Automatic Econometric Selection using PcGets,* 4th Edition, Timberlake Consultants Press.

Jagannathan, R. and Z. Wang (1996). "The conditional CAPM and the Cross-Section of Expected Returns", *Journal of Finance*, 51, 3-53.

Jansen, W.J. and N.J. Nahuis (2003). "The Stock Market and Consumer Confidence: European Evidence", *Economic Letters*, 79, 89-98.

Jegadeesh, N. and S. Titman (1993). "Returns to Buying winners and Selling Losers: Implications for Stock Market Efficiency", *Journal of Finance*, 48, 65-91.

Joliffe, I.T. (1986). *Principal Components Analysis*, Springer-Verlag, New York.

Koopman, S.J., A.C. Harvey, J.A. Doornik, and N. Shephard (2000). *STAMP 6.0: Structural Time Series Analyser, Modeller and Predictor*, Timberlake Consultants Press.

Mei, J. (1993). Explaining the Cross-Section of Returns via a Multi-Factor APT Model", *Journal of Financial and Quantitative Analysis*, 28, 331-345.

Otoo, M.W. (1999). "Consumer Sentiment and the Stock Market", *Finance and Economics Discussion Paper 60* (November), Federal Reserve Board.

Poon, S. and S.J. Taylor (1991). "Macroeconomic Factors and the UK Stock Market", *Journal of Business Finance and Accounting*, 18, 619-636.

Priestley, R. (1996). "The Arbitrage Pricing Theory, Macroeconomic and Financial Factors, and Expectations Generating Process", *Journal of Banking and Finance*, 20, pp 869-890.

Priestley, R. (1997). "Seasonality, Stock Returns and the Macroeconomy", *The Economic Journal*, 107, 1742-1750.

Roll, R and S.A. Ross (1980). "An Empirical Examination of the Arbitrage Pricing Theory", *Journal of Finance*, 35, 1073-1103.

Ross, S.A. (1977). "Risk, Return and Arbitrage", in J. Bicksler and I. Friend, *Risk and Return in Finance*, Ballinger, Cambridge MA.

Sadorsky, P. (2003). "The Macroeconomic Determinants of Technology Stock Price Volatility", *Review of Financial Economics*, 12, 191-205.

Shukla, R. and C. Trzcinka (1990). "Sequential Tests of the Arbitrage Pricing Theory: a comparison of Principal Components and Maximum Likelihood Factors", *Journal of Finance*, 45, 1541-1564.

Stevens, J. (1996). *Applied Multivariate Statistics for the Social Sciences*, 3rd edition, Lawrence Erlbaum Associates, Mahwah, New Jersey.

---

[1] Mei (1993) finds that size and dividend yield effects are captured by his PCA model but that book-to-market and E/P ratio effects are not. More recently, Avramov and Chordia (2005) report that a conditional version of the FF model captures the impact of firm-level size and book-to-market variables but that liquidity and momentum effects (Jegadeesh and Titman, 1993) remain unexplained in all models tested.

[2] While many have valuably asked how risk evolves over time, such research has focused on the time-series properties of the returns themselves (Bollerslev *et al.*, 1992; Fraser, 1996), on conditional asset pricing models with pre-specified risk factors or conditioning variables (Jagannathan and Wang, 1996; Avramov and Chordia, 2005), the time-varying properties of given risk factors or risk premia (Priestley, 1997) or the links between economic variables and the conditional volatility of stock prices (Sadorsky, 2003) rather than on shifts between different sources of risk.

[3] Notwithstanding the finding of Booth and Booth (2003) that the 'presidential cycle' in stock returns may be captured by business cycle variables.

[4] Jollife (1986) asserts that the presence of heavy-zero data influences the correlation structure and the results of the PCA, but there seems to be no criterion by which to define 'heavy-zero'.

[5] The R^2 values reported here are somewhat lower than are often found for the market model, and considerably lower than those found by Fama and French (1993) for their time-series regressions. However, the variables to be explained by our regressions are not portfolio returns but component scores (the loadings do not sum to unity) capturing only part of common risk.