Bachelor thesis

Kees ter Brugge

November 30, 2011

1 introduction and abstract

2 Stock price models

An often used model to valuate assets is a simplified version of the model Lucas (1978) suggests. In this model, there are a large number of infinitely lived and identical agents. We assume these agents to be risk neutral. This gives us the following stochastic difference equation for equilibrium prices

$$P_t = e^{-r} E_t(D_t + P_{t+1}) \tag{1}$$

where P_t is the real stock price at the beginning of period t, D_t are the real dividends per share paid over period t and E_t is the mathematical expectation conditioned on information available at the beginning of period t. e^{-r} is the discount rate, where r is the constant, continuous real rate of interest.

Equation (1) has multiple solutions. A particular one is the present-value solution P_t^{pv}, which is obtained by iterating and using the law of iterated expectation:

$$
\begin{aligned}
P_t^{pv} &= e^{-r} E_t(D_t) + e^{-r} E_t(P_{t+1}) = e^{-r} E_t(D_t) + e^{-r} E_t(e^{-r} E_{t+1}(D_{t+1} + P_{t+2})) \\
&= e^{-r} E_t(D_t) + e^{-2r} E_t(D_{t+1}) + e^{-2r} E_t(e^{-r} E_{t+2}(D_{t+2} + P_{t+3})) \\
&= e^{-r} E_t(D_t) + e^{-2r} E_t(D_{t+1}) + e^{-3r} E_t(D_{t+2}) + \cdots \\
&= \sum_{s=0}^{\infty} e^{-(s+1)r} E_t(D_{t+s})
\end{aligned}
\tag{2}
$$

if we assume

$$\lim_{s \to \infty} e^{-rs} E_t(P_s) = 0 \tag{3}$$

Condition (3) is known as the transversality condition and can be interpreted as assuming that the present value of an asset at time infinity is zero. This rules out the possibility of infinitely lived agents holding on to their stocks for infinity. The present-value solution equates a stock's price to its discounted expected future payments. We assume the continuously compounded growth rate of expected dividends to be smaller than r, so that the present-value solution does not diverge and therefore always exists.

Another group of solutions to (1) is obtained by adding a rational bubble component

$$P_t = P_t^{pv} + B_t \tag{4}$$

where

$$B_t = e^{-r} E_t(B_{t+1}) \tag{5}$$

These solutions violate condition (3) unless the bubble component is equal to zero. Condition (5) excludes arbitrage oppurtunities.

2.1 Intrinsic bubbles

Froot and Obstfeld (1991) claim that the component of prices left unexplained by the present-value model is highly positively correlated with dividends. They suggest an 'intrinsic' bubble that only depends on fundamentals and not on extraneous factors. To find the function of this

3

bubble, assumptions about the stochastic process for dividends D_t have to be made. A common assumption is that log dividends d_t are a random walk [1]with constant drift μ

$$d_{t+1} = \mu + d_t + \xi_{t+1} \tag{6}$$

where ξ_{t+1} is a normal random variable with conditional mean zero and variance σ^2. If we additionally assume D_t to be known at the beginning of period t and

$$
\begin{aligned}
P_t^{pv} &= \sum_{s=0}^{\infty} e^{-(s+1)r} E_t(D_{t+s}) = \sum_{s=0}^{\infty} e^{-(s+1)r} E_t(e^{d_{t+s}}) = \sum_{s=0}^{\infty} e^{-(s+1)r} E_t(e^{d_t + s\mu + \sum_{i=1}^{s} \xi_{t+i}}) \\
&= D_t e^{-r} \sum_{s=0}^{\infty} e^{-sr + s\mu + \frac{s\sigma^2}{2}} = D_t e^{-r} \sum_{s=0}^{\infty} \left(e^{-r + \mu + \frac{\sigma^2}{2}} \right)^s = D_t e^{-r} \frac{1}{1 - e^{-r + \mu + \frac{\sigma^2}{2}}} \\
&= D_t \frac{1}{e^r - e^{\mu + \frac{\sigma^2}{2}}} = D_t \kappa
\end{aligned}
\tag{7}
$$

since $E_t(e^{\xi_{t+1}}) = e^{\frac{\sigma^2}{2}}$. κ represents the inverse of the required rate of return on stocks less the expected rate of growth of dividends and is equal to the ratio of the present-value price and dividends. It is defined as follows

$$\kappa = \frac{1}{e^r - e^{\mu + \frac{\sigma^2}{2}}} \tag{8}$$

The sum in (7)converges because we have assumed $r > \mu + \frac{\sigma^2}{2}$ to ensure the existence of the present-value solution.

The bubble proposed by Froot and Obstfeld is a nonlinear function of current dividends

$$B(D_t) = cD_t^{\lambda} \tag{9}$$

which satisfies condition (5) if the following condition holds

$$\lambda = \frac{\sqrt{\mu^2 + 2r\sigma^2} - \mu}{\sigma^2} \tag{10}$$

because then

$$
\begin{aligned}
e^{-r} E_t(B(D_{t+1})) &= e^{-r} E_t(cD_{t+1}^{\lambda}) = e^{-r} E_t(ce^{d_{t+1}\lambda}) = e^{-r} E_t(ce^{(\mu + d_t + \xi_{t+1})\lambda}) \\
&= ce^{d_t \lambda} e^{-r + \lambda\mu} E_t(e^{\xi_{t+1}\lambda} = cD_t^{\lambda} e^{-r + \lambda\mu + \frac{\lambda^2 \sigma^2}{2}} = cD_t^{\lambda} = B(D_t)
\end{aligned}
$$

We choose λ to be the positive root of the equation, because we want the size of the bubble to go to zero as dividends go to zero. This implies $\lambda > 1$. c is a nonnegative constant, because stock prices cannot be negative. It would violate free disposability. Now we can rewrite (4) so that it includes an intrinsic bubble, obtaining our intrinsic bubble model.

$$P_t = P_t^{pv} + B(D_t) = \kappa D_t + cD_t^{\lambda} \tag{11}$$

When λ is near to one this model can suffer from colinearity. Therefore it is better to use the price-dividend ratio model

$$\frac{P_t}{D_t} = \kappa + cD_t^{\lambda - 1} \tag{12}$$

While the present-value solution suggests a constant price-dividend ratio κ, the inclusion of an intrinsic bubble implies that the ratio is a function of dividends.

---

[1]Froot and Obstfeld (1991) argue that a random walk is a plausible approximation to the mechanism the market uses to forecast dividends (see appendix A, p. 1209)

2.2 Nonlinear least squares data fitting

If we want to estimate coefficients of a model with a nonlinear specification, we use the method of nonlinear least squares (NLS) estimation. We consider the nonlinear specification

$$y = f(\mathbf{x}; \theta) + e(\theta) \tag{13}$$

where y is the dependent variable, f a nonlinear function, $\mathbf{x} \in \mathbb{R}^k$ a vector of explanatory variables, $\theta \in \mathbb{R}^l$ a vector of parameters and e the specification error. Given T observation of $(y, \mathbf{x})$, let

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_T \end{bmatrix}, f(\mathbf{X}; \theta) = \begin{bmatrix} f(\mathbf{x_1}; \theta) \\ f(\mathbf{x_2}; \theta) \\ \vdots \\ f(\mathbf{x_T}; \theta) \end{bmatrix}, \mathbf{e}(\theta) = \begin{bmatrix} e_1(\theta) \\ e_2(\theta) \\ \vdots \\ e_T(\theta) \end{bmatrix} \tag{14}$$

We want to find the parameter θ for our function that best fits the data $(\mathbf{y}, \mathbf{X})$. How well our choice of parameters is, is measured by the following formula

$$S(\theta) = [y - f(\mathbf{X}; \theta]'[y - f(\mathbf{X}; \theta] \tag{15}$$

We want to minimize $S(\theta)$ with respect to θ. A solution $\theta^* \in \mathbb{R}^l$ to this problem must satisfy the first and second order condition of the minimization problem. The first order condition is that the the gradient of (13) evaluated at θ^* must be equal to zero

$$\frac{dS(\theta)}{d\theta}(\theta^*) = -2\frac{df(\mathbf{X}; \theta)}{d\theta}(\theta^*)[y - f(\mathbf{X}; \theta^*] = \mathbf{0} \tag{16}$$

where

$$\frac{df(\mathbf{X}; \theta)}{d\theta} = \begin{bmatrix} \frac{df(\mathbf{x_1}; \theta)}{d\theta} & \frac{df(\mathbf{x_2}; \theta)}{d\theta} & \cdots & \frac{df(\mathbf{x_T}; \theta)}{d\theta} \end{bmatrix}$$

Furthermore, to ensure that $S(\theta^*)$ is a minimum and not a maximum, the second order condition of a minimization problem has to be satisfied. The hessian matrix of (13), denoted below, evaluated at point θ^* must be a positive definite matrix.

$$\frac{d^2 S(\theta)}{d\theta d\theta'} = -2\frac{d^2 f(\mathbf{X}; \theta)}{d\theta d\theta'} \cdot [y - f(\mathbf{X}; \theta] + 2\frac{df(\mathbf{X}; \theta)}{d\theta} \cdot \frac{df(\mathbf{X}; \theta)}{d\theta'} \tag{17}$$

The problem with f being nonlinear is that there can be multiple $\theta \in \mathbb{R}$ that satisfy these two conditions. If f were linear, (14) would be a system with l equations and l unknowns. This would give us a unique solution. Furthermore, $\frac{d^2 f(\mathbf{X}; \theta)}{d\theta d\theta'}$ would be equal to zero and thus $\frac{d^2 S(\theta)}{d\theta d\theta'} = 2\frac{df(\mathbf{X}; \theta)}{d\theta} \cdot \frac{df(\mathbf{X}; \theta)}{d\theta'}$. This is a sufficient condition for positive definitness. There is no garantee that there exists a unique solution to this problem if f is nonlinear. This means (13) can have multiple local minima. We cannot solve this problem analytically, but we can with an iterative process.

There are a lot of iterative procedures that can lead to a fairly good approximation of the optimal parameter choice. An often used method is the Newton-Raphson algorithm. It uses the second order Taylor expansion of (13) around some chosen initial value θ_0

$$S(\theta) \approx S(\theta_0) + (\frac{dS(\theta)}{d\theta}(\theta_0))'(\theta - \theta_0) + \frac{1}{2}(\theta - \theta_0)'\frac{d^2 S(\theta)}{d\theta d\theta'}(\theta_0)(\theta - \theta_0) \tag{18}$$

The first order condition of this expansion with respect to (θ) is

$$\frac{dS(\theta)}{d\theta}(\theta_0) + \frac{d^2S(\theta)}{d\theta d\theta'}(\theta_0)(\theta - \theta_0) = 0$$

This can be rewritten as

$$\theta = \theta_0 - (\frac{d^2S(\theta)}{d\theta d\theta'}(\theta_0))^{-1}\frac{dS(\theta)}{d\theta}(\theta_0) \tag{19}$$

So we now know the value θ that minimizes the taylor expansion . If we now replace θ_0 and θ with θ_i and θ_{i+1} respectively, we get the following algorithm

$$\theta_{i+1} = \theta_i - (\frac{d^2S(\theta)}{d\theta d\theta'}(\theta_i))^{-1}\frac{dS(\theta)}{d\theta}(\theta_i) \tag{20}$$

The $(i+1)^{th}$ iterated value θ_{i+1} is obtained from the value of the previous iteration. The downside of this algorithm is that for every iteration the hessian matrix must be positive definite, otherwise it might not be invertible of point in the wrong direction. From (16)(18) we can deduce that

$$S(\theta_{i+1}) - S(\theta_i) \approx -\frac{1}{2}(\frac{dS(\theta)}{d\theta}(\theta_i))'\frac{d^2S(\theta)}{d\theta d\theta'}(\theta_i)\frac{dS(\theta)}{d\theta}(\theta_i) \tag{21}$$

Succesive iterations of S only become smaller if the hessian matrix is positive definite.
Sometimes a step length s is added to (13)

$$\theta_{i+1} = \theta_i - s_i(\frac{d^2S(\theta)}{d\theta d\theta'}(\theta_i))^{-1}\frac{dS(\theta)}{d\theta}(\theta_i) \tag{22}$$

The step length could be determined by minimizing $S(\theta_{i+1})$ with respect to s_i.
An iterative algorithm needs a beginning and an end. The beginning may be some initial value set of values. To minimize the risk of no iteration converging to the global minimum we can generate a lot of initial values from some distribution and choose the initial value that gives the best result. An algorithm stops when some sort of convergence criteria is met. For example, when

$$(\theta i + 1 - \theta_i)'(\theta i + 1 - \theta_i) < c \tag{23}$$

where c is some small positive pre-determined number.
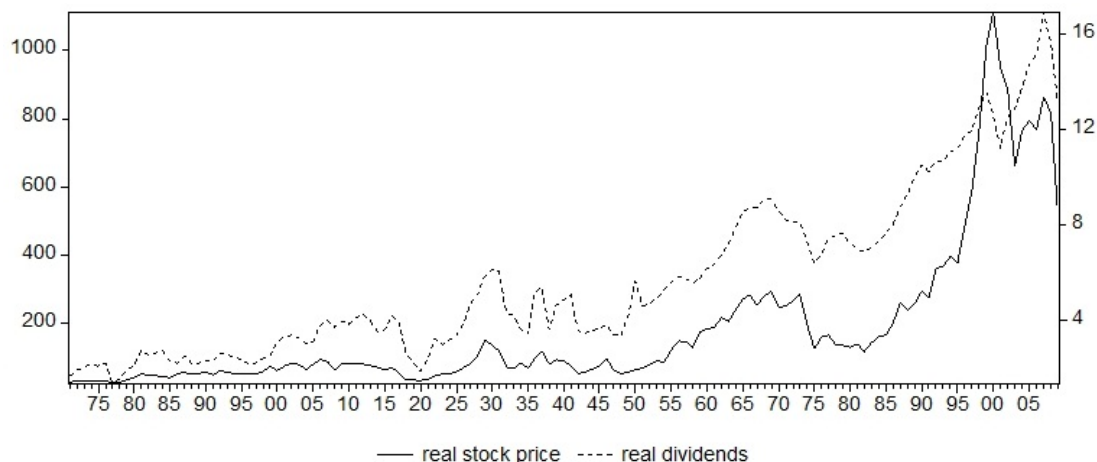
3   Application to Standard &Poor's 500

Figure 1: Plot of real dividends D_t and real prices P_t over 1871 - 2009. The left axis scales the prices and the right axis scales the dividends.

To examine the validity of the models we defined before, we use data on the U.S. stock market from 1871 untill 2009. The used stockprice index is the January values of the Standard and Poor's Composite annual stock price index. The time series is an updated version of the often used data from Shiller (1989). Shiller used data till 1988, which gives us a oppurtunity to see if recent events, such as the dot-com bubble, change the results of the analysis of this model. Since data on January values of the dividends index are not available, we use annual averages for the calendar year, following Froot and Obstfeld (1991). They found that this does not affect results statistically[2]. Stock prices and dividends are deflated by the Produces Price Index of 1982. We use data till 2006 to estimate the coefficients, so that we can perform out of sample forecasts to compare with the last three realizations. The Standard and Poor's composite index is considered as one of the best benchmarks of U.S. stock market performance. It is a market value weighted index that includes 500 stocks that trade on either the New York Stock Exchange and the NASDAQ. The selection of stocks is mostly based on market size, liquidity and sector. The S&P monthly index series was created in 1957 and is extended back to January 1871 by Cowles.

In figure (1) we see that dividends and prices seem to have an exponential trend. To assess the nature of these processes, we want to produce stationary series. In figure (2) the log transformation of both series is shown. The series seem to have a linear trend now. The series move together, which indicates correlation of the series. Furthermore, their trend also seems equally steep, which suggests that these processes might be cointegrated.

In figure (3) the first difference of the log transformation of prices and dividends is plotted. The series appear to be stationary. This suggests d_t and p_t are integrated of order one. It is common to assume that a series is integrated of order one, when its log transformation is integrated of order one.

If we look at figure (3) we see that after 1955 Δd_t becomes much less volatile. Since dividends is the explanatory variable of our model it is a good idea to analyse a subsample with 1955 as start

---

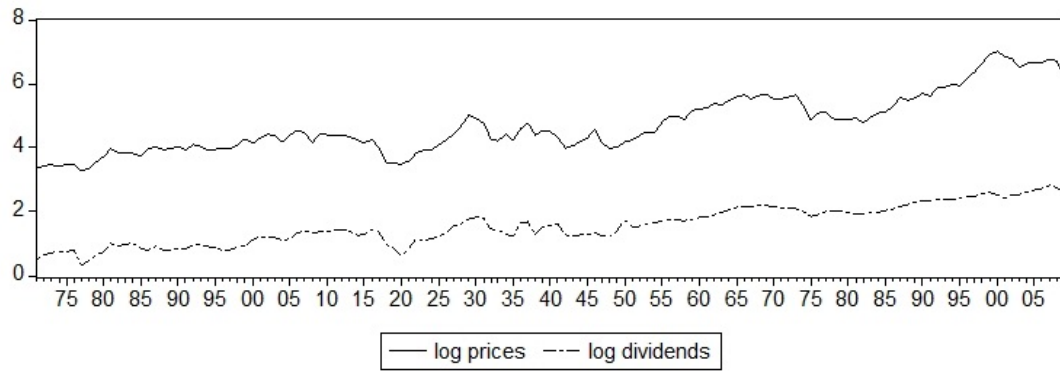[2]see Froot and Obstfeld (1991) footnote 18, p. 1198

Figure 2: Plot of log prices and log dividends over 1871 - 2009

Figure 3: Plot of Δd_t and Δp_t over 1872 - 2009
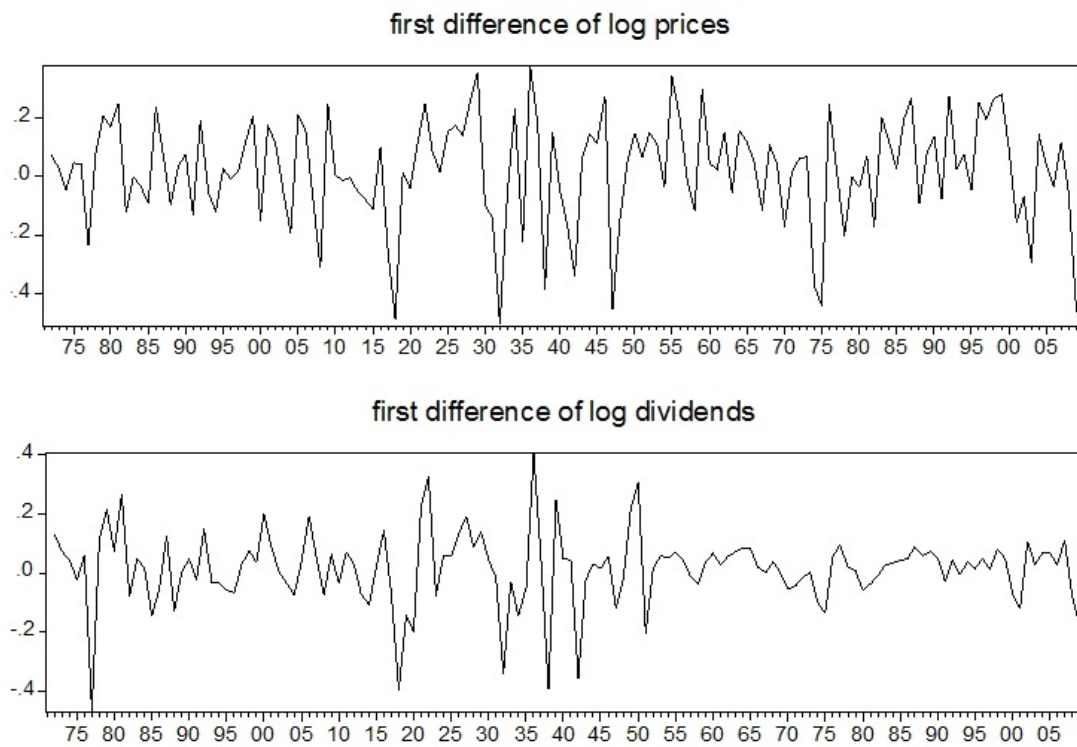
date in addition to the whole sample. This could give us better estimates, which in turn leads to a better assessment of the statistical significance of the model and more accurate forecasts.

To be able to execute some tests on the intrinsic bubble model we have to assume that the error term η_t we add to (10) to get

$$\frac{P_t}{D_t} = c_0 + cD_t^{\lambda-1} + \eta_t \tag{24}$$

is well-behaved, meaning that η_t is statistically independent of dividends at all leads and lags and has unconditional mean zero[3]. Otherwise the statistical tools we use do not lead to unbiased estimates. η_t can be interpreted as any shock to the price-dividend ratio that has no predictive power with respect to future dividens, such as fads.

Our null hypothesis is that there is no bubble, that the present-value model is true, so that the level of dividends has no impact one the price-dividend ratio and therefore $c_0 = \kappa$ and $c = 0$. Our alternative hypothesis is that there is a bubble involved, a relation between dividends and the price-dividend ratio, implying $c_0 = \kappa$ and $c > 0$.

To obtain the estimates of κ and λ implied by the intrinsic bubble model, which we from now on denote as $\overline{\kappa}$ and $\overline{\lambda}$ respectively, we first have to obtain the point estimates of μ, σ and r. The estimates are reported in table (??). We obtain the continuous interest rate as follows $\overline{r} = log(\frac{\sum_{t=m}^{2005} \frac{P_{t+1}+D_t}{P_t}}{2005-m})$, where m is 1871 for the estimate based one the whole sample and 1955 based on the subsample. In figure (??) we clearly see that log dividends has a positive trend. Under the assumption that log dividends has no trend, the p-value of the estimate of μ is equal to 0.14 for the whole sample and 0.01 for the subsample. $\overline{\kappa}$ and $\overline{\lambda}$ are obtained by using the point estimates in equation (??) and (??) respectively.

parameter	whole sample 1871 - 2006	subsample 1955 - 2006
r	0.0904	0.1009
μ	0.0162	0.0204
σ	0.1257	0.0544
κ	14.2411	11.8984
λ	2.5094	3.8634

Table 1: Estimates of parameters for the two samples

The model estimates the coefficient λ that indicates the explosiveness of the bubble much higher for the subsample than for the whole sample.

3.1 Present-value model

The present-value model (2) predicts that the following regression

$$\mathrm{P}_t = \beta_0 + \beta D_t + v_t \tag{25}$$

gives us an estimate of β that is close to $\overline{\kappa}$. When we perform an regression on the log transformations of prices and dividends instead

$$\mathrm{p}_t = \beta_0 + \beta d_t + v_t \tag{26}$$

---

[3] Froot and Obstfeld (1991) also made these assumptions. For a justification see footnote 16, p. 1198

we expect to find an estimate of β near one, since the linear relation between prices and dividends implied by the present-value model has an elasticity equal to one.

If (22) is true, the estimate of β we expect to find from (??) is much bigger than $\overline{\kappa}$, since the derivative of prices with respect to dividends predicted by the intrinsic model adds a positive nonlinear component to the derivative predicted by the present-value model

$$\frac{dP_t}{dD_t} = c_0 + \lambda c D_t^{\lambda-1} > c_0 = \frac{dP_t^{pv}}{dD_t}$$

Furthermore, under (22) implies a nonlinear relation between P_t and D_t and since λ is bigger than one, this relation has an explosive nature. The estimate of β that follows from (??) the log transformation of the variables should be bigger than one. Prices would appear to overreact to changes in dividends.

regression equation	β whole sample	β subsample
$P_t = \beta_0 + \beta D_t + v_t$	56.0023	89.2800
$p_t = \beta_0 + \beta d_t + v_t$	1.4475	2.1273

Table 2: OLS regression of real prices on a constant and real dividends.

In table (1) we see that the estimates of β obtained by (??) are much higher for both samples than the values in table (??) predicted by the present-value model. Furthermore, the coefficients obtained from (??) are also much higher than the value predicted by the present-value model, one.

If two processes appear to have a long-run relationship, it is usefull to test if they are cointegrated. Finding that two processes are cointegrated increases forecasting powers. If two time series are integrated of some order and a linear combination of these processes yields a time series that is integrated of a lower order, these time series are cointegrated. Prices and dividends are generally thought of as processes which are integrated of order one. If they are, then under the present-value model we would expect to find that a linear function, with cointegrating coefficient κ, of prices and dividends gives us a stationary time series. This linear function is called the spread by Campbell and Shiller's (1987). Equivalently, if we use the log prices and log dividends instead, we expect to find cointegration with cointegrating coefficient one. The price-dividend ratio is also expected to be stationary. Under (22) we expect the null hypothesises of nonstationarity of these various measures of the price-dividend ratio to not be rejected.

As we can see in table (2) seven out of eight tests do not reject the hypothesis that prices, dividends or their log transformations contain a unit root and are therefore nonstationary. Furthermore, for none of the measures of the price-dividend ratio was the null hypothesis of a unit root rejected. They were not able to produce a series that had a lower order of integration.

There are a few problems with the present-value model. It can not explain the high price-dividend ratios given the rate of return $\overline{r}$, prices are much more sensitive to dividends than predicted by the present-value model. When we test for a linear long-run relationship between prices and dividends, which under the present-value model should exist, we do not find evidence to support this relationship.

Furthermore, the result of the regression of log prices on log dividends can be interpreted as evidence that the relation between prices and dividends is a nonlinear one.

10

Variable X_t	β_1 with trend	β_1 without trend
P_t	-0.0573	-0.0214
	(-2.27)	(-1.27)
p_t	-0.0871	-0.0176
	(-2.58)	(-1.09)
D_t	-0.0757	-0.0043
	(-2.34)	(-0.06)
d_t	-0.2084	-0.0233
	(-4.10)*	(-1.01)
$P_t - \overline{\kappa} D_t$	-0.0619	-0.0318
	(-2.40)	(-1.65)
$\frac{P_t}{D_t}$	-0.0802	-0.0488
	(-2.65)	(-1.80)
$p_t - d_t$	-0.1027	-0.0590
	(-2.72)	(-1.93)

Table 3: The reported values are the estimates of β_1 in the regression $\Delta x_{t+1} = \beta_0 + \beta_1 x_t + \beta_2 t + v_{t+1}$ where β_2 is equal to zero in the regression without trend. The sample is 1871 - 2009. In parentheses under the estimates the Phillips-Perron test statistics are reported, where * denotes statistical significance at the 1 % level.

3.2 Intrinsic bubble model

To see if a nonlinear function of dividends describes the variation in price-dividend ratios more accurately than the present-value model, we estimate (22). To test for absence of an intrinsic bubble we first have to know the distribution of the t statistic under the null hypothesis. So, additional assumptions about the error term η_t have to be made. Besides η_t being identically distributed, independent of dividends at all leads and lags and having unconditional mean zero, we also assume that they have a normal distribution. The nonnegativity of price-dividend ratios is not likely to give any problems since the mean of the price-dividend ratios is more than three times the standard error of η_t. Even though (22) contains the explosive regressor $D_t^{\lambda-1}$, the standard t statistic approximates a normal distribution under these assumptions[4].
Since we have not assumed that η_t is distributed independently, the residuals may be serially correlated. This can lead to incorrect estimations of the standard errors of coefficients. To counteract this, we correct the residuals using Newey and West's (1987) covariance matrix estimator.

In table (3) the results of the estimation of (22) are reported. The model is estimated with and without the restriction that λ and κ are equal to their point estimates obtained earlier from the log dividend process and the mean return on stock. When we restrict $\lambda = \overline{\lambda}$, the bubble coefficient $\widehat{c}$ is statistically very significant. When we do not restrict the exponent λ, $\widehat{c}$ gets smaller and $\widehat{\lambda}$ is estimated larger than its point estimate. The bubble coefficient becomes statistically insignificant. To test whether the nonlinear component is statistically insignificant, We use a F-test. We test the restricted no bubble model where $c = 0$ and $\lambda = \widehat{\lambda}$ versus the alternative bubble model where $c \neq 0$ and $\lambda = \widehat{\lambda}$. The no bubble model is rejected at the 1% level[5]. We can use this test, since the no bubble model is nested in the bubble model. The

---

[4]see Froot and Obstfeld (1991), Appendix B

[5]The F statistic is defined as follows: $F = \frac{RSS_1 - RSS_2}{p_2 - p_1} \cdot \frac{n - p_2}{RSS_2}$ where RSS_1, p_1, RSS_2 and p_2 denote the residual sum of squares and the amount of parameters of model 1 and 2 respectively. n is the sample size.

estimated model	$\widehat{c_0}$	$\widehat{c}$	$\widehat{\lambda}$	F stat.$(c=0)$	R^2	d.f.
$\frac{P_t}{D_t} = c_0 + cD_t^{\lambda-1} + \eta_t$	16.8904 (8.06)**	0.1824 (0.69)	3.0152 (5.43)**	131.72 **	0.66	136
$\frac{P_t}{D_t} = c_0 + cD_t^{\overline{\lambda}-1} + \eta_t$	14.2736 (9.58)**	0.6863 (5.25)**			0.65	137
$\frac{P_t}{D_t} = \overline{\kappa} + cD_t^{\lambda-1} + \eta_t$		0.4284 (1.93)*	2.7120 (7.19)**	39113.20 **	0.65	137
$\frac{P_t}{D_t} = \overline{\kappa} + cD_t^{\overline{\lambda}-1} + \eta_t$		0.6770 (7.14)**			0.65	138

Table 4: nonlinear OLS regression. Sample is 1871 - 2009. Standard errors are corrected by the Newey-West covariance matrix, allowing for fourth-order serial correlation and conditional heteroskedasticity. In parentheses under the estimates the standart t test statistics are reported, where *(**) denotes statistical significance at the 10(1)% level.

estimates of κ are almost equal to $\overline{\kappa}$. If we apply the restriction $\kappa = \overline{\kappa}$, $\widehat{\lambda}$ is almost equal to $\overline{\lambda}$. Another way to check the validity of the intrinsic bubble model is by testing for cointegration. Under (22) there should exist a long-run relationship between the variation in the price-dividend ratio left unexplained by the present-value model and the nonlinearly transformed dividend series. To assess if such a relation exists, we have to perform a nonlinear cointegration test. We use the method desribed by Granger and Hallman (1991). We test if a linear combination of $\frac{P_t}{D_t} - \overline{\kappa}$ and the nonlinearly transformated dividends series $D_t^{\lambda-1}$ can produce a series with a lower order of integration. The result are reported in table (4). The cointegration coefficient are obtained from table (3)

Variable X_t	β_1 with trend	β_1 without trend
$\frac{P_t}{D_t} - \overline{\kappa} - 0.4284D_t^{1.7120}$	-0.1620 (-3.94)*	-0.1556 (-3.84)**
$\frac{P_t}{D_t} - \overline{\kappa} - 0.6770D_t^{\overline{\lambda}-1}$	-0.1536 (-3.82)*	-0.1501 (-3.78)**

Table 5: The reported values are the estimates of β_1 in the regression $\Delta x_{t+1} = \beta_0 + \beta_1 x_t + \beta_2 t + v_{t+1}$ where β_2 is equal to zero in the regression without trend. The sample is 1871 - 2009. In parentheses under the estimates the Phillips-Perron test statistics are reported, where *(**) denotes statistical significance at the 5(1) % level.

The results of a Phillips-Perron unit root test on the price-dividend ratio and on dividends are reported in table 2. For both these series, the unit root hypothesis cannot be rejected. For the sake of robustness we test with and without time trend, with and without restricting λ for a unit root in variable X_t. In table (4) we see that in all four tests the unit root hypothesis is rejected. This means that the price-dividend ratio and the nonlinear transformation D_t^{λ} are cointegrated.

The estimation of model (22) confirmed the explosive nature and the statistical significance of the bubble component in stock prices. The bubble component absorbs almost all the oversensitivity of prices to dividends, because when we include a nonlinear regressor in the regression, the estimates of κ are close to its point estimate. This means model (22) does a good job of explaining the price-dividend ratio given the rate of return $\overline{r}$. This can also be seen in figure (4). The positive correlation between the price-dividend ratio and the nonlinear bubble component implies that as dividends grow, the overvaluation of prices relative to the present-value price, grows exponentially. This is hardly noticeable when dividends are low, but when dividends increase, prices explode. In figure (5) the size of the bubble component in prices can be estimated by comparing the stock price with the present-value price. Figure (5) and (4) show that after the 1950's the economic relevance of the bubble component in stock prices and price-dividend ratios has grown. The nonlinear cointegration test showed us (22) produces stationary residuals, which means that in the long-run we have found a fairly good approximation of the underlying data generating process.
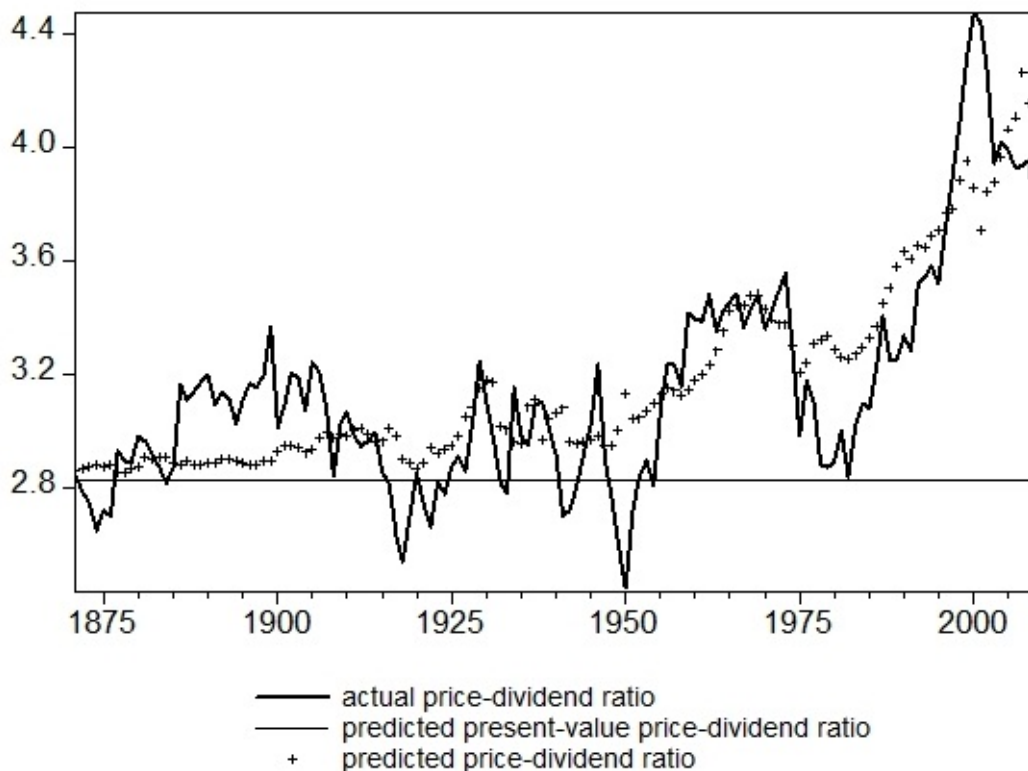

Figure 4: This figure shows the log transformation of the following ratios: the actual price-dividend ratio $\frac{P_t}{D_t}$, the predicted price-dividend ratio under the intrinsic bubble model $\frac{\widehat{P}_t}{D_t} = \widehat{c}_0 + \widehat{c} D_t^{\widehat{\lambda}-1}$ and its predicted present-value price-dividend ratio $\frac{\widehat{P}_t^{pv}}{D_t} = \widehat{c}_0$. The coefficients that are used are from the unrestricted regression found in table (3).
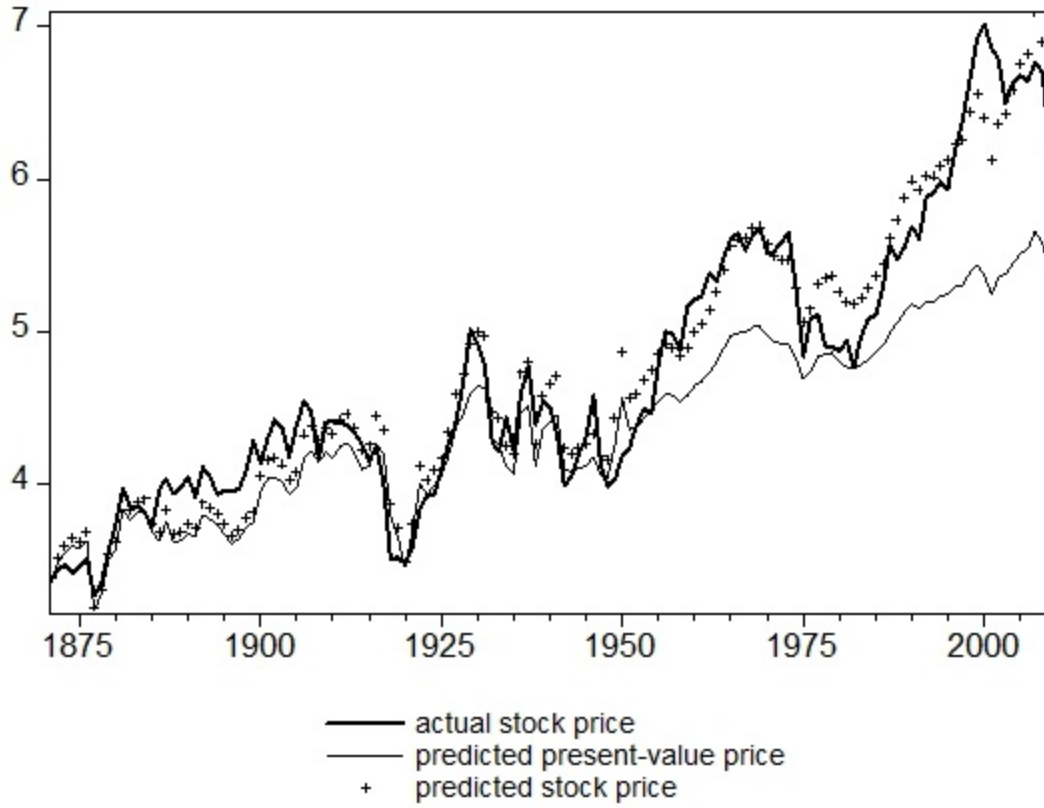
13

Figure 5: This figure shows the log transformation of the following prices: the actual stock price P_t, the predicted stock price under the intrinsic bubble model $\widehat{P}_t = \widehat{c}_0 D_t + \widehat{c} D_t^{\widehat{\lambda}}$ and its predicted present-value stock price $\widehat{P}_t^{pv} = \widehat{c}_0 D_t$. The coefficients that are use are from the unrestricted regression found in table (3).

4   Conclusion

5   References