ISSN 1988-088X

fae2
Departamento de Fundamentos del Análisis Económico II

Department of Foundations of Economic Analysis II
University of the Basque Country UPV/EHU
Avda. Lehendakari Aguirre 83
48015 Bilbao (SPAIN)
http://www.dfaeii.ehu.es

DFAE-II WP Series

2014-05

Juan-Miguel Londoño, Marta Regúlez & Jesús Vázquez

An Alternative View of the US Price-Dividend Ratio Dynamics

An Alternative View of the US Price-Dividend Ratio Dynamics*

Juan-Miguel Londoño†, Marta Regúlez‡, Jesús Vázquez§

October, 2014

Abstract

As a necessary condition for the validity of the present value model, the price-dividend ratio must be stationary. However, significant market episodes seem to provide evidence of prices significantly drifting apart from dividends while other episodes show prices anchoring back to dividends. This paper investigates the stationarity of this ratio in the context of a Markov-switching model à la Hamilton (1989) where an asymmetric speed of adjustment towards a unique attractor is introduced. A three-regime model displays the best regime identification and reveals that the first part of the 90's boom (1985-1995) and the post-war period are characterized by a stationary state featuring a slow reverting process to a relatively high attractor. Interestingly, the latter part of the 90's boom (1996-2000), characterized by a growing price-dividend ratio, is entirely attributed to a stationary regime featuring a highly reverting process to the attractor. Finally, the post-Lehman Brothers episode of the subprime crisis can be classified into a temporary nonstationary regime.

JEL Classification: C32, G12

Keywords: Markov regime-switching, price-dividend ratio stationarity

*We are grateful for comments from Lieven Baele, Guillermo Llorente, Alfonso Novales, Eliseo Navarro, Stefano Siviero, Miguel Martinez, Trino Ñiguez, Susan Orbe and participants in seminars at Universidad Complutense de Madrid, Universidad del País Vasco, Foro de Finanzas and EFMA meeting. Financial support from Ministerio de Ciencia e Innovación (Spain) through projects SEJ2004-04811/ECON, SEJ2007-66592-C03-01/ECON, ECO2010/16970 is also acknowledged.

†Division of International Finance. Mailstop 43, Board of Governors of the Federal Reserve System, 20th Street and Constitution Avenue N.W., Washington, D.C. 20551 (USA). E-mail: Juan-miguel.londono-yarce@frb.gov

‡Departamento de Economía Aplicada III, Universidad del País Vasco (UPV/EHU). E-mail: marta.regulez@ehu.es

§Departamento de Fundamentos del Análisis Económico II, Universidad del País Vasco (UPV/EHU). E-mail: jesus.vazquez@ehu.es

1 Introduction

According to traditional asset pricing literature, prices should not drift apart permanently from dividends. In other words, the price-dividend (PD hereafter) ratio should show a reverting behavior towards an attractor. However, around 1985, the US stock market started its latest remarkable episode of apparent divergence between equity prices and dividends. As a consequence of this, the PD ratio showed a sustained increase during the so-called 90's boom. During this particular episode, dividends seem to have become less important as the key payout mechanism.[1] A few years later, this episode of sustained increase was followed by the episode of significant drops in equity prices experienced during the latest recession or subprime crisis. Around the subprime crisis, equity prices seem to anchor back to dividends. The dynamics of the ratio around these two episodes reopened the debate on the relation between equity prices and their fundamentals around recessions.

This paper provides an alternative view of the PD ratio dynamics consistent with the existence of a high, unique attractor and the presence of alternative regimes. The idea of a unique attractor is closely linked to that of a long-run equilibrium relationship between prices and dividends whereas the existence of a high attractor is motivated by the rapid increase in the PD ratio during the late 20th century.[2] This alternative view is based on the empirical evidence obtained from the estimation of a Markov-switching (MS) model à la Hamilton (1989) where neither the number of states nor the parameters driving the PD ratio dynamics in each state are restricted. Applying this methodology to US stock market data from 1871 to 2009, our empirical results show evidence of three alternative regimes in the dynamics of the PD ratio. The characteristics of the states identified differ significantly from those previously identified in the related literature. In particular, our results suggest that the PD ratio has experienced transitory episodes where it has drifted apart from its attractor.

The relationship between equity prices and fundamentals is a basic concern for market participants and researchers since it determines the uniqueness of prices, the rationality

[1] Fama and French (2001) show that firms, regardless of their characteristics, have become less likely to pay dividends. Other papers like Baker and Wurgler (2004) and Hoberg and Prabhala (2009) investigate the phenomenon of disappearing dividends. Some changes to the dividend policy regulation may partially explain this decreasing dividend payment trend. In particular, the enactment of U.S. Securities and Exchange Commission (SEC) rule 10b-18 in 1982 (see, for instance, Grullon and Michaely, 2002; Boudoukh, Michaely, Richardson and Roberts,2007), and the 2003 tax cut (as investigated by Chetty and Saez, 2005).

[2] The idea of a transition to a relatively high attractor advocated in this paper is not completely new, although somewhat controversial. This idea follows the initial beliefs of a "permanently high plateau" of stock prices proposed by Irving Fisher just before the 1929 crash. Even more, it is in line with the hypothesis in Glassman and Hasset (1999) that stock prices might keep rising in relation to dividend payments in order to reach what they believe is an appropriate risk premium. By no means are we arguing that our results support the hypothesis in Glassman and Hasset (1999) that the Dow Jones will someday hit 36,000 or their beliefs about the risk premium. Our approach simply considers the possibility that the long-run attractor for the ratio could be higher than what historical data initially suggest, and of course, different from the historical mean.

of agents, and the forecastability of returns (Cochrane, 2008). Since the seminal paper by Shiller (1981), the dynamic features of the relationship between equity prices and dividends has been investigated by many authors. The papers in this area of research can be classified into two strands of literature. There is a first strand that considers that the reversion process of the PD ratio exhibits linear dynamics. Several studies have found nonconclusive empirical evidence for the cointegration relation linking stock prices and dividends as shown in Cochrane (1992, 2001) and Lettau and Ludvigson (2005), among others. In particular, the evidence on cointegration between equity prices and dividend series turns out to be highly sensitive to the sample considered. This sample-dependence is not surprising given the different characteristics of alternative episodes in the equity market. These episodes clearly suggest the presence of an asymmetric behavior in the dynamics of the ratio, which implies that the reversion process may not be linear. The second strand of literature, where this paper can be classified into, is somewhat more flexible and allows precisely for the possibility of an asymmetric reverting process for the PD ratio. In this line of research, the long-run relationship implied by the present value model holds while the short run dynamics of the ratio can be affected by certain episode-specific characteristics such as the existence of transaction costs, noise traders and swings in market sentiment. Most papers in this line of research propose different versions of threshold autoregressive (TAR) models where the speed of adjustment to an attractor is assumed to depend on a threshold variable driving regime-switches.[3]

The MS approach followed in this paper allows for the possibility of a state-dependant speed of adjustment towards a unique long-run equilibrium as in the TAR-related papers. However, the MS method provides several unique features to go deeper and investigate further the dynamics of the PD ratio during particular market episodes. First of all, this method makes simpler to consider additional regimes. This feature allows us to deviate from the two-state (bull and bear markets or recession and non recession episodes) framework traditionally investigated in the previous literature. Moreover, within this method we do not impose any predetermined characteristics for the dynamics of the ratio within each regime. The lack of restrictions of the MS method stands in sharp contrast to the approach suggested by the TAR method. We argue that the MS approach is more flexible since the estimation results from TAR models may depend on the definition of the threshold variable chosen by the researcher. The definition of the threshold variable imposes *a priori* features identifying the alternative regimes. In contrast, the switching process characterized by an MS approach is governed by a latent variable that is not predefined by the researcher. In this sense, we believe that the MS approach allows the data to speak more freely than a TAR approach because the variable governing regime-switches is not defined *a priori* under the MS approach.[4] Finally, the MS approach

[3]See, for instance, Bohl and Siklos (2004), Coakley and Fuertes (2006) and McMillan (2006, 2007). This branch of literature stands in sharp contrast with papers that consider switches in the attractor (see for instance Lettau and Van Nieuwerburgh, 2008).

[4]The flexibility of our methodology as in the number of states, their characteristics and the possibility of nonstationary states has of course a trade off. In particular, this methodology might be prompt to deliver nonidentifiable states. For example, the states obtained might be hard to interpret and not related to any characteristic of either the stock market or the business cycle. In order to avoid this potential drawback

allows for the possibility of transitory episodes where the PD ratio shows no evidence of convergence to the attractor. The possibility of temporary non-stationary regimes is motivated, as mentioned above, by the high sensitivity of the evidence on cointegration between stock prices and dividends to the sample period considered found in previous studies.[5]

Within the three-regime MS specification, we investigate the particular characteristics of some relevant historical episodes such as the post-war period (up to 1975), the 90's boom and the subprime crisis. We find that the post-war period and the first part of the subprime crisis are characterized by a stationary behavior, but with a low speed of adjustment to the attractor. Moreover, two sub-periods can be identified in the 90's boom. The first sub-period (1985-1995) shows a slow reversion to the attractor whereas the second sub-period, characterized by a fast growing PD ratio, features a strong reversion regime. Interestingly, this last result suggests that the period 1996-2000 is characterized by a stationary regime. A result that stands in sharp contrast to the conclusions reached by previous papers which rely on a TAR approach, especially to those papers that identify the long-run attractor with the historical mean of the PD ratio. The contrasting regime characterization is determined by another important estimation result found: even if the attractor is poorly identified, by using alternative samples and MS specifications, we robustly find larger estimated values of this parameter than those reported in the previous related literature. More important, the remaining parameter estimates and the regime features identified by the MS approach are not affected whenever the attractor parameter is fixed to a particular value as long as this value is close to the historical maximum reached by the PD ratio. This high attractor in turn explains why a fast growing PD ratio is linked to a stationary state where the PD ratio is catching up the high long-run equilibrium. Finally, we also find that a short-lived nonstationary regime characterizes several remarkable recession episodes. In particular, recession episodes involving significant drops in equity prices such as the 1929 crisis and the post-Lehman Brothers episode during the subprime crisis belong to this regime.

The rest of the paper is organized as follows. Section 2 summarizes the related literature on the analysis of PD ratio stationarity within the PV framework. Section 3 describes the data and presents a preliminary investigation on the stability of the PD ratio reversion process. Section 4 presents the MS model considered in this paper and discusses the empirical results found using the benchmark three-regime model. This section also assesses the importance of attractor estimation. Section 5 provides a robustness analysis along several dimensions. Finally, Section 6 concludes.

---

and keep the state identification interpretable, we carry out an analysis of the relationship between the states identified using the MS methodology and several business cycle and stock market characteristics such as inflation, GDP growth and equity market volatility.

[5] Put differently, the presence of transitory non-reverting regimes may lead to the non-conclusive evidence of cointegration in a linear framework found in the literature. By contrast, the MS approach accommodates the possibility of detecting temporary non-stationary regimes together with at least one stationary regime featuring a strong reversion to an attractor, which implies a strong cointegration relationship.

2 Nonlinear Processes and the PD Ratio Dynamics

In this section, we introduce the present value (PV) model as well as its implications on the stationarity of the PD ratio. We then summarize the results found in the related literature with respect to the PD ratio stationarity. We review first those papers considering linear processes in the dynamics of the ratio. Then, we summarize the literature considering nonlinear processes in the dynamics of the ratio.

Campbell and Shiller (1988a, 1988b) develop a log-linear approximation to the PV framework that can be used to study stock price behavior under any model of expected returns. Their approach leads to the following PV equation:

$$p_t = \frac{k}{(1-\phi)} + E_t \left\{ \sum_{j=0}^{\infty} \phi^j [(1-\phi) d_{t+j} - r_{t+j}] \right\} + \lim_{j \to \infty} E_t(\phi^j p_{t+j}), \tag{1}$$

where p_t is the logged value of the stock price at the beginning of period t, d_t is the logged value of the dividend accruing to the stock paid out throughout period t, and r_t is the log return associated with stocks at time t (i.e. $r_t = \ln(1 + R_{t+1})$). Finally, k and ϕ are constants obtained from the log-linear approximation. Equation (1) can be written in terms of the (log of the) PD ratio as follows:

$$p_t - d_{t-1} = \frac{k}{(1-\phi)} + E_t \left\{ \sum_{j=0}^{\infty} \phi^j [\Delta d_{t+j} - r_{t+j}] \right\} + \lim_{j \to \infty} E_t[\phi^j (p_{t+j} - d_{t-1+j})]. \tag{2}$$

The last term on the RHS of equation (2) drops out under the transversality condition $\lim_{j\to\infty} E_t[\phi^j (p_{t+j} - d_{t-1+j}) = 0$,[6] which implies that

$$p_t - d_{t-1} = \frac{k}{(1-\phi)} + E_t \left\{ \sum_{j=0}^{\infty} \phi^j [\Delta d_{t+j} - r_{t+j}] \right\}. \tag{3}$$

In addition, if dividends are assumed to be I(1) and returns are stationary, equation (3) implies that the stationarity of the PD ratio can be viewed as a necessary condition for the validity of the PV model, and the logged values of prices and dividends are then cointegrated, with a cointegration vector given by $(1, -1)$. To understand this cointegration relationship, one may intuitively think that if current stock prices are high in relation to current dividends (the stock is overpriced with respect to actual/lagged dividends, or there are expectations of future high dividend payments), dividends are expected to grow. That is, if agents are fully rational under this model, prices and dividends cannot drift apart forever and the ratio will show a reverting behavior towards an attractor. In other words, in the presence of rational agents assigning unique prices to stocks in relation to

[6] Imposing the transversality condition ensures the uniqueness of the solution for stock prices obtained from the PV model.

their dividend payments, the stationarity of the PD ratio is a necessary condition for the PV model.

Previous studies such as Cochrane (1992, 2001), and Lettau and Ludvigson (2005), among others consider linear dynamics for the analysis of the PD ratio reversion process. These papers have investigated different samples as well as alternative model specifications and have found nonconclusive evidence for the cointegration relationship implied by the PV model. In particular, the evidence on cointegration reported in these studies depends highly on the sample period considered. This nonconclusive evidence as well as the different characteristics of alternative equity market episodes suggest that the reversion process of the PD ratio may not be linear, and then, any inference based on a linear framework might be at least misleading.

The key difference among papers following a nonlinear approach comes from the alternative driving forces assumed for the asymmetric reversion process.[7] There is a branch of literature in which the non-linearity features come from stock price fundamentals (for instance, dividends). In Shiller (1989), for example, there are different types of agent who react differently to historical events, macroeconomic news or just fads. These agents can be long-run investors who show a more stable behavior or noise traders who tend to react to fads or overreact to news. Alternatively, Froot and Obstfeld (1991) introduce the possibility of an "intrinsic bubble" which depends exclusively on dividends. Extending the intrinsic bubble specification, Drifill and Sola (1998) include structural breaks in the dividend process. The possibility of having structural breaks in the dividend series is in turn motivated by the empirical evidence on unstable dividend processes. This evidence on structural breaks paved the way for other regime-switching specifications as in Evans (1998) and Gutiérrez and Vázquez (2004).[8]

Another branch of literature focuses entirely on the stationarity of the PD ratio implied by the PV model and considers that stock prices are driven by non-fundamental components. In particular, Bohl and Siklos (2004), Coakley and Fuertes (2006), Kapetanios, Shin and Shell (2006) and McMillan (2006, 2007), introduce a non-fundamental term u_t in equation (3) as in

$$p_t - d_{t-1} = \frac{k}{(1-\phi)} + E_t \left\{ \sum_{j=0}^{\infty} \phi^j [\Delta d_{t+j} - r_{t+j}] \right\} + u_t. \tag{4}$$

There are different interpretations for the error term u_t in Eq. (4). For example, Bohl and Siklos (2004) argue that u_t is a bubble term that captures run-ups in stock prices

[7] In this paper, we focus our attention on the literature that considers nonlinearities towards a unique long-run attractor. A widely analyzed alternative considers that the nonlinearity comes from shifts in the attractor as analyzed by Lettau and Van Nieuwerburgh (2008) and papers cited therein. In their paper, the hypothesis analyzed is that deviations around different steady-states are stationary.

[8] Related papers considering alternative specifications including regime changes not directly linked with the dividend process can be found in Cecchetti, Lang and Mark (1990), Veronessi (1999), Timmermann (2001), and Bonomo and Garcia (1994), among others.

before a crash, suggesting the presence of asymmetries in the PD ratio reverting process. Kapetanios et al. (2006) interpret this term as capturing the presence of transaction costs such that small, uncorrected deviations may arise but larger deviations would be arbitraged away. McMillan (2006, 2007) and Coakley and Fuertes (2006) link this mispricing term to market sentiment as in behavioral finance models or as in one of the hypotheses explaining disappearing dividends (Baker and Wurgler, 2004). In those behavioral models, the existence of noise traders in the market who react differently to the arrival of good or bad fundamental news provides a possible source for asymmetries since the trend-chasing behavior of such traders after the arrival of positive news leads to a market over-reaction such that the price change is greater than required by the news. In contrast, these noise traders would be more conservative in bear markets thus anchoring prices to dividends. Therefore, the reversion process of the PD ratio could be more persistent in bull markets and more rapidly reverting in bear markets.

The evidence found by Bohl and Siklos (2004), Coakley and Fuertes (2006) and McMillan (2007) is based on a two-regime framework for the speed of adjustment. In general, these papers consider alternative TAR specifications for the dynamics of the PD ratio that build on the model of Enders and Granger (1998), i.e., using a model similar to the augmented Dickey-Fuller (ADF) regression specification such as:

$$\Delta pd_t = I_t \rho_1 (pd_{t-1} - \mu) + (1 - I_t)\rho_2 (pd_{t-1} - \mu) + \sum_{j=1}^{l} \beta_j (\Delta pd_{t-j}) + \varepsilon_t, \tag{5}$$

where pd_t denotes the (log of the) PD ratio at time t; ρ_j $(j = 1, 2)$ is the speed of adjustment in each regime to the attractor, μ; ε_t is an i.i.d. shock; and I_t is an indicator function that takes the value of one if $q_t \geq 0$, and zero otherwise. Finally, q_t is the threshold variable that predetermines the regimes.

Following the TAR setting, Coakley and Fuertes (2006) propose *a priori* a two-regime framework (called bull and bear regimes) for the speed of adjustment around the same long-run equilibrium. In their paper, the threshold variable is highly persistent with respect to dividend growth and is defined as $q_t(w, d) = w_1 \Delta pd_{t-1} + ... w_d \Delta pd_{t-d}$, where $w' = (w_1, ... w_d) > 0$ is a vector of predefined weights, and d is the number of lags to be selected from the data. The regime changes work as follows: if $q_t > 0$, the stock market is in a bull episode with speed of adjustment ρ_1 and if $q_t < 0$ the market is in a bear episode with speed of adjustment ρ_2. By contrast, Bohl and Siklos (2004) assume a threshold variable showing lower persistence simply defined as $q_t = \Delta pd_{t-1} - \tau$, where τ is a threshold parameter to be estimated. Recently, McMillan (2006, 2007) considers an exponential smooth transition model specification for the dividend-price ratio. His model implies that the dynamics of the middle ground differ from the dynamics associated with large deviations. McMillan (2006,2007) also introduces asymmetries between regimes of rising and falling prices. His model falls into the STAR family of models, where a continuous transition function $G(q_t)$ between 0 and 1 is used instead of the indicator function I_t.[9]

[9] For a generalization of TAR models and their reversion analysis see, for instance, Tong (1993) and

The empirical evidence found in this last branch of literature is also dependent on the sample and the particular specification considered. For instance, the evidence in Coakley and Fuertes (2006) is based on monthly data from January 1871 to September 2001 for the Standard and Poors PD ratio. Based on their two-regime TAR framework, they find support for the hypothesis that the PD ratio regularly behaves as a random walk with an upward drift where stock prices drift away from fundamentals during bull market episodes. In particular, they find that the 90's boom falls into this category. However, in bear markets, the adjustment of the ratio towards the equilibrium level is reinstated. Their conclusion remains the same if observations from 1993:01 onwards are excluded. A result that could be interpreted either as quite robust or as driven by the small weight of the observations corresponding to the last eight years relative to the whole sample. Moreover, Coakley and Fuertes (2006) report an estimate of the attractor smaller than the historical mean, with no associated measure of precision, so its significance cannot be assessed.

Bohl and Siklos (2004) investigate the (demeaned and detrended) US log dividend-price ratio from January 1871 to September 2001. For this sample, they find evidence of a stationary ratio and bubble-like asymmetric short-run adjustments such that stock prices increase relative to fundamentals followed by a crash. The exception to these dynamics is the episode between 1947 and 1982. They suggest that the different pattern observed for this period is due to the absence of bull market periods followed by crashes which makes this an atypical period relative to the rest of the sample. Moreover, they also find strong differences between two non-overlapping periods: 1871-1936 and 1937-2001. In the first period they find no evidence of either a unit root in the log dividend-price ratio nor asymmetric effects in its dynamics, whereas the opposite is true for the second period.

Finally, McMillan (2006) finds that for the period 1980-1995 the cointegration relationship between stock prices and dividends gets stronger. This episode is followed by an increase in real dividends from the beginning of 1995 that is in turn followed by an increase in real prices. Given the dynamics of prices and dividends from 1995, the strength of the stationary relationship falls quite significantly for the period 1995-1999, when it enters a slow transition from a reverting regime to a random walk regime. However, for the period 2000-2004 he finds that the adjustment becomes stronger again. The results in McMillan (2007) also support the presence of nonlinearities in the reversion process of the PD ratio for the period February 1965 to May 2004. For this sample, he obtains that the speed of transition is lower when stock prices rise relative to dividends than when prices are below the level supported by dividends. In particular, he finds that the recent dynamics of the PD ratio (in the episode related to the technological boom) falls into a random walk regime where stock prices seem to diverge from fundamentals. It is important to remark that all conclusions reached by McMillan (2006, 2007) are affected by the fact that he considers demeaned time series, so the historical mean is taken *a priori* as the long-run attractor.

Enders and Granger (1998).

The MS method proposed in this paper adds to this literature by allowing for more flexibility in the investigation of the dynamics of the PD ratio. First, we do not impose any number of regimes that may have been present during the sample period. The flexibility in the number of regimes allows us to deviate from the two-states framework traditionally considered in the literature. In contrast with the TAR-related papers summarized above, we do not impose any predetermined characteristics for the dynamics of the ratio within each regime. In particular, we argue that the possibility of a long-run equilibrium different from the historical mean of the PD ratio may play a key role as a reference point in the identification of regimes as well as in the interpretation of their characteristics. Thus, in this paper we investigate the asymmetric reversion hypothesis together with the estimation of a unique attractor. Finally, the MS method allows for the possibility of transitory episodes where the PD ratio shows no evidence of convergence to the attractor. That is, again in contrast to the related literature, our approach allows for the possibility of identifying temporary nonstationary regimes for which the cointegration relationship between stock prices and dividends breaks down temporary until the stock market switches to a stationary state.

3 Data and Preliminary Stability Analysis for the Dynamics of the PD Ratio

Before introducing the MS model, this section presents the data and performs a preliminary analysis. This analysis motivates empirically the presence of switches in the dynamics of the ratio during the sample considered as well as the possibility that the attractor is not well captured by the sample mean of the PD ratio.

In this paper, we consider annual and quarterly data for the Standard and Poor's index price as well as for the dividends and earnings accruing to this index. These series are available at Robert Shiller's web site for the period 1971 to 2009. For this sample, we calculate the PD and PE ratios. The PD ratio is calculated as $pd_t = p_t - d_{t-1}$, where p_t is the log of the (closing) price of the SP500 index and d_{t-1} is the log of the dividends accruing to the index paid out throughout period $t-1$. Now, the price-earnings (PE hereafter) ratio used to investigate the robustness of the results to alternative proxies for payouts is calculated as $pe_t = p_t - \sum_{k=1}^{10} e_{t-k}$, where e_{t-k} is the log of the earnings reported in period $t-k$.

Throughout the paper, we mainly focus our attention on the annual frequency. The use of this frequency is justified by the main interest of the paper in relating switching regimes to business cycle characteristics. We believe that this relationship is in principle well captured by using annual data that ignore the noise associated with higher frequency data.[10] Nevertheless, we also consider quarterly data below as an additional robustness check.

Figure 1 shows the annual pd_t and pe_t for the full sample. This figure already provides some preliminary motivation for the nonlinear analysis of the dynamics of these ratios. It is interesting for example to identify several episodes where the magnitude of the ratio remains above or below the historical mean for several years. In particular, there is one such episode of high prices related to dividends and earnings (ratios above the mean for long episodes) between around 1955 and 1975, and a more recent one starting around 1990 and running at least to 2009. The latter episode includes the maximum values of both the PD and PE ratios for the whole sample, the trend shift after the technology boom and the large drop in stock prices related to the recent crisis episode or subprime crisis.[11]

Table 1 shows a summary of descriptive statistics for the PD ratio for several alternative samples.[12] The information in this table includes a commonly used test for cointegration in a non-state-dependent context namely the ADF test. The values of the statistics for this test suggest that only for the pre-1993 sample can the null hypothesis of nonstationarity of the PD ratio be rejected for any standard significance level. This preliminary result implies that if one does not consider the possibility of nonlinear reversion, the hypothesis of the PD ratio being stationary might be rejected when the last part of the sample (1993-2009) is included in the analysis. In sum, the preliminary evidence suggests that the PD ratio varies considerably through time, shows episodes of sustained increase, large drops and significant peaks related to particular business cycle episodes. The evidence also suggests that the stationarity of the PD ratio implied by the PV model is not decisively supported by the data for different subsamples.

The information in Table 1 and Figure 1 summarizes in a simple way part of the evidence already found in the relevant literature considering a linear framework for the analysis of PD ratio stationarity. In short, the empirical evidence presented here as well as that found in this literature is nonconclusive when testing the stationarity implications of the PV model. These results also show how difficult it may prove to reach a conclusion about the stationarity of the PD ratio in a non-state-dependent context, especially after the significant increase in the PD ratio that took place at the end of the millennium and the subsequent drop in prices related to the subprime crisis.

[10] Several related papers rely on the annual frequency. See, for instance, Froot and Obstfeld (1991), Driffill and Sola (1998), Gutiérrez and Vázquez (2004), Cochrane (2008) and Lettau and Van Nieuwerburgh (2008), among others. As for the monthly frequency considered in other papers such as Coakley and Fuertes (2006) and McMillan (2007), we also believe that since monthly data on dividends are actually an interpolation of quarterly data, the noise contained in these data may potentially affect the results.

[11] Throughout the paper, we focus our attention mainly on the PD ratio at the annual frequency. The dynamics of the PE ratio as well as the discussion for the quarterly frequency are left to be discussed in Section 5.

[12] We investigate the pre-1993 sample mostly for comparison with previous papers such as Coakley and Fuertes (2006), and Lettau and Van Nieuwerburgh (2008).

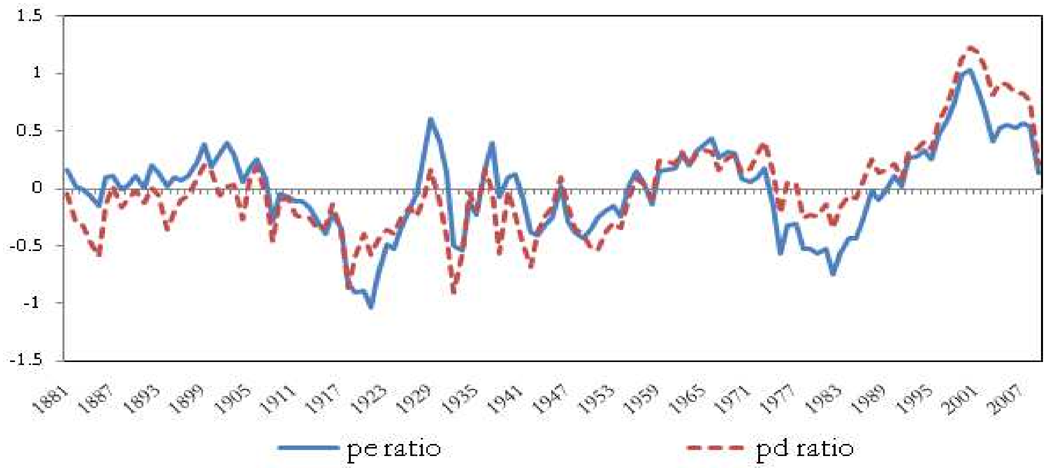


Figure 1: Standard and Poors Composite Index PD and PE ratios. Full sample.

Note: This figure plots the PD and PE ratios calculated as in Shiller (1989). That is, the PD ratio is calculated as $pd_t = p_t - d_{t-1}$, where p_t is the log of the (closing) price of the SP500 index and d_{t-1} is the log of the dividends accruing to the index paid out throughout period $t-1$. The price-earnings (PE hereafter) ratio is calculated as $pe_t = p_t - \sum_{k=1}^{10} e_{t-k}$, where e_{t-k} is the log of the earnings reported in period $t-k$. In order to simplify the preliminary analysis and make the ratios comparable, the figure shows the demeaned series.

	Mean	Median	St. Dev	Min.	Max.	Skew.	ADF	5% Crit.
Full sample	3.21	3.16	0.41	2.32	4.45	0.84	−2.95	−2.91
Pre 1993 sample	3.09	3.12	0.27	2.32	3.63	−0.33	−5.11	−2.92
Pre 2001 sample	3.15	3.13	0.35	2.32	4.45	0.74	−3.04	−2.92
Post 1950 sample	3.46	3.42	0.43	2.67	4.45	0.50	−1.73	−2.94

Table 1: Summary Statistics for the PD ratio.

Note: The results under the column ADF represent the Augmented Dickey Fuller test statistic (for the equation with an intercept and one-year lag for the PD ratio). The null hypothesis of this test is that the speed of adjustment of the ratio towards an attractor is null. The column named 5% Crit. reports the 5% critical value for the rejection of the null hypothesis.

Given the nonconclusive nature of the evidence based on a linear framework, we now introduce the possibility of nonlinearities in the dynamics of the ratio. Figure 2 provides a preliminary, but highly intuitive, analysis for considering the possibility of an asymmetric speed of adjustment around a constant attractor. More precisely, this figure shows the rolling estimates for the parameters α and the speed of adjustment ρ in a non state-dependent framework based on a Dickey-Fuller-type equation for the PD ratio such as:

$$\Delta x_t = \alpha + \rho x_{(dmd)t-1} + u_t, \tag{6}$$

where $x_{(dmd)t}$ is the demeaned value of the PD ratio using the sample mean $\overline{x}_t$ for the associated window of data.[13] In order to interpret the figure, it is useful to consider first the following equation for the ratio

$$x_t = \eta + \gamma_t x_{t-1} + u_t. \tag{7}$$

If we subtract x_{t-1} from both sides of the equation, we can write (7) as follows,

$$\Delta x_t = \eta + \rho_t \mu + \rho_t (x_{t-1} - \mu) + u_t, \tag{8}$$

where $\rho_t = (\gamma_t - 1)$ and μ is the constant attractor. Then, we can write the following identity

$$\mu \equiv \overline{x}_t + a_t,$$

where a_t is just the deviation of the sample mean associated with the window from the attractor. Substituting this identity into (8), we obtain a time-varying parameter version of (6):

$$\Delta x_t = \alpha_t + \rho_t (x_{t-1} - \overline{x}_t) + u_t, \tag{9}$$

where $\alpha_t = \eta + \rho_t \overline{x}_t$.

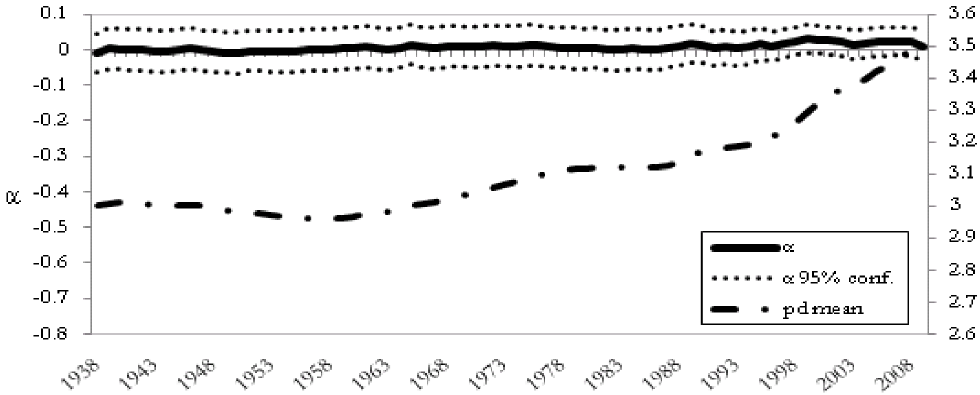
The information in Figure 2 and the implications of equations (7) to (9) can be used as preliminary evidence that the dynamics of the ratio might show asymmetry in the speed of adjustment towards an attractor. This information also challenges the usual estimate of the attractor as the historical mean since the rolling sample mean, $\overline{x}_t$, appears to follow an upward trend in the sample considered in this paper.

4 An MS model for adjusting the PD ratio

In the previous sections we have reviewed the literature that considers a nonlinear framework to investigate the stationarity of the PD ratio in light of the implications of the PV

[13] In this rolling-regression, we use a constant data size of 57 observations for each window. The choice of this number is determined by the number of observations available until the 1929 crash. The first window corresponds to an episode where the PD ratio seems to follow a stationary process fluctuating around the sample mean associated with this pre-crash period.

Panel A: α stability.



Panel B: ρ stability.

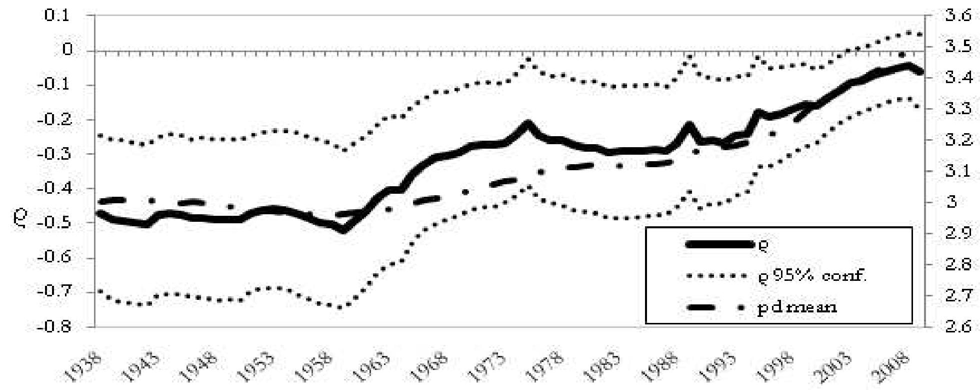


Figure 2: Rolling regression stability analysis.

Notes: This figure analyzes the stability of parameters α and ρ in equation (9). We report the estimated parameters and their respective 95% confidence intervals (in dashed lines) for rolling windows with a length of 57 years. The mean PD ratio or $\overline{x}_t$ denotes the rolling sample mean (roll.mean) and its scale is depicted on the right-hand side vertical axe of both panels.

model. We have also provided preliminary evidence on the possibility of an asymmetric speed of adjustment of the PD ratio towards a long-run attractor and how this attractor is not necessarily the average PD ratio. In this section, we explain the unique features of the MS method to explain the relation between prices and dividends in particular stock market episodes. In sum, the MS approach (Hamilton, 1989) provides a more flexible non-linear framework for the dynamics of the PD ratio than the TAR approach. In the TAR models reviewed in Section 2, the characterization of the alternative regimes is driven by the choice of a particular threshold variable and the transition function between regimes. However, on the one hand, the definition of the threshold variable chosen by the researcher clearly determines the features and the asymmetric behavior of the PD ratio associated with each regime. On the other hand, the number of regimes considered is limited by the features assumed by the researcher *a priori*, such as growing and decreasing markets, run-ups in prices and crashes, or bubble episodes, etc.

4.1 The MS setting

The MS approach for the dynamics of the PD ratio can be seen as a generalization of Eq. (5) given by the following model:

$$\Delta pd_t = \alpha + \rho_{s_t}(pd_{t-1} - \mu) + \sum_{j=1}^{l} \beta_j(\Delta pd_{t-j}) + \varepsilon_t, \tag{10}$$

where ε_t is assumed to be i.i.d. as a normal with mean zero and standard deviation σ.[14] In this framework, the variable characterizing the transition between regimes is not defined by the researcher. Instead, it is driven by an unobserved variable s_t that describes the state or regime of the process at time t. The latent variable s_t is the outcome of a k-regime Markov chain with s_t being independent of ε_t.[15,16]

As stated above, the basic difference between imposing a threshold variable as in TAR models and the k-regime MS model is that the latter does not impose any particular

[14] We will also briefly discuss the possibility of regime-dependent volatility σ_{s_t}. However, in order to maintain the parsimony of the model, and given the minor gains in fitting the actual data for this specification, we consider the model with only regime-dependent speed of reversion as the benchmark specification.

[15] The MS estimation methodology for a three-regime specification is briefly described in Appendix 1.

[16] As for an alternative specification with a regime-dependent attractor, it must be clear that the MS approach allows for transitory shifts in the characteristics of the dynamics which may have a relatively short duration. Frequent attractor switches that might be captured by an alternative MS specification clearly go against the concept of a long-run equilibrium motivated in the introduction. We believe that structural break approaches might be more suited for dealing with the hypothesis of a switching attractor, whereas TAR and MS approaches may be more suitable for capturing temporary switches in the speed of adjustment.

characteristic on each regime. Thus, this method allows us to analyze the unique characteristics of alternative stock market episodes in our sample. In particular, we investigate the possibility of having two and three regimes to capture these market episodes. Now, for each particular specification and for each state, we estimate the parameters characterizing the dynamics of the ratio (Eq. 10). These estimation results will allow us to (i) identify which episodes belong to each regime; (ii) assess which episodes exhibit a reverting behavior and which ones do not; and (iii) link those regimes to particular business cycle and stock market episodes previously discussed in the relevant literature.

By no means are we arguing that an MS approach is better than a TAR approach under all circumstances. There is always a trade-off between imposing more restrictions that might help to interpret the different regimes and imposing less restrictions giving more freedom to the data and trying later to match the characteristics of the regimes obtained. The MS approach followed in this paper should be viewed rather as a way of assessing, and perhaps challenging, some of the results and interpretations obtained in the recent literature by following alternative TAR approaches.

4.2 Evidence from a k-Regime MS Model

Table 2 shows the estimation results for the three-regime MS model in Eq. (10) for the full and pre-1993 samples. In order to evaluate the state identification, the characteristics of each regime and their relation to the business cycle are analyzed below. Figure 3 shows the estimated smoothed probabilities of being in each regime for the two alternative samples investigated under the benchmark specification (columns 2 and 7 of Table 2). Table 3 shows the contemporaneous correlation between the probabilities in Figure 3 and some business cycle and stock market characteristics. More precisely, the left-hand panel shows the correlations between the smoothed probabilities of each state with inflation, GDP growth and realized volatility of stock market excess returns, respectively. The middle panel shows the correlations between the smoothed probabilities of each regime with *moderate* and *extreme* events related to inflation, GDP growth and stock market volatility. Thus, an extreme event of high (low) inflation is defined as a period where inflation is above (below) the 90% (10%) percentile. A moderate event of high (low) inflation is defined as a period where inflation is above (below) the 70% (30%) percentile. Similarly, moderate and extreme events are defined for GDP growth and realized market volatility. Finally, the last panel completes the characterization of the alternative regimes identified by showing the correlations of each regime with different sets of event co-exceedances.

We find that the regimes are clearly identified by the three-regime MS benchmark specification. That is, at least one of the smoothed probabilities in each period is close to 1, which is confirmed by the low (adjusted) regime classification measure (RCM) reported in Table 2 and described in Appendix 2. Thus, regime 1 is a state that occurs occasionally

	Full Sample	Pre 1993
Mean Log-lik.	0.84	0.82
Adj RCM	7.21	9.55
Param.		
α	-0.77	0.01
	(0.67)	(0.09)
ρ_1	-0.08	0.15
	(0.05)	(0.10)
ρ_2	-0.14	-0.02
	0.03	0.05
ρ_3	-0.18	-0.18
	(0.03)	(0.06)
μ	9.07	4.88
	(4.34)	(0.67)
β_1	0.18	0.16
	(0.05)	(0.05)
σ_1	0.08	0.08
	(0.01)	(0.01)
$P_{\mathbf{11}}$	1.00	0.99
$P_{\mathbf{12}}$	0.00	0.01
$P_{\mathbf{21}}$	0.02	0.06
$P_{\mathbf{22}}$	0.95	0.91
$P_{\mathbf{31}}$	0.01	0.00
$P_{\mathbf{33}}$	0.96	0.97

Table 2: Estimated parameters and stationarity analysis results for the three-regime model.

Notes: This table reports the estimated parameters in equation (10), as well as their standard deviations (in parenthesis) for the full and the pre-1993 samples. The table also reports the adjusted RCM diagnostic tool briefly described in Appendix 2. P_{ij} are the components of the transition matrix P, defined as $P_{ij} = p(s_t = j, s_{t-1} = i)$, for $i, j = 1, 2, 3$ (see Appendix 1).

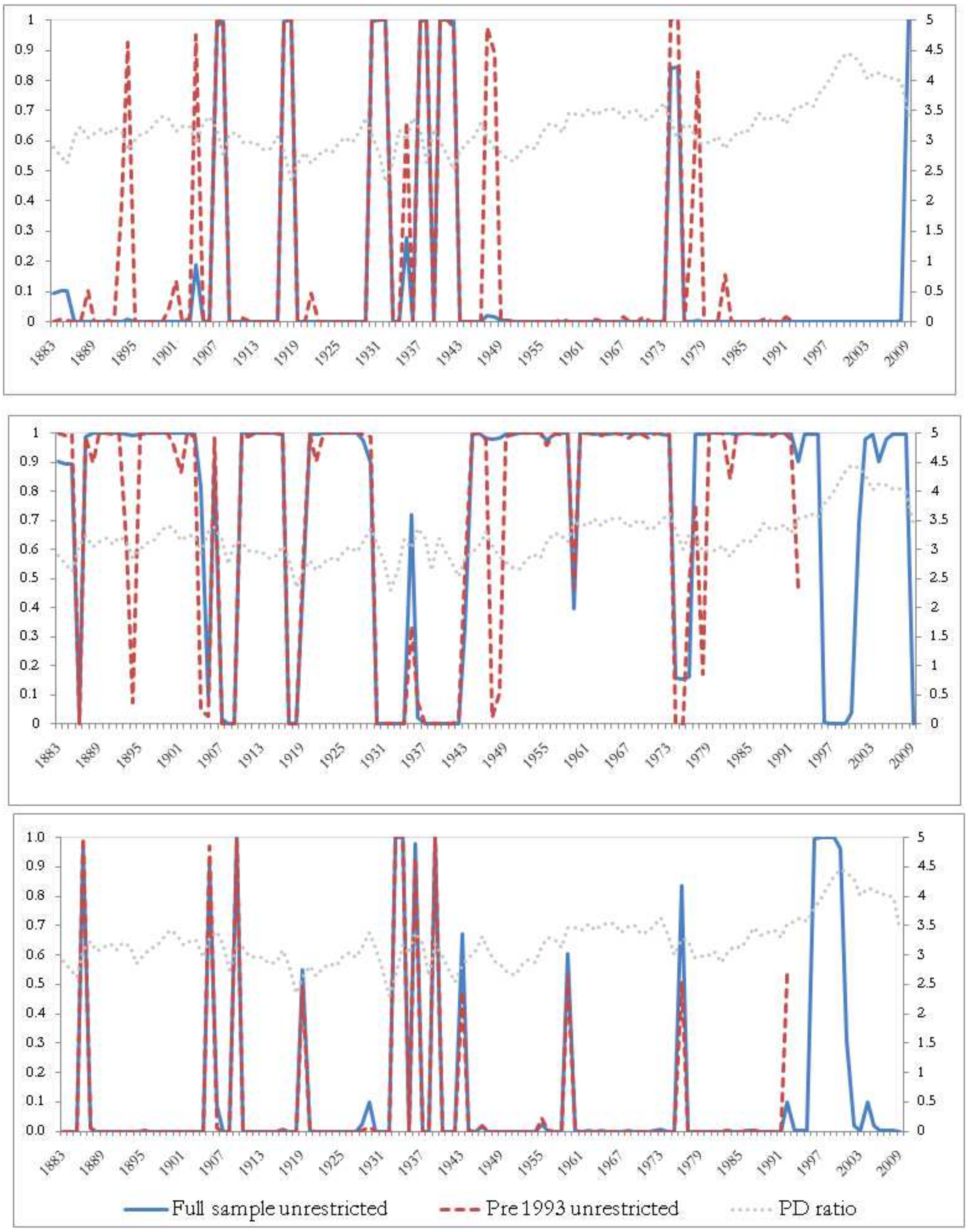


Figure 3: Smoothed probabilities for the three-regime benchmark specification.
Notes: This figure shows the smoothed probabilities of states 1, 2 and 3 for the PD ratio for the full sample and the pre-1993 sample for the model in Eq. (10) (benchmark model). The smoothed probabilities are calculated as in Hamilton (1989) and described in Appendix 1.

in very short episodes which never last more than 3 years (its average duration is 1.47 years), most of them before 1950, and most recently in 2009. This regime is a near (non-) stationary state, i.e., the speed of adjustment ρ_1, is either positive or negative depending on the sample considered, and in any case non-significant. This regime is associated with a few relevant market episodes involving large drops in the PD ratio that took place just after a strong run-up in stock prices relative to dividends, such as the 1929 crash and the subsequent crises (1929-1932) as well as the first oil crisis (1974-1975) and more recently, the post-Lehman Brothers episode during the latest recession. As can be seen in Table 3, this regime is correlated with episodes of high inflation, low economic activity and high realized volatility of excess stock returns. The co-exceedances analysis shows that this regime is also related to episodes of stagflation and high volatility in the stock market.[17,18]

Regime 2 is the most likely state in over two thirds of the sample.[19] In particular, this regime (with an average duration of 5.55 years) identifies two key historical episodes: the post-war period up to the mid 70's and the first part of the 90's boom from 1980 to 1995. When compared to the stock market or business cycle characteristics, this regime turns out to be in general not related to periods where any moderate or extreme events occur.

Finally, regime 3 is a very interesting state. Before 1995, this regime only occurs occasionally in very short episodes, which never last more than two years. All these episodes are related to temporary run-ups in prices with respect to dividends. Interestingly, it is clearly the most likely regime in the second part of the 90's boom, from 1996 to 2000. That is, this state identifies the episode with the highest slope of the PD ratio. Intuitively, one might think that if the estimated attractor is higher than the maximum value, as the upward drift of the PD ratio suggests, episodes of large and persistent growth of the PD ratio are related to a high speed of adjustment consistent with a highly reverting process to a high attractor. Interestingly, this highly reverting process associated with the second part of the 90's boom stands in sharp contrast with McMillan (2006), who identifies almost the same subsample period with a significant fall in the strength of the stationary relationship by considering the sample mean as an attractor.

Regime 3 exhibits quite distinctive features when compared to the particular business cycle characteristics around the episode 1996-2000. Table 3 shows that this regime is related to low inflation and high GDP growth. In addition, this third regime differs substantially from the other two regimes since it is also associated (i.e. positively correlated)

[17] There are two exceptions for the probability of being in regime 1 around episodes of stagflation. These exceptions are the 1929 crash and the year 2009, which are assigned to regime 1 but which are also characterized by extreme low inflation or even deflation.

[18] It is important to highlight that p_{11} (the probability of staying in regime 1) is very close to 1 suggesting an absorbing regime. Nevertheless, in unreported results, we find that regime characterization holds if p_{11} is restricted to be strictly lower than 1.

[19] This percentage is obtained following a rule of thumb by associating any time period with a particular regime only when the smoothed probability of being in that regime is higher than 80%.

	Business Cycle Characteristics			Extreme Events							
	Inflation	GDP growth	Market Volat.	Low infl.	Mid-lo infl.	High infl.	Mid-hi infl.	Low Δ GDP	Mid-lo Δ GDP	High Volat.	Mid-hi Volat.
Sample 1948:2009											
State 1	0.25**	−0.41***		0.18*	0.03	0.35***	0.16***	0.56***	0.34**		
State 2	0.04	0.07		−0.23**	−0.31**	−0.05	0.04	−0.22*	−0.01		
State 3	−0.20*	0.19*		0.14	0.33***	−0.18	−0.15	−0.13	−0.22**		
Sample 1964:2009											
State 1	0.24**	−0.52***	0.30**	0.42***	0.10	0.35***	0.16**	0.56***	0.44***	0.22*	0.22
State 2	0.11	0.11	−0.32**	−0.46***	−0.48***	−0.05	0.07	−0.22*	−0.09	−0.19	−0.43***
State 3	−0.29**	0.22*	0.16	0.24*	0.48***	−0.18*	−0.19*	−0.13	−0.20*	0.06	0.34***

Table 3: Contemporaneous correlations between the smoothed probabilities and stock market/business cycle characteristics.

Notes: This table shows the contemporaneous correlation coefficients between the smoothed probabilities of each state for the estimated parameters reported in column 1 of Table 2 and several stock market/business cycle characteristics. *,**, and *** represent significance at the 1, 5 and 10% levels, respectively. The correlations are reported for the business cycle characteristics, and the extreme and moderate events are defined as those values of each characteristic above or below the 10th (high, low) and 30th (Mid-Hi, Mid-lo) percentiles, respectively. In addition, we also report the correlations for the co-exceedances of events. The co-exceedances of events is defined as the coincidence of certain events in the same period of time. For example, the column called "stagflation+M-hi RV" reports the correlation between the smoothed probabilities of each state and episodes where the following three characteristics occur simultaneously: inflation is above the 70th percentile, GDP growth is below the 30th percentile and the realized market volatility is above the 70th percentile. The inflation is calculated as the annual change in the GDP deflator. The GDP growth is calculated as the annual change in GDP. The data source for these macroeconomic indicators is the Federal Reserve Bank in St. Louis (only available from 1948 onwards). The equity market realized volatility is calculated as the square root of the realized variance RV, where $RV = \sum_{i=1}^{N}(r_{m,i} - r_{f,i})^2$, where N is the number of days in each year, $r_{m,i} - r_{f,i}$ is the daily excess market return available on Kenneth French's web site. The data for the RV measure is only available from 1964 onwards. Given the low occurrence of co-exceedances at the 10th percentile level, this analysis is only reported for the 30th percentile level.

	Co-exceedances (of Extreme Events)			
	Stagflation (M-hi infl+ M-lo Δ GDP)	M-lo infl +M-hi Δ GDP	Stagflation +M-hi Volat.	M-lo infl. +M-hi Volat.
Sample 1948:2009				
State 1	0.40***	−0.08		
State 2	−0.12	−0.23*		
State 3	−0.13	0.32***		
Sample 1964:2009				
State 1	0.39***	−0.08	0.26**	0.16
State 2	−0.09	−0.25*	−0.06	−0.48***
State 3	−0.17	0.34***	−0.11	0.44***

Table 3: Continued. Contemporaneous correlations between the smoothed probabilities and stock market/business cycle characteristics

with episodes featuring low inflation, high economic growth and high realized volatility simultaneously. It is interesting to see how these characteristics (high real growth and increasing market risk) are clearly related to the technological boom.

In sum, although our estimation results support the evidence of nonlinearities in the dynamics of the ratio in Bohl and Siklos (2004), Coakley and Fuertes (2006) and McMillan (2006, 2007), these results show significant differences on the reversion features displayed by the alternative regimes. Thus, regime 1 is hardly identified with a stationary state, i.e., the estimated speed of adjustment for this regime is statistically non-significant for most specifications considered. Regime 2 is a stationary state where the reversion parameter ρ_2 is negative and significant. Finally, regime 3 is a stationary state featuring a high, significant speed of adjustment that identifies the episode 1996-2000. These differences in regime characterization crucially depend on the attractor considered. This comes as no surprise since we estimate this parameter while most of the related literature relies on the historical PD ratio as the long-run attractor. Thus, we obtain a (relatively) high and imprecise estimation of the attractor that leads to a characterization of the PD ratio dynamics that stands in sharp contrast with the one provided in this literature. Given the relevance of these results, next subsection assesses the importance of estimating the attractor for regime identification.

4.3 The importance of estimating the attractor

The different characteristics of the states identified with the MS approach depend highly on the estimated attractor. The attractor that we obtain turns out to be very large although imprecisely estimated for every specification (and sample) considered. Given that the estimated attractor is very high, the reader may worry that it really just functions as a time trend, and that the changing speeds of adjustment just capture changes in the trend. This is an easy hypothesis to investigate; we have simply added linear trends to each of their three regimes and (i) test the exclusion restrictions on the three trends, (ii) test the exclusion restrictions on the three adjustment terms. Our estimation results show that the linear trend coefficients are all close to zero and non-significant, whereas the adjustment coefficient estimates for states 2 and 3 remain significant as in the benchmark model.[20]

To understand the lack of precision in the estimation of the attractor parameter, it is useful to note that its estimated value is much larger than the maximum value for the PD ratio (4.45) and its sample mean (3.21). Then, under the view that the PD ratio is reverting to a high level of the attractor, it is reasonable to obtain an imprecise estimation of an attractor that has not yet been reached. However, a large value of the attractor makes sense if we consider that the attractor might not be directly linked to the historical mean when the PD ratio starts from a low level and the transition to the

[20] These results are not reported in order to save space and are available from the authors upon request.

long-run equilibrium or attractor is not symmetric. This hypothesis is supported by the long-run upward drift followed by the PD and PE ratios around the technological boom episode as shown in Section 3. This upward drift in the PD ratio may be the outcome of several forces such as the fact that firms have become less likely to pay dividends, the changes introduced in legislation such as the enactment of new SEC rules, a more favorable treatment for corporate taxes than for personal income tax, new participants in the market, and changing risk profiles as previously investigated in the literature.[21]

We further assess the importance of a high attractor estimate in the regime characterization of the PD ratio by estimating the MS model under three alternative upper-bound restrictions on the attractor. The first restriction imposes that the estimated attractor must be less or equal than 6 (i.e. a value which is much lower than the unrestricted estimate of μ and yet higher than the maximum value of the PD ratio). The second restriction imposes that μ must be lower than the historical maximum (i.e. $\mu \leq 4.45$). Finally, the third alternative restriction ($\mu \leq 4$) imposes that the attractor cannot be much higher than the historical mean as often assumed in the relevant literature. Table 4 and Figure 4 show the estimation results and state identification of the MS model under these three alternative restrictions on the attractor, respectively. Two important conclusions emerge from this analysis. First, the estimated value of the attractor reaches the upper bound imposed unless the upper-bound for the attractor is restricted to be close to the historical mean. Second, regime features and state identification are mostly similar to those obtained in the unrestricted MS model as long as the estimated attractor is allowed to be sufficiently different from the historical mean. In particular, the stationary regime 3 characterizes (almost uniquely) the second part of the 90's boom when the attractor is allowed to be equal or larger than the maximum value for the PD ratio as in the benchmark specification. However, when the attractor is restricted to be lower than 4, its point estimate is 3.66, close to the historical mean. Moreover, the state identification implied by this estimated attractor becomes similar to that obtained in the previous literature where the 90's boom is characterized by a non-stationary ($\rho_1 = 0.23$) state.

In sum, our analysis shows that the size of the attractor parameter is crucial in our understanding of the PD ratio dynamics. In particular, our investigation highlights the limitation of regime characterization in previous related literature. That is, by relying on the assumption that the historical mean of the PD ratio is a good proxy for this parameter previous literature provides, at least, a limited view of the PD ratio dynamics.

[21] See, for instance, Fama and French (2001), Grullon and Michaely (2002), Baker and Wurgler (2004), Chetty and Saez (2005), Boudoukh, Michaely, Richardson and Roberts (2007), and Hoberg and Prabhala (2009).

	$\mu \leq 6$	$\mu \leq 4.45$	$\mu \leq 4$
Mean Log-lik.	0.829	0.747	0.703
RCM	8.973	10.611	12.826
Param.			
α	−0.367	−0.095	−0.110
	(0.100)	(0.058)	(0.017)
ρ_1	−0.034	0.139	0.233
	(0.033)	(0.044)	(0.052)
ρ_2	−0.139	−0.092	−0.251
	(0.034)	(0.043)	(0.032)
ρ_3	−0.235	−0.293	−0.636
	(0.035)	(0.043)	(0.064)
μ	6.000	4.450	3.658
			(0.047)
β_1	0.185	0.207	0.058
	(0.049)	(0.053)	(0.052)
σ	0.084	0.089	0.092
	(0.006)	(0.008)	(0.006)

Table 4: Estimated parameters for the three-regime model. Upper bounds for the attractor
Notes: This table reports the estimated parameters in equation (10), as well as their standard deviations (in parenthesis) when an upper bound for parameter μ is imposed. The following bounded models are considered $\mu \leq 6$, $\mu \leq 4.45$ (the maximum value of the PD ratio for the full sample), and $\mu \leq 4$. The table also reports the adjusted RCM diagnostic tool briefly described in Appendix 2.

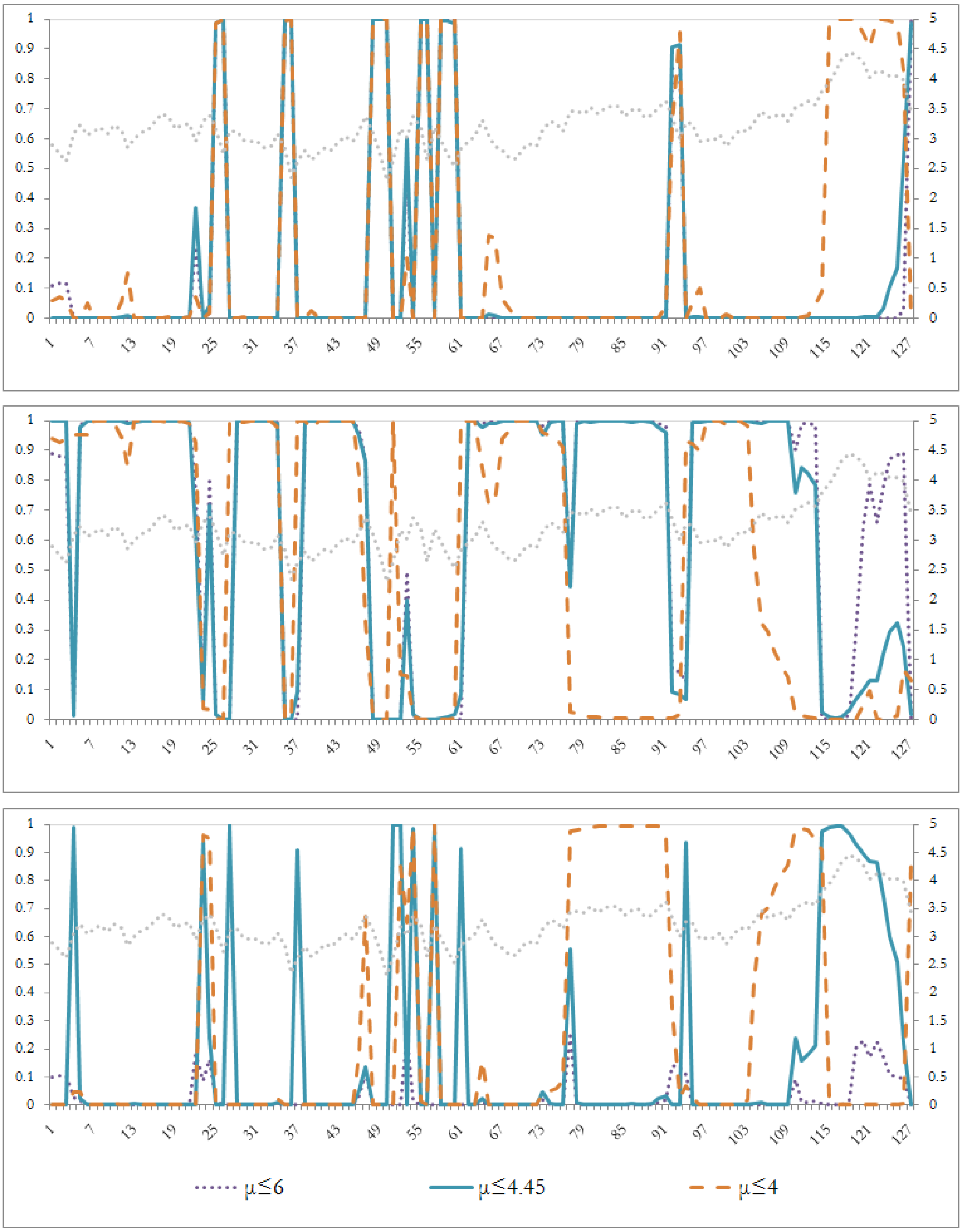


Figure 4: Smoothed probabilities for the three-regime model. Upper bounds for the attractor

Notes: This figure shows the smoothed probabilities of states 1, 2 and 3 for the PD ratio for the full sample and the pre-1993 sample for the model in Eq. (10) with upper bouds for the parameter μ. The smoothed probabilities are calculated as in Hamilton (1989) and described in Appendix 1.

5 Robustness Checks

In this section, we carry out an evaluation of the robustness of our results in four important dimensions. First, we investigate the possibility that the volatility of the ratio varies through time. Second, we investigate the more traditional two-regime approach for the asymmetry of the PD ratio reversion process in the context of an MS model. Third, we analyze the robustness of results to considering PD ratio quarterly data. Finally, we move on to analyze the estimated results in the case where the PE ratio is considered instead of the PD ratio.

5.1 Time-varying volatility

Table 5 shows the estimation results for the regime-dependent volatility specification. The estimated characteristics within the states for this alternative MS specification are quite similar to those under the benchmark specification. Under the regime-dependent volatility specification, regime 1 turns out to be a stationary state, but it is again characterized by the lowest speed of reversion and the highest volatility of the PD ratio, while regime 3 is characterized by the highest speed of reversion and the lowest volatility as under the benchmark specification.[22] Figure 5 shows the smoothed probabilities for the state-dependent volatility specification. A comparison of Figures 3 and 5 shows that the main difference among these specifications occurs in the identification of regime 3. Interestingly, regime 3 is the most likely state for a much longer period, from 1996 to 2008, under the regime-dependent volatility specification which reinforces the importance of considering a third regime in the analysis.

5.2 Two-state MS model

We also investigate the more traditional two-regime approach for the asymmetry of the PD ratio reversion process in the context of an MS model. The related literature has mainly focused on this two-regime specification due to the link established by researchers between regime identification and the specific characteristics attributed to each regime (for instance, the market episodes defined as bull and bear markets in Coakley and Fuertes, 2006; and the outer -reverting- and inner -random walk- regimes defined by McMillan,

[22] A simple likelihood ratio test shows that the augmented model with regime-dependent volatility hardly increases the estimation fit. The p-value for this test is 0.027 (formally, the null hypothesis of this test is $\sigma_1 = \sigma_2 = \sigma_3$). A likelihood ratio test is a limited, but rather useful test in this context since the number of regimes both under the null and under the alternative hypotheses is identical. A much more reliable comparison between alternative specifications is performed below using the Hamilton specification tests described in Appendix 2.

	$\rho_{s_t};\sigma_{s_t}$
Mean Log-lik.	0.87
Adj RCM	11.14
Param.	
$\boldsymbol{\alpha}$	-1.57
	(1.34)
$\boldsymbol{\rho}_1$	-0.21
	(0.06)
$\boldsymbol{\rho}_2$	-0.26
	0.03
$\boldsymbol{\rho}_3$	-0.31
	(0.02)
$\boldsymbol{\mu}$	9.39
	(4.67)
$\boldsymbol{\beta}_1$	0.15
	(0.04)
$\boldsymbol{\sigma}_1$	0.10
	(0.02)
$\boldsymbol{\sigma}_2$	0.08
	(0.01)
$\boldsymbol{\sigma}_3$	0.07
	(0.01)
$P_{\mathbf{11}}$	0.97
$P_{\mathbf{12}}$	0.03
$P_{\mathbf{21}}$	0.04
$P_{\mathbf{22}}$	0.95
$P_{\mathbf{31}}$	0.00
$P_{\mathbf{33}}$	0.98

Table 5: Estimated parameters and stationarity analysis results for the three-regime model with regime-dependent volatilities

Notes: This table reports the estimated parameters in equation (10), as well as their standard deviations (in parenthesis) for a model with regime-dependent volatility (σ_{st}). The table also reports the adjusted RCM diagnostic tool briefly described in Appendix 2. P_{ij} are the components of the transition matrix P, defined as $P_{ij} = p(s_t = j, s_{t-1} = i)$, for $i, j = 1, 2, 3$ (see Appendix 1).

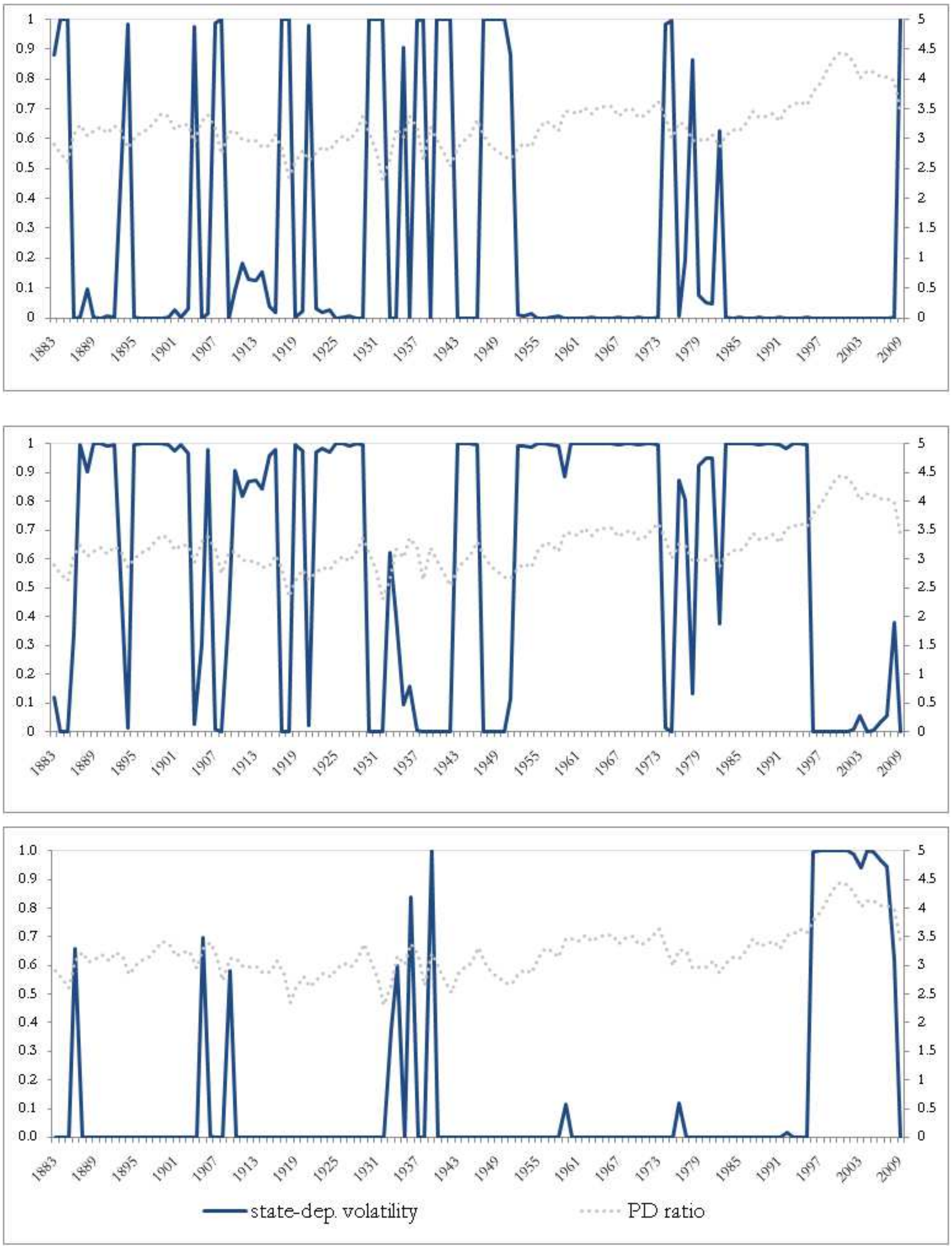


Figure 5: Smoothed probabilities for the three-regime model with regime-dependent volatility.

Notes: This figure shows the smoothed probabilities of states 1, 2 and 3 for the PD ratio for the full sample for a specification of model (10) with regime dependent volatility σ_{st}. The smoothed probabilities are calculated as in Hamilton (1989) and are described in Appendix 1.

	Full Sample	Pre 1993
Mean Log-lik.	0.55	0.58
Adj RCM	15.42	16.60
Param.		
$\boldsymbol{\alpha}$	-0.80	-0.38
	(1.05)	(0.23)
$\boldsymbol{\rho}_1$	-0.10	-0.11
	(0.07)	0.09
$\boldsymbol{\rho}_2$	-0.16	-0.30
	(0.04)	0.07
$\boldsymbol{\mu}$	8.97	4.82
	(6.57)	0.75
$\boldsymbol{\beta}_1$	0.02	0.03
	(0.07)	0.10
σ	0.11	0.11
	(0.01)	0.01
P_{11}	0.97	0.97
P_{12}	0.93	0.93

Table 6: Estimation results for the two-regime model.

Notes: This table reports the estimated parameters in equation (10), as well as their standard deviations (in parenthesis) for the full and pre-1993 samples for a two-regime model. The table also reports the Regime Classification Measure (RCM) diagnostic tool briefly described in Appendix 2. P_{ij} are the components of the transition matrix P, defined as $P_{ij} = p(s_t = j, s_{t-1} = i)$, for $i, j = 1, 2$ (see Appendix 1).

2006). However, we maintain that one of the relevant features of the MS method is that we are not assigning *a priori* features neither to the states identified nor to the transition mechanism from one state to another. In other words, our approach does not impose, prior to estimating, any of the alternative definitions of states and market episodes proposed in the related literature.

Table 6 shows the results for the two-regime MS model. The results for this specification as for the three-regime specification show a high estimated attractor, especially if the full sample is considered. This estimated attractor suffers again from a lack of precision. Figure 6 shows the smoothed probability of being in state 1 for the full and the pre-1993 samples under the two-regime specification. As can be seen from this figure, the two-regime specification displays a less clear state identification, so we do not go further analyzing the implications of this specification.[23]

[23] The estimation results as well as the state identification for the two-regime specification are also not robust to considering a regime-dependent volatility. Results for this specification are available upon request.

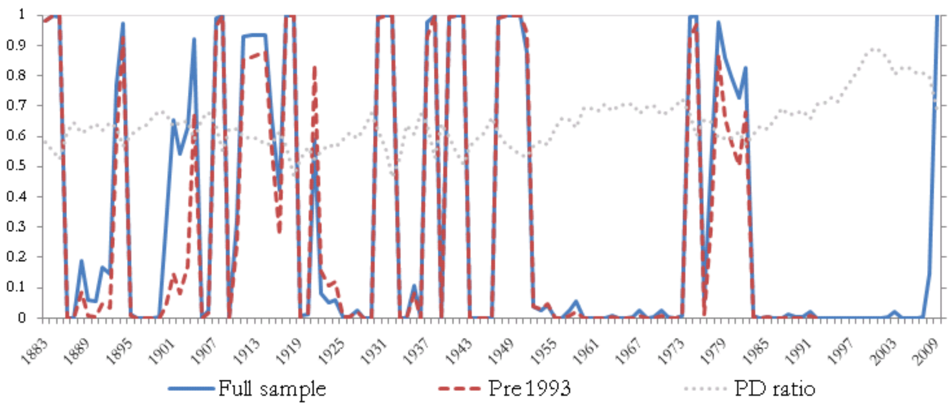


Figure 6: Two-regime model. Smoothed probability of state 1 for PD ratio.
Note: This figure compares the results for a two-regime model (Eq. 10) for the full and pre-1993 samples at the annual frequency.

In order to evaluate the correct specification and compare the alternative MS specifications, Table 7 includes the Lagrange multiplier tests proposed in Hamilton (1996). These tests are explained in more detail in Appendix 2. As can be seen from this table, the MS assumptions (absence of autocorrelation and heteroskedasticity in the form of ARCH structure within regimes) hold for almost all cases. There are two exceptions though. On the one hand, there is evidence of ARCH effects in regime 2 for the three-regime specification and on the other hand, there is evidence of autocorrelation for regime 2 in the two-regime specification at the 10% confidence level. In general, the results for the specification tests suggest that although the three-regime specification provides the most accurate and robust state identification, there is evidence of autocorrelation and heteroskedasticity structure in the residuals across states.[24]

5.3 Quarterly frequency

Table 8 shows the results for the three-regime specification using quarterly data and Figure 7 displays the smoothed probabilities for this data frequency. The overall effect of using a higher frequency data on the state identification appears clear from this figure. That is, it is harder to identify episodes of several quarters that are linked to one of the regimes since on the one hand regime switching is likely to be a low frequency feature that is harder to identify when quarterly data is considered. On the other hand, the regime switching is determined by the dynamics of a latent (unobservable) variable and its identification

[24] As pointed out by Hamilton (1996), "all of the tests are more difficult for a correctly specified model to pass than one would have anticipated based on an asymptotic distribution." However, the specification tests proposed in his paper are, as far as we know, the only tool specifically designed to test for the correct specification in the residuals for this type of models.

	3 States			2 States
	Benchmark	$\mu = 4.45$	ρ_{s_t}, σ_{s_t}	Benchmark
Autocorr. In Regime 1	0.07	0.00	0.24	2.32
	0.80	*0.98*	*0.63*	*0.13*
Autocorr. In Regime 2	1.48	1.22	0.07	0.97
	0.23	*0.27*	*0.79*	*0.33*
Autocorr. In Regime 3	0.05	0.11	0.68	
	0.82	*0.74*	*0.41*	
Autocorr. Across regimes	5.27	1.88	0.01	0.75
	0.02	*0.17*	*0.92*	*0.39*
ARCH effects in regime 1	0.03	0.19	0.03	1.89
	0.86	*0.66*	*0.86*	*0.17*
ARCH effects in regime 2	11.77	2.87	2.24	1.02
	0.00	*0.09*	*0.14*	*0.32*
ARCH effects in regime 3	0.01	0.37	0.51	
	0.92	*0.54*	*0.48*	
ARCH across regimes	11.19	2.83	3.06	1.73
	0.00	*0.10*	*0.08*	*0.19*

Table 7: Lagrange Multipler Specification tests. Annual Data. Full sample. Two and three regimes.

Notes: This table reports the Lagrange Multiplier test statistics for the specification of MS models proposed by Hamilton (1996) as well as the p-values (in italics). The p-values are reported for the F distributions with $(1, T-m+1)$ degrees of freedom, where T is the total sample size and m is the number of parameters estimated in each case. The Lagrange Multiplier Tests are performed with modified versions of the programs used in Hamilton (1996). These specification tests investigate the MS specification model in equation (10) for the two- and three-regime specifications against an alternative specification that allows for autocorrelation and ARCH structure in the residuals from one period to the next one. The statistics and the small sample correction used for these tests can be found in Hamilton (1996) and are described in more detail in Appendix 2.

	Full sample	Pre 1993
Mean Log-lik.	1.31	1.36
Adj RCM	16.08	17.89
Param.		
α	-0.11	-0.25
	(0.08)	(0.13)
ρ_1	0.01	-0.03
	(0.04)	(0.05)
ρ_2	-0.06	-0.12
	(0.03)	(0.03)
ρ_3	-0.13	-0.19
	(0.04)	(0.05)
μ	5.30	4.89
	(0.61)	(0.97)
β_1	-0.07	-0.10
	(0.10)	(0.06)
σ	0.05	0.05
	(0.01)	(0.00)
P_{11}	0.96	0.97
P_{12}	0.03	0.02
P_{21}	0.05	0.03
P_{22}	0.93	0.91
P_{31}	0.02	0.00
P_{33}	0.96	0.95

Table 8: Estimation results for quarterly data.

Notes: This table reports the estimated parameters in equation (10), as well as their standard deviations (in parenthesis) for the full and the pre-1993 samples for the quarterly frequency. The table also reports the adjusted RCM diagnostic tool briefly described in Appendix 2. P_{ij} are the components of the transition matrix P, defined as $P_{ij} = p(s_t = j, s_{t-1} = i)$, for $i, j = 1, 2, 3$ (see Appendix 1).

becomes harder when using quarterly data, which are noisier than annual data. Nevertheless, a third regime with a clearly significant speed of adjustment parameter is again attributed to the 1996-2000 episode when data at the quarterly frequency are considered. For the data at this frequency, the estimated attractor is again much higher than the historical mean, but close to the maximum value of the PD ratio.

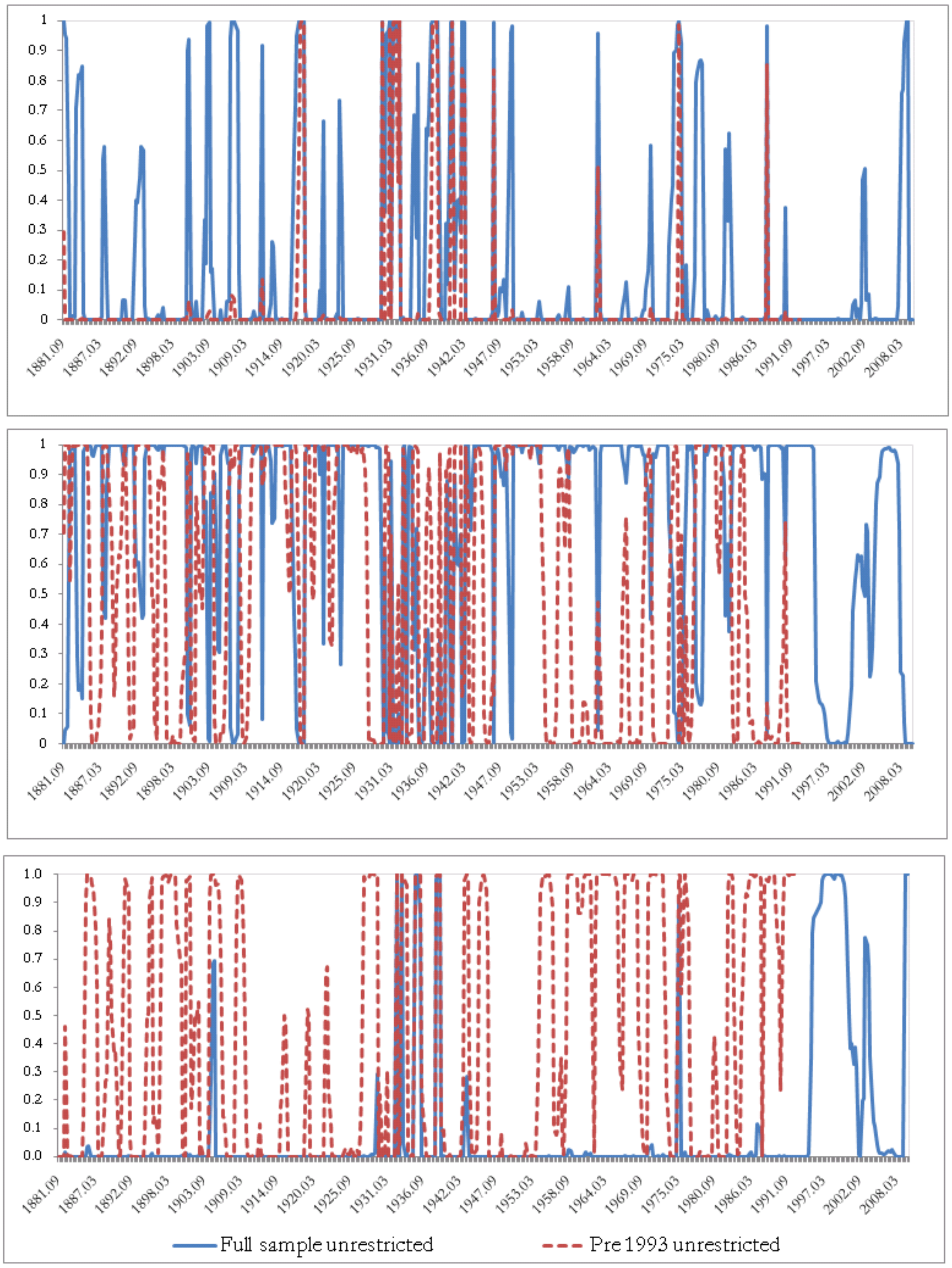


Figure 7: Smoothed probabilities for the benchmark specification using PD ratio quarterly data.

Notes: This figure shows the smoothed probabilities of states 1, 2 and 3 for the full sample and the pre-1993 sample for the benchmark model (10) using quarterly data for the PD ratio.

5.4 PE ratio

Finally, we also consider earnings as a proxy for payouts, and investigate an MS specification for the PE ratio. In the PE ratio, the proxy for payouts is an average of the earnings paid out in the last 10 years. By including 10-year average earnings we also incorporate a more stable measure of payouts that could be less affected by particular short-lived market episodes. Table 9 shows the estimated parameters for this ratio and Figure 8 shows the smoothed probabilities in comparison to those obtained for the PD ratio. The results suggest that some parameter estimates and the presence of a third regime associated with the 90's boom for the PE ratio are qualitatively similar features to those found for the PD ratio estimated under both the benchmark and the regime-dependent volatility specifications. In particular, the estimated attractor is much higher than the historical mean of the PE ratio, as was the case when considering the PD ratio above. Moreover, it is also fair to say that the state identification and the characteristics of the states are not robust across ratios. This comes as no surprise since, as can be observed from Figure 1, the two ratios exhibit quite different short-term dynamics. Nonetheless, as shown by the bottom graph in Figure 8, the two ratios identify the 90's boom with a stationary, highly reverting state, although that state is also related to other sample periods when considering the PE ratio. Thus, this highly reverting state is also present in other episodes of run-up in prices relative to earnings with relatively long duration.

6 Conclusions

Previous research related to the present value model and its implications for the stationarity of the price-dividend (PD) ratio has been nonconclusive to say the least when analyzing the reversion process of the PD ratio in a linear framework. In this paper, we find empirical evidence that the speed of adjustment of the PD ratio has not been constant over time. More precisely, that there are several transitory episodes in the US stock market characterized by a nonsignificant reversion to a long-run equilibrium or attractor. A finding that might explain the nonconclusive evidence on the stationarity of the PD ratio found in the related literature. Moreover, our empirical results show major changing episodes closely related to historical events in the US stock market. The nonlinear analysis of the reversion process of the PD ratio based on a three-regime Markov-switching (MS) model à la Hamilton (1989) carried out in this paper shows robust empirical evidence of switching regimes in the parameters characterizing the speed of adjustment of the PD ratio around a constant long-run attractor.

We find evidence of an asymmetric speed of adjustment identifying at least three relevant market episodes: the post-war period (up to 1975), the so called "90's boom" and the subprime crisis. A three-regime MS model shows a sharp regime classification.

		ρ_{s_t}, σ_{s_t}
Mean Log-likelihood	0.89	0.97
Adj RCM	11.55	14.11
Param.		
α	0.07	-0.02
	(0.16)	(0.09)
ρ_1	0.20	0.10
	(0.04)	(0.04)
ρ_2	0.04	0.00
	(0.05)	(0.03)
ρ_3	-0.06	-0.08
	(0.07)	(0.04)
μ	4.94	5.54
	(0.54)	(0.87)
β_1	0.03	0.09
	(0.07)	(0.04)
σ_1	0.08	0.11
	(0.01)	(0.02)
σ_2		0.05
		(0.00)
σ_3		0.07
		(0.01)
P_{11}	1.00	0.96
P_{12}	0.00	0.03
P_{21}	0.02	0.02
P_{22}	0.94	0.95
P_{31}	0.02	0.02
P_{33}	0.95	0.94

Table 9: Estimation results for the PE ratio based on annual data.

Notes: This table reports the estimated parameters in equation (10), as well as their standard deviations (in parenthesis) for the full sample for the PE ratio. The table shows the results for the benchmark model and for the alternative specification with regime-dependent volatility (σ_{st}). It also reports the adjusted RCM diagnostic tool briefly described in Appendix 2. P_{ij} are the components of the transition matrix P, defined as $P_{ij} = p(s_t = j, s_{t-1} = i)$, for $i, j = 1, 2, 3$ (See Appendix 1).

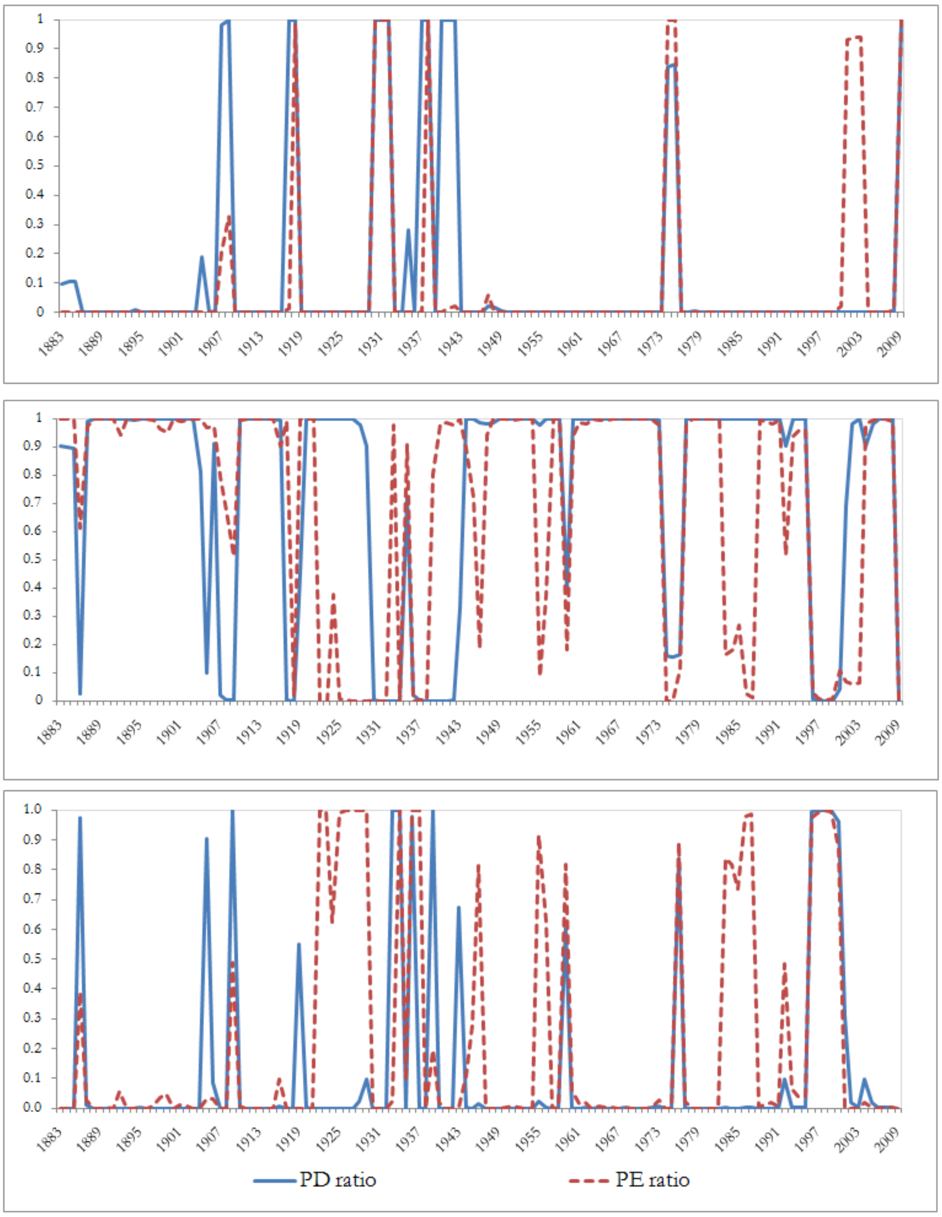


Figure 8: Three-regime model. Smoothed probabilities. PE ratio. Annual data

Notes: This figure shows the smoothed probabilities of states 1, 2 and 3 for PE ratio in comparison with those estimated for the PD ratio for the full sample. To save space, we only report the results for the benchmark model. However, the state identification for this ratio remains robust when regime-dependent volatility is considered.

Moreover, the three-regime model suggests that the post-war period (up to 1975) and the 90's boom episodes do not share the same characteristics and that the additional third state is needed to properly model the PD ratio dynamics, as suggested by the unique characteristics of the market in the last part of the sample. The post-war period is characterized by a stationary regime featuring low speed of adjustment to the attractor, and the 90's boom is divided into two parts. The first part exhibits features similar to the post-war period, whereas the second part (1996-2000), when the PD ratio grows faster and the apparent divergence between prices and dividends becomes greater, is characterized by a new regime with a stronger reversion to the large estimated attractor. Finally, the results from this specification suggest that the post-Lehman Brothers episode of the subprime crisis can be classified into a temporary nonstationary regime. This state identification is related to another interesting finding on the value of the attractor. Even when the attractor is poorly identified, by using alternative samples and MS specifications we robustly find higher estimated values of this parameter than those estimated in the previous related literature. More important, the remaining parameter estimates and the regime features identified are not affected whenever the attractor parameter is restricted by an upper-bound value as long as this value is close to the historical maximum reached by the PD ratio.

The empirical evidence of a high estimated attractor suggests that the apparent divergence between prices and dividends reflects the transition to a long-run equilibrium (attractor) that has not yet been reached. The evidence then suggests that the high increase of the PD ratio during the 90's boom is consistent with a higher speed of adjustment to the long-run equilibrium. This interpretation stands in sharp contrast to alternative interpretations of this episode suggested in the previous literature considering a regime-dependent speed of adjustment specification.

The evidence found for the three-regime MS model supports the idea that the stationarity hypothesis implied by the present value model is better understood as a long-run concept. While there are occasional episodes of non-stationary behavior of the PD ratio (as in regime 1), those episodes are followed with higher probability by stationary regimes (regime 2 in almost the entire sample and regime 3 in the last part of it). Non-stationary periods, such as the one related to the recent recession episode, can therefore be understood as temporary episodes if we consider a unique long-run equilibrium.

APPENDIX 1

State-dependent models: the Markov-switching approach

This appendix briefly describes the MS framework for estimating nonlinear models. Following Hamilton (1989, 1990), in a three-regime MS model, for estimating an equation such as (10), a transition matrix for s_t (the latent variable governing the switching-regime process) must be defined as:

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{21} & p_{31} \\ p_{12} & p_{22} & 1 - p_{31} - p_{33} \\ 1 - p_{11} - p_{12} & 1 - p_{21} - p_{22} & p_{33} \end{bmatrix},$$

where $p_{ij} = P(s_t = j \mid s_{t-1} = i, x_{t-1})$, and x_{t-1} is a vector containing all observations for the PD ratio obtained through date $t-1$. If at time t, $s_t = j$, the conditional density of Δx_t is given by:

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

where Θ is a vector containing the estimated parameters (depending on each case considered). 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 three-regime model, the conditional densities are gathered together on a vector denoted by $\boldsymbol{\eta}_t$

$$\boldsymbol{\eta}_t = \begin{bmatrix} f(\Delta x_t \mid x_{t-1}, s_t = 1; \Theta) \\ f(\Delta x_t \mid x_{t-1}, s_t = 2; \Theta) \\ f(\Delta x_t \mid x_{t-1}, s_t = 3; \Theta) \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ \frac{-(\Delta x_t - \alpha - \rho_1 (x_{t-1} - \mu) - \beta_1 (\Delta x_{t-1}))^2}{2\sigma^2} \right\} \\ \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ \frac{-(\Delta x_t - \alpha - \rho_2 (x_{t-1} - \mu) - \beta_1 (\Delta x_{t-1}))^2}{2\sigma^2} \right\} \\ \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ \frac{-(\Delta x_t - \alpha - \rho_3 (x_{t-1} - \mu) - \beta_1 (\Delta x_{t-1}))^2}{2\sigma^2} \right\} \end{bmatrix}.$$

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

$$\mathcal{L}(\Theta) = \sum_{t=1}^{T} \log f(\Delta x_t \mid x_{t-1}; \Theta), \tag{11}$$

where

$$f(\Delta x_t \mid x_{t-1}; \Theta) = \mathbf{1}'(\widehat{\xi}_{t|t-1} \odot \boldsymbol{\eta}_t),$$

1 is a (3x1) vector of ones, and $\widehat{\xi}_{t|t-1}$ are the filtered probabilities defined as

$$\widehat{\xi}_{t|t-1} = \mathbf{P} \cdot \widehat{\xi}_{t-1|t-1}, \tag{12}$$

where

$$\widehat{\xi}_{t-1|t-1} = \frac{(\widehat{\xi}_{t-1|t-2} \odot \boldsymbol{\eta}_{t-1})}{\mathbf{1}'(\widehat{\xi}_{t-1|t-2} \odot \boldsymbol{\eta}_{t-1})}. \tag{13}$$

The optimization algorithm works as follows. Given an initial value $\widehat{\xi}_{1|0}$, equations (13) and (12) can be used to calculate $\widehat{\xi}_{t|t-1}$ and $\widehat{\xi}_{t|t}$ for any t. Following Hamilton (1989), we choose set $\xi_{1|0}$ equal to the vector of unconditional probabilities, $\boldsymbol{\pi}$, determined by $\pi = (A'A)^{-1}A'e_4$, where $A = [I_4 - P, 1']'$ and e_4 denotes the fourth column of I_4 (i.e. the 4x4 identity matrix). The value of $\widehat{\xi}_{t|t-1}$ is introduced in (11) and the procedure iterates until Θ^* is found according to a predefined convergence criterion.

In addition to the filtered probabilities previously obtained for each t, as a by-product the procedure also finds the probability of being in each state given the information from the whole sample considered. These probabilities are called smoothed probabilities

$$p_{i,t} = P(s_t = i \mid x_T; \Theta).$$

Kim and Nelson (1999) suggest the following algorithm to compute the smoothed probabilities:

$$\widehat{\xi}_{t|T} = \widehat{\xi}_{t|t} \odot \left\{ \mathbf{P}' \cdot [\widehat{\xi}_{t+1|T} (\div) \widehat{\xi}_{t+1|t}] \right\},$$

where $(\div)$ denotes element-by-element division. From the filtered probabilities, one can obtain the vector $\widehat{\xi}_{T|T}$ and iterate backward to obtain the smoothed probabilities for each t.

APPENDIX 2

Model diagnostics

Two types of test are commonly used in the relevant literature for evaluating performance when estimating MS models. The first type assesses the correct specification of the model. The Lagrange Multiplier test used in this paper compares the MS specification against one where the errors potentially show autocorrelation or heteroskedastic structure. The second type of test assesses the ability of the model to correctly identify states. More precise, we consider regime classification measures (RCM) to evaluate whether the model is clearly able to attribute a regime to each period of time.

- **Lagrange Multiplier Tests**

  Lagrange multiplier tests for MS models were proposed by Hamilton (1996). Of all these tests, we apply only those for testing for autocorrelation in each regime and across regimes, and those for ARCH effects in residuals (within and across regimes). In all tests displayed in Table 7, the null hypothesis is the model in equation (10). The alternative hypotheses tested are:

  - For autocorrelation within regimes:

  $H_A : (\varepsilon_t \mid x_{t-1}, s_t = j, s_{t-1} = i, s_{t-2} = k, ...; \Theta) \sim N[\delta_{[s_t=i, s_{t-1}=i]}\phi_i \varepsilon_{t-1}, \sigma]$,

  $H_0 : \phi_i = 0$.

  - For autocorrelation across regimes:

  $H_A : (\varepsilon_t \mid x_{t-1}, s_t = j, s_{t-1} = i, s_{t-2} = k, ...; \Theta) \sim N[\phi\varepsilon_{t-1}, \sigma]$,

  $H_0 : \phi = 0$.

  Similarly, for the possible structure in the variance of the residuals or ARCH test:

  - For ARCH effects (across regimes)

  $H_A : (\varepsilon_t \mid x_{t-1}, s_t = j, s_{t-1} = i, s_{t-2} = k, ...; \Theta) \sim N[0, h_t]$, where $h_t = \sigma[1 + \frac{\xi(\varepsilon_{t-1})^2}{\sigma}]$,

  $H_0 : \xi = 0$.

  The statistics obtained by Hamilton (1996) are distributed as a $\chi^2(1)$. We also use a small sample correction for these tests suggested by Hamilton (1996). With the correction proposed, the statistics are distributed as $F(1, (T - m + 1))$, where m is the number of parameters estimated in each specification and T is the total sample size.[25]

[25] Although Hamilton (1996) points out that the bias in small samples is less significant for samples of size 100 or more, which is our case, we still perform the correction for small sample bias. Inference based on χ^2 distributions is in any case close to the inference based on F given our sample size.

- **Regime classification measures**

  Ang and Bekaert (2002) suggest a summary statistic that assesses the regime classification quality provided by a k-regime MS model. The RCM is defined as

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

  where $p_{i,t}$ is the smoothed probability of being in regime i in period t as defined in Appendix 1. This measure captures the fact that if at least one of the smoothed probabilities in t is close to 0 for every t, the RCM will also be close to 0. In this case, regimes are properly identified and the model provides a good regime classification. If regimes are not well identified, the probabilities of being in a particular regime will be far from 1, and close to $1/k$ in the worst possible scenario. Thus, the RCM will be close to 100 in this case. The RCM measure is not always useful. For instance, it is not useful for comparing MS models with different numbers of regimes. Moreover, for $k > 2$, the RCM does not punish the fact that as more regimes are included the probability of at least one of those regimes being close to zero is always higher, but this does not necessarily mean that the model correctly identifies at least one of the regimes. Baele (2005) proposes an extended RCM for k-regime models that is equivalent to Ang and Bekaert's (2002) measure when $k = 2$. His measure allows for a comparison between models with different numbers of regimes and correctly captures the case where the model is clearly identifying one state in each period. The RCM2 proposed by Baele (2002) is defined as:

$$RCM2 = 100(1 - \frac{k}{k-1} \frac{1}{T} \sum_{t=1}^{T} \sum_{i=1}^{k} (p_{i,t} - \frac{1}{k})^2).$$

  The RCM reported in the tables is a modified version of RCM and RCM2. It is actually equivalent to RCM and RCM2 when $k = 2$. This new measure also shares with RCM2 two desired characteristics for an RCM. First, it is useful for comparing models with different numbers of regimes. Second, it provides a measure closer to 0 only when the model correctly identifies one regime in each period and a measure closer to 100 when no information about the regime identification is obtained. The adjusted RCM is defined as:

$$Adj\ RCM = 100(\frac{k}{k-1})^k \frac{1}{T}(\sum_{t=1}^{T} \prod_{i=1}^{k} (1 - p_{i,t})).$$

References

[1] Ang, A. and Bekaert G. (2002) "Regime switches in interest rates", *Journal of Business and Economic Statistics* 20, 163-182.

[2] Baele, L. (2005) "Volatility spillover effects in European equity markets: Evidence from a regime-switching model", *Journal of Financial and Quantitative Analysis* 40.

[3] Baker, M., and Wurgler, J. (2004) "A Catering Theory of Dividends," *Journal of Finance* 59(3), 1125-1165.

[4] Bohl, M.T., and Siklos, P.L. (2004) "The present value model of US stock prices redux: a new testing strategy and some evidence", *The Quarterly Review of Economics and Finance* 44, 208-223.

[5] Bonomo, M., and Garcia, R. (1994) "Can a well-fitted equilibrium asset-pricing model produce mean-reversion?" *Journal of Applied Econometrics* 9, 19-29.

[6] Boudoukh, J., M., Michaely, R., Richardson, M., and Roberts, M. (2007) "On the importance of measuring payout yield: implications for empirical asset pricing", *Journal of Finance* 62, 877-915.

[7] Campbell, J.J. and Shiller, R.J.(1988a) "The dividend-price ratio and expectations of future dividends and discount factors", *Review of Financial Studies* 1, 195-228.

[8] Campbell, J.J. and Shiller, R.J.(1988b) "Stock prices, earnings, and expected dividends", *Journal of Finance* 48, 661-676.

[9] Cechetti, S.G., Lang, P.S., and Mark, N.C. (1990) "Mean reversion in equilibrium asset prices", *American Economic Review* 80, 398-418.

[10] Chetty, R., and Saez, E. (2005) "Dividend Taxes and Corporate Behavior: Evidence from the 2003 Dividend Tax Cut", *The Quarterly Journal of Economics*, 120(3), 791-833.

[11] Coakley, J. and Fuertes, A. (2006) "Valuation ratios and price deviations from fundamentals", *Journal of Banking and Finance* 30, 2325-2346.

[12] Cochrane, J.H. (1992) "Explaining the variance of price-dividend ratios ", *Review of Financial Studies* 5, 243-280.

[13] Cochrane, J.H. (2001) *Asset Pricing Theory.* Princeton University Press, Princeton.

[14] Cochrane, J.H. (2008) "The Dog That Did Not Bark: A Defense of Return Predictability", *Review of financial Studies* 21, 1533-1575.

[15] Drifill, J., and Sola, M. (1998) "Intrinsic bubbles and regime-switching", *Journal of Monetary Economics* 42, 357-373.

[16] Enders, W., and Granger, C.W.J. (1998) "Unit-root tests and asymmetric adjustment with an example using the term structure of interest rates", *Journal of Business and Economic Statistics* 16, 304-311.

[17] Evans, M.D.D. (1998) "Dividend variability and stock market swings", *Review of Economic Studies* 65, 711-740.

[18] Fama, E.F., and French, K.R. (2001) "Disappearing dividends: changing firm characteristics or lower propensity to pay", *Journal of Financial Economics* 60, 3-43.

[19] Froot, K.A., and Obstfeld, M. (1991) "Intrinsic bubbles: The case of stock prices", *American Economic Review* 81, 1189-1214.

[20] Glassman, J.K., and Hasset, K.A. (1999). *Dow 36,000: The New Strategy for Profiting From the Coming Rise in the Stock Market.* New York: Times Books.

[21] Grullon, G. and Michaely, R. (2002) "Dividends, Share Repurchases, and the Substitution Hypothesis," *Journal of Finance*, 57(4), 1649-1684.

[22] Gutiérrez, M.J., and Vázquez, J. (2004) "Switching equilibria: The present value model for stock prices revisited", *Journal of Economics Dynamics and Control* 28, 2297 – 2325

[23] Hamilton, J.D. (1989) "A new approach to the economic analysis of nonstationary time series and the business cycle", *Econometrica* 57, 331-339.

[24] Hamilton, J.D. (1990) "Analysis of time series subject to changes in regime", *Journal of Econometrics* 45, 39-70.

[25] Hamilton, J.D. (1996) "Specification testing in Markov-switching time-series models", *Journal of Econometrics* 70, 127-157.

[26] Hoberg, G., and Prabhala, N.R. (2009) "Disappearing Dividends, Catering, and Risk", *Review of Financial Studies* 22, 79-116.

[27] Kapetanios, G. , Shin, Y. Snell, A. (2006) "Testing for cointegration in nonlinear smooth transition error-correction models", *Econometric Theory* 22, 279-303.

[28] Kim, C.J., and Nelson C.R. (1999) *State-Space Models with Regime-Switching: Classical and Gibbs-Sampling Approaches with Applications.* The MIT Press, Cambridge.

[29] Lettau, M. and Ludvigson, S. (2005) "Expected returns and expected dividend growth", *Journal of Financial Economics* 76, 583-626.

[30] Lettau, M., and Van Nieuwerburgh, S. (2008) "Reconciling the Return Predictability Evidence", *Review of Financial Studies* 21,1607-1652.

[31] McMillan, D. G. (2006) "The price-dividend ratio and limits to arbitrage: Evidence from a time-varying ESTR model," *Economic Letters* 91, 408-412.

[32] McMillan, D. G. (2007) "Bubbles in the dividend-price ratio? Evidence from an asymmetric exponential smooth-transition model", *Journal of Banking and Finance* 31, 787-804.

[33] Shiller, R.J. (1981) "Do stock prices move too much to be justified by subsequent changes in dividends?", *American Economic Review* 71, 421-436.

[34] Shiller, R. (1989) *Stock Market Volatility.* MIT Press, Cambridge.

[35] Timmermann, A. (2001) "Structural breaks, incomplete information, and stock prices", *Journal of Business and Economic Statistics* 19, 299-314.

[36] Tong, H. (1993) *Non-Linear Time Series*. Oxford University Press, Oxford.

[37] Veronessi, P. (1999) "Stock market overreaction to bad news in good times: A rational expectations equilibrium model", *Review of Financial Studies* 5, 975-1007.