Another look to the Price-Dividend ratio: the Markov-switching approach

Juan-Miguel Londoño, Marta Regúlez, Jesús Vázquez

Universidad del País Vasco, Bilbao 48015, Spain

Preliminary draft. This version October 2007

Abstract

Since the mispricing effect from a rational valuation model has to be stationary and insignificant; the price-dividend ratio in the present value model must be mean-reverting; however, relevant market episodes in the last part of the past century seem to have shown non-conclusive evidence about this fact. In this paper, we analyze the stationarity of this ratio in the context of a Markov-switching model à la Hamilton. This particular specification robustly supports the asymmetrical mean-reversion hypothesis of the ratio and identify two important episodes: the post World War II and the 90’s "boom". A model with 3 states reports the best state identification and reveals that only the first part of the 90’s "boom" (1985-1995) and the post war period are near-nonstationary states, but the las part of the 90´s "boom" is entirely characterized by a new stationary state.

JEL Classification: C32, G12

Keywords: Markov processes, Regime-Switching, Stationarity, Valuation ratios, price-dividend ratio.

*Corresponding author. E-mail juanchoyoarc@hotmail.com
†Financial support from Ministerio de Ciencia y Tecnología (Spain) through project SEJ2004-04811.
1 INTRODUCTION

The last clear episode of persistent deviations of US stock prices from the prices predicted by traditional present value (PV) models starting by the end of the last century has motivated lots of research in the last years trying to explain whether this so-called by many authors "boom" or "run up in prices" has shown that traditional fundamentals like dividends can no longer explain stock prices. Traditional payout measures like dividends seem to be less important at least in the US since 1978 when 66.5% of the firms were paying cash dividends, but only 20.8% of the firms were paying them by 1999. Fama and French (2001) showed that regardless of their characteristics, firms have become less likely to pay dividends. In their paper, they analyze how there is an increasing number of firms with characteristics typical of firms that have never paid dividends: small size, low profitability and strong growth opportunities. Ordinary dividends might then be a bad measure of total payouts at least for the last two decades, specially since the enactment of the U.S. Securities and Exchange Commission (SEC) rule 10b-18 in 1982. This rule provided a legal safe harbor for firms repurchasing their shares, and did the repurchase of stocks a very important way of payout mainly because of their tax benefits or the presence of agency problems between regular shareholders and managers shareholders as mentioned by Boudokh, Michaely, Richardson and Roberts (2004).

Since Shiller’s (1981) seminal paper, some research works trying to explain the dynamic features of the relation between prices and dividends have included time varying discount factors, noise traders, and fads (Shiller, 1989); the introduction of an "intrinsic bubble" like in Froot and Obstfeld (1991), and in Drifill and Sola (1998), and the possibility of having structural breaks in the dividends motivated by the presence of instabilities in their processes, making room for regime-switching models, like in Cecchetti, Lam, and Mark (1990), Veronessi (1999), Timmermann (2001), Bonomo and Garcia (1994), Drifill and Sola (1998), Evans (1998) and Gutiérrez and Vázquez (2004) (GV hereafter) among others. Some recent studies have even analyzed the relation between stock prices and real activity like in Binswanger (2000, 2003) and have introduced alternative valuation ratios, like the price-output in Rangvid (2006). The motivation for this relation comes from the evidence that some macroeconomic variables may contain information about expected returns that could not be explained by dividends or earnings.

Since the PV model implies that the price-dividend (PD) ratio must have a mean-reverting behavior, so the unexpected shock of the model is stationary and in-

\footnote{Evans (1998) and GV focus their analysis in regime-switching models for the dividends process.}
significant, the first step in analyzing the performance of this ratio in valuing stocks should be a mean reversion analysis. Relative to this particular topic, previous studies have found non-conclusive evidence for the cointegration relation using different model specifications and samples in a linear context, like in Cochrane (1992, 2001), Lettau and Ludvigson (2005), among others. Another school of thought has suggested that the cointegration relation implied by the model does not necessarily mean that the reversion process exhibits a linear dynamics as in Gabriel et.al (2002), GV, Coakley and Fuertes (2006) (CF hereafter) and McMillan (2007). GV analyze the possibility of having alternative equilibria in the stock market depending on the dividend process and the feedback from prices to dividends, identifying different episodes for the market. CF seriously concentrate on the mean-reversion analysis and propose a two-regime framework for the PD and the price-earnings (PE) ratios. In this framework they define a bull and a bear regime that may show an asymmetrical speed of adjustment around the same long-run equilibrium. Their analysis is focused on a threshold autoregressive (TAR) model estimation of a reduced-form similar to the augmented Dickey-Fuller (ADF) regression form. McMillan (2007) estimates an equation similar to the one in CF, but they choose a threshold variable with an economical interpretation. In their paper, the non-linear dynamics is a consequence of transaction costs and the presence of noise traders in the market.

The aim of this paper is to analyze if the divergence between prices and dividends can be explained considering that the ratio may show an asymmetrical speed of adjustment among different episodes around the same long-run equilibrium, without necessarily breaking the stationarity hypothesis implied by the PV model. We also analyze whether the last "boom" episode in the late 90's has the same characteristics as previous boom episodes when we assume that the long-run equilibrium must hold for the entire sample. We then estimate a reduced-form similar to the one proposed by CF, but without a priori giving any characteristics to the regimes that may have been present during the sample period, and allowing the data to identify the existence of switching regimes à la Hamilton (1989) characterizing the mean-reverting behavior of the PD ratio. We consider two samples for this ratio in order to analyze the stability of the stationarity analysis and the identification of states; one considers annual data from 1871 to 2004, called full sample, and another one considers only annual data previous to 1993 and it is called pre-1993 sample. We also analyze the robustness of the estimated results by considering two cases for the model, one is a 3-state model, and another one is a 2-state model.

The paper is organized as follows. Section 2 explains the methodology, and introduces the PV model and the mean-reversion analysis proposed. Section 3 presents the data and the empirical evidence including a preliminary stationarity analysis and the robustness analysis. Finally, section 4 summarizes the main conclusions.
2 METHODOLOGY

2.1 The present value model and the PD ratio

Following Campbell and Shiller (1988), the PV model or "dynamic Gordon model" for finding assets prices through fundamentals can be derived as follows: first, from the definition of continuously compounded one-period returns, we have that

\[ r_{t+1} = \ln(1 + R_{t+1}) = \ln(P_{t+1} + D_t) - \ln(P_t), \]

(1)

where in this case, \( P_t \) is the Standard and Poors (S&P) index price at the beginning of period \( t \) and \( D_t \) is the dividend accruing to the index paid out throughout period \( t \) (dividends paid sometime within period \( t \)). We can rewrite this log-return as:

\[ r_{t+1} = p_{t+1} - p_t + \ln[1 + e^{d_t - p_{t+1}}], \]

(2)

where lower case letters denote logged variables.

Now, using a first-order Taylor approximation of equation (2) around \( \bar{d} - \bar{p} \) (e.g. the average PD ratio), we have that

\[ r_{t+1} \approx k_d + \rho_d p_{t+1} + (1 - \rho_d)d_t - p_t, \]

(3)

where \( k_d = -\ln \rho_d - (1 - \rho_d) \ln(\frac{1}{\rho_d} - 1) \) and \( \rho_d = \frac{1}{[1+e^{d_t-p_t}]} \).

Finally, solving forward this term and taking expectations we have

\[ p_t = \frac{k_d}{(1 - \rho_d)} + E_t \left\{ \sum_{j=0}^{\infty} \rho_d^j[(1 - \rho_d)d_{t+j} - r_{t+1+j}] \right\} + \lim_{j \to \infty} E_t(p_d^j p_{t+j}). \]

(4)

If we want to write this equation in terms of the PD ratio, we can take the first term in the summatory and rewrite (4) as
\[ p_t - d_{t-1} = \frac{k_d}{(1 - \rho_d)} + E_t \left\{ \sum_{j=0}^{\infty} \rho_d^j [\Delta d_{t+j} - r_{t+1+j}] \right\} + \lim_{j \to \infty} E_t[\rho_d^j (p_{t+j} - d_{t-1+j})] \], \quad (5) \]

where the last term drops out under the transversality condition \( \lim_{j \to \infty} E_t[\rho_d^j (p_{t+j} - d_{t-1+j})] = 0 \). Finally, we have

\[ p_t - d_{t-1} = \frac{k_d}{(1 - \rho_d)} + E_t \left\{ \sum_{j=0}^{\infty} \rho_d^j [\Delta d_{t+j} - r_{t+1+j}] \right\} + u_{dt}, \quad (6) \]

where \( u_{dt} \) is a zero-mean stationary term, since the relation implied by the model between the future expected returns and the actual ratio is not deterministic, and the desired residual term should be stationary, reflecting the rational behavior of economic agents under this model.

### 2.2 Mean-reversion analysis in a N-state Markov context

In the PV model, testing if a valuation ratio is stationary implies that there exists a relation between future returns and the current ratio value so the actual index price will equal the expected PV of the future payouts whatever the proxy we use to measure them. For the PD ratio, we can intuitively think that if current stock prices are high in relation to current dividends (investors are willing to pay more or the stock is overpriced), dividends are expected to grow. The zero-mean stationary term \( u_{dt} \) can be seen as a mispricing term; and if agents are fully rational under this model, prices and fundamentals cannot always drift apart and the ratio will evidence a mean-reverting behavior.

The basic explanation for the mean-reverting behavior can also be extended to a non-linear framework, in particular, to a multiple state context with an asymmetrical speed of adjustment as explained, among others, by CF and previously discussed by Lee (1998). In this case, the speed of adjustment of the ratios can vary over time appearing very persistent in some periods and rapidly mean-reverting in other periods usually related to market episodes such as bear and bull markets; however, identifying periods with particular characteristics requires the definition of a state variable that tell us when the markets are in a bull or bear period for example. This
state variable is the threshold or transition variable that specify the regime-switching in this kind of models.²

To study the implications of the mean-reverting behavior of the ratio in equation (6), we propose a general non-linear framework for the dynamics of the ratio that will allow us to test stationarity including regime changes. This framework can be seen as a representation similar to the augmented Dickey-Fuller equation (ADF) given by

$$\Delta x_t = \alpha + \rho_{s_t}(x_{t-1} - \mu) + \sum_{j=1}^{k} \beta_j(\Delta x_{t-j}) + \varepsilon_t,$$

where $x_t$ is the PD ratio, $\alpha$ is a non-zero parameter, $\mu$ is the long-run equilibrium or attractor, and $\varepsilon_t \sim N(0, \sigma)$ is an iid shock.³⁴ In this context, instead of having the possibility of two predefined states as in CF, we have an unobserved regime $s_t$ that is the outcome of an unobserved N-state Markov chain with $s_t$ independent of $\varepsilon_t$. The basic difference between imposing a threshold variable as in TAR models and the N-state Markov context is that in the latter, we do not give any particular a priori characteristic to each state. Instead, we only impose the possibility of having N states and leave the data set available in each period to tell us the estimated parameters and the smoothed probability for each state.⁵ For each case we test some hypotheses related to the mean-reversion analysis:

²For example, in CF, the threshold variable is defined as

$$q_t(w, d) = w_1 \Delta x_{t-1} + \cdots + w_d \Delta x_{t-d},$$

where $w = (w_1, \ldots, w_d) > 0$ are predefined weights and $d$ are predefined lags. If $q_t > 0$, the stock market is in a bull episode with speed of adjustment $\rho^c$, and if $q_t < 0$ there is a bear episode with $\rho^b$. For a generalization on TAR models and their mean-reversion analysis see for example Tong (1993) and Enders and Granger (1998).

³For the optimal number of lags ($k$) in equation (7), we analyze the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) for the non-state dependant form of the equation. Previously knowing that these kind of models suffer from overparametrization, and that it is possible to find flat zones for the objective function for one or some of the parameters that can derive in local optima, it is convenient to keep the number of estimated parameters as low as possible. We find appropriate to keep only one lag ($k = 1$).

⁴In a 2-states model where $\sigma$ is also state-dependant, the empirical evidence suggests that we cannot reject the hypothesis that $\sigma_1 = \sigma_2$. Results for models where both $\rho$ and $\sigma$ are state-dependant are available upon request.

⁵Appendix A provides a more brief description of the Markov-switching model estimation algorithm.
**Hypothesis 1:** Symmetrical or asymmetrical mean reversion

$H_{01}$: all $\rho_i$ are equal.

$H_{a1}$: at least one $\rho_i$ is different from the others (asymmetrical mean reversion).

For this test, since we use a maximum-likelihood algorithm to estimate the parameters, a Likelihood-ratio (LR) test seems to be appropriate. The LR statistic in this case is distributed as a $\chi^2_{(1)}$ for the 2-state model and $\chi^2_{(2)}$ for the 3-state model.

**Hypothesis 2:** Individual state unit-root persistence or mean reversion

$H_{02}: \rho_i = 0; \rho_j \neq 0, i \neq j$

$H_{a2}: \rho_i \neq 0$;

### 3 EMPIRICAL EVIDENCE

#### 3.1 Data and preliminary analysis

This paper considers annual data for the S&P index price (January data), and annual data for the dividends accruing to this index in each year available at Robert Shiller’s web site. The PD ratio is calculated as $PD_t = p_t - d_{t-1}$. We use data for the period 1871-2004.\(^6\)

Table 1 shows a summary of descriptive statistics for the ratio for both samples. It includes the two most commonly used test for cointegration in a linear context: the Dickey-Fuller test (ADF) and the Phillips-Perron test (PP). Figure 1 shows the annual demeaned PD ratio for the full sample. In this Figure we can preliminary identify some relevant episodes. We can see for example a sustained increase starting

---

\(^6\)Even though the dividends and earnings data are related to payouts during each year considered, some of the literature contrasted here takes into account other frequency data, as in CF or McMillan (2007) who use monthly data.
around 1950, a local maximum around 1970, and then again, a considerable increase starting around 1990. The last increase clearly corresponds with the 90’s "boom". The ratio has a maximum value of 4.44 in year 2000, almost twice its minimum value. This maximum shows a sharp contrast with an average PD of 3.021 for the period 1871-1985. From the Unit-root tests values in the table, we can also conclude for the full sample that we cannot reject the null hypothesis of overall nonstationarity of the ratio, while for the pre-1993 sample, this hypothesis is rejected.

The information in Table 1 and Figure 1 basically summarize the non-conclusive evidence found in the relevant previous literature. It shows how difficult may result to conclude about the valuation ratios stationarity in a linear context specially after the divergence between prices and fundamentals found by the end of the last century.

Figures 2 and 3 show a preliminary stability analysis for the parameters $\alpha$ and the speed of adjustment $\rho$ in a linear context such as in a Dickey-Fuller-type equation for the ratio like $\Delta x_t = \alpha + \rho x_{(demeaned)} t-1 + u_t$, for samples starting in 1930 to the end of the full sample for moving cumulative windows and moving windows respectively. Both figures motivate the possibility that for the demeaned series, neither the estimated $\alpha$, nor the estimated speed of adjustment remain constant for different samples, and that the estimated parameters have shown different episodes. In Figure 2, we can see that as we include more recent data in the sample, the estimated $\rho$ shows important episodes of increase (getting closer to 0). At least two important episodes of sustained increase are remarkable, one starting around 1955 to around 1975 and another one starting around 1990 to the end of the full sample. In Figure 3, the possibility that both parameters are not stable is even more clear, specially when in this case, from 1965 to 1971 and then again starting in 1996 to the end of the sample, the 95% confidence intervals for $\rho$ includes the 0.

3.2 Evidence for a N-state model

Tables 2 and 3 show the estimated parameters for the full and the pre-1993 samples as well as their standard deviations, the LR test values for hypothesis 1 and the RCM values for a 3-state and 2-state model respectively. Figures 4 and 5 show the smoothed probabilities of being in each state for each case considered. We

---

7 The moving cumulative windows analysis begins with 57 data in 1930 and includes new data until the end of the sample. The moving windows analysis uses a constant data size of 57.

8 Standard deviations are calculated as the square-root of the main diagonal elements of the quasi-maximum likelihood covariance matrix.

9 The smoothed probabilities of being in each state are calculated as suggested by Hamilton (1990).
can see in these figures, that while in a 3-state model, we obtain robust results for both samples, in a 2-state model, the state identification is considerably different among samples. Evidence then suggests that a 3-state model robustly shows a better state identification. A model with 2 states not only shows a poor state identification, in particular, for the full sample in the pre-world war II period, but also, the conclusions about the location of the states are not robust among samples. The RCM is lower and almost 0 for the 3-state model and we obtain a slightly better state identification for the full sample in this case. In a 2-state model, the RCM is lower for the pre-1993 sample, confirming that in this case, we get a better state identification for the pre-1993 sample, as previously suggested by figure 5.

We can also see in the tables that the evidence supports the asymmetrical mean-reversion hypothesis, since the LR test value is high for both, the 2 and the 3-state models. The reversion process seems to exhibit a non-linear dynamics as theoretically proposed by Gabriel et. al (2002) and GV; and empirically tested by GV, CF and McMillan (2007). For the 3-state model, even when the attractor is poorly estimated for both samples, we can conclude that (i) \( \rho_1 \) is statistically insignificant in both samples (\( \hat{\rho}_1 \) is even positive for the full sample). State 1 is then a nonstationary state and occurs punctually in some years only for very short periods that does not clearly correspond with relevant historical episodes in the stock market (except maybe for the period 1930-1932 that corresponds to the post-1929 crash). Evidence suggests that state 1 only characterizes the dynamics of outlayer data; (ii) even when \( \hat{\rho}_3 < \hat{\rho}_2 < 0 \), only for the full sample, both parameters are statistically significant for a 5% confidence level. For the pre-1993 sample, these two parameters are more poorly estimated suggesting that both states are near-nonstationary. In almost the entire sample, the ratio dynamics can be characterized by state 2. In particular, the probability of being in state 2 is very close to 1 in the period 1949-1973, corresponding with the post WWII period, and from 1979 to 1995, corresponding at least with the first part of the 90’s "boom", and finally (iii) state 3, the state with the higher speed of adjustment is clearly a stationary state. This last state basically characterize only the dynamics of the last part of the 90´s boom (from around 1996 to 2000).

Until now, previous related research involving non-linear dynamics in the reversion process has supported the hypothesis of asymmetrical speed of adjustment. They have found evidence of this asymmetry allowing only for two regimes, usually related to bull and bear markets as in CF or episodes where noise traders tend to overreact to news as in McMillan (2007). Evidence for our 2-state model basically supports previous evidence only if we use the pre-1993 sample at least in the state identification of remarkable episodes. From table 3, for the pre-1993 sample, we can conclude that both \( \rho_1 \) and \( \rho_2 \) are statistically negative (with \( \rho_2 \) statistically significant at a critical value just above 5%). For this sample, we can then conclude that
state 2 is a near-nonstationary state while state 1 is a stationary state. Analyzing the probabilities of being in each state in Figure 5 we can identify some important episodes. From around 1930 to 1932 the probability of being in state 1 is almost 0, which corresponds to a bear episode related to the Wall Street crash, previously analyzed by CF. From 1952 to around 1973 the probability of being in state 1 is almost 1, also supporting GV that had previously identified a different state for this ratio in the period 1955-1975. According to CF, there is evidence of asymmetrical mean reversion $\rho^c \neq \rho^b$, with $\rho^c$ (bull markets) closer to 0 (in fact, $\rho^c$ is statistically insignificant) than $\rho^b$ (bear markets), while in this context, for the pre-1993 sample $\hat{\rho}_2$ is closer to 0 than $\hat{\rho}_1$, with state 2 being a near-nonstationary state. The results for the pre-1993 seems to be really affected by the poor estimation of the long-run equilibrium $\mu$.

4 CONCLUSIONS

Previous research in the stationarity analysis of the PD ratio has been non-conclusive at least in a linear context. We find support to the evidence that the speed of adjustment of the ratio has not been constant among samples and that it has shown important changing episodes closely related to historical episodes in the stock market. The non-linear analysis of the stationarity hypotheses based on a Markov-switching model shows robust evidence of switching regimes characterizing the speed of adjustment of these ratios around the same long-run equilibrium.

The poor state identification in a 2-state context when we include the part of the sample corresponding to the last part of the past century and the robust conclusions obtained with a 3-state model suggest that both "booms" does not share the same characteristics and that a third state is necessary to properly model the ratio dynamics. In this context, the first "boom" (post WWII) is characterized by a near-nonstationary ratio, and the second "boom" is divided into two parts, the first one has the same characteristics as the first "boom", and the second part, where the slope of the ratio is higher, and the divergence between prices and dividends becomes higher is characterized by a new state with a stronger pull from the attractor, implying that it is more probable that the market is again in a stationary state with respect to the estimated attractor.

The evidence found for the 3-state model supports the PV model and suggests that the overall rational behavior of the agents is explained when punctual episodes of nonstationary behavior (as in state 1) are followed by a higher probability of
changing to a stationary state (state 2 in almost the entire sample and state 3 in the last part of it), and then, nonstationarity is only a temporary state.
APPENDIX A

This appendix briefly describes the Markov-switching framework for estimating non-linear models as well as a regime classification measure. Following Hamilton (1989, 1990), in a 2-state Markov-switching model, for estimating an equation such as (7), a transition matrix for $s_t$ (the latent variable governing the switching-regime process) has to be defined as

$$\mathbf{P} = P(S_t = s_t \mid S_{t-1} = s_{t-1}, x_{t-1}) = \begin{bmatrix} p_{11} & 1 - p_{22} \\ 1 - p_{11} & p_{22} \end{bmatrix},$$

where $x_{t-1}$ is a vector containing all observations for a valuation ratio obtained through date $t - 1$. If by time $t$, $s_t = j$, the conditional density of $\Delta x_t$ will be given by

$$f(\Delta x_t \mid x_{t-1}, s_t = j, s_{t-1} = i, s_{t-2} = k, \ldots; \Theta),$$

where $\Theta$ is a vector containing the estimated parameters (depending on each case considered), and it is assumed that the conditional density depends only on the current regime $s_t$, so the conditional density is given by

$$f(\Delta x_t \mid x_{t-1}, s_t = j; \Theta).$$

For instance, in the two regime model, the conditional densities will be collected in a vector called $\eta_t$

$$\eta_t = \begin{bmatrix} f(\Delta x_t \mid x_{t-1}, s_t = 1; \Theta) \\ f(\Delta x_t \mid x_{t-1}, s_t = 2; \Theta) \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{(\Delta x_t - \alpha_1(x_{t-1} - \mu) - \beta_1(\Delta x_{t-1}))^2}{2\sigma^2} \right\} \\ \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{(\Delta x_t - \alpha_2(x_{t-1} - \mu) - \beta_1(\Delta x_{t-1}))^2}{2\sigma^2} \right\} \end{bmatrix}.$$

The maximum-likelihood algorithm seeks to find a vector $\Theta^*$ that maximizes the log-likelihood function $\mathbb{L}(\Theta)$ for the observed data $x_t$. $\mathbb{L}(\Theta)$ is given by

$$\mathbb{L}(\Theta) = \sum_{t=1}^{T} \log f(\Delta x_t \mid x_{t-1}; \Theta),$$

12
where

\[ f(\Delta x_t \mid x_{t-1}; \Theta) = 1(\xi_{t|t-1} \odot \eta_t), \]  

(11)

where 1 is a (2x1) vector of ones, and

\[ \xi_{t|t-1} = P \cdot \xi_{t-1|t-1}, \]  

(12)

where

\[ \xi_{t-1|t-1} = \frac{(\xi_{t-1|t-2} \odot \eta_{t-1})}{1(\xi_{t-1|t-2} \odot \eta_{t-1})}. \]

The optimization algorithm works as follows: given an initial value \(\xi_{1|0}\) equations (11) and (12) can be used to calculate \(\xi_{t|t}\) for any \(t\). These values can be replaced in (10) and iterate until finding \(\Theta^*\) depending on a predefined convergence criterion.

To analyze the ability of the Markov-switching model to identify different states, we use the regime classification measure (RCM) proposed by Ang and Bekaert (2002). The RCM is defined as

\[ RCM = 100k^2 \frac{1}{T} \left( \sum_{i=1}^{T} \prod_{t=1}^{k} p_{i,t} \right), \]

where \(k\) is the number of regimes, \(p_{i,t}\) is the smoothed probability of being in state \(i\) at period \(t\), and the constant (100) normalize the RCM between 0 and 100. If the probabilities of being in one state are close to 1 or to 0 in every period, the states would be properly identified, we have then a good regime classification, and the RCM statistic will be close to zero; but if the states are not well identified, the probabilities of being in a particular state will be far from 0 and 1, implying a large value for the RCM statistic.

\(^{10}\)Following Hamilton (1989) for \(\xi_{1|0}\) we choose the vector of unconditional probabilities \(\pi\)

\[ \pi = (A' A)^{-1} A e_3, \]

where

\[ A = \begin{bmatrix} I_3 - P \\ 1' \end{bmatrix}, \]

and \(e_3\) denotes the third column of \(I_3\), the (3x3) identity matrix.
References


Table 1. Summary Statistics and Unit-root tests.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Median</th>
<th>Min.</th>
<th>Max.</th>
<th>St. Dev</th>
<th>SK</th>
<th>Exc. Ku</th>
<th>ADF Test</th>
<th>PP Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>PD</td>
<td>3.1580</td>
<td>3.1262</td>
<td>2.3147</td>
<td>4.4475</td>
<td>0.3928</td>
<td>0.9425</td>
<td>1.6052</td>
<td>-2.3606</td>
<td>-2.5039</td>
</tr>
<tr>
<td>Pre-1993 sample</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PD</td>
<td>3.0721</td>
<td>3.0772</td>
<td>2.3147</td>
<td>3.6268</td>
<td>0.2806</td>
<td>-0.2587</td>
<td>-0.4407</td>
<td>-4.0340</td>
<td>-4.2498</td>
</tr>
</tbody>
</table>
Table 2. PD ratio. Estimated parameters and stationarity analysis results for a 3-state model.

<table>
<thead>
<tr>
<th>PD ratio</th>
<th></th>
<th>Pre 1993</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>RCM</td>
<td>0.0119</td>
<td>0.0168</td>
<td></td>
</tr>
<tr>
<td>LR for H</td>
<td>178.79</td>
<td>163.92</td>
<td></td>
</tr>
<tr>
<td>Param.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>α</td>
<td>-0.2310</td>
<td>-0.3342</td>
<td></td>
</tr>
<tr>
<td>st. dev</td>
<td>0.2302</td>
<td>0.2747</td>
<td></td>
</tr>
<tr>
<td>ρ1</td>
<td>0.0053</td>
<td>-0.0263</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0705</td>
<td>0.0748</td>
<td></td>
</tr>
<tr>
<td>ρ2</td>
<td>-0.0880</td>
<td>-0.1200</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0419</td>
<td>0.0709</td>
<td></td>
</tr>
<tr>
<td>ρ3</td>
<td>-0.1753</td>
<td>-0.2052</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0348</td>
<td>0.1112</td>
<td></td>
</tr>
<tr>
<td>μ</td>
<td>6.1761</td>
<td>6.1769</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.2097</td>
<td>2.6136</td>
<td></td>
</tr>
<tr>
<td>β1</td>
<td>0.1312</td>
<td>0.1476</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0431</td>
<td>0.0511</td>
<td></td>
</tr>
<tr>
<td>σ1</td>
<td>0.0763</td>
<td>0.0793</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0065</td>
<td>0.0060</td>
<td></td>
</tr>
<tr>
<td>P11</td>
<td>0.9759</td>
<td>0.9723</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0137</td>
<td>0.0154</td>
<td></td>
</tr>
<tr>
<td>P12</td>
<td>0.0206</td>
<td>0.0222</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0123</td>
<td>0.0142</td>
<td></td>
</tr>
<tr>
<td>P21</td>
<td>0.0442</td>
<td>0.0333</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0377</td>
<td>0.0201</td>
<td></td>
</tr>
<tr>
<td>P22</td>
<td>0.9243</td>
<td>0.9355</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0429</td>
<td>0.0262</td>
<td></td>
</tr>
<tr>
<td>P31</td>
<td>0.0085</td>
<td>0.0187</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0213</td>
<td>0.0360</td>
<td></td>
</tr>
<tr>
<td>P33</td>
<td>0.9557</td>
<td>0.9493</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0197</td>
<td>0.0419</td>
<td></td>
</tr>
</tbody>
</table>
Table 3. PD ratio. Estimated parameters and stationarity analysis results for a 2-state model.

<table>
<thead>
<tr>
<th></th>
<th>Full sample</th>
<th>Pre 1993</th>
</tr>
</thead>
<tbody>
<tr>
<td>PD ratio</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RCM</td>
<td>38.8478</td>
<td>16.3206</td>
</tr>
<tr>
<td>LR for H1</td>
<td>44.23</td>
<td>92.57</td>
</tr>
<tr>
<td>Param.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha$</td>
<td>-0.0266</td>
<td>-0.4991</td>
</tr>
<tr>
<td>st. dev</td>
<td>0.0784</td>
<td>0.3211</td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>0.0923</td>
<td>-0.3078</td>
</tr>
<tr>
<td></td>
<td>0.1357</td>
<td>0.0580</td>
</tr>
<tr>
<td>$\rho_2$</td>
<td>-0.7521</td>
<td>-0.1491</td>
</tr>
<tr>
<td></td>
<td>0.1724</td>
<td>0.0884</td>
</tr>
<tr>
<td>$\mu$</td>
<td>3.0763</td>
<td>5.1512</td>
</tr>
<tr>
<td></td>
<td>0.1483</td>
<td>1.0245</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.1646</td>
<td>0.0230</td>
</tr>
<tr>
<td></td>
<td>0.1097</td>
<td>0.0793</td>
</tr>
<tr>
<td>$\sigma_1$</td>
<td>0.1499</td>
<td>0.1055</td>
</tr>
<tr>
<td></td>
<td>0.0105</td>
<td>0.0101</td>
</tr>
<tr>
<td>$P_1$</td>
<td>0.9658</td>
<td>0.9310</td>
</tr>
<tr>
<td></td>
<td>0.0206</td>
<td>0.0182</td>
</tr>
<tr>
<td>$P_2$</td>
<td>0.9539</td>
<td>0.9613</td>
</tr>
<tr>
<td></td>
<td>0.0192</td>
<td>0.0167</td>
</tr>
</tbody>
</table>
Figure 1. S&P Composite demeaned PD ratio. Full sample.

Note: the series are normalized to 1 in 1970.
Figure 2. Moving cumulative windows stability analysis for parameters $\alpha$ and $\rho$ in equation $\Delta x_t = \alpha + \rho x_{(demeaned)} t-1 + u_t$, for the PD ratio and 95% confidence intervals.
Figure 3. Moving windows stability analysis for parameters $\alpha$ and $\rho$ in equation
$\Delta x_t = \alpha + px_{(demeaned)_{t-1}} + u_t$, for the PD ratio and 95% confidence intervals.
Figure 4. 3-state model. Smoothed probability of state 1, 2 and 3 for PD ratio. Full and Pre-1993 sample.
**Figure 5.** 2-state model. Smoothed probability of state 1 for PD ratio. Full and Pre-1993 sample.