



Kasimir Kaliva – Lasse Koskinen

MODELLING BUBBLES AND CRASHES ON THE STOCK MARKET

Kasimir Kaliva – Lasse Koskinen

MODELLING BUBBLES AND CRASHES ON
THE STOCK MARKET

Quantitative Methods in Economics and
Management Science

July
2004

HELSINGIN KAUPPAKORKEAKOULU
HELSINKI SCHOOL OF ECONOMICS
WORKING PAPERS
W-374

HELSINGIN KAUPPAKORKEAKOULU
HELSINKI SCHOOL OF ECONOMICS
PL 1210
FIN-00101 HELSINKI
FINLAND

© Kasimir Kaliva, Lasse Koskinen and
Helsinki School of Economics

ISSN 1795-1828
ISBN 951-791-865-8 (Electronic working paper)

Helsinki School of Economics -
HeSE print 2004

Modelling Bubbles and Crashes on the Stock Market

by

Kasimir Kaliva and Lasse Koskinen

Insurance Supervisory Authority
P.O. Box 449 Helsinki
Finland

ABSTRACT. The aim of this paper is to model the risk caused by stock market bubbles. The framework of the modelling is the present value approach with time-varying expected returns. A logistic mixture model of stock price dynamics is derived in which the process behaves like a random walk in one regime while it is like an error-correcting process in the other. The probability of regime switch depends on exogenous inflation. The developed bubble model is applied to the U.S. and UK stock market data. It's goodness-of-fit is found to be superior to the linear alternatives. Usually stock return models assume that the Central Limit Theorem applies and drives returns towards normality in a few months. However, the proposed bubble model may produce nonnormal predictive return distributions also over a horizon of several years. Simulation experiments are conducted with the result that risk depends heavily on the holding period, price-dividend ratio and inflation. The model produces skewed and fat-tailed short-term returns matching with the empirical observations. Also long term predictability exists.

KEYWORDS. Dividend-Price Ratio, LMARX Model, Stock Return Distribution, Time-Varying Expected Returns, Value-at-Risk. JEL C53, G10

1. Introduction

The standard efficient market model with a constant discount factor claims that stock returns are unpredictable and Random Walk is the appropriate model for stock prices. However, there is compelling empirical evidence against the standard efficient market model. The historical stock market bubbles, which find extensive description in Shiller's *Irrational Exuberance* (Shiller (2000)), show that in some time periods the behaviour of stock prices cannot be reduced to market fundamentals. More precisely, empirical studies of Shiller (1981) and West (1988) imply that stock price volatility in comparison to volatility of dividends is too large with respect to the standard efficient market models. This implies that in the long run the Random Walk model for the stock returns leads to a too broad prediction interval, given that the variability of the stock price reverts to the variability of dividends over some time horizon. Moreover, empirical studies indicate that the Random Walk model for price dynamics fails to capture extreme price movements. An extensive treatment of extremal models is eg. Embrechts et al (2001).

The starting point of this paper is the present value approach with time-varying expected returns (Campbell and Shiller 1988a-b). A detailed treatment of the approach can be found in Campbell, Lo, and MacKinlay (1997) pp. 253 – 289. Campbell and Shiller have introduced the so-called *dynamic Gordon growth model* (Campbell and Shiller 1988a-b), which allows that expected stock returns and expected growth rate of dividends are time-varying. This model is based on log-linear approximation of the dividend-price ratio. When the log price, the log dividends and the log returns are denoted by letters p_t , d_t and r_t , the log dividend-price ratio can be expressed in the form

$$(1) \quad d_t - p_t = c + E_t \sum_{j=0}^{\infty} \rho^j [-\Delta d_{t+1+j} + r_{t+1+j}],$$

where c is a constant, $\rho = 1/(1 + \exp(\overline{d-p}))$, and $\overline{d-p}$ is the average log dividend-price ratio. Formula (1) presumes that

$$P(\lim_{j \rightarrow \infty} \rho^j p_{t+j} = 0) = 1.$$

On the basis of formula (1) the log dividend-price ratio is stationary if the log difference of dividend Δd_t and return r_t are stationary. In other words

log dividend and log price are cointegrated under the above conditions. This relationship is very important in long horizon modelling. It means that

$$(2) \quad \lim_{j \rightarrow \infty} [Var_t(p_{t+j} - p_t) / Var_t(d_{t+j} - d_t)] = 1.$$

Hence long-term variability of the stock price reverts to variability of the dividend. On the basis of formula (1) either log dividend growth or stock returns are predictable. The simple form of the efficient market theory based on the constant discount factor claims that returns in the future are unforecastable. According to the empirical study of Campbell and Shiller (1988*a - b*), the log dividend-price ratio has only a weak power to forecast the growth of dividends over either a one-year or ten-year horizon. This means that stock market fluctuations are not based on superior knowledge about future dividends. On the other hand, the log dividend-price rate has no power to forecast short run log real returns, but there is a significant positive relation between the dividend-price ratio and long run log real stock returns. These results have some statistical pitfalls (Campbell et al. (1997) pp. 273 - 274) or (Campbell and Yogo (2003)). But Campbell and Yogo (2003) have found evidence for long run predictability of stock returns, when those pitfalls are taken into account.

The empirical results of stock prices excess volatility and long run predictability of stock returns have many interpretations. We interpret those so that part of the stock price fluctuation is due to stock market bubbles. In this context bubbles are time periods when stock price fluctuations cannot reduce rational expectations of future cash flows. Our definition of bubbles is quite different from "rational bubble" (e.g Campbell et al. (1997) pp. 258 - 260), which imply that the log price-dividend ratio is a unit root process. There are lots of theoretical and empirical arguments against rational bubbles (Campbell et al. (1997) pp. 258 - 260) and (Santos & Woodford (1997)). Bubbles are well-know empirical phenomena in many kinds of markets and also experimental studies are made eg. Smith et al. (1988) and Lei et al. (2001). Typical characteristics for bubbles are rapidly increasing stock prices, followed by a sudden crash.

There is no consensus about the mechanisms behind bubbles. Smith et al. (1988) argue that bubbles are caused by the lack of common knowledge of rationality. In that case some individuals understand the relationship between stock price and dividends but assume that some other individuals do

not. An experimental study of Lei et al. (2001) shows that this explanation is not sufficient. They find that some participants actually do not understand the relationship between stock price and dividends. Shiller (2000) pp. 96 – 134 has studied bubbles in a real stock market. Based on the history of stock markets he argues that stock market bubbles are usually caused by so called *new era* stories about the economy. This term refers to common beliefs about fundamental change of the economy. Those beliefs are usually closely connected to a breakthrough of a new technology.

The aim of this paper is to study what implications the existence of bubble periods has for the risk and the investment management. For this purpose we introduce a regime switching process for modelling bubbles and crashes in the above context. Here a new model that is intuitively appealing, having a "bubble" regime and a "crash" regime, is proposed. The model produces outstanding crashes in the stock market and takes into account the large fluctuation of prices. In a mathematical sense the model is a special case of the recently proposed logistic mixture autoregressive model (LMARX) (Wong and Li (2001)), which allows exogenous variables and time-varying predictive distribution. Market crashes have been modelled earlier, but these studies have adopted other regime-switching mechanisms (cf. eg. Van Norden and Schaller (1999)). The proposed bubble model is applied to the US and UK stock market data and compared with linear alternatives. Finally the risk characteristics in different holding periods and initial values are studied.

2. The bubble model

Based on the bubble phenomenon and the dynamic Gordon model, we arrive at a regime-switching model, where in regime 1 the changes of stock price are independent of dividends and in regime 2 the stock price change depends on the log price-dividend ratio (we use price-dividend ratio here for positivity). Following the principle of parsimony, we assume that in regime 1 the log stock price follows standard Random Walk

$$(3) \quad \Delta p_t = a_1 + \sigma_1 \varepsilon_t$$

and in regime 2 the change of the log stock price depends on the price-dividend ratio through simple linear regression

$$(4) \quad \Delta p_t = a_2 - by_{t-1} + \sigma_2 \varepsilon_t$$

$$= -b(y_{t-1} - \frac{a_2}{b}) + \sigma_2 \varepsilon_t$$

and

$$\varepsilon_t \sim NID(0, 1).$$

The model can be equivalently expressed in the form

$$y_t = (1 - z_t)(a_1 + y_{t-1} + \sigma_1 \varepsilon_t) + z_t(a_2 + (1 - b)y_{t-1} + \sigma_2 \varepsilon_t) - \Delta d_t,$$

where z_t is an unobservable indicator function, which is one in regime 2 and zero otherwise.

Our starting point is that the behaviour of the stock price should be almost unpredictable in the short run. This means that there is no easy and quick way to make large profits with low risk in the stock market. This assumption poses some restrictions on our model. Because a switch from regime 1 to regime 2 can cause a level shift in the stock price, the indicator function z_t has to be unobservable. Otherwise this switch can give an easy profit opportunity for an informative investor. This restriction excludes all TAR models (Tong (1990)) where regime is determined by some observable variables.

When seeking for the regime-switching mechanism we build on the following observations. Modigliani and Cohn have advanced a hypothesis that people suffer from so-called "inflation illusion" (Modigliani (1979)). According to this hypothesis, market participants discount real dividends by a nominal interest rate, which has a strong dependence on inflation. Recently Ritter and Warr (2002) have found evidence supporting that hypothesis. Secondly, Shiller has studied public attitudes towards inflation by a survey method (Shiller (1997)). He concluded that people associate a high inflation rate with economic disarray and lower purchasing power. A low inflation rate is associated with economic prosperity and social justice. Thirdly, Sharpe (2002) has found that the market expectations of real earnings growth are negatively related to expected inflation. These studies imply that inflation is negatively related to the log dividend-price ratio and a low inflation rate has a tendency to create optimistic expectations and bubble behavior in the stock market.

The above findings indicate that inflation would be a suitable explanatory variable for regime switch. Since it is desirable that regime switching is difficult to predict, we confine ourselves to dealing only with stochastic

regime-shifting processes. A final requirement for the model is that there exists in practice a working estimation method. Hence we assume that the probability of regime 1 is given by

$$(5) \quad 0 \leq \pi_t = f(i_t, i_{t-1}, \dots, i_{t-n}) \leq 1$$

where f is a function of the past log inflation values $i_t, i_{t-1}, \dots, i_{t-n}$. Now the model defined by formulas (3 – 5) is a special case of the logistic mixture autoregressive model with an exogenous variable (LMARX model) proposed by Wong and Li (Wong and Li (2001)). It can be estimated directly via the log-likelihood function, or estimation can be carried out by the EM algorithm.

When assessing investment risk it is important to take into account that stock returns are often nonnormally distributed. The return distribution has fat tails and it is asymmetric. The general practice in financial modelling is to use a nonnormal distribution for innovations. Here we have chosen another approach. We think that fat tails and asymmetry of an empirical probability distribution are an indication of nonlinearity in the data generating process defined by (3 – 5). This approach includes the assumption that the risk of the stock investment is not constant over time.

3. Estimation

Data

We apply the model defined by formulas (3 – 5) to the U.S. quarterly stock data until 1995. The data which we use is from Standard & Poor and it is available on the home page of Shiller. The stock price and dividend indices are value-weighted indices of the 500 largest companies of the USA. The inflation rate is calculated from the Consumer Price Index of the USA.

One problem, which affects the statistical inference of the log price-dividend ratio, is share repurchases. Share repurchases is a tax-favoured alternative to transfer cash to shareholders. Share repurchases by S & P 500 companies have risen sharply in recent years. Liang and Sharpe (1999) have studied the effect of share repurchases and emission. Their sample includes the 144 largest firms of the S & P 500 index. They concluded that share repurchases have a significant influence on the dividend price ratio in the end of 1990. Because we have not data from share repurchases of companies, we do not include observations after 1995 in the estimation period.

Dividends

The empirical result of Campbell and Shiller (1988a) implies that the log price-dividend ratio has no power to forecast the growth of dividends. Furthermore, the empirical cross-correlation function (Figure 1) shows that lagged change of log dividends has better explanatory power for future inflation than vice versa. Based on those observations we use the change of log dividends as the leading factor of our model. Here dividend is the running sum of dividends that are distributed during the previous 12 months. Figure 2 indicates that there seems to be a structural break in the variance of the dividend series at the turn of the 1950s and 1960s. Hence we limited the estimation period from 1959 up to 1994. We take the standard Box-Jenkins approach and model the log difference of the dividend by an autoregressive model. Since the residuals were clearly nonnormal, the innovation terms are assumed to be t -distributed. The resulting model is

$$(6) \quad \Delta d_t = 0.0031 + 0.3945\Delta d_{t-1} + 0.3711\Delta d_{t-2} + 0.0065\varepsilon_t \\ \varepsilon_t \sim t(5)$$

whose mean $E(\Delta d_t)$ is 0.013. The diagnostic study of residuals does not find evidence for serial correlation and conditional heteroskedasticity. The test results of all models are given in Appendix A.

Inflation

We model the log inflation by the AR(4) model using the log difference of the dividend of the previous quarter as explanatory variable. The resulting process is

$$(7) \quad i_t = 0.335i_{t-1} + 0.310i_{t-3} + 0.189i_{t-4} + 0.133\Delta d_{t-1} + \varepsilon_t \\ \varepsilon_t \sim N(0, 0.005^2).$$

Using the same test as for the dividends we concluded that the serial correlations of residuals are insignificant at the 5% level. Normality of standardized residuals was tested by the Jargue Bera test. It was not rejected at the 5% level.

Linear models for the price-dividend ratio

For the sake of comparison we model the log price-dividend ratio by ARMAX(1,0) with explanatory variable log inflation. We assumed the following linear symmetric model for the log price

$$\Delta p_t = \alpha + (\beta_1 - 1)y_{t-1} + (1 - \theta_1)\Delta d_t - \theta_2 i_t + \varepsilon_t.$$

By the conditional maximum likelihood technique (we assume that the first observation is fixed) we got the following parameters

$$(8) \Delta p_t = 0.215 - 0.056y_{t-1} - 1.545i_t + \varepsilon_t, \\ \varepsilon_t \sim N(0, 0.068^2).$$

This resembles linear models which are widely used in insurance business (cf. Wilkie (1995)). The serial correlation of residuals is insignificant at the 5% level, excluding lag 20. Then the p-value is 0.0281. We tested the conditional heteroskedasticity of residuals by Lagrange's multiplier test. It implies that the conditional heteroskedasticity of residuals is insignificant at the 5% level. The normality of residuals is rejected clearly. The skewness of residuals is -0.916 and the kurtosis is 5.690 . Also the normal probability plot (see figure 3) indicates that the residuals of the symmetric linear model include several divergent observations.

Then the following asymmetric linear model whose error distribution is a mixture of two normal distributions is estimated

$$\Delta p_t = \alpha_1 - \beta y_{t-1} + \theta i_t + \sigma_1 \varepsilon_t | s_t = 1 \\ \Delta p_t = \alpha_2 - \beta y_{t-1} + \theta i_t + \sigma_2 \varepsilon_t | s_t = 2 \\ \varepsilon_t \sim N(0, 1).$$

$$(9) \Delta p_t = -0.054y_{t-1} - 1.383i_t + (1 - b_t) \\ * \{0.222 + 0.045\varepsilon_t + b_t * (0.172 + 0.097\varepsilon_t)\}$$

$$\varepsilon_t \sim N(0, 1), b_t \sim BER(0, 278).$$

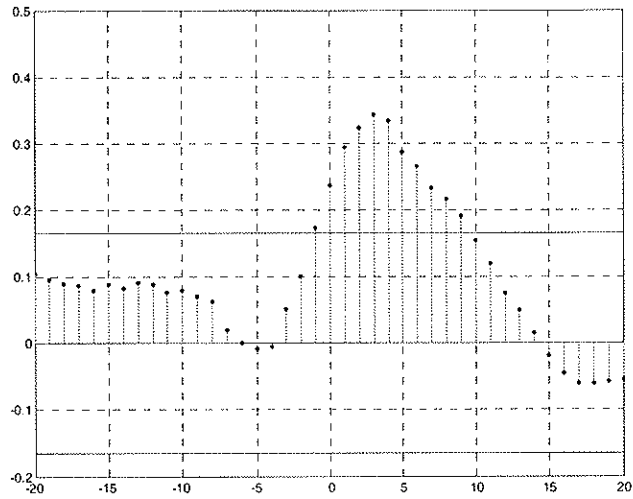


Figure 1: The cross-correlation between the change of the log dividends and the inflation rate

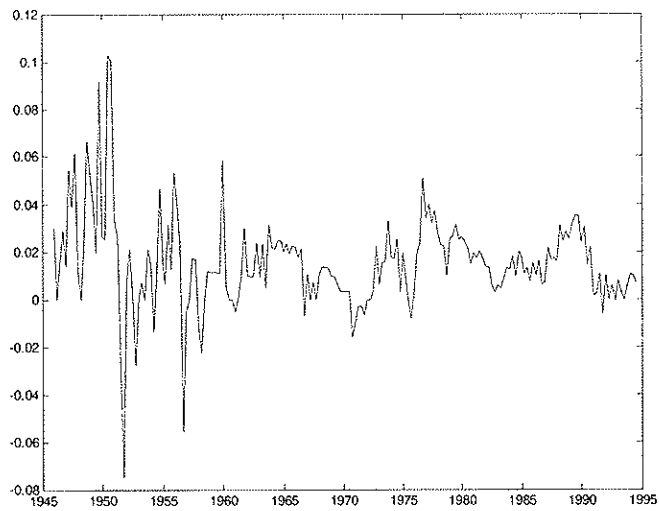


Figure 2: The change of log dividends 1946 – 1994

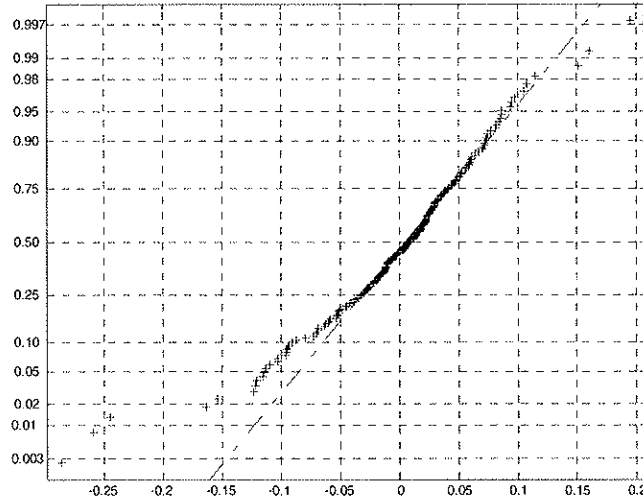


Figure 3: The normal probability plot of the symmetric linear model

The residuals of the symmetric linear model have several divergent observations. The three most divergent values are 2/1962, 3/1974 and 4/1987. The divergent observation of 1962 is related to the so-called *tronic boom* in 1959 – 1962 (Malkiel (1999)). At that time people invested in companies whose name sounded as electronics, even if the company had nothing to do with the electronics industry. A consequence of this bubble was a sharp decline in stock prices in the second quarter of 1962. The second divergent observation in 1974 is related to the oil crises. The divergent observation in 1987 is a consequence of the stock market crash in October 1987.

The bubble model

Finally, we study the bubble model defined by formulas (3 – 5). We estimate it directly via the log-likelihood function using standard numerical estimation techniques. The resulting parameter values of the model are

Regime 1:

$$(10) \quad \Delta p_t = 0.027 + 0.52\varepsilon_t$$

Regime 2:

$$(11) \quad \begin{aligned} \Delta p_t &= 1.078 - 0.357y_{t-1} + 0.077\varepsilon_t \\ &= 0.357(y_{t-1} - 3.020) + 0.077\varepsilon_t \end{aligned}$$

When quarterly inflation is denoted by i_t the probability of regime 2 is

$$(12) \quad \pi_t = \Phi(-2.120 + 333.44(i_t + i_{t-1} + i_{t-2} + i_{t-3})^2).$$

Regime probabilities are illustrated in figure 4. It shows that the probability for the process being in regime 2 is much larger during high inflation periods. By the likelihood ratio test the influence of inflation is very significant. When we test the influence of inflation the p -value is $8.2 * 10^{-6}$.

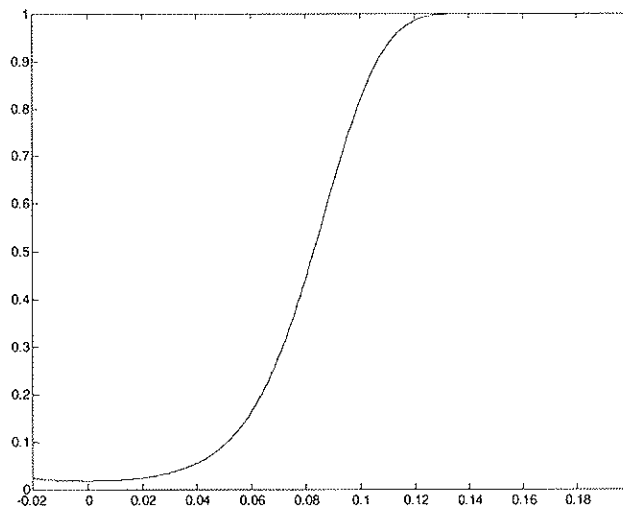


Figure 4: The probability of regime 2 with respect to inflation rate

4. Testing and diagnostics

Comparison with the linear model

Unfortunately there does not exist any straightforward way to test the null hypothesis of a linear model against two regime alternatives. The reason for this is that when the null hypothesis holds and parameters are the same in both regimes, nuisance parameters c_1 and c_2 are not identified. In other words, under the null hypothesis, values of the likelihood function are independent of these parameters. The consequence of this is that standard asymptotic properties of a likelihood-ratio test do not hold. (Andrews & Ploberger (1994))

We compared the linear model and the bubble model by model selection criteria AIC and BIC (e.g. Box et al (1994) pp. 200 – 201). Akaiken’s information criterion AIC is defined by

$$\text{AIC}(\beta) = -2 \log(L(y; \beta)) + 2p,$$

where $L(y; \beta)$ is the likelihood function of observation vector y in respect of parameter β , and p is the length of parameter vector β . The Bayesian (also called Schwarz-Rissanen) information criterion BIC is defined by

$$\text{BIC}(\beta) = -2 \log(L(y; \beta)) + p \log(n),$$

where n is the number of observations. This model selection criterion is a conservative criterion in the sense that it punishes more the complexity of the model than AIC in large samples. Table 1 shows that the regime-switching model is clearly better than the linear alternatives by both criteria.

Model	AIC	BIC
Linear symmetric	-844.5	-831.4
Linear asymmetric	-862.9	-840.1
Bubble	-875.7	-852.8

Table 1. The values of AIC and BIC for the linear models and the bubble model.

Testing against a more general alternative

To test the sufficiency of the bubble model we use the likelihood ratio test against an alternative, where the change of the log stock price depends on the log price dividend ratio in both regimes:

Regime 1:

$$\Delta p_t = a_1 + b_1 y_{t-1} + \sigma_1 \varepsilon_t$$

Regime 2:

$$\Delta p_t = a_2 + b_2 y_{t-1} + \sigma_2 \varepsilon_t$$

The value of the test statistics is 2.74, which is lower than the critical value 3.84 of the $\chi^2(1)$ -distribution at the 5% level. So the null hypothesis of the simpler model cannot be rejected at the 5% level.

Model diagnostics

For diagnostics we studied the quantiles v_t of the conditional distribution. If the bubble model (10 – 12) is true, v_t are independent and approximately standard uniform. In the case of the linear model (8) the most divergent observation is the market crash in 1987, whose studentized value is -4.24 . If the linear model (8) is true, at least as severe an observation occurs on average every 22 800 years. In the case of the bubble model the lowest quantile is 0.0052 in 2/1962. Hence, if the true data generating process is the nonlinear model, S&P 500 collapses so dramatically once in every 48 years. The market crash in 1987 has the quantile value 0.0122 and at least as severe an observation occurs on average every 20 years. These calculations show that the bubble model is much better in modelling the short-term risk of share investment than the linear model or an estimation period including extremely unusual observations.

Quantile residuals (Dunn and Smyth 1996) u_t are based on the fact that the inverse normal distribution transformations of standard uniform variables

$$u_t = \Phi^{-1}(v_t)$$

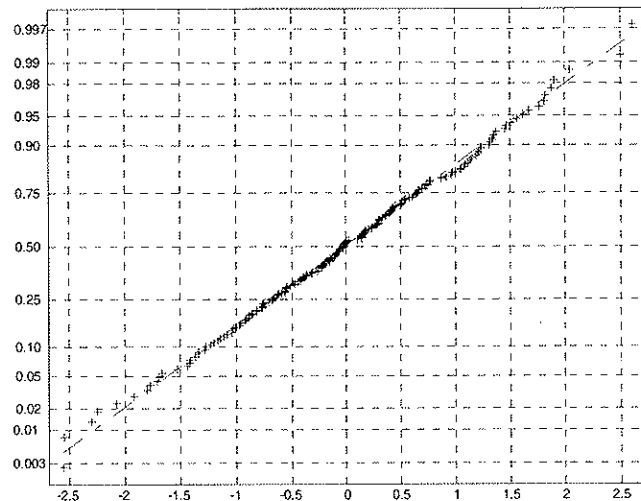


Figure 5: The normal probability plot of the quantile residuals of the bubble model

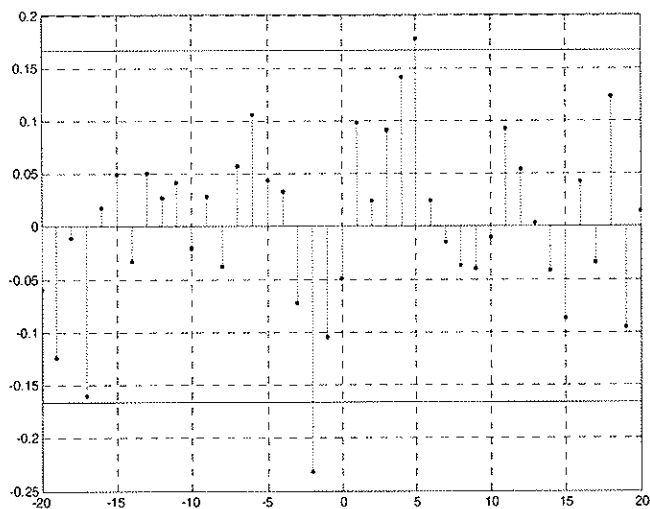


Figure 6: The cross-correlations between the quantile residuals and the log change of dividends

are standard normal variables. Figure 5 demonstrates how well this assumption holds compared with the linear model. Because many regression diagnostic tests are based on the assumption of normality, it is reasonable to use quantile residuals in diagnostic study. We tested serial autocorrelation and conditional heteroskedasticity of quantile residuals in the same way as for the linear model. The conclusion of these tests is also the same as in the case of the linear model: there is no clear evidence of serial correlation or conditional heteroskedasticity.

One basic assumption in our model is that dividend is an exogenous variable. We study this assumption by the cross-correlation function between the quantile residuals of the share price and the change of the log dividends. The cross-correlation function (see figure 6) does not show clear cross-correlations between these variables.

5. Properties of the bubble model

Interpretation of parameters

The bubble model (10 – 12) operates much more often in regime 1 than in regime 2. This means that most of the time the stock price does not react to the information on dividends. This is consistent with the short-run unforecastability of the stock price. A regime switch from regime 1 to regime 2 can cause a jump in the stock price. The sign and magnitude of the jump is mainly determined by the log price-dividend ratio of the previous quarter y_{t-1} . Formula (10) implies that if $y_{t-1} < 3.02$, a jump is more likely positive than negative and vice versa. For parameter a_1 in regime 1 it holds that $a_1 = 0.027 > E(\Delta d_t) = 0.014$. Hence, the stock price grows faster than the dividend most of the time. It means that the jump of the process after a switch from regime 1 to regime 2 is more likely negative (causing market crash or bubble burst) than positive.

The error-correction regime 2 has a crucial role in our model. It keeps the price-dividend ratio stable in the long run. The price-dividend ratio may reach very high values, but after some time period it reverts back to its normal level.

The relationship between inflation and stock price in the bubble model is totally different from that in the linear model (8). In both models inflation is statistically strongly significant. In a low-inflation period the probability

of a regime switch is approximately constant. Inflation raises the regime-switching probability substantially only when it is clearly higher than average. In a very high-inflation period the probability of regime 2 is almost 1. In other words, only a large change in inflation has a clear effect on the stock price.

A clear difference between the regime-switching model and the linear model is how the stock price behaves under a high-deflation (negative inflation) period. The used data does not include any high-deflation period. According to the linear model, the price-dividend ratio is high during a deflation period. Under the regime-switching model this relationship is negative, because the probability of regime 2 depends on the square of inflation. It is well known that deflation is often related to general economic depression. It is unreasonable to assume that the stock price is overvalued under that kind of circumstances. It is an interesting further question to check how well the nonlinear model fits the data that includes deflation periods (e.g. in Japan in the 1990s).

Time-varying moments

Traditionally risk management concentrates on the changes of the conditional volatility. The bubble model implies that also higher central moments are time-varying. It has many interesting implications from the risk and investment management perspective. One consequence is that the risk management which concentrates only on the two lowest moments might underestimate the investment risk in some cases.

Let $\Omega_t = \{i_{t-j}, y_{t-k} \mid 0 \leq j < \infty, 1 \leq k < \infty\}$ be the information set including inflation until time t and price-dividend ratio until time $t - 1$ and $\gamma(t)$ the probability of regime 2 at time t . For the bubble model (10-12) the conditional density of price difference Δp_t is

$$f(y_t|\Omega_t) = (1 - \pi(t)) \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left(-\frac{(y-\mu_1)^2}{2\sigma_1^2}\right) + \pi(t) \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left(-\frac{(y-\mu_2(t))^2}{2\sigma_2^2}\right),$$

where $\sigma_1^2 = 0.052^2$, $\sigma_2^2 = 0.077^2$, $\mu_1 = 0.027$, $\mu_2(t) = 1.078 - 0.357y_{t-1}$, and $\pi(t) = \Phi(-2.120 + 333.44(\sum_{k=0}^3 \hat{i}_{t-k})^2)$.

Hence, the first four conditional moments of price differences Δp_t based on the information set Ω_t are

$$E(\Delta p_t | \Omega_t) = 0.027 + 1.051\pi(t) - 0.357\pi(t)y_{t-1}$$

$$\begin{aligned} & Var(\Delta p_t | \Omega_t) \\ &= (1 - \pi(t))0.052^2 + \pi(t)0.077^2 + 2\pi(t)(1 - \pi(t))0.357^2(y_{t-1} - 2.944)^2 \end{aligned}$$

$$\begin{aligned} & E\{(\Delta p_t - E(\Delta p_t | \Omega_t))^3 | \Omega_t\} / \{Var(\Delta p_t | \Omega_t)\}^{3/2} \\ &= \pi(t)(1 - \pi(t))0.357(y_{t-1} - 2.944) \times \{-0.0097 + (\pi(t))^2 \\ & - (1 - \pi(t))^2\}0.357^2(y_{t-1} - 2.944)^2 / \{Var(\Delta p_t | \Omega_t)\}^{3/2} \end{aligned}$$

$$\begin{aligned} & E\{(\Delta p_t - E(\Delta p_t | \Omega_t))^4 | \Omega_t\} / \{Var(\Delta p_t | \Omega_t)\}^2 \\ &= 3(1 - \pi(t))0.052^4 + 3\pi(t)0.077^2 + (1 - \pi(t))\pi(t) \\ & \times 0.357^2(y_{t-1} - 2.944)^2 [6(0.052^2\pi(t) + (1 - \pi(t))0.077^2) \\ & + (\pi(t))^3 + (1 - \pi(t))^3]0.357^2(y_{t-1} - 2.944)^2 / \{Var(\Delta p_t | \Omega_t)\}^2 \end{aligned}$$

These formulae tell that the moments vary with inflation and price-dividend ratio. This is a feature where the proposed bubble model differs fundamentally from other stock return models. It results from the time-varying marginal distributions of the LMARX model (see Wong and Li (2001)). For interpretation let us assume that inflation is constant. Then we obtain immediately conclusions concerning the first two moments. First, increasing price-dividend ratio results in decreasing conditional expectation. Second, the smaller the quantity $|y_{t-1} - 2.9440|$ is the smaller is the conditional variance. In other words low dividend-price ratio is related to low conditional variance. The conditional skewness and kurtosis have a rather complex dependence structure. The higher the moment is the more sensitive it is to the change of price-dividend ratio y_{t-1} .

6. Testing with the UK-data

Here we do not model dividends and inflation. The only aim is to test the bubble model (3-5) with the UK monthly FTA index between 1966 and 1995. This data is available in the appendix of Terence Mills's book *Econometric Analysis of Financial Time Series* (Mills (2000)).

Nonlinear equations (3-5) and the linear alternatives (8-9) were estimated. According to the model selection criteria the bubble model is again superior (see Table 2). This model has one parameter more in the case of the UK than the USA. In the regime 1 model is heteroskedastic with respect to annual inflation rate. The resulting bubble model is:

Regime 1:

$$(13) \quad \Delta p_t = 0.0077 + (0.0417 + 0.1227(\sum_{k=0}^{11} i_{t-k})^2)\varepsilon_t$$

Regime 2:

$$(14) \quad \begin{aligned} \Delta p_t &= 1.5000 - 0.5194y_{t-1} + 0.0139\varepsilon_t \\ &= -0.5194(y_{t-1} - 2.888) + 0.0139\varepsilon_t \end{aligned}$$

where $\varepsilon_t \sim NID(0, 1)$ and the probability of state 2 is

$$(15) \quad \gamma_t = \Phi(-2.9836 + 12.7522\sum_{k=0}^{11} i_{t-k}).$$

The p -value of the test for the impact of inflation on function γ_t is $5.5 * 10^{-6}$. So the influence of inflation is statistically very significant. In addition that the influence of inflation on the stock price through the function γ_t in the UK, the conditional variance in regime 1 depends on the inflation rate. This effect is also very significant. When we test it by the likelihood-ratio test we get p -value $5.9 * 10^{-6}$. We also test the bubble model against a more general alternative, where the log dividend-price ratio has influence on the stock price in both regimes. The conclusion of that test is the same as in the case of the USA. This model does not fit the data statistically significantly better than the bubble model. The linear alternative is

$$\Delta p_t - i_t = 0.0665 - 0.0213y_{t-1} + \varepsilon_{yt}$$

where $\varepsilon_{yt} \sim N(0, 0.0608)$. Residuals of the linear model diverge very significantly from the normal distribution. The skewness is 0.645 and the kurtosis is 14.359. In table 2 we compare three models - the linear model with normal error distribution, the linear model with asymmetric error distribution (a mixture of two normal distributions) and the bubble model - by the model selection criteria. The bubble model is much better than linear models in this comparison.

Model	AIC	BIC
Linear symmetric	-1644.7	-1633.0
Linear asymmetric	-1654.9	-1631.6
Bubble	-1748.5	-1721.3

Table 2. The values of AIC and BIC for the linear and the bubble model

A comparison between the UK and the US is very illustrative. According to the linear model these markets seem to be quite different. The regression coefficient of the inflation rate is positive for the US data and negative for the UK data. The skewness of the empirical residuals has also a different sign in both data. However, the structure of the bubble model is quite similar in both markets. This observation indicates that the bubble model is less sensitive to the selection of estimated sample although it is nonlinear. Moreover, by the model selection criteria the bubble model is clearly better than models that assume higher central moments as time-invariant in both the US and the UK markets.

7. Risk evaluation

Albrecht et al (2001) have analytically quantified the short- term and long-term risks of a stock investment in the traditional Random Walk context. This paper gives a different picture of some aspects of risks, because of the unequal underlying model. In particular the risk imposed by changing predictive distributions is studied by conducting simulation experiments by the US bubble model (10 – 12). Here we consider merely real returns including dividends, which are reinvested into the index at the end of each year. The quartely returns are calculated by assuming that the dividends of each year are equally distributed. The factors whose effect we study are holding period, inflation and price-dividend ratio. Holding periods that we study are 1/4, 5 and 20 years. All simulations were repeated 10^4 times.

Predictability

First we study the linear predictability of real returns in different holding periods. For that purpose we simulate our model 10^4 times 250 years (1000 quarters). Our initial values for the log price-dividend ratio is 3.2 and 0.04 for the annual inflation rate. We assume that 250 years is a sufficiently long time to guarantee that the last observation of each scenario is in practice independent of the initial value. We use the last observation of each scenario as the initial value for the annual inflation rate and the log dividend-price ratio of the investment holding periods. Then we simulate each scenario 20 years further and calculate cumulative log real returns in selected holding periods for each initial value of the investment holding period.

The results of the simulation experiment are represented in table 3 and 4. Table 3 reports the correlation coefficient between the previous year's annual

inflation rate and the log real returns in holdings periods. There is a weak negative correlation between the inflation rate in the short run and a little stronger positive correlation between the inflation rate and the twenty-year log real returns. These results support the argument that stocks form at least a partial hedge against inflation in the long run.

Horizon	1/4	5	20
Correlation	-0.123	-0.049	0.244
R^2	0.015	0.002	0.060

Table 3. Inflation rate and log real returns

Table 4 reports the correlation coefficient between the log price-dividend ratio at the begin of the holding period and the log real returns in holdings periods. Results are consistent with Campbell and Shiller (1988), who conclude that the log price-dividend ratio has strong power to predict stock returns in the long run but only weak power to predict stock returns in the short run. The R^2 statistics grows from 0.8 per cent to 48.7 percent when the length of the holding period grows from a quarter to twenty years.

Horizon	1/4	5	20
Correlation	-0.090	-0.466	-0.698
R^2	0.008	0.217	0.487

Table 4. The log price-dividend ratio and log real returns

An important problem in investment management is the question about mean reversion. When stock returns are mean-reverting the variance ratio

$$Var(\sum_{i=1}^h r_{t+i}) / hVar(r_{t+1})$$

is lower than unit, when h is sufficiently large. Our simulation experiment implies weak mean reversion in the twenty-year holding period with respect to quarterly returns. The standard deviation of the log real quartely returns is 0.0832 and the standard deviation of the twenty-year log real returns is 0.6805. The variance ratio is then $\frac{0.6805^2}{80 \cdot 0.0832^2} = 0.8362$. The standard deviation of the five-year log real returns is 0.4109 and the variance ratio with respect

to quarterly returns is 1.2195. So in our model the stock returns are weakly mean-reverting in twenty-year holding period and weakly mean-averting in a five-year holding period.

Predictive distributions, moments and value-at-risk

In order to assess the impact of the distributional characteristics on the risk we compute conditional moments and value-at-risk (quantile) values from simulations. The results are shown in Appendix B. Three initial values (0, 0.4, and 0.1) for inflation and four initial values (2.8, 3.2, 3.6, 4.0) for log price-dividend ratio are considered. Log price-dividend ratio 3.2 is a typical historical level and 4.0 a very high one.

The shape of the predictive distribution depends heavily on the investment horizon. This is illustrated in figures (7 – 8) where forecasting horizon varies from one quarter to twenty years with the initial value of inflation 0 percent and log price-dividend ratio 4.0. The predictive distribution over a one-quarter horizon is negatively skewed, the five-year distribution is rather symmetric, but the twenty-year predictive distribution is positively skewed. The outcomes of the other simulations are fairly similar. These observations are important since a negatively skewed distribution imposes much higher risk than a symmetric one and the other way around. Value-at-risk simulations show that a five-year investment period possesses the highest risk when the log price-dividend ratio is high (3.6 or 4.0) and one-quarter is the riskiest when the log price-dividend ratio is low. The twenty-year horizon is quite safe unless the log price-dividend ratio is very high 4.0. Luckily, simultaneously high inflation and price-dividend ratio is not a likely occurrence in the real economy since it would be the most dangerous combination.

The simulation results can be summarized as follows. The risk depends to a considerable extent on the present price-dividend ratio and the risk persists over a very long period. High inflation is an important risk factor, but it plays a much lesser role over a long horizon.

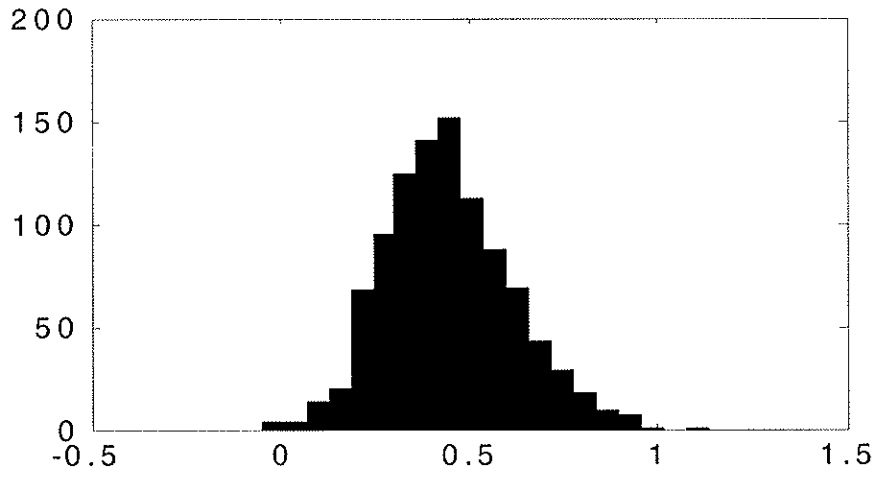


Figure 7: Five-year predictive distribution.

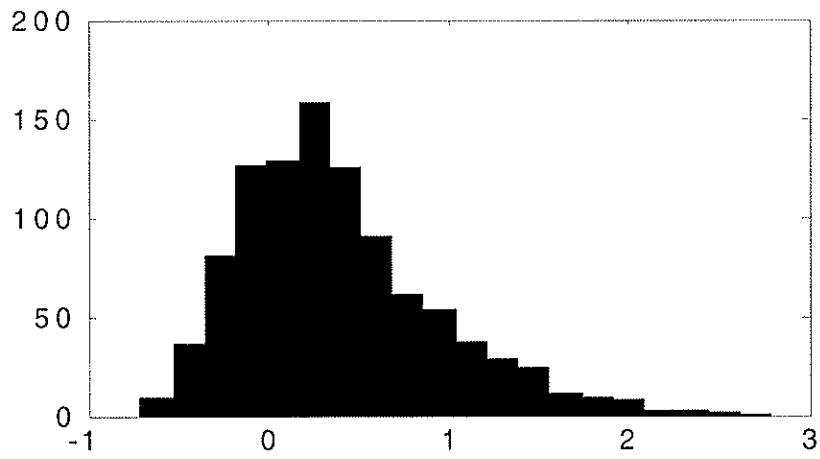


Figure 8: Twenty-year predictive distribution.

Conclusions

We have introduced a new statistical model for stock price bubbles, which is based on the recently published regime-switching technique (Wong & Li (2001)). The proposed bubble model is based on the assumption that all fluctuations of the stock price cannot be reduced to the rational expectations of cash flows. The model has two regimes: the bubble regime and the error-correction regime. The regimes have a useful interpretation - regime one produces bubble periods and regime two crashes which bring prices back to fundamentals whose indicators are valuation ratios. According to the used statistical criteria the bubble model is a superior alternative for both the U.S. and the UK data. Bubble periods are most often associated with a low inflation rate. The relationship is statistically very significant. This finding is consistent with the survey study that claims that low inflation is associated with positive economic expectations (Shiller (1997)).

The simulations show that the short run returns are almost unforecastable linearly and the bubble model does not include quick profit opportunities with a low risk for a well informed investor. Despite weak short-run forecastability the bubble model includes clear long-run forecastability. The log price-dividend ratio explained almost 50% of the variance of the twenty-year log real returns. These results are consistent with Shleifer and Vishny (1997), who show that existence of arbitrageurs does not necessarily eliminate market inefficiencies. Moreover the model implies that the stock returns are weakly mean-reverting in the twenty-year holding period and mean-averting in the five-year holding period.

According to the proposed bubble model, stock investment risk is much higher when the price-dividend ratio is high. The inflation works as an exogenous trigger variable. Usually a stock return model assumes that the nonnormality exists merely in the short run. However, the proposed bubble model may produce nonnormal predictive log real return distributions also over a horizon of several years.

Campbell and Shiller (1988a) argue that the conditional expectation of stock return depends on the price-dividend ratio. We have found evidence for higher moments also. Return simulations show that risk management which concentrates only on the time dependence of the two lowest moments might underestimate the investment risk.

The stock market's downturn at the beginning of this century has been a heavy burden on the solvency of insurance companies and corporate pension

funds. Scherer (2003) argues that there has been too much focus on the long horizon when the real problem has been short-term solvency. The proposed bubble model reveals time-varying short-term and long-term risks. Thus it might be useful for assessing the solvency of an insurance company and for ALM purposes.

REFERENCES:

Albrecht, P., R. Mayurer and U. Ruckpaul: "The shortfall-risks of stocks in the long run," *Financial Markets and Portfolio Management* 15, 481 - 499, (2001).

Andrews, D. W. K and W. Ploberger: "Optimal tests when a nuisance parameter present only under alternative," *Econometrica* 62, 683 – 734, (1994)

Box, G. E. P., G. M. Jenkins and G. C. Reinsel: *Time Series Analysis Forecasting and Control*, 3rd edition, Prentice Hall.(1994)

Campbell, J.Y. and R. J. Shiller: "The dividend -Price Ratio and Expectations of Future Dividends and Discount Factors." *Review of Financial Studies*, 1, 195-227. (1988a)

Campbell, J.Y. and R. J. Shiller: "Stock Prices, Earnings, and Expected Dividends," *Journal of Finance* 43, 661-676. (1988b)

Campbell, J.Y., A.W. Lo, and A.C. Mackinlay: *The Econometrics of Financial Markets*. Princeton University Press. (1997).

Campbell, J.Y. and M. Yogo : "Efficient Tests of Stock Return Predictability," NBER Working Papers 10026 (2003).

Dunn, P.K. and G. K. Smyth: "Randomized Quantile Residuals," *Journal of Computational and Graphical Statistics* 5, 236 – 244 (1996)

Embrechts, P., C. Klueppelberg, and T. Mikosch: *Modelling Extremal Events for Insurance and Finance*. 3rd edition, Springer-Verlag Berlin Heidelberg New York. (2001)

Engle, R. F: "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of UK Inflation," *Econometrica* 50, 987 – 1008 (1982).

Lei, V., C. Noussair and C. Plott: "Non-Speculative Bubbles in Experimental Asset Markets: Lack of Common Knowledge of Rationality vs. Actual Irrationality," *Econometrica* 69, 831 – 859.(2001)

Liang, J. N. and S. A. Sharpe: "Share repurchases and employee stock options and their implications for S&P 500 share retirements and expected returns," *Finance and Economics Discussion Series 1999–59*, Board of Governors of the Federal Reserve System (U.S.) (1999).

Malkiel, B.: *A Random Walk Down Wall Street*. 7rd edition, W. W. Norton & Company. (1999)

Mills, T. C.: *Econometric Analysis of Financial Time Series*. Cambridge University Press. (2000)

Modigliani, F. and R. Cohn : "Inflation, Rational Valuation, and the Market," *Financial Analysts Journal* 35, 24 – 44 (1979).

Ritter, J. R., and R.S. Warr: 'The Decline of Inflation and the Bull Market,' *Journal of Financial and Quantative Analysis* 37, 29 – 61. (2002)

Santos, M. S. and M. Woodford : "Rational Asset Pricing Bubbles," *Econometrica* 65, 19 – 57.(1997)

Scherer, B. (Edited): *Asset and Liability Management Tools, A Handbook of Best Practice*, Risk Books, London (2003).

Sharpe, S. A. : "Reexamining Stock Valuation and Inflation: The Implications of Analysts' Earnings Forecasts," *The Review of Economics and Statistics* 84, 632–648, (2002)

Shiller, R. J. : "Do Stock Prices Move Too Much to Be Justified by Subsequent Changes in Dividends?," *American Economic Review* 71, 421 – 435.(1981)

Shiller, R. J.: "Why Do People Dislike Inflation?," Romer, C. D ja Romer, D. H: *Reducing Inflation: Motivation and Strategy*. University of Chicago Press. (1997)

Shleifer, A. and R. Vishny : "The limits of arbitrage," *Journal of Finance* 52, 35 – 55 (1997).

Smith, V., G. Suchanek and A. Williams : 'Bubbles, Crashes and Endogenous Expectations in Experimental Spot Asset Markets,' *Econometrica* 56, 1119 – 1151 (1988).

Tong, H. : *Non-linear Time Series: A Dynamic System Approach* Oxford University Press (1990).

van Norden, S. and H. Schaller : "Speculative Behaviour, Regime-Switching and Stock Market Fundamentals," *Nonlinear Time Series Analysis of Economic and Financial Data*, (Philip Rothman ed.) Kluwer Academic Publishers, Boston (1999).

West, D. K.: "Dividend innovations and stock price volatility," *Econometrica* 56, 37 – 61. (1988)

Wilkie, A. D.: "More on a stochastic asset model for actuarial use," *British Actuarial Journal*, 1,777 – 964 (1995).

Wong, C. S. and W.K. Li : "On a logistic mixture autoregressive model," *Biometrika* 88 : 833 – 846. (2001).

Appendix A

We test serial correlation of residuals by the Ljung-Box Q-test (Box (1994) pp. 314 – 317) and conditional heteroskedasticity by Engle's test statistics (Engle (1982)). We start by testing in one lag, second including the first four lags, third including the first ten lags, fourth including the first twenty lags. Results of these tests are represented in the tables below.

Lags	1	4	8	20
Q-statistics	0.25(0.62)	0.62(0.96)	8.23(0.41)	14.95(0.78)
Engle's statistics	5.89(0.015)	12.88(0.012)	17.10(0.029)	28.83(0.091)

Residuals of the dividend model.

Lags	1	4	8	20
Q-statistics	0.003(0.96)	4.6(0.33)	5.4(0.72)	19.5(0.49)
Engle's statistics	2.7(0.10)	11.2(0.024)	15.3(0.054)	25.5(0.18)

Residuals of the inflation model.

Lags	1	4	8	20
Q-statistics	2.6(0.11)	6.7(0.15)	12.3(0.15)	34.2(0.025)
Engle's statistics	0.051(0.82)	3.6(0.46)	5.7(0.68)	12.7(0.89)

Quantile residuals of the bubble model (USA).

Lags	1	4	8	20
Q-statistics	3.5(0.062)	8.0(0.091)	17.3(0.14)	28.4(0.10)
Engle's statistics	7.5($6.0 * 10^{-3}$)	14.7($5.4 * 10^{-3}$)	32.4($1.2 * 10^{-3}$)	34.9(0.89)

Quantile residuals of the bubble model (UK).

Appendix B

Simulated conditional moment and VaR for the US bubble model (10-12). Simulations are repeated 10^4 times over each horizon.

p_0-d_0	2.8	2.8	2.8	3.2	3.2	3.2
Inflation	0	0.04	0.10	0	0.04	0.10
mean	0.071	0.048	0.012	0.035	0.023	-0.067
st. deviation	0.068	0.067	0.068	0.054	0.058	0.081
skewness	-0.008	-0.030	0.013	-0.190	-0.419	-0.074
kurtosis	2.978	3.024	3.040	3.520	4.149	2.869

The conditional moments of the quarterly log real returns.

p_0-d_0	3.6	3.6	3.6	4.0	4.0	4.0
Inflation	0	0.04	0.10	0	0.04	0.10
mean	0.030	0.010	-0.191	0.025	-0.002	-0.319
st. deviation	0.059	0.076	0.109	0.071	0.105	0.146
skewness	-0.951	-1.641	0.651	-2.464	-2.495	1.290
kurtosis	6.999	7.700	3.217	16.175	10.331	4.152

The conditional moments of the quarterly log real returns.

$p_0-d_0 \backslash$ Inflation	0	0.04	0.10
2.8	-0.077(92.62)	-0.087(91.70)	-0.107(89.89)
3.2	-0.010(90.53)	-0.137(87.18)	-0.254(77.54)
3.6	-0.159(85.32)	-0.284(75.30)	-0.400(67.01)
4.0	-0.332(71.47)	-0.432(64.92)	-0.544(58.24)

The 1% VaR of the quarterly a) log real returns
b) total return (in parenthesis).

p_0-d_0	2.8	2.8	2.8	3.2	3.2	3.2
Inflation	0	0.04	0.10	0	0.04	0.10
mean	0.688	0.652	0.619	0.510	0.464	0.393
st. deviation	0.203	0.198	0.191	0.253	0.255	0.257
skewness	0.217	0.269	0.235	0.115	0.164	0.272
kurtosis	3.134	3.120	3.130	2.768	2.740	2.827

The conditional moments of the five-year log real returns.

p_0-d_0	3.6	3.6	3.6	4.0	4.0	4.0
Inflation	0	0.04	0.10	0	0.04	0.10
mean	0.353	0.291	0.182	0.219	0.134	-0.006
st. deviation	0.338	0.344	0.352	0.426	0.444	0.463
skewness	-0.111	-0.061	0.112	-0.288	-0.190	0.024
kurtosis	2.559	2.515	2.388	2.480	2.352	2.146

The conditional moments of the five-year log real returns.

$p_0-d_0 \backslash$ Inflation	0	0.04	0.10
2.8	0.234(126.36)	0.224(125.11)	0.203(122.49)
3.2	-0.036(96.50)	-0.075(92.81)	-0.124(88.30)
3.6	-0.406(66.60)	-0.458(63.26)	-0.526(59.12)
4.0	-0.770(46.28)	-0.834(43.18)	-0.905(40.44)

The 1% VaR of the five-year log real a) log real returns
b) total return (in parenthesis).

$p_0 - d_0$	2.8	2.8	2.8	3.2	3.2	3.2
Inflation	0	0.04	0.10	0	0.04	0.10
mean	1.735	1.713	1.668	1.261	1.232	1.189
st. deviation	0.435	0.438	0.436	0.475	0.468	0.462
skewness	0.420	0.461	0.465	0.681	0.626	0.628
kurtosis	3.192	3.296	3.406	3.688	3.652	3.630

The conditional moments of the twenty-year log real returns.

$p_0 - d_0$	3.6	3.6	3.6	4.0	4.0	4.0
Inflation	0	0.04	0.10	0	0.04	0.10
mean	0.810	0.779	0.748	0.398	0.358	0.325
st. deviation	0.529	0.508	0.500	0.587	0.570	0.546
skewness	0.866	0.772	0.825	1.010	1.109	1.011
kurtosis	4.109	3.833	4.061	4.444	4.762	4.577

The conditional moments of the twenty-year log real returns.

$p_0 - d_0 \backslash$ Inflation	0	0.04	0.10
2.8	0.866(237.77)	0.826(228.46)	0.780(218.04)
3.2	0.364(143.97)	0.303(135.32)	0.294(134.16)
3.6	-0.132(87.66)	-0.132(87.62)	-0.158(85.41)
4.0	-0.613(54.18)	-0.591(55.39)	-0.610(54.32)

The 1% VaR of the twenty-year log real a) log real returns
b) total return (in parenthesis).