

# Learning, structural instability and present value calculations

Hashem Pesaran

(University of Cambridge and USC)

Davide Pettenuzzo

(University of Bocconi and Bates and White)

Allan Timmermann

(University of California San Diego)



Discussion Paper  
Series 1: Economic Studies  
No 27/2006

**Editorial Board:**

Heinz Herrmann  
Thilo Liebig  
Karl-Heinz Tödter

Deutsche Bundesbank, Wilhelm-Epstein-Strasse 14, 60431 Frankfurt am Main,  
Postfach 10 06 02, 60006 Frankfurt am Main

Tel +49 69 9566-1

Telex within Germany 41227, telex from abroad 414431, fax +49 69 5601071

Please address all orders in writing to: Deutsche Bundesbank,  
Press and Public Relations Division, at the above address or via fax +49 69 9566-3077

Reproduction permitted only if source is stated.

ISBN 3-86558-189-7 (Printversion)

ISBN 3-86558-190-0 (Internetversion)

This paper was presented at the 8th Bundesbank spring conference (May 2006) on “New Developments in Economic Forecasting”. The views expressed in the paper are those of the authors and not necessarily those of the Bundesbank.

## **New Developments in Economic Forecasting** 8th Bundesbank Spring Conference

(Preliminary programme as at 10 March 2006)

Organisers: Michael Binder, Jörg Breitung, Heinz Herrmann

Eltville, 05/06 May 2006

### Friday, 5 May 2006

9:00 - 9:30      **Introductory remarks by Axel A Weber** (*Deutsche Bundesbank*)

9:30 - 10:45    **Multiple structural breaks, forecasting and present value calculations**  
Hashem Pesaran (*University of Cambridge*)

Discussant:                      Uwe Hassler (*Johann Wolfgang Goethe-  
University*)  
Siem Jan Koopman (*Free University  
Amsterdam*)

10:45 – 11:00    *Coffee break*

11:00 - 12:15    **Testing predictive ability using estimated co integration  
relationships: Business as usual?**  
Lutz Kilian (*University of Michigan*)

Discussant:                      Hans-Eggert Reimers (*Hochschule Wismar*)  
Andreas Beyer (*European Central Bank*)

12:15 – 14:00

*Lunch*

14:00 – 15:00

**Global yield curve dynamics and interactions**

Francis X Diebold (*University of Pennsylvania*)

Discussant:

Stefan Mittnik (*Ludwig-Maximilians-University  
Munich*)

Joachim Grammig (*Eberhard-Karls-University  
Tübingen*)

15:00 - 15:15

*Coffee break*

15:15 - 16:30

**A benchmark for models of growth and inflation**

Massimiliano Marcellino (*Università Bocconi*)

Discussant:

George Kapetanios (*University of London*)

Todd Clark (*Federal Reserve Bank of Kansas  
City*)

16:30 - 17:45

**Forecasting with panel data**

Badi Baltagi (*Texas A&M University*)

Discussant:

Helmut Herwartz (*Christian-Albrechts-University  
Kiel*)

Jean-Pierre Urbain (*University Maastricht*)

20:00

*Dinner hosted by Axel A Weber (Deutsche Bundesbank)*

Speaker:

Zdenek Tuma (*Czech National Bank*)

Saturday, 6 May 2006

9:30 - 10:45

**Did the ECB make a difference for Euro area business cycle?**

Fabio Canova\* (*Universitat Pompeu Fabra*)

Matteo Ciccarelli (*European Central Bank*)

Eva Ortega (*Bank of Spain*)

Discussant:

Sandra Eickmeier (*Deutsche Bundesbank*)

Domenico Giannone (*Université Libre de  
Bruxelles*)

10:45 - 11:00

*Coffee break*

11:00 - 12:15

**Bayesian density forecasting with best performing priors determined  
by entropic tilting**

Charles Whiteman\* (*University of Iowa*)

Kurt Lewis (*University of Iowa*)

Discussant:

John Geweke (*University of Iowa*)

Stephane Adjemian (*CEPREMAP*)

12:15 - 13:30

*Lunch*

13:30 - 14:45

**Shrinkage methods for forecasting using many predictors**

Mark Watson\* (*Princeton University*)

James Stock (*Harvard University*)

Discussant:

Gary Koop (*University of Leicester*)

Carlo Favero (*Universitat Pompeu Fabra*)



14:45 - 16:00

**Forecasting using a large number of predictors**

Lucrezia Reichlin (*European Central Bank*)

Discussant:

Christian Schumacher (*Deutsche Bundesbank*)

Peter Vlaar (*Dutch Central Bank*)

16:00

**Concluding remarks by** Hermann Remsperger (*Deutsche Bundesbank*)

# Learning, Structural Instability and Present Value Calculations\*

Hashem Pesaran

University of Cambridge and USC

Davide Pettenuzzo

University of Bocconi and Bates and White

Allan Timmermann

University of California San Diego

January 2006

## Abstract

Present value calculations require predictions of cash flows both at near and distant future points in time. Such predictions are generally surrounded by considerable uncertainty and may critically depend on assumptions about parameter values as well as the form and stability of the data generating process underlying the cash flows. This paper presents new theoretical results for the existence of the infinite sum of discounted expected future values under uncertainty about the parameters characterizing the growth rate of the cash flow process. Furthermore, we explore the consequences for present values of relaxing the stability assumption in a way that allows for past and future breaks to the underlying cash flow process. We find that such breaks can lead to considerable changes in present values.

JEL Classifications: C11, G12, G22

Keywords: present value, stock prices, structural breaks, Bayesian learning

---

\*We would like to thank Ron Smith and Martin Weitzman for their comments on a preliminary version of this paper.



## Non-technical summary

Present value relations play a key role in economics and finance and are used in testing the permanent income hypothesis, in standard inventory models and to calculate the present value of assets such as stocks and bonds. Computing present values requires forecasting a stream of future values of the variable of interest at horizons that can be long, but finite (as in the case of bonds) or even infinitely long (as in the case of stocks). It is customary in such calculations to assume that the underlying driving process follows a simple ARMA process with stable and known parameter values. This assumption is a gross oversimplification in almost any realistic economic context and so it becomes important both to dispense with the assumption of known parameters and to consider the possibility of past and future breaks in the data generating process of the driving variable. As an example of a present value relationship, we study in this paper long-lived assets such as stocks. Stocks are claims on unknown future dividends and so the stock price at any point in time must reflect the present value of the expected future dividends. A key question that investors are faced with is therefore how to compute expected values of future dividends in the presence of the considerable uncertainty surrounding not just dividends in the near future but dividends at very distant future points in time. In particular, how high is the growth rate of future dividends likely to be and how much does it vary through time? These are key issues that investors must answer when pricing long-lived assets with unknown future payoffs. We shall consider the problem of present value calculations under a variety of circumstances. To begin with we assume that the process of the driving variable is known with stable parameters but consider the implications of incomplete learning and parameter uncertainty that arises when dividends or incomes are predicted into an infinite future from a finite past. In practice, dividends or labour income processes are unlikely to remain stable and may be subject to structural breaks. Therefore we explore how to forecast future dividends and compute the present value of dividends in the context of a model where the dividend growth process is subject to occasional structural breaks. Such breaks give rise to considerable uncertainty about the stock price when compared to a model that ignores breaks although the latter, as we argue in this paper, is clearly mis-specified. Our analysis uses the hierarchical hidden Markov chain model introduced in Pesaran, Pettenuzzo and Timmermann (Review of Economic Studies 2006, forthcoming) for the purpose of forecasting time-series that are subject to multiple breaks. Using US dividend data we find that the present value stock price is very sensitive to the underlying modelling assumptions for the dividend process. In particular, it depends on whether the possibility of past breaks during the historical sample is considered and also whether future breaks are allowed for.

## Nicht technische Zusammenfassung

Barwertfunktionen spielen eine wichtige Rolle im Wirtschafts- und Finanzwesen und werden zur Überprüfung der permanenten Einkommenshypothese, im Rahmen standardisierter Bestandsmodelle und zur Berechnung des inneren Wertes etwa von Aktien und Anleihen herangezogen. Zur Berechnung von Barwerten ist es erforderlich, die zukünftigen Werte der Renditevariablen auf lange, aber zeitlich begrenzte Sicht (z. B. bei Anleihen) oder sogar auf unbegrenzte Dauer (z. B. bei Aktien) zu prognostizieren. In der Regel wird bei solchen Berechnungen unterstellt, dass die zugrunde liegenden Wirkungsmechanismen einem einfachen ARMA-Prozess mit stabilen und bekannten Parameterwerten folgen. Diese Annahme stellt in fast allen Fällen eine grobe Vereinfachung dar, weshalb es wichtig ist, auf die Annahme bekannter Parameter zu verzichten und die Möglichkeit zurückliegender und zukünftiger Brüche im Datengenerierungsprozess der entscheidenden Variablen in Betracht zu ziehen. Im vorliegenden Papier werden als Beispiel langfristige Vermögenswerte wie Aktien untersucht. Aktien beinhalten Ansprüche auf unbekannte zukünftige Dividendenzahlungen, weshalb sich in den Aktienkursen zu jedem Zeitpunkt der Barwert der in Zukunft erwarteten Dividenden widerspiegeln muss. Eine wichtige Frage, mit der sich Anleger beschäftigen, besteht also darin, wie der Wert erwarteter künftiger Dividendenausschüttungen berechnet werden kann, wenn nicht nur die Dividenden in naher Zukunft, sondern vor allem jene, die in ferner Zukunft fällig werden, mit erheblicher Unsicherheit behaftet sind. So stellt sich vor allem die Frage nach der Höhe der Wachstumsrate künftiger Dividendenzahlungen und wie sehr die Rate im Zeitverlauf variiert. Diese zentralen Fragen sind von den Anlegern zu klären, wenn sie langfristige Vermögenswerte mit unbekanntem zukünftigen Dividendenausschüttungen bewerten. In diesem Papier wird die Berechnung des Barwerts unter verschiedenen Voraussetzungen betrachtet. Zunächst wird unterstellt, dass der Prozess, dem die entscheidenden Variablen unterliegen bekannt ist; es werden jedoch auch die Implikationen unvollständiger Erkenntnisse und unsicherer Parameter, welche sich aus der Prognose von Dividenden oder Erträgen aus der endlichen Vergangenheit in eine unendliche Zukunft ergeben, betrachtet. In der Praxis dürften Dividenden oder Einkommen Schwankungen unterliegen und können Strukturbrüchen ausgesetzt sein. Aus diesem Grund wird hier untersucht, wie zukünftige Dividenden prognostiziert und der Barwert von Dividenden im Rahmen eines Modells berechnet werden können, bei dem das Dividendenwachstum gelegentlichen Strukturbrüchen unterliegt. Derartige Brüche lassen erhebliche Unsicherheit über den Aktienkurs entstehen, wenn man das Modell mit einem, das solche Brüche ignoriert, vergleicht. Allerdings wäre letztgenanntes Modell, so die Argumentation in diesem Papier,

eindeutig fehlspezifiziert. Um Zeitreihen zu prognostizieren, die mehrfachen Brüchen unterliegen, zieht die hier vorgenommene Analyse das von Pesaran, Pettenuzzo und Timmermann (Review of Economic Studies 2006, im Erscheinen) eingeführte hierarchische „versteckte“ Markov-Ketten-Modell heran. Bei Verwendung von US-Dividendendaten zeigt sich, dass der innere Wert des Aktienkurses stark von den unterstellten Modellannahmen für den Dividendenprozess abhängig ist – insbesondere davon, ob die Möglichkeit zurückliegender Brüche während des Beobachtungszeitraums in der Vergangenheit berücksichtigt wird, und auch, ob zukünftige Brüche einbezogen werden.



## 1. Introduction

Present value relations play a key role in economics and finance and are used in testing the permanent income hypothesis, in standard inventory models and to calculate the present value of assets such as stocks and bonds. Computing present values requires forecasting a stream of future values of the variable of interest at horizons that can be long, but finite (as in the case of bonds) or even infinitely long (as in the case of stocks). It is customary in such calculations to assume that the underlying driving process follows a simple ARMA process with *stable* and *known* parameter values. This assumption is a gross oversimplification in almost any realistic economic context and so it becomes important both to dispense with the assumption of known parameters and to consider the possibility of past and future breaks in the data generating process of the driving variable.

As an example of a present value relationship, we study in this paper long-lived assets such as stocks. Stocks are claims on unknown future dividends and so the stock price at any point in time must reflect the present value of the expected future dividends. A key question that investors are faced with is therefore how to compute expected values of future dividends in the presence of the considerable uncertainty surrounding not just dividends in the near future but dividends at very distant future points in time. In particular, how high is the growth rate of future dividends likely to be and how much does it vary through time? These are key issues that investors must answer when pricing long-lived assets with unknown future payoffs.

We shall consider the problem of present value calculations under a variety of circumstances. To begin with we assume that the process of the driving variable is known with stable parameters but consider the implications of incomplete learning and parameter uncertainty that arises when dividends or incomes are predicted into an infinite future from a finite past. In the case of geometric random walks with normally distributed innovations we show that expected present value can be divergent even if the parameter uncertainty is confined to the mean of the dividend process. This is a new finding and differs from similar results by Geweke (2001) and Weitzmann (2005) who show that the expected utility does not exist in the case of power utility functions where the consumption growth is normally distributed but with unknown mean and variance. In their set up expected utility is well defined when consumption growth is normally distributed with a known innovation variance. The non-convergence of the discounted sum of expected future values arises because parameter uncertainty increases at a faster rate than the discounting of future outcomes.

In practice, dividends or labour income processes are unlikely to remain stable and may be subject to structural breaks. Indeed, empirical studies have increasingly found evidence of incomplete learning and instability in a range of macroeconomic and financial time series processes that are likely to be related to the determinants of asset payoffs. Stock and Watson (1996) document evidence of breaks in the univariate time-series representation of a wide variety of financial and macroeconomic variables.<sup>1</sup> Similarly, Clements and Hendry (1998, 1999) emphasize the importance

---

<sup>1</sup>Other studies finding evidence of breaks in such time series include Alogoskoufis and Smith (1991), Banerjee et al (1992), Garcia and Perron (1996), Koop and Potter (2004a,b), Pastor and Stambaugh (2001), Paye and Timmermann (2005), Pesaran and Timmermann (2002), Pesaran, Pettenuzzo and Timmermann (2005) and Timmermann (2001).

of breaks to forecasting performance.

Such evidence opens up the possibility that the dividend process underlying common stock portfolios is also subject to breaks, a point verified by Timmermann (2001). Our paper presents new evidence of breaks in US dividends over the period 1872 - 2003. We find evidence of five breaks, three of which cluster over the period 1911-1930 (before the Great Depression), with the other two occurring in 1952 (around the Korean War), and in 1960 (the start of the Golden Age). The parameter estimates in the associated regimes differ significantly both in economic and statistical terms in a way that suggests that the dividend process has become less volatile but also more persistent through time.

Building on this evidence, we next explore how to forecast future dividends and compute the present value of dividends in the context of a model where the dividend growth process is subject to occasional structural breaks. Such breaks give rise to considerable uncertainty about the stock price when compared to a model that ignores breaks although the latter, as we argue in this paper, is clearly mis-specified. Our analysis uses the hierarchical hidden Markov chain model introduced in Pesaran, Pettenuzzo and Timmermann (2005) for the purpose of forecasting time-series that are subject to multiple breaks. Building on work by Chib (1998), this approach introduces a meta distribution that characterizes the distribution from which parameters within each dividend growth regime are drawn following a new break. Without this approach, forecasting future values of dividends is infeasible unless, of course, the possibility of future breaks to the parameters of the dividend process is ruled out. Using Gibbs sampling techniques we draw values from the parameter distribution within the regime that is in effect at the time of the forecast. To allow for possible breaks, we next draw new values of the discrete state indicator that characterizes how future states evolve. In the event that a future break occurs, new values of the parameters of the subsequent regime are drawn from the meta distribution.

Using the parameter estimates for the break point process fitted to US dividends, we find that the present value stock price is very sensitive to the underlying modeling assumptions for the dividend process. In particular, it depends on whether the possibility of past breaks during the historical sample is considered and also whether future breaks are allowed for. Since the regimes identified for the dividend process are typically quite persistent, there is no particular ranking of the present value stock price computed under no (historical or future) breaks, under historical breaks only or under past and future break scenarios. Instead, the ranking will reflect the value of the dividend growth rate in the current state relative to its historical average computed across different regimes.

The outline of the paper is as follows. Section 2 presents new theoretical results on the existence of present values under parameter estimation uncertainty. Section 3 discusses the role of structural breaks and presents empirical results for a model with multiple break points fitted to US dividend data. Section 4 shows how the present value stock price can be computed under different assumptions concerning parameter instability and reports empirical results for US data. Finally, Section 5 concludes.

## 2. Parameter Uncertainty and Present Value Calculations

Many intertemporal optimization problems result in rational expectations models with future expectations whose solution involves calculation of the discounted value of an infinite sum of forecasts formed at time  $T$ , for many periods ahead into the future. A simple example is given by

$$y_T = \lim_{H \rightarrow \infty} \left\{ \sum_{h=1}^H \delta^h E(x_{T+h} | \mathcal{I}_T) \right\}, \quad (1)$$

where  $\mathcal{I}_T$  is the forecaster's information at time  $T$ ,  $r > 0$  is the (known) discount rate so the discount factor,  $\delta = (1+r)^{-1}$ , lies in  $(0, 1)$ , and  $E(x_{T+h} | \mathcal{I}_T)$  is the conditional expectation taken with respect to the probability distribution(s) assumed for the driving process,  $\{x_t\}$ , over the past  $m$  periods ( $t = T - m + 1, T - m + 2, \dots, T$ ), and the future ( $t = T + 1, T + 2, \dots, T + h$ ), where  $m$  is the length of the estimation window while  $h$  is the forecast horizon. In general, the driving process need not be known or stable. In many applications in finance and economics it is assumed that  $\{x_t\}$  follows the geometric random walk model

$$\Delta \ln x_{t+1} = \mu + \sigma \varepsilon_{t+1}, \quad (2)$$

where  $\mu$  and  $\sigma$  are fixed constants, and  $\varepsilon_{t+1}$  is identically and independently distributed with zero means and unit variances. For given (known) values of  $\mu$  and  $\sigma$ , and assuming that these values apply to the past as well as to the indefinite future, we have

$$E(x_{T+h} | \mathcal{I}_T; \mu, \sigma) = \left( e^{h\mu} \right) [M_\varepsilon(\sigma)]^h, \quad (3)$$

where  $M_\varepsilon(\sigma)$  is the moment generating function of  $\varepsilon_t$ , assuming that it exists. Under the above assumptions the present value,  $y_T$ , is convergent and is given by

$$y_T = \frac{\delta e^\mu M_\varepsilon(\sigma) x_T}{1 - \delta e^\mu M_\varepsilon(\sigma)}, \quad (4)$$

so long as  $\lambda = \delta e^\mu M_\varepsilon(\sigma) < 1$ . In the case of normally distributed errors  $M_\varepsilon(\sigma)$  exists and is given by  $\exp(0.5\sigma^2)$ . This yields the familiar result in the literature,  $\lambda = \delta \exp(\mu + 0.5\sigma^2)$ .<sup>2</sup>

### 2.1. Unknown $\mu$ with a Known $\sigma^2$

Consider next the case where  $\sigma$  is known but  $\mu$  is unknown and estimated based on the past observations,  $\mathbf{X}_{m,T} = (x_{T-m+1}, x_{T-m+2}, \dots, x_T)'$ , with a Gaussian prior:

$$\mu \sim N(\underline{\mu}, \underline{\sigma}_\mu^2), \quad \underline{\sigma}_\mu^2 > 0. \quad (5)$$

Assuming that  $\varepsilon_{t+1} \sim N(0, 1)$ , the posterior distribution of  $\mu$  will also be Gaussian and is given by

$$\mu | \mathbf{X}_{m,T}, \sigma, \underline{\mu}, \underline{\sigma}_\mu^2 \sim N(\bar{\mu}, \bar{\sigma}_\mu^2),$$

---

<sup>2</sup>Notice, however, that even when  $\mu$  and  $\sigma$  are known, the present value calculations are quite fragile in the case of geometric random walk models, where non-convergent outcomes will follow if the innovations,  $\varepsilon_t$ , are distributed as a Student- $t$ . A similar result has also been pointed out by Geweke (2001) in the case of expected utility optimization where the consumption growth follows a geometric random walk model and the utility function is of the CRRA variety.

where (see, for example, Geweke (2005, pp. 25-27))

$$\bar{\mu} = \bar{\sigma}_\mu^2 \left( \frac{\underline{\mu}}{\underline{\sigma}_\mu^2} + \frac{m \bar{x}_{m,T}}{\sigma^2} \right), \quad (6)$$

$$\bar{x}_{m,T} = m^{-1} \sum_{t=T-m+1}^T x_t, \text{ and}$$

$$\bar{\sigma}_\mu^2 = \left( \frac{1}{\underline{\sigma}_\mu^2} + \frac{m}{\sigma^2} \right)^{-1} \quad (7)$$

In this case

$$\begin{aligned} E(x_{T+h} | \mathcal{I}_T; \sigma, \underline{\mu}, \underline{\sigma}_\mu^2) &= x_T \left[ E(e^{h\underline{\mu}} | \mathcal{I}_T; \sigma, \underline{\mu}, \underline{\sigma}_\mu^2) \right] [M_\varepsilon(\sigma)]^h \\ &= x_T E(e^{h\underline{\mu}} | \mathcal{I}_T; \sigma, \underline{\mu}, \underline{\sigma}_\mu^2) e^{\frac{1}{2}h\sigma^2}, \end{aligned} \quad (8)$$

where

$$E(e^{h\underline{\mu}} | \mathcal{I}_T; \sigma, \underline{\mu}, \underline{\sigma}_\mu^2) = e^{h\bar{\mu} + \frac{1}{2}h^2\bar{\sigma}_\mu^2},$$

and the individual elements in the infinite sum, (1), exist and are given by

$$E(x_{T+h} | \mathcal{I}_T; \sigma) = x_T e^{h\bar{\mu} + \frac{1}{2}h^2\bar{\sigma}_\mu^2 + \frac{1}{2}h\sigma^2} = x_T \left( e^{\bar{\mu} + \frac{1}{2}h\bar{\sigma}_\mu^2 + \frac{1}{2}\sigma^2} \right)^h. \quad (9)$$

Finally, the present value, truncated at forecast horizon  $H$ , becomes

$$y_{T:T+H} = x_T \sum_{h=1}^H \left( \delta e^{\bar{\mu} + \frac{1}{2}h\bar{\sigma}_\mu^2 + \frac{1}{2}\sigma^2} \right)^h = x_T \sum_{h=1}^H [\rho(h, m)]^h, \quad (10)$$

where

$$\rho(h, m) = e^{-\ln(1+r) + \bar{\mu} + \frac{1}{2}h\bar{\sigma}_\mu^2 + \frac{1}{2}\sigma^2}. \quad (11)$$

To check if this is convergent as  $H \rightarrow \infty$ , we first note that

$$\begin{aligned} \bar{\mu} &= \frac{\bar{x}_{m,T} + \left( \frac{\sigma^2}{m} \right) \left( \frac{\underline{\mu}}{\underline{\sigma}_\mu^2} \right)}{1 + \left( \frac{\sigma^2}{m} \right) \left( \frac{1}{\underline{\sigma}_\mu^2} \right)} = \bar{x}_{m,T} \left[ 1 - \left( \frac{\sigma^2}{m} \right) \left( \frac{1}{\underline{\sigma}_\mu^2} \right) \right] + \left( \frac{\sigma^2}{m} \right) \left( \frac{\underline{\mu}}{\underline{\sigma}_\mu^2} \right) + O\left( \frac{1}{m^2} \right), \\ \bar{\sigma}_\mu^2 &= \frac{\sigma^2}{m} \left( \frac{1}{1 + \frac{1}{m} \left( \frac{\sigma^2}{\underline{\sigma}_\mu^2} \right)} \right) = \frac{\sigma^2}{m} + O\left( \frac{1}{m^2} \right). \end{aligned} \quad (12)$$

Using these results in (11) yields  $\rho(h, m) = \delta e^{\frac{1}{2}\sigma^2} e^{g(h, m)}$ , where

$$g(h, m) = \bar{x}_{m,T} \left[ 1 - \left( \frac{\sigma^2}{m} \right) \left( \frac{1}{\underline{\sigma}_\mu^2} \right) \right] + \left( \frac{\sigma^2}{m} \right) \left( \frac{\underline{\mu}}{\underline{\sigma}_\mu^2} \right) + \frac{\sigma^2}{2} \left( \frac{h}{m} \right) + O\left( \frac{h}{m^2} \right),$$

or after some algebra

$$g(h, m) = \bar{x}_{m,T} + \left( \frac{\sigma^2}{m} \right) \left[ \frac{\underline{\mu} - \bar{x}_{m,T}}{\underline{\sigma}_\mu^2} \right] + \frac{\sigma^2}{2} \left( \frac{h}{m} \right) + O\left( \frac{h}{m^2} \right).$$



Typically one expects  $(\underline{\mu} - \bar{x}_{m,T}) / (m\sigma_\mu^2)$  to be quite small and the present value would be dominated by the term  $\frac{\sigma^2}{2} \left(\frac{h}{m}\right)$ . Therefore,

$$y_{T:T+H} \approx x_T \sum_{h=1}^H \left( e^{-\ln(1+r) + \frac{1}{2}\sigma^2(1+h/m) + \bar{x}_{m,T}} \right)^h, \quad (13)$$

and even if  $\hat{\lambda}_{m,T} = \exp(-\ln(1+r) + \bar{x}_{m,T} + 0.5\sigma^2) < 1$  (the estimated certainty equivalence convergence condition),  $y_{T:T+H}$  will be divergent as  $H \rightarrow \infty$ .<sup>3</sup>

The only case where the present value convergences in the presence of the estimation uncertainty will be if the limit of  $h/m$  as  $h \rightarrow \infty$  is bounded from above by some constant,  $\kappa < 2[\bar{x}_{m,T} - \ln(1+r)]/\sigma^2 - 1$ . This corresponds to a thought experiment in which the length of the estimation sample ( $m$ ) somehow grows at a sufficiently fast rate with the forecast horizon,  $h$ , so that  $\lim(h/m)$  tends to  $\kappa < 2[\mu - \ln(1+r)]/\sigma^2 - 1$ , as  $h$  and  $m$  tend to infinity jointly. In practice, of course, such thought experiments have little relevance since the estimation window (often determined by the historical data set available) and the forecast horizon are separate concepts and there is no reason why they should be linked in any particular way. Furthermore, in cases with evidence of instability in the parameters of the dividend process, an unbiased estimator of the parameters of the dividend generating process can only be obtained by restricting the estimation sample to the post-break data. This provides another reason for why  $m$  is finite in practice. In the following we shall, without loss of generality, assume  $m = T$  and suppress the  $m$  notation for simplicity.

## 2.2. Unknown $\mu$ and $\sigma^2$

The non-convergence problem of the present value will be accentuated if we also assume that  $\sigma$  is unknown and is estimated from the past data,  $\mathbf{X}_{m,T}$ . For example, using conjugate priors for  $\mu$  and  $\sigma^2$  the posterior distribution of  $\mu$  will be  $t$ -distributed and  $E(e^{h\mu} | \mathcal{I}_T; \underline{\mu}, \underline{\sigma}_\mu^2, \underline{\sigma}^2, \underline{\nu})$  ceases to exist for any  $h > 0$ , where  $\underline{\sigma}^2$  and  $\underline{\nu}$  are the parameters of the gamma prior density assumed for  $\sigma^2$  which can be written conveniently as

$$\frac{\underline{\sigma}^2}{\sigma^2} | \underline{\sigma}^2, \underline{\nu} \sim \chi^2(\underline{\nu}).$$

As pointed out by Geweke (2001), the use of non-conjugate priors for  $\mu$  and  $\sigma^2$  does help in resolving the non-existence of  $E(e^{h\mu} | \mathcal{I}_T; \underline{\mu}, \underline{\sigma}_\mu^2, \underline{\sigma}^2, \underline{\nu})$ . However, it does not resolve the non-convergence of the infinite sums that are involved in present value calculations.

Non-Bayesian approaches to dealing with the uncertainty of  $\mu$  and  $\sigma^2$  are unlikely to help either. One possible approach would be to bootstrap the present values. This involves (i) drawing  $\mu^{(b)}$  and  $\sigma^{(b)}$  from the observed empirical distribution of the estimators of  $\mu$  and  $\sigma^2$  (say  $\hat{\mu}_{m,T}$  and  $\hat{\sigma}_{m,T}^2$ ), (ii) computing present values for each choice of  $\mu^{(b)}$  and  $\sigma^{(b)}$  denoted as  $y_T^{(b)}$ , and (iii) obtaining the bootstrap present value as  $B^{-1} \sum_{b=1}^B y_T^{(b)}$ , where  $B$  is the total number of bootstraps. However, for this procedure to yield a convergent outcome it will be required that  $y_T^{(b)}$  is convergent for each

---

<sup>3</sup>The above analysis also shows the danger of letting  $m \rightarrow \infty$  first before computing the limit of the present value with  $H \rightarrow \infty$ . In reality  $m$  could be quite large but still finite as  $H \rightarrow \infty$ .

$b$ , which is extremely unlikely unless the empirical distribution of  $\hat{\mu}_{m,T}$  and  $\hat{\sigma}_{m,T}^2$  is constrained so that all draws from that distribution satisfy the conditions  $\delta e^{\mu^{(b)}} M_\varepsilon(\sigma^{(b)}) < 1$  for all  $b$ .

### 2.3. Trend Stationary Log-linear Driving Processes

The non-convergence problem continues to be present if the unit root process in (2) is replaced by the following trend stationary process:

$$\Delta [\ln x_{t+1} - a - \mu(t+1)] = -(1-\rho)(\ln x_t - a - \mu t) + \sigma \varepsilon_{t+1}, \quad (14)$$

where  $|\rho| < 1$ , and as before  $\mu$  represents the average growth of the logarithm of the driving process,  $x_t$ . In the case of this process

$$\ln(x_{T+h}/x_T) = -(1-\rho^h)(\ln x_T - a - \mu T) + \mu h + \sigma \sum_{j=1}^h \rho^j \varepsilon_{T+j},$$

and

$$E(x_{T+h} | \mathcal{I}_T; a, \mu, \rho, \sigma) = e^{-(1-\rho^h)(\ln x_T - a - \mu T)} \left( e^{h\mu} \right) \left[ \prod_{j=1}^h M_\varepsilon(\sigma \rho^j) \right], \quad (15)$$

which is a direct generalization of (8) and reduces to it for  $\rho = 1$ . It is clear that the various issues discussed for the unit root case readily apply here. Even if  $\varepsilon_{t+1}$  has a moment generating function, the present value is unlikely to exist if  $\mu$  is not known with certainty. For example, suppose  $a, \rho$  and  $\sigma$  are known and  $\mu$  is estimated based on the regression of  $\ln x_t - \rho \ln x_{t-1} - a(1-\rho)$  on  $(1-\rho)t + \rho$ . Assuming, as before, that conditional on  $a, \rho$  and  $\sigma$  the prior probability distribution of  $\mu$  is Gaussian and given by (5), then the posterior distribution of  $\mu$  will be given by

$$\mu | \mathbf{X}_T, a, \rho, \sigma, \underline{\mu}, \underline{\sigma}_\mu^2 \sim N(\bar{\mu}, \bar{\sigma}_\mu^2),$$

where

$$\begin{aligned} \bar{\mu} &= \bar{\sigma}_\mu^2 \left( \frac{\underline{\mu}}{\underline{\sigma}_\mu^2} + \frac{\hat{\mu}_T}{\hat{\sigma}_T^2} \right), \\ \hat{\mu}_T &= \frac{\sum_{t=1}^T [\ln x_t - \rho \ln x_{t-1} - a(1-\rho)] [(1-\rho)t + \rho]}{\sum_{t=1}^T [(1-\rho)t + \rho]^2}, \\ \hat{\sigma}_T^2 &= \frac{\sigma^2}{\sum_{t=1}^T [(1-\rho)t + \rho]^2}, \text{ and } \bar{\sigma}_\mu^2 = \left( \frac{1}{\underline{\sigma}_\mu^2} + \frac{1}{\hat{\sigma}_T^2} \right)^{-1}. \end{aligned}$$

Hence

$$\begin{aligned} E(x_{T+h} | \mathcal{I}_T; a, \rho, \sigma) &= \left[ e^{-(1-\rho^h)(\ln x_T - a)} \prod_{j=1}^h M_\varepsilon(\sigma \rho^j) \right] E_\mu \left( e^{[h+(1-\rho^h)T]\mu} \right), \\ &= \left[ e^{-(1-\rho^h)(\ln x_T - a)} \prod_{j=1}^h M_\varepsilon(\sigma \rho^j) \right] e^{[h+(1-\rho^h)T]\bar{\mu} + \frac{1}{2}[h+(1-\rho^h)T]^2 \bar{\sigma}_\mu^2}, \end{aligned}$$

and for a fixed  $T$  its rate of expansion is governed by the term  $\exp(.5\bar{\sigma}_\mu^2 h^2)$ . Therefore, as far as the existence of the present value is concerned the outcomes are very similar irrespective of whether the logarithm of the driving variable has a unit root or is trend stationary. The differences between the two cases is a matter of degree and differs only due to the differences in the precision with which  $\mu$  is estimated under the two cases. Under the unit root process the precision of  $\mu$  is of order  $T^{-1}$ , while when  $\ln x_t$  is trend stationary it is given by  $T^{-3/2}$ .

#### 2.4. Present Values with a Stochastic Discount Factor: The Lucas Tree Model

Normally the discount rate is formed as the risk-free rate plus some risk-premium to reflect the uncertain nature of future payoffs and correct for correlations between dividend shocks and variations in the stochastic discount factor. It is clearly of interest to relate the discount rate—taken to be fixed and strictly exogenous so far—to the growth rate in dividends, using equilibrium consumption based asset pricing models. In the case of consumption based asset pricing models the expression for the present value is more complicated and depends on a stochastic discount factor that varies with  $x$ . In the context of a representative agent model with the utility function,  $u(c_t)$ , we have<sup>4</sup>

$$y_T = \lim_{H \rightarrow \infty} \left\{ \sum_{h=1}^H \delta^h E \left( \frac{u'(c_{t+h})}{u'(c_t)} x_{T+h} \mid \mathcal{I}_T \right) \right\}, \quad (16)$$

which reduces to the present value expression (1) in the risk neutral case where  $u(c)$  is linear. But for a general specification of  $u(c)$  the analysis of convergence of the present value depends on the form of the utility function and the nature of the dependence of  $c_{t+h}$  and  $x_{t+h}$  in a general equilibrium context. Although such a general analysis is beyond the scope of the present paper, certain analytical results can be obtained for the Lucas's tree model (Lucas, 1978) where consumption is equal to dividends ( $c_{T+h} = x_{T+h}$ ) and the utility function is specified to have the power form  $u(c) = (1 - \gamma)^{-1}(c^{1-\gamma} - 1)$  ( $\gamma \neq 1$ ). In this case,

$$y_T = \lim_{H \rightarrow \infty} \left\{ \sum_{h=1}^H E \left( e^{-h \ln(1+r) + (1-\gamma)(\ln x_{T+h} - \ln x_T)} \mid \mathcal{I}_T \right) \right\},$$

and under the geometric random walk model (2) with a known mean and variance we have

$$E(y_T \mid \mu, \sigma^2) = x_T \lim_{H \rightarrow \infty} \left\{ \sum_{h=1}^H E \left( e^{-h \ln(1+r) + (1-\gamma)\mu h + 0.5(1-\gamma)^2 \sigma^2 h} \mid \mu, \sigma^2, \mathcal{I}_T \right) \right\},$$

which is convergent for given values of  $\mu$  and  $\sigma^2$  so long as  $-\ln(1+r) + (1-\gamma)\mu + 0.5(1-\gamma)^2 \sigma^2 < 0$ .

---

<sup>4</sup>See, for example, Cochrane (2005, p.24). Deriving (16) from the first order inter-temporal optimization conditions also requires that the transversality condition

$$\lim_{h \rightarrow \infty} E \left( \frac{u'(c_{T+h})}{u'(c_T)} y_{T+h} \mid \mathcal{I}_T \right) = 0,$$

is satisfied.

Consider now the case where  $\mu$  is unknown and continue to assume that  $\sigma^2$  is given (known). Then using the above results we have

$$E\left(e^{(1-\gamma)\mu h + 0.5(1-\gamma)^2\sigma^2 h} \mid \sigma^2, \mathcal{I}_T\right) = e^{(1-\gamma)\bar{\mu}h + 0.5(1-\gamma)^2\sigma^2 h + 0.5(1-\gamma)^2 h^2 \bar{\sigma}_\mu^2}, \quad (17)$$

where  $\bar{\mu}$  and  $\bar{\sigma}_\mu^2$  are the posterior mean and variance of  $\mu$  given by (6) and (7) where  $m = T$  and that do not vary with  $h$ . Hence, the elements in the series expansion of  $E(y_T/x_T \mid \sigma^2)$  eventually will be dominated by terms  $e^{0.5(1-\gamma)^2 h^2 \bar{\sigma}_\mu^2}$ ,  $h = 1, 2, \dots$  and the present value expression will be divergent unless  $\gamma = 1$ , corresponding to the special case of log-utility.

### 2.5. Possible Solutions to the Non-Convergence Problem

The reason for the fragility of the present value under the geometric random walk model with constant but unknown parameters can be illustrated using a simple discrete state process for  $\mu$ . Suppose that over the forecast horizon  $T + 1, T + 2, \dots, T + H$ ,  $\mu$  can take any one of the values  $\mu_1, \mu_2, \dots, \mu_m$  with probabilities  $\pi_1, \pi_2, \dots, \pi_m$  where  $\sum_{i=1}^m \pi_i = 1$ , and  $1 > \pi_i > 0$ . To simplify the analysis also assume that  $\sigma^2$ ,  $\mu_i$  and  $\pi_i$  are known at time  $T$ . Under this example, the present value is given by

$$y_T = x_T \left\{ \sum_{i=1}^m \pi_i \lim_{H \rightarrow \infty} \sum_{h=1}^H e^{-h \ln(1+r) + h\mu_i} [M_\varepsilon(\sigma)]^h \right\}. \quad (18)$$

Since  $1 > \pi_i > 0$ ,  $y_T$  exists if  $\delta e^{\mu_i} M_\varepsilon(\sigma) < 1$  for all  $i$ . Contrast this result with the associated certainty equivalent expression that accounts for uncertainty about the value of the underlying mean parameter (but disregards uncertainty about future dividend innovations):

$$y_T^{CE} = x_T \left\{ \lim_{H \rightarrow \infty} \sum_{h=1}^H e^{-h \ln(1+r) + h\bar{\mu}_\pi} [M_\varepsilon(\sigma)]^h \right\}, \quad (19)$$

where  $\bar{\mu}_\pi = \sum_{i=1}^m \pi_i \mu_i$ . The condition for  $y_T^{CE}$  to exist is given by  $\delta e^{\bar{\mu}_\pi} M_\varepsilon(\sigma) < 1$ . Clearly, it is possible for the latter to be satisfied without  $\delta e^{\mu_i} M_\varepsilon(\sigma) < 1$  being satisfied for all  $i$ . A sufficiently large  $\mu_i$ , even if it is extremely unlikely (with  $\pi_i$  very close to zero), can result in divergence of  $y_T$ , although for all other outcomes that are much more likely the associated infinite sums could be convergent.

It is clear from this example that the non-convergence of the present value arises from the particular combinations of (i) a geometric random walk driving process, (ii) an infinite horizon and (iii) constant, but unknown parameters drawn from a Gaussian posterior distribution. One could consider relaxing any one or all of these elements. We discuss the first two assumptions below and then deal with the parameter stability assumption in more detail in the next section.

#### 2.5.1. Use of Linear Driving Processes

The non-convergence problem can be avoided altogether if the geometric random walk model is abandoned in favour of a linear driving process. Consider for example the simple random walk

model with drift

$$x_t = \mu + x_{t-1} + \sigma \varepsilon_t. \quad (20)$$

Then  $E(x_{T+h} | \mathcal{I}_T) = x_T + \mu h$ , and

$$\begin{aligned} y_T &= \lim_{H \rightarrow \infty} \left\{ \sum_{h=1}^H \delta^h E(x_{T+h} | \mathcal{I}_T) \right\} \\ &= \frac{x_T}{1-\delta} + \mu \sum_{\tau=1}^{\infty} h \delta^h, \end{aligned}$$

or

$$y_T = \frac{x_T}{1-\delta} + \frac{\mu \delta}{(1-\delta)^2}. \quad (21)$$

Uncertainty surrounding  $\mu$  can easily be dealt with in a way that does not cause non-convergence problems. Generalizing the process to higher order models with possible serial correlation in the innovations,  $\varepsilon_t$ , would not alter the main conclusion. For example, for the  $p^{\text{th}}$  order driving process

$$x_t = \left(1 - \sum_{j=1}^p \rho_j\right) \mu + \sum_{j=1}^p \rho_j x_{t-j} + \sigma \varepsilon_t,$$

conditional on  $\boldsymbol{\rho} = (\rho_1, \rho_2, \dots, \rho_p)'$  we have

$$y_T | \boldsymbol{\rho} = \frac{\sum_{j=0}^{p-1} \phi_j(\delta) x_{T-j}}{1 - \phi_0(\delta)} + \frac{\delta \mu (1 - \sum_{j=1}^p \rho_j)}{(1-\delta)(1 - \phi_0(\delta))}, \quad (22)$$

where

$$\phi_j(\delta) = \sum_{i=j+1}^p \rho_i \delta^{i-j}, \text{ for } j = 0, 1, \dots, p-1.$$

Under parameter uncertainty

$$y_T = \sum_{j=0}^{p-1} E \left[ \frac{\phi_j(\delta)}{1 - \phi_0(\delta)} | \mathbf{X}_T \right] x_{T-j} + \frac{\delta}{1-\delta} E \left( \frac{\mu (1 - \sum_{j=1}^p \rho_j)}{1 - \phi_0(\delta)} | \mathbf{X}_T \right), \quad (23)$$

where expectations are taken with respect to the posterior distribution of  $\mu$  and  $\boldsymbol{\rho}$ . These expectations are likely to exist for a sufficiently large  $T$ , and do not depend on the forecast horizon.

The problem with this approach is, as pointed out by Campbell, Lo and MacKinlay (1997, p. 258), that linear models for real dividends, consumption or labour income do not fit the data well since these series tend to grow exponentially over time. This means that linear models are usually dominated by log-linear specifications.

### 2.5.2. Use of Finite Horizons

A simple, but rather *ad hoc*, solution would be to define the present values over a given finite future, say  $\tilde{H}$ , and then write the solution as

$$y_T(\tilde{H}) = \sum_{h=1}^{\tilde{H}} \delta^h E(x_{T+h} | \mathcal{I}_T), \quad (24)$$

which is finite so long as  $E(x_{T+h} | \mathcal{I}_T)$  exists for  $h = 1, \dots, \tilde{H}$ . In practice the choice of  $\tilde{H}$  could be problematic. In the case of the life cycle consumption model,  $\tilde{H}$  can be viewed as the life of the household and treated as a truncated random variable. For example, we could assume that (for  $0 < \theta < 1$ )

$$\begin{aligned} \Pr(\tilde{H} = s) &= \frac{(1-\theta)\theta^s}{\theta(1-\theta^{\bar{H}})}, \text{ for } s = 1, 2, \dots, \bar{H} \\ &= 0, \text{ for } s > \bar{H}, \end{aligned} \quad (25)$$

where  $\bar{H}$  is an arbitrarily large but finite value. The non-truncated case where  $\bar{H} \rightarrow \infty$  yields the familiar geometric distribution used by Yaari (1965), Cass and Yaari (1967) and Blanchard and Fischer (1989) to model uncertain life times in models of household consumption. Integrating out the uncertainty of  $\tilde{H}$  we have

$$E_{\tilde{H}} [y_T(\tilde{H})] = \frac{(1-\theta)}{\theta(1-\theta^{\bar{H}})} \sum_{s=1}^{\bar{H}} \sum_{h=1}^s \theta^s \delta^h E(x_{T+h} | \mathcal{I}_T),$$

which can be written more compactly as

$$E_{\tilde{H}} [y_T(\tilde{H})] = \sum_{s=1}^{\bar{H}} \frac{1-\theta^{\bar{H}-s+1}}{\theta(1-\theta^{\bar{H}})} (\theta\delta)^s E(x_{T+s} | \mathcal{I}_T). \quad (26)$$

In this set up the choice of  $\bar{H}$  is of secondary importance. However, it is worth noting that for  $\bar{H} \rightarrow \infty$  the uncertain life time present value problem reduces to

$$\lim_{\bar{H} \rightarrow \infty} E_{\tilde{H}} [y_T(\tilde{H})] = \theta^{-1} \sum_{s=1}^{\infty} (\theta\delta)^s E(x_{T+s} | \mathcal{I}_T), \quad (27)$$

which is the infinite horizon problem with a lower discount factor given by  $\theta\delta$ . By increasing the discount rate the stochastic life time assumption will help towards achieving convergence, but does not resolve the problem altogether. A finite  $\bar{H}$  would still be required in general.

### 3. Present Value Models with Structural Breaks

Perhaps a more appealing way to handle the non-convergence problem is to relax the assumption that the parameters of the underlying growth process are constant through time. This assumption clearly goes to the root of the non-convergence: As long as there is even an infinitesimal probability of drawing a set of parameters for which the (constant) growth rate exceeds the discount rate, the present value will not exist. Conversely, if the parameters of the growth rate are subject to structural breaks, there are cases where the growth rate temporarily exceeds the discount rate, yet the present value continues to exist. This happens provided that the underlying driving process most of the time grows at a slower rate than the discount rate. The condition for the existence of the present value is now the rather weaker one that paths leading to an unbounded present value

have zero probability. This will trivially hold if the distribution of the maximum growth rate is truncated so that it always falls below the discount rate, but can also hold in the absence of this assumption.

Breaks to the cash flow process will not, however, in and of itself, resolve the problem. This can easily be seen as follows. Suppose that the mean of the increment to the logarithm of the first-differenced future dividend process falls in different regimes during the period  $T + 1, \dots, T + h$ . Denote the number of these regimes by  $N_h$  and let their duration be  $h_1, \dots, h_{N_h}$ , so that  $\sum_{i=1}^{N_h} h_i = h$ . Equivalently, the fraction of the time spent in regime  $i$  is given by  $\pi_i = h_i/h$ ,  $0 < \pi_i < 1$ . Suppose that  $\mu_i \sim N(\bar{\mu}, \sigma_\mu^2)$ . We then get the present value as follows:

$$\begin{aligned} y_T &= E_T \left[ x_T \sum_{h=1}^H \exp \left[ -\ln(1+r)h + h \sum_{i=1}^{N_h} \pi_i \mu_i + \sigma \sum_{j=1}^h \varepsilon_j \right] \right] \\ &= x_T \sum_{h=1}^H \exp \left[ -\ln(1+r)h + h(\bar{\mu} + \frac{1}{2}\sigma^2) + \frac{1}{2}\sigma_\mu^2 h^2 \sum_{i=1}^{N_h} \pi_i^2 \right]. \end{aligned} \quad (28)$$

Notice that the last term is not convergent as  $H \rightarrow \infty$ . Hence while the possibility of breaks adds some flexibility to the model, one has to be careful to ensure convergence of the present value either by using a finite  $H$ , by truncating the distribution from which the future growth rate is drawn, or through some other means.

### 3.1. A Dividend Model with Breaks

To illustrate the above issues, we next consider the empirical evidence of breaks in the dividend process underlying US stocks. Real dividends underlying broadly diversified stock market indices are often assumed to follow a simple process of the form (2) with  $\varepsilon_{t+1} \sim N(0, 1)$ . Depending on the frequency at which dividends are modelled, autoregressive dynamics may also be present, in which case the process can be generalized to

$$\Delta \ln(x_{t+1}) - \mu = \sum_{i=1}^p \beta_i (\Delta \ln(x_{t+1-i}) - \mu) + \sigma \varepsilon_{t+1}, \quad (29)$$

where  $\beta_i$  ( $i = 1, \dots, p$ ) are autoregressive parameters and  $\mu$  reflects the long-run mean of the dividend growth rate, whereas  $\tilde{\mu} = \mu(1 - \sum_{i=1}^p \beta_i)$  is the intercept for the AR(p) process in  $\Delta \ln(x_{t+1})$ . Both specifications (2) and (29) assume that the parameters of the dividend growth process remain constant through time—an assumption that, in view of the significant shocks to economic growth observed throughout the twentieth century, is unlikely to be satisfied over the long sample periods typically used for estimation of the parameters of the dividend growth process, see Timmermann (2001).

To capture the possibility of structural shifts in the parameters of the dividend growth process, we adopt the change-point process proposed by Chib (1998).<sup>5</sup> This approach assumes that shifts to

---

<sup>5</sup>McCulloch and Tsay (1993) is another prominent example of breakpoint analysis in a Bayesian setting.

the parameters of a time-series process are determined by the evolution in a discrete state variable,  $S_t$ , initialized so that  $S_1 = 1$  and increasing over time to capture the possibility of breaks. Every time the state variable increases by one unit, the parameters of the process shift. Hence, if the break point indicator  $S_T = K + 1$ , this means that there were  $K$  breaks between time 1 and time  $T$ . For example, assuming that the dividend process can be characterized as an AR(1) model whose parameters are subject to breaks, we have

$$\begin{aligned} \Delta \ln(x_{t+1}) - \mu_1 &= \beta_1(\Delta \ln(x_t) - \mu_1) + \sigma_1 \varepsilon_{t+1}, & \tau_0 \leq t \leq \tau_1 \\ \Delta \ln(x_{t+1}) - \mu_2 &= \beta_2(\Delta \ln(x_t) - \mu_2) + \sigma_2 \varepsilon_{t+1}, & \tau_1 + 1 \leq t \leq \tau_2 \\ \vdots & & \vdots \\ \Delta \ln(x_{t+1}) - \mu_{K+1} &= \beta_{K+1}(\Delta \ln(x_t) - \mu_{K+1}) + \sigma_{K+1} \varepsilon_{t+1}, & \tau_K + 1 \leq t \leq T \end{aligned} \quad (30)$$

where  $\tau_1, \dots, \tau_K$  are the breakpoints and  $\{\mu_{S_{t+1}}, \beta_{S_{t+1}}, \sigma_{S_{t+1}}^2\}$  are the parameters associated with the dividend process regime that is in effect at time  $t + 1$ . Our other assumptions follow Pesaran, Pettenuzzo and Timmermann (2005) which we next briefly review. The state variable,  $S_{t+1}$ , can either remain in the  $k^{\text{th}}$  regime, which happens with probability  $p_{kk}$  or move on to the next regime, which happens with probability  $p_{k,k+1} = 1 - p_{kk}$ . These probabilities are assumed to be drawn independently across regimes from a beta distribution with prior parameters  $\underline{a}$  and  $\underline{b}$ :

$$p_{ii} \sim \text{Beta}(\underline{a}, \underline{b}), \text{ for } i = 1, 2, \dots, K. \quad (31)$$

For the AR(1) specification the parameters determining the conditional mean of the dividend growth process,  $\beta_i = (\tilde{\mu}_i, \beta_i)$  for  $i = 1, 2, \dots, K + 1$ , are drawn from a Gaussian distribution,  $\beta_{S_{t+1}} \sim N(\mathbf{b}_0, \mathbf{B}_0)$ , while the error term precision parameters,  $\sigma_{S_{t+1}}^{-2}$ , are identically, independently distributed (IID) draws from a Gamma distribution,  $\sigma_j^{-2} \sim G(v_0, d_0)$ . At the level of the meta distribution, we make the following distributional assumptions:

$$\mathbf{b}_0 \sim N(\underline{\mu}_\beta, \underline{\Sigma}_\beta) \quad (32)$$

$$\mathbf{B}_0^{-1} \sim W(\underline{v}_\beta, \underline{\mathbf{V}}_\beta^{-1}), \quad (33)$$

where  $W(\cdot)$  is a Wishart distribution.  $\underline{\mu}_\beta$ ,  $\underline{\Sigma}_\beta$ ,  $\underline{v}_\beta$  and  $\underline{\mathbf{V}}_\beta^{-1}$  are hyperparameters that are specified *a priori*. Finally, the error term precision,  $v_0$  and  $d_0$  are assumed to follow an exponential and Gamma distribution, respectively, with hyperparameters  $\underline{\rho}_0$ ,  $\underline{c}_0$  and  $\underline{d}_0$ :

$$v_0 \sim \text{Exp}(\underline{\rho}_0) \quad (34)$$

$$d_0 \sim \text{Gamma}(\underline{c}_0, \underline{d}_0). \quad (35)$$

We refer to this specification as the composite-meta model. All prior parameters are under-scored.

### 3.2. Estimation Results

We use data from Shiller (2000) available at <http://www.econ.yale.edu/~shiller/data.htm>. This provides monthly dividends paid by the firms included in a broad index of US firms. The data



runs from 1871:03 to 2003:09. Nominal dividends are divided by the consumer price index to get a time-series of real dividends. Statistics on the first-differenced values of this series are provided in Table 1. Monthly dividend growth rates are serially correlated and left-skewed with fat tails. Figure 1 plots the associated time-series and indicates strong evidence of instability in the underlying process as it has evolved over the sample. Most notably, the volatility was very high in the early parts of the sample, and around World War II, and has become much lower after this period. In addition, the persistence of the series appears to have shifted over time as reflected in a more volatile (less persistent) time-series plot in the early parts of the sample. These are of course only visual impressions and must be confirmed by a more formal econometric analysis.

To this end we next document the presence of breaks in the dividend process. Table 2 shows estimates for a variety of models with different numbers of breaks. Bayes factors based on the ratios of the marginal likelihoods for models with different numbers of breaks suggest selecting a model with five breaks, i.e. a break occurring roughly every 25 years. In fact, assuming equal prior probabilities on the models with between zero and six breaks, almost all of the posterior probability mass goes to the model with five breaks.

Based on the posterior modes for the probability of a shift in the state variable,  $S_t$ , the five breaks are estimated to have occurred in 1911, 1922, 1930 (at the beginning of the Great Depression), 1952 (around the Korean War) and in 1960 (the start of the Golden Age of Capitalism). Figure 2 shows that the date of the first break is very poorly determined with probabilities of a break in individual months well below 6% and spread out between 1900 and 1920. The remaining break dates are more precisely determined with modal probabilities varying from 0.15 to 0.35.

Table 3 reports parameter estimates for the model with five breaks (six regimes). As one might expect from a sample period as heterogenous as the twentieth century, there is considerable variation in the parameters across regimes. The intercept parameter varies from -0.02 to 0.16, while the AR(1) parameter varies from a low point of 0.37 (between 1871 and 1911) to a high point of 0.73 between 1930 and 1952. Confirming the visual impression from Figure 1, the standard deviation of the dividend process has varied considerably from a peak of 0.78 prior to 1911 to its value of 0.15 after 1960. Clearly the dividend process has become less volatile but also more persistent through time. Finally, the mean value of the ‘stayer’ probability parameter that characterizes the duration of the various states has varied from a high of 0.9997 in the regime prior to 1911 to 0.985 in the regime over the period 1952 to 1960.

Consistent with the large variation across regimes in the parameters of the dividend growth process, the mean value of the standard errors of the meta distribution parameters ( $b_0(1), b_0(2)$ ) are quite large at 0.04 and 0.08, respectively. In fact, Table 4 shows that the 95% confidence interval for  $b_0(1)$ , the parameter in the meta distribution characterizing the mean intercept across regimes, goes from -0.021 to 0.104 and from 0.43 to 0.68 for  $b_0(2)$ , the parameter in the meta distribution characterizing the mean persistence across regimes. Following a future break, the parameters of the dividend process will be drawn from the meta distribution so these values indicate that there is considerable uncertainty about the process driving future dividend growth.

Figure 3 shows that parameter instability of the dividend growth process has a large effect on

the predictive density. It shows the predictive density under three different models at horizons of 1, 12, 24, 36, 48 and 60 months. The first model is the full-sample, no-break model, while the other two models allow for past breaks and past and future breaks. The model that allows for past breaks effectively bases predictions of future values on the parameters from the last regime—the regime prevailing from 1960-2003 in this case—while the model that allows for future breaks starts off from this regime but lets new parameters following a future break be drawn from the meta distribution. This also explains why the predictive densities are more concentrated under the break regimes in Figure 3 since the standard deviation of innovations to dividends was very low in this regime compared with the full-sample average value. While the dispersions of the densities are quite different, however, in this regime the centering of the dividend growth process is not greatly affected by the presence of breaks.

The relationship between the predictive density under breaks and under no breaks can be quite different depending on the parameter values in the regime from which the forecast is calculated. To see this, we plot in Figure 4 the same three graphs but now for the case where the parameters in the last regime are based on their values in the regime prevailing during 1922-1930. It now becomes clear that, particularly at the longer horizons, the three predictive densities are very different once breaks are considered. In this case breaks shift the mean growth rate to the right compared with the no-break case. The reason why the difference is largest at the longer horizons is due to the cumulated effect of having different mean and persistence parameters under the three scenarios. Since the forecasts from the AR(1) model are computed based on the same initial value of the dividend process, differences in the parameters have a relatively smaller effect at short horizons. This observation is by no means unique to the regime from 1922-1930 and—as shown in Figure 5—also holds for the regime that was in effect from 1952-1960.

Differences between predictive densities under the two breakpoint models can be explained with reference to Figure 6. This figure plots the weight on the current regime (thus assuming that the parameter values for the final regime remain in effect) as a function of the forecast horizon. This is similar to a survival plot for the current state and shows how the probability of a break (computed as one minus the ‘stayer’ probability plotted in Figure 6) increases to more than 50% as the forecast horizon extends beyond five years.

#### 4. Present Value Stock Price under Breaks

As we showed in Section 2, to compute the present value of future dividends, we need to evaluate an expression of the form

$$\lim_{H \rightarrow \infty} y_T = \lim_{H \rightarrow \infty} \sum_{h=1}^H \delta^h E_T[x_{T+h}] = \lim_{H \rightarrow \infty} \sum_{h=1}^H \exp(-h \ln(1+r)) E_T[x_{T+h}]. \quad (36)$$

Notice that  $E_T[\cdot]$  is calculated not just conditional on current dividends,  $x_T$ , but on the entire past sequence  $\{x_t\}_{t=1}^T$ . Hence the complete historical track record of dividends matters when forecasting future dividends.

To compute the future stream of dividends, we use the Gibbs sampler to generate draws from the predictive distribution of  $\Delta \ln(x_{T+1}), \Delta \ln(x_{T+2}), \dots, \Delta \ln(x_{T+H})$ . To see how this works, consider the sum of log first-differences of dividends between period  $T + 1$  and period  $T + H$ :

$$\sum_{h=1}^H \Delta \ln(x_{T+h}) = \ln(x_{T+H}/x_T), \quad (H \geq 1) \quad (37)$$

so

$$x_{T+H} = x_T \exp\left(\sum_{h=1}^H \Delta \ln(x_{T+h})\right),$$

or, in terms of present values,

$$\delta^H x_{T+H} = x_T \exp\left(\sum_{h=1}^H \Delta \ln(x_{T+h}) - H \ln(\delta)\right). \quad (38)$$

First assume that there are no breaks between period  $T$  and  $T + h$ . When dividends follow the AR(1) process (30), the value of  $\Delta \ln(x_{T+h})$  is given by

$$\Delta \ln(x_{T+h}) - \mu_{S_T} = \beta_{S_T}^h (\Delta \ln(x_T) - \mu_{S_T}) + \sum_{i=1}^h \beta_{S_T}^{h-i} \varepsilon_{T+1+h-i}. \quad (39)$$

Using (37), we see after some algebra that

$$\ln(x_{T+h}/x_T) = h\mu_{S_T} + \frac{\beta_{S_T}(1 - \beta_{S_T}^h)}{1 - \beta_{S_T}} [\Delta \ln(x_T) - \mu_{S_T}] + \varepsilon_{T+1:T+h}, \quad (40)$$

where  $\varepsilon_{T+1:T+h}$  denotes the weighted shocks to the present value of dividends between period  $T$  and  $T + h$ , defined by

$$\varepsilon_{T+1:T+h} = \varepsilon_{T+h} + \left(\frac{1 - \beta_{S_T}^2}{1 - \beta_{S_T}}\right) \varepsilon_{T+h-1} + \left(\frac{1 - \beta_{S_T}^3}{1 - \beta_{S_T}}\right) \varepsilon_{T+h-2} \dots + \left(\frac{1 - \beta_{S_T}^h}{1 - \beta_{S_T}}\right) \varepsilon_{T+1}.$$

It is now easily seen that

$$\text{Var}(\varepsilon_{T+1:T+h}) = \frac{h + \frac{\beta_{S_T}^2(1 - \beta_{S_T}^{2h})}{1 - \beta_{S_T}^2} - \frac{2\beta_{S_T}(1 - \beta_{S_T}^h)}{1 - \beta_{S_T}}}{(1 - \beta_{S_T})^2}. \quad (41)$$

Future dividends can therefore be simulated by drawing a set of parameters,  $\{\mu_{S_T}, \beta_{S_T}\}$ , and, for these parameters, compute

$$x_{T+h} = x_T \exp\left\{h\mu_{S_T} + \frac{\beta_{S_T}(1 - \beta_{S_T}^h)}{1 - \beta_{S_T}} [\Delta \ln(x_T) - \mu_{S_T}] + \varepsilon_{T+1:T+h}\right\}, \quad (42)$$

where  $\varepsilon_{T+1:T+h}$  has mean zero and variance as given in (41).

Turning to the case that allows for breaks between periods  $T + 1$  and  $T + h$ , and recalling that  $\tilde{\mu}_{S_{T+i}} = \mu_{S_{T+i}}(1 - \beta_{S_{T+i}})$ , we have

$$\begin{aligned}
\Delta \ln(x_{T+1}) &= \tilde{\mu}_{S_{T+1}} + \beta_{S_{T+1}} \Delta \ln(x_T) + \varepsilon_{T+1} \\
\Delta \ln(x_{T+2}) &= \tilde{\mu}_{S_{T+2}} + \beta_{S_{T+2}} \tilde{\mu}_{S_{T+1}} + \beta_{S_{T+2}} \beta_{S_{T+1}} \Delta \ln(x_T) + \varepsilon_{T+2} + \beta_{S_{T+2}} \varepsilon_{T+1} \\
\Delta \ln(x_{T+3}) &= \tilde{\mu}_{S_{T+3}} + \beta_{S_{T+3}} \tilde{\mu}_{S_{T+2}} + \beta_{S_{T+3}} \beta_{S_{T+2}} \tilde{\mu}_{S_{T+1}} + \beta_{S_{T+3}} \beta_{S_{T+2}} \beta_{S_{T+1}} \Delta \ln(x_T) \\
&\quad + \varepsilon_{T+3} + \beta_{S_{T+3}} \varepsilon_{T+2} + \beta_{S_{T+3}} \beta_{S_{T+2}} \varepsilon_{T+1} \\
&\quad \vdots \\
\Delta \ln(x_{T+h}) &= \tilde{\mu}_{S_{T+h}} + \varepsilon_{T+h} + \sum_{j=1}^{h-1} \left( \prod_{i=j+1}^h \beta_{S_{T+i}} \right) (\tilde{\mu}_{S_{T+j}} + \varepsilon_{T+j}) + \left( \prod_{i=j+1}^h \beta_{S_{T+i}} \right) \Delta \ln(x_T).
\end{aligned} \tag{43}$$

Comparing this expression to (39), clearly, the presence of breaks complicates calculations of future expected dividends very considerably and numerical methods are required to compute the present value.

#### 4.1. Computing Present Values

In practice we calculate the present value of the stock price as follows:

$$\lim_{H \rightarrow \infty} y_T = \lim_{H \rightarrow \infty} \sum_{h=1}^H \exp(-h \ln(1+r)) \int x_{T+h} p(x_{T+h} | \mathcal{I}_T) dx_{T+h}, \tag{44}$$

where  $\mathcal{I}_t$  is again the forecaster's information set at time  $T$ , which we shall assume comprises past dividends only, i.e.  $\mathcal{I}_T = \{x_1, \dots, x_T\}$ .  $p(x_{T+h} | \mathcal{I}_T)$  is the predictive density of dividends at time  $T + h$  conditional on  $\mathcal{I}_T$ .

The expectation is computed under three different scenarios capturing different assumptions about the forecaster's beliefs:

1. A model that accounts for parameter estimation uncertainty but ignores past and future breaks to dividends by using the predictive density:

$$p(x_{T+h} | S_{T+h} = 1, \mathcal{I}_T) = \int p(x_{T+h} | \Theta_t, S_{T+h} = 1, \mathcal{I}_T) \pi(\Theta_T | S_{T+h} = 1, \mathcal{I}_T) d\Theta_T, \tag{45}$$

where  $\Theta$  are the constant model parameters whose posterior distribution given the data at time  $T$  is  $\pi(\Theta | S_{T+h} = 1, \mathcal{I}_T)$ .

2. A model that accounts for historical breaks to the dividend process but ignores the possibility of future breaks (and hence assumes that the last regime stays in effect forever but with uncertain parameters) by using the predictive density:

$$p(x_{T+h} | S_{T+h} = K+1, \mathcal{I}_T) = \int p(x_{T+h} | \Theta_{K+1}, S_{T+h} = S_T = K+1, \mathcal{I}_T) \pi(\Theta_{K+1} | \mathcal{H}, p, \mathcal{I}_T) d\Theta_{K+1}, \tag{46}$$

where  $\Theta_{K+1}$  are the parameters in the last regime (labelled  $K+1$ ), while  $\mathcal{H}$  is the set of hyper parameters.

3. A model that accounts for parameter estimation uncertainty as well as past and future breaks to the dividend process by using the following predictive density:

$$\int p(x_{T+h} | S_{T+h} = K + n_b + 1, S_T = K + 1, \mathcal{I}_T), \quad (47)$$

where  $n_b$  is the maximum number of out-of-sample breaks so the predictive density can be calculated (integrating out uncertainty about the dates of the breaks,  $\tau_{K+1} = T + j_1, \dots, \tau_{K+n_b} = j_{n_b}$ ) as

$$\begin{aligned} p(x_{T+h} | S_{T+h} = K + n_b + 1, S_T = K + 1, \mathcal{I}_T) &= \sum_{j_1=1}^{h-n_b+1} \dots \sum_{j_{n_b}=j_{n_b-1}+1}^h \int \dots \int \\ p(x_{T+h}, \Theta_{K+2}, \dots, \Theta_{K+n_b+1}, \mathcal{H}, S_{T+h} = K + n_b + 1, \\ \tau_{K+1} = T + j_1, \dots, \tau_{K+n_b} = T + j_{n_b}, S_T = K + 1, \mathcal{I}_T) \\ \times \pi(\tau_{K+1} = T + j_1, \dots, \tau_{K+n_b} = t + j_{n_b} | S_{T+h} = K + n_b + 1, S_T = K + 1) \\ \times \pi(\Theta_{K+2}, \dots, \Theta_{K+n_b+1}, \mathcal{H} | \mathcal{I}_T) d\Theta_{K+2} \dots d\Theta_{K+n_b+1} d\mathcal{H}. \end{aligned} \quad (48)$$

To get a more complete picture of the possible impact of breaks on the present value price, we compute the stock price under the three scenarios based on different terminal regimes. One thousand draws from the Gibbs sampler were used (after discarding the first 500 draws) to forecast dividends and the present value of dividends. In computing the monthly dividends, we set the forecast horizon at  $H = 1000$  and assumed an annualized discount rate of 10%. The parameters of the prior were as follows:  $p_{ii} \sim \text{Beta}(\underline{a}, \underline{b})$  with  $\underline{a} = \underline{b} = 0.5$ . We assume an uninformative prior for the parameters of the conditional mean of the dividend process by setting  $\underline{\mu}_\beta = \mathbf{0}_{2 \times 1}$ ,  $\underline{\mathbf{V}}_\beta = 1000\mathbf{I}_2$  (recall that  $\beta_i = (\tilde{\mu}_i, \beta_i)'$  in the  $i^{\text{th}}$  regime). The hyperparameters determining the error term precision are  $\underline{c}_0 = 1; \underline{d}_0 = 1/100; \underline{\rho}_0 = 100$ , while the prior for the transition probability matrix is assumed to be drawn from a  $\text{Gamma}(\underline{a}_0, \underline{b}_0)$  distribution with  $\underline{a}_0 = 1; \underline{b}_0 = 1/10$ .

To shed light on the practical importance of our choice of  $H$ , the terminal value at which dividends are computed, Figure 6 plots the present value of the expected dividend as a function of the forecast horizon,  $H$ . The sum of expected discounted dividends stabilizes rapidly under all three models, suggesting that, in this parameterization and for our choice of discount rate  $r = 10\%$  per annum, the present value is not very sensitive to our choice of  $H$ . Our earlier theoretical results suggest, however, that if we were to let  $H \rightarrow \infty$ , the results could be quite different. Consistent with this, when we chose a smaller value for the discount rate of  $r = 5\%$  per annum, the present value series failed to converge, blowing up in the process. These results suggest that in the presence of parameter estimation uncertainty and model instability the present value stock price can be very sensitive to modeling assumptions.

#### 4.2. Empirical Results

Table 5 reports the stock price computed under these assumptions relative to the stock price from the model that ignores model instability which we normalize at 100.<sup>6</sup> As indicated by the parameter

<sup>6</sup>Since we are using a finite horizon in these calculations, normalizing the present values in this way is innocuous.

estimates in Table 3, the mean value of the log first-differenced dividends in the regime prevailing during 1922-1930 was unusually high. As a result, in this regime the stock price computed under the assumption that the last regime stays in effect and no future breaks occur (the middle column in Table 5) was nearly 40% higher than the full-sample, no-break stock price. In this regime, the stock price computed under the composite-meta model that allows for both past and future breaks lies in the middle of the full-sample and current regime values, 20% above the no-break price. Stock prices under past and/or present breaks that exceed the value under the no-break assumption are also observed under the parameters based on the regime that was in effect during the period 1952-1960.

Conversely, the stock price based on dividend growth parameter values from the regime prevailing from 1930-1952 falls below the full-sample value by 7% since dividend growth was quite low in this regime as indicated by the negative intercept for this state shown in Table 3. Furthermore, in this regime the stock price computed under the composite meta distribution, at 101, is only marginally above the full-sample value. A similar set of results is obtained on the basis of the parameters from the last regime prevailing during 1960-2003.

The reason for these rankings is easy to understand from Figure 7: At short investment horizons, the weight on the current state tends to be very high, but this weight declines gradually as the horizon is expanded and the weight on draws from the meta distribution increases. Consequently, the stock price under the current regime lies above both the full-sample value and the price computed under the composite meta distribution whenever dividend growth is very high in the current regime, i.e. the parameters associated with the current regime are drawn from the right tail of the meta distribution. In this situation, the stock price under the composite-meta distribution is also likely to be considerably higher than its full-sample counterpart, but it falls below the value conditioned on remaining in the last regime since dividend growth after a future break is likely to be below the growth rate in the current regime.

These results also demonstrate that, in general, the stock price under the composite meta distribution will tend to be above the full-sample value due to the convexity of the mapping from the dividend growth rate to the stock price implied by the present value relation (see Timmermann (2001)). This explains why we see higher stock prices as a result of accounting for parameter uncertainty and model instability. Furthermore, model instability generally increases the effect of parameter uncertainty. The intuition for this finding is that under breaks fewer observations are effectively used to estimate the model parameters in the last regime, so the standard errors of the parameters tend to increase under breaks compared with full-sample estimates.

However, in a given regime, any ranking between stock prices under the three scenarios is in fact possible. For example, if the current state experiences a sufficiently low dividend growth rate and the state is highly persistent, then the effect of conditioning the stock price on the dividend growth parameters from the current state will dominate the convexity effect and hence the stock price under the composite meta distribution (as well as under the assumption that the current state remains in effect) will be smaller than the full-sample, no break price.

Furthermore, stock prices under the model that accounts for breaks need not exceed prices under a no-break assumption in models where shocks to the dividend growth process are correlated

with a stochastic discount factor and uncertainty surrounding future dividend growth leads to a higher risk premium. This will occur, for example, in consumption asset pricing models where dividends have a large positive correlation with consumption growth. But our discussion of the Lucas model in Section 2.4 suggests that taking account of possible correlations between dividends and consumption growth might not be sufficient to resolve the non-convergence problem so long as there are important uncertainties surrounding the future mean dividend growth rates.

## 5. Conclusion

This paper showed how to compute stock prices as the present value of future dividends when we do not assume that the dividend process is stable through time and that its parameters are known. We showed that stock prices can be quite sensitive to the nature of the assumptions concerning uncertainty and instability of the parameters of the dividend process. These findings suggest that our understanding of the dynamics in stock prices can be improved by focusing on the uncertainty surrounding the underlying fundamentals process.

Our emphasis on the sensitivity of present values to uncertainty about the growth rate in different ‘regimes’ or states of nature is closely related to the literature on how investors’ learning about the dividend growth process can give rise to the ‘excess volatility’ patterns observed for asset prices, c.f. Timmermann (1993). Even with a finite horizon, present values can be very sensitive to small changes in the estimated growth rate, particularly as this gets close to the discount rate. It is also related to recent work on asset pricing puzzles by Weitzmann (2005) who, following earlier insights by Geweke (2001), points out the sensitivity of equilibrium asset prices and returns to assumptions concerning the precision of the parameters characterizing the distribution from which fundamentals are generated, questions the ergodicity assumption made in much of the rational expectations literature and proposes modifications to this. For example, Weitzmann (2005) writes that “the unobservable nature of structural growth parameters adds to expectation beliefs a permanent thick-tailed background layer of uncertainty that never converges to a stationary-ergodic rational expectations equilibrium.”

The empirical results presented here clearly have implications for the equity premium puzzle, although we chose not to address this issue here. Recent papers by Barro (2005), de Santis (2005) and Weitzmann (2005) emphasize the importance of parameter uncertainty, instability and rare events as potential explanations of the historically large equity premium. Indeed, through their large effect on the present value stock price, persistent shifts in the dividend growth rate tend to increase the uncertainty about future returns which may be a reason why a larger equity premium is required compared to the standard model that ignores such effects.<sup>7</sup>

---

<sup>7</sup>Allowing for estimation uncertainty could also be important in resolving the so called Deaton’s paradox, namely the excess smoothness of observed consumption growth to changes in labour income growth. See, for example, Deaton (1992).

## References

- [1] Alogoskoufis, G.S. and R. Smith, 1991, The Phillips Curve, the Persistence of Inflation, and the Lucas Critique: Evidence from Exchange Rate Regimes. *American Economic Review* 81, 1254-1275.
- [2] Banerjee, A., R. Lumsdaine and J.H. Stock, 1992, Recursive and Sequential Tests of the Unit-Root and Trend-Break Hypotheses: Theory and International Evidence. *Journal of Business and Economic Statistics* 10, 271-287.
- [3] Barro, R.J., 2005, Rare Events and the Equity Premium. Mimeo, Harvard.
- [4] Blanchard, Oliver J. and Stanley Fischer (1989), *Lectures on Macroeconomics*, MA: MIT Press.
- [5] Campbell, J.Y., A. Lo and C. MacKinlay, 1997, *The Econometrics of Financial Markets*. Princeton University Press: Princeton.
- [6] Cass, David and Menahem E. Yaari (1967), "Individual Saving, Aggregate Capital Accumulation, and Efficient Growth." In K. Shell (ed.), *Essays on the Theory of Optimal Economic Growth*. Cambridge, MA: MIT Press.
- [7] Chib, S., 1998, Estimation and Comparison of Multiple Change Point Models, *Journal of Econometrics*, 86, 221-241.
- [8] Clements, M.P. and D.F. Hendry, 1998, *Forecasting Economic Time Series*, Cambridge University Press.
- [9] Clements, M.P. and D.F. Hendry, 1999, *Forecasting Non-stationary Economic Time Series*, The MIT Press.
- [10] Cochrane, J., 2005, *Asset Pricing*. Revised Edition. Princeton University Press: Princeton.
- [11] de Santis, M., 2005, Movements in the Equity Premium: Evidence from Time-Varying VAR. Unpublished manuscript.
- [12] Deaton, A., (1992), *Understanding Consumption*, Clarendon Press, Oxford.
- [13] Garcia, R. and P. Perron, 1996, An Analysis of the Real Interest Rate under Regime Shifts. *Review of Economics and Statistics* 78, 111-125.
- [14] Geweke, J., 2001, A Note on Some Limitations of CRRA Utility. *Economics Letters* 71, 341-346.
- [15] Geweke, J., 2005, *Contemporary Bayesian Econometrics and Statistics*. New York: Wiley.
- [16] Koop, G. and S. Potter, 2004a, Forecasting and Estimating Multiple Change-point Models with an Unknown Number of Change-points. Mimeo, University of Leicester and Federal Reserve Bank of New York.



- [17] Koop, G. and S. Potter, 2004b, Prior Elicitation in Multiple Change-point Models. Mimeo, University of Leicester and Federal Reserve Bank of New York.
- [18] Lucas, R.E., 1978, Asset Prices in an Exchange Economy. *Econometrica* 46, 1429-1445.
- [19] McCulloch, R.E. and R. Tsay, 1993, Bayesian Inference and Prediction for Mean and Variance Shifts in Autoregressive Time Series. *Journal of the American Statistical Association* 88, 965-978.
- [20] Pastor, L. and R.F. Stambaugh, 2001, The Equity Premium and Structural Breaks. *Journal of Finance*, 56, 1207-1239.
- [21] Paye, B. and A. Timmermann, 2005, Instability of Return Prediction Models. Forthcoming in *Journal of Empirical Finance*.
- [22] Pesaran, M.H., D. Pettenuzzo and A. Timmermann, 2005, Forecasting Time Series Subject to Multiple Structural Breaks. Forthcoming in *Review of Economic Studies*.
- [23] Pesaran, M.H. and A. Timmermann, 2002, Market timing and return prediction under model instability. *Journal of Empirical Finance*, 9, 495-510.
- [24] Shiller, R.J., 2000, *Irrational Exuberance*. Princeton: Princeton University Press.
- [25] Stock, J.H. and M.W. Watson, 1996, Evidence on Structural Instability in Macroeconomic Time Series Relations. *Journal of Business and Economic Statistics* 14, 11-30.
- [26] Timmermann, A., 1993, How Learning in Financial Markets Generates Excess Volatility and Predictability in Stock Prices, *Quarterly Journal of Economics* 108, 1135-1145.
- [27] Timmermann, A., 2001, Structural Breaks, Incomplete Information, and Stock Prices. *Journal of Business and Economic Statistics* 19, 299-315.
- [28] Weitzmann, M.L., 2005, Risk, Uncertainty, and Asset Pricing ‘Puzzles’. Mimeo, Harvard University.
- [29] Yaari, M.E. (1965): Uncertain Lifetime, Life Insurance, and the Theory of the Consumer, *Review of Economic Studies*, 32, 137-150.

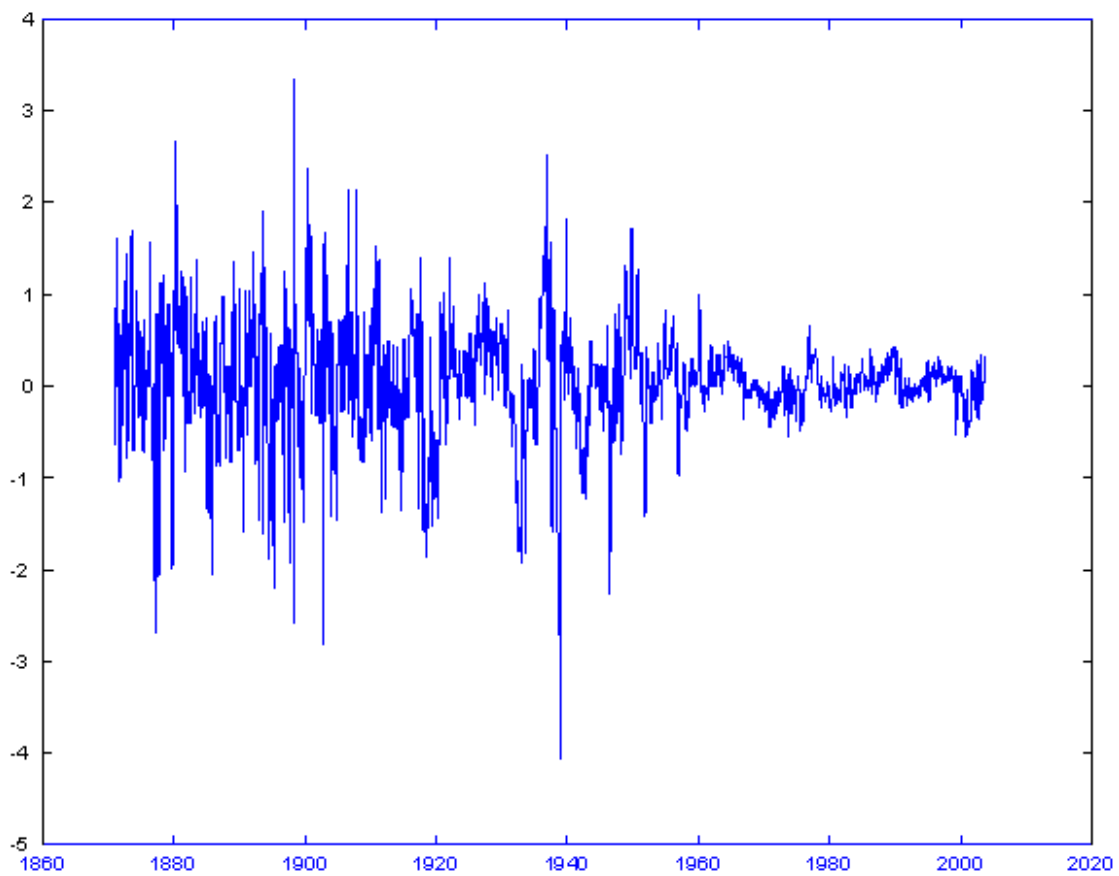


Figure 1: Monthly real dividend growth rates, 1871:3 - 2003:9

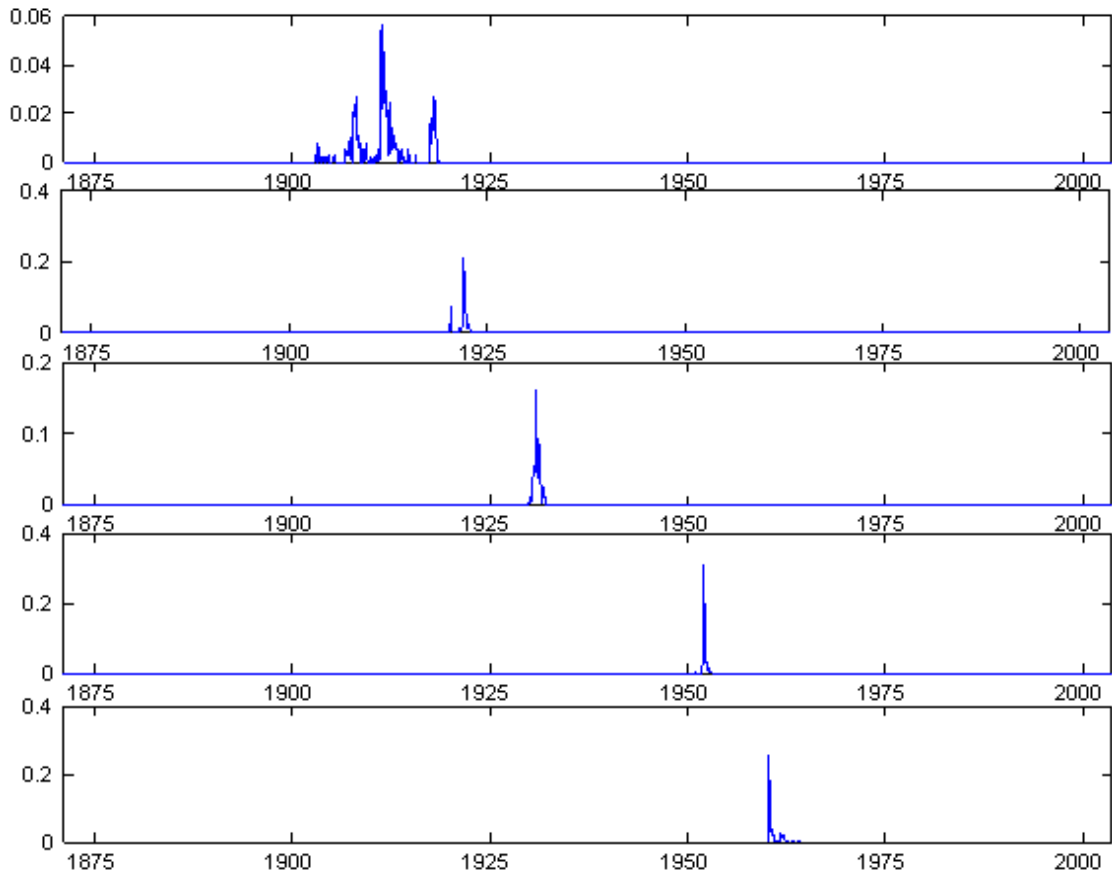


Figure 2: Posterior probability of break occurrence in the AR(1) model for the real dividend growth rate, assuming  $K = 5$ .

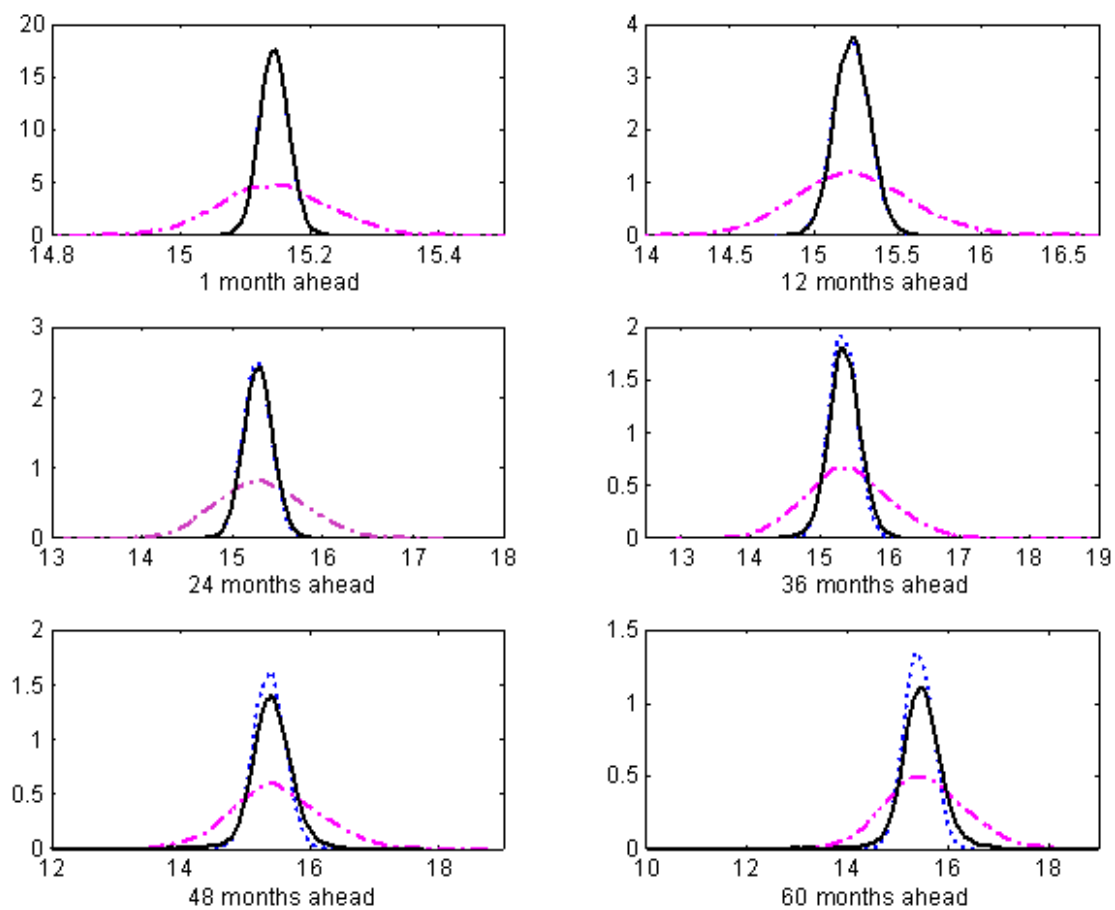


Figure 3: The graphs show the predictive distributions for the dividend series under various forecast horizons. The solid line represents the predictive density from the composite-meta model (assuming  $K=5$  breaks) while the dotted line represents the predictive density under the last regime (1960-2003). Finally, the dashed/dotted line represents the predictive density from the full sample/no breaks model.

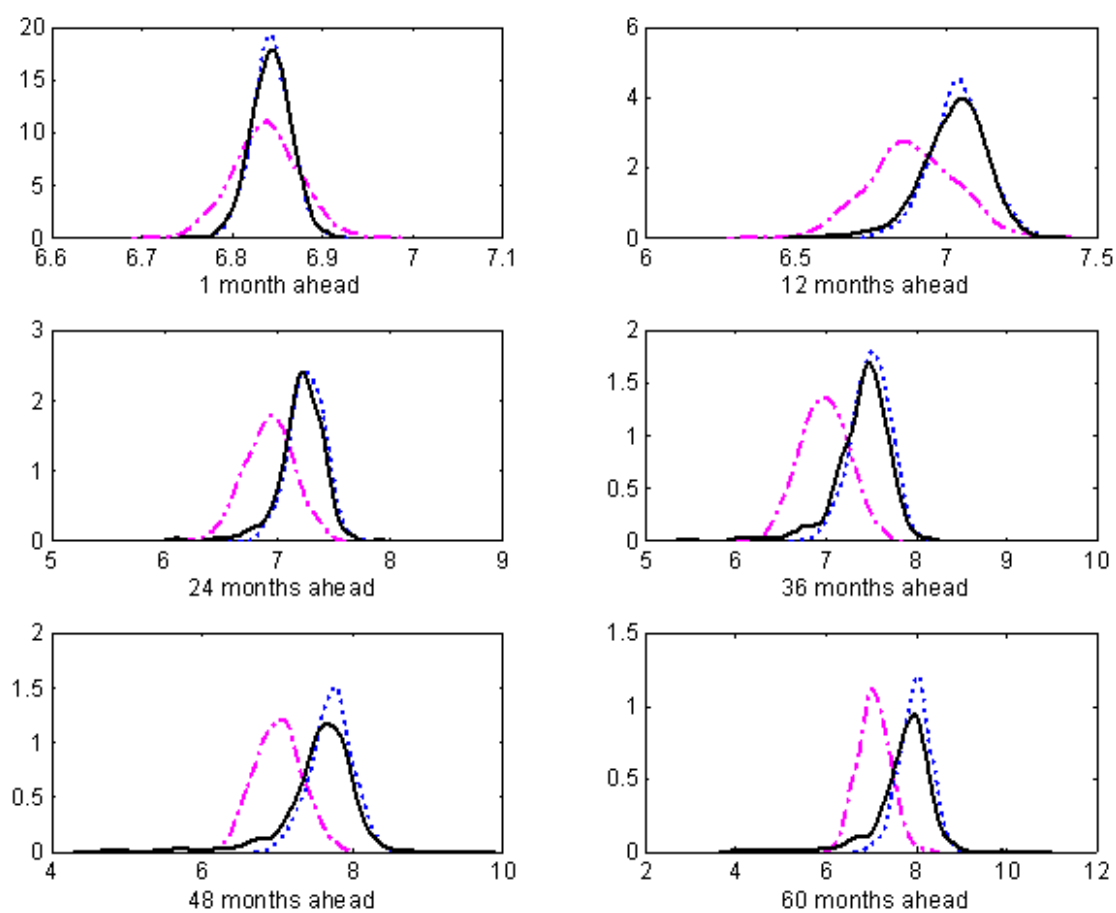


Figure 4: The graphs show the predictive distributions for the dividend series under various forecast horizons. The solid line represents the predictive density from the composite-meta model (assuming  $K=5$  breaks) while the dotted line represents the predictive density under regime 3 (1922-1930). Finally, the dashed/dotted line represents the predictive density from the full sample/no breaks model.

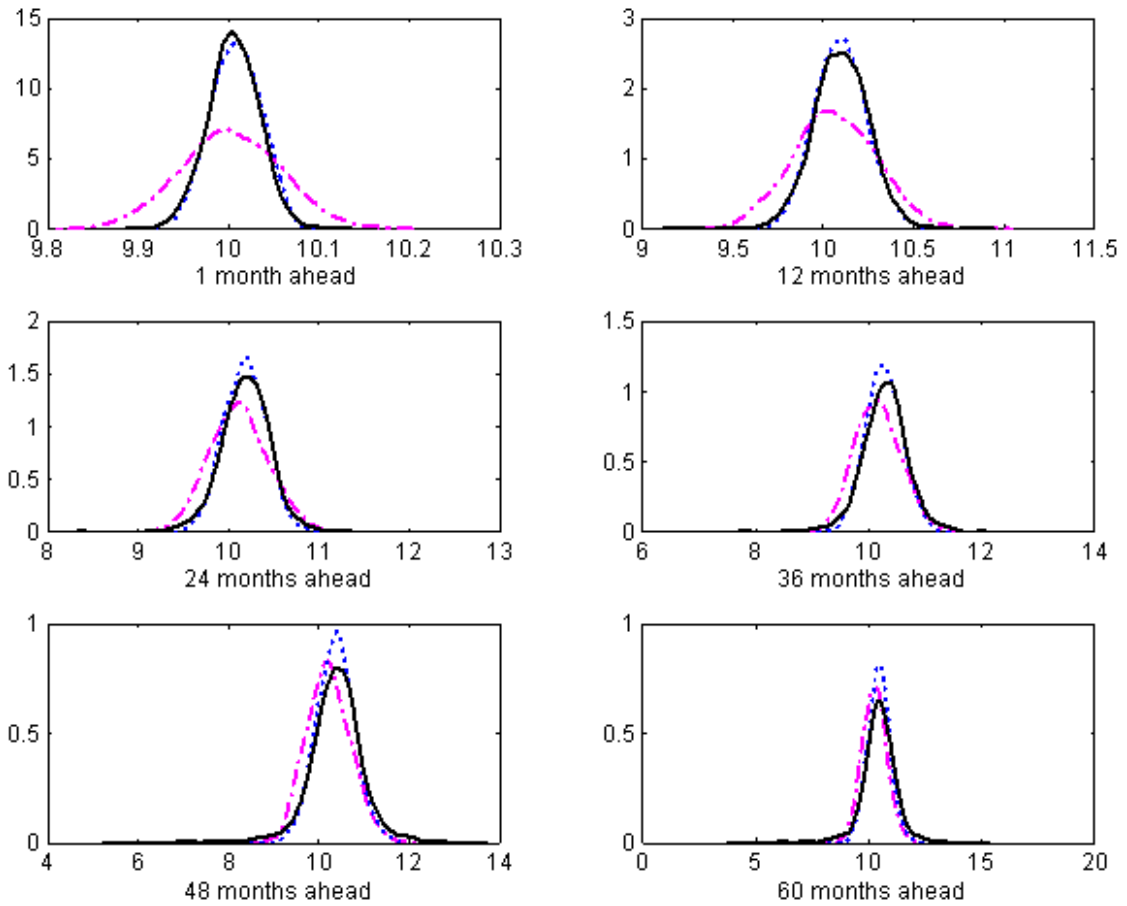


Figure 5: The graphs show the predictive distributions for the dividend series under various forecast horizons. The solid line represents the predictive density from the composite-meta model (assuming  $K=5$  breaks) while the dotted line represents the predictive density under regime 5 (1952-1960). Finally, the dashed/dotted line represents the predictive density from the full sample/no breaks model.

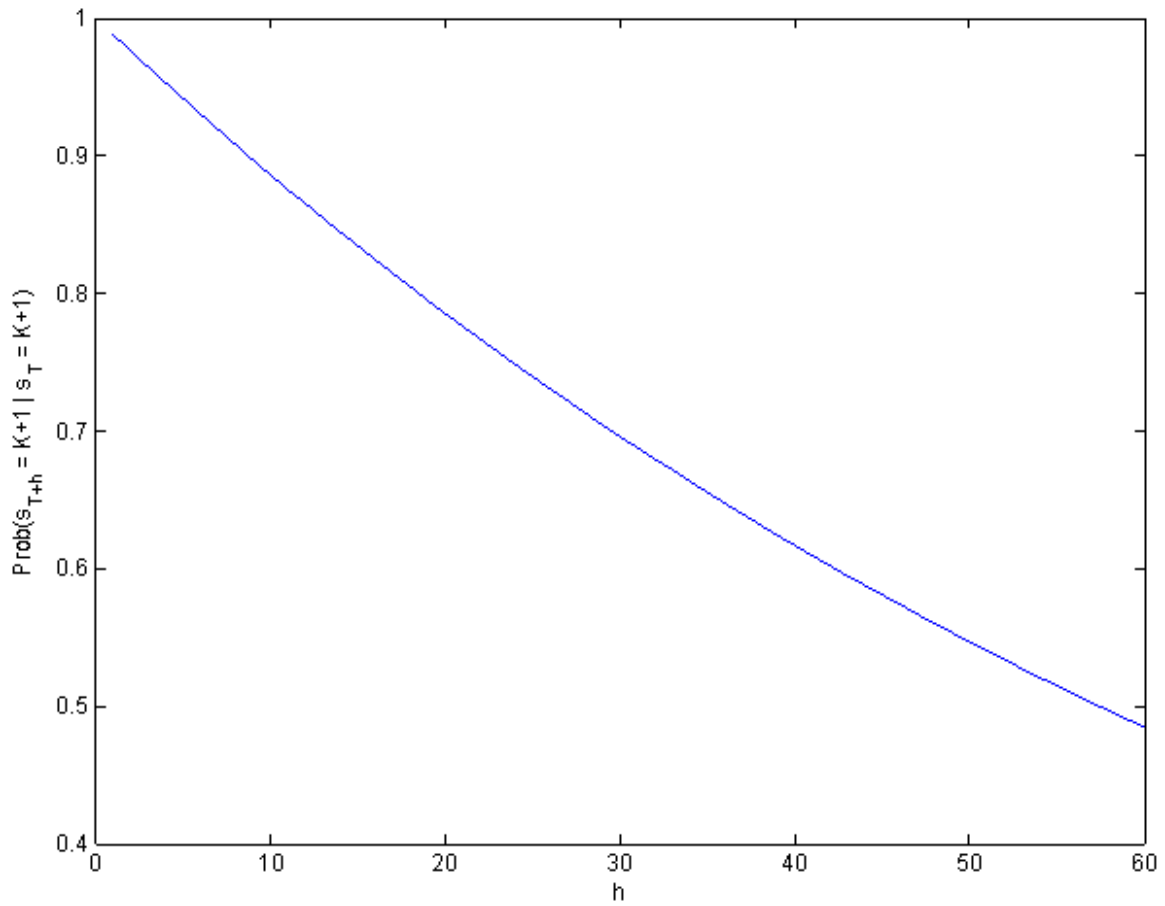


Figure 6: Posterior probability of staying in regime  $K + 1$  at time  $T + h$ ,  $\Pr(s_{T+h} = K + 1 | s_T = K + 1)$ .  $h$  is the forecast horizon that ranges from 1 to 60 months.

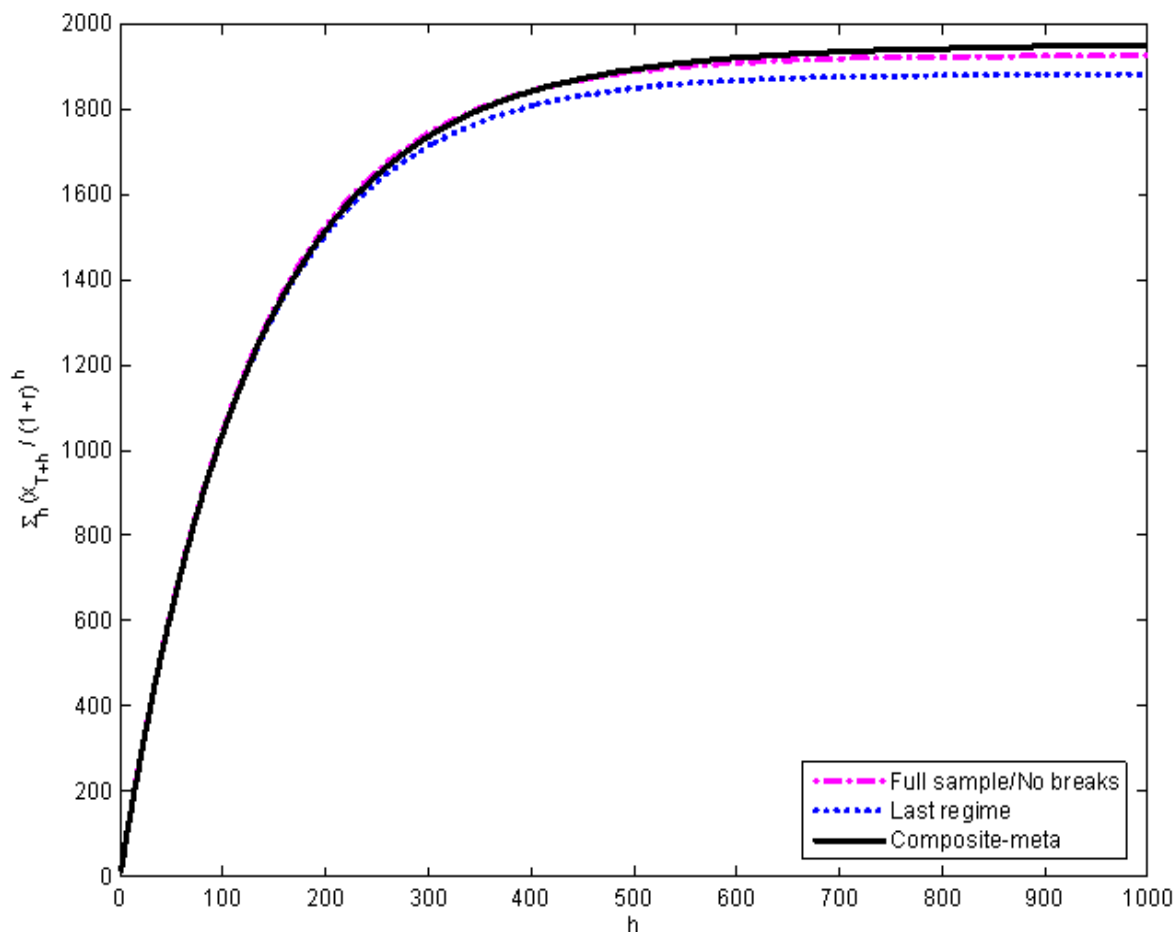


Figure 7: Cumulative discounted dividend  $\sum_{j=1}^h \frac{x_{T+j}}{(1+r)^j}$  under different scenarios, where  $1 \leq h \leq 1000$  months and  $r = 10\%$  per annum. The solid line represents the predictive density from the composite-meta model (assuming  $K = 5$  breaks) while the dotted line represents the predictive density under the last regime. Finally, the dashed/dotted line represents the predictive density from the full sample/no breaks model. The dividend is set at the end of sample value,  $x_T = 15.12$ .



$\Delta \log(x_t)$ Sample Statistics					
Mean	Std. dev.	Skewness	Kurtosis	1st order AC	12th order AC
0.0004	0.0064	-0.5133	7.5251	0.5220	0.0480

Table 1: Sample statistics for the monthly growth rate of the real dividends,  $\Delta \log(x_t)$ . The sample period is 1871:3-2003:9.

No. of breaks	Log lik.(LL)	Marginal LL	Break dates			
0	-1298.05	-1329.4062				
1	-806.482	-875.8323	Feb-52			
2	-749.527	-820.8898	Sep-11	Apr-60		
3	-723.086	-795.1873	May-11	Feb-52	Apr-60	
4	-705.76	-778.9393	Jan-11	Nov-30	Jan-52	Apr-60
5	-693.636	-769.1268	Jan-11	Jan-22	Nov-30	Jan-52
			Apr-60			
6	-692.465	N.A.	Nov-17	Jun-20	Feb-31	Jan-52
			Sep-60	Oct-60		

Table 2: Model comparison. This table shows log likelihood and marginal log likelihood values for first-order autoregressive models with different numbers of breaks. Also reported are posterior modes of the time of the breaks.

Parameters estimates						
Regimes						
	1	2	3	4	5	6
date	71-11	11-22	22-30	30-52	52-60	60-04
Constant						
Mean	0.071	-0.022	0.155	-0.009	0.030	0.009
s.e.	0.035	0.044	0.052	0.033	0.025	0.006
AR(1) coefficient						
Mean	0.372	0.606	0.398	0.729	0.621	0.658
s.e.	0.043	0.073	0.104	0.042	0.069	0.032
Standard deviation						
Mean	0.781	0.589	0.310	0.597	0.270	0.148
s.e.	0.028	0.039	0.024	0.027	0.020	0.005
Transition Probability matrix						
Mean	0.997	0.989	0.986	0.994	0.985	1
s.e.	0.003	0.010	0.011	0.005	0.012	0

Table 3: Posterior parameter estimates for the AR(1) hierarchical Hidden Markov Chain model with five break points in the monthly growth rate of real dividends.

Mean Parameters				
	Mean	s.e.	95% conf interval	
$b_0(1)$	0.039	0.040	-0.021	0.104
$b_0(2)$	0.563	0.078	0.430	0.684
Variance Parameters				
	Mean	s.e.		
$B_0(1, 1)$	0.008	0.008		
$B_0(2, 2)$	0.032	0.037		
Error term precision				
	Mean	s.e.	95% conf interval	
$v_0$	1.281	0.555	0.531	2.336
$d_0$	0.113	0.059	0.035	0.217

Table 4: Posterior estimates of the hyperparameters of the meta distribution for the AR(1) hierarchical Hidden Markov Chain model with five break points for the monthly growth rate of real dividends.

$y_T$			
Regime	Full sample/No breaks	Last regime	Composite-Meta
1960-2003	100	97.736	101.272
1922-1930	100	136.427	120.306
1930-1952	100	93.012	101.036
1952-1960	100	104.677	107.656

Table 5: Stock price  $y_T$  under different scenarios when the interest rate is set at 10 % per annum.

## **The following Discussion Papers have been published since 2005:**

### **Series 1: Economic Studies**

1	2005	Financial constraints and capacity adjustment in the United Kingdom – Evidence from a large panel of survey data	Ulf von Kalckreuth Emma Murphy
2	2005	Common stationary and non-stationary factors in the euro area analyzed in a large-scale factor model	Sandra Eickmeier
3	2005	Financial intermediaries, markets, and growth	F. Fecht, K. Huang, A. Martin
4	2005	The New Keynesian Phillips Curve in Europe: does it fit or does it fail?	Peter Tillmann
5	2005	Taxes and the financial structure of German inward FDI	Fred Ramb A. J. Weichenrieder
6	2005	International diversification at home and abroad	Fang Cai Francis E. Warnock
7	2005	Multinational enterprises, international trade, and productivity growth: Firm-level evidence from the United States	Wolfgang Keller Steven R. Yeaple
8	2005	Location choice and employment decisions: a comparison of German and Swedish multinationals	S. O. Becker, K. Ekholm, R. Jäckle, M.-A. Muendler
9	2005	Business cycles and FDI: evidence from German sectoral data	Claudia M. Buch Alexander Lipponer
10	2005	Multinational firms, exclusivity, and the degree of backward linkages	Ping Lin Kamal Saggi

11	2005	Firm-level evidence on international stock market comovement	Robin Brooks Marco Del Negro
12	2005	The determinants of intra-firm trade: in search for export-import magnification effects	Peter Egger Michael Pfaffermayr
13	2005	Foreign direct investment, spillovers and absorptive capacity: evidence from quantile regressions	Sourafel Girma Holger Görg
14	2005	Learning on the quick and cheap: gains from trade through imported expertise	James R. Markusen Thomas F. Rutherford
15	2005	Discriminatory auctions with seller discretion: evidence from German treasury auctions	Jörg Rocholl
16	2005	Consumption, wealth and business cycles: why is Germany different?	B. Hamburg, M. Hoffmann, J. Keller
17	2005	Tax incentives and the location of FDI: evidence from a panel of German multinationals	Thiess Buettner Martin Ruf
18	2005	Monetary Disequilibria and the Euro/Dollar Exchange Rate	Dieter Nautz Karsten Ruth
19	2005	Berechnung trendbereinigter Indikatoren für Deutschland mit Hilfe von Filterverfahren	Stefan Stamford
20	2005	How synchronized are central and east European economies with the euro area? Evidence from a structural factor model	Sandra Eickmeier Jörg Breitung
21	2005	Asymptotic distribution of linear unbiased estimators in the presence of heavy-tailed stochastic regressors and residuals	J.-R. Kurz-Kim S.T. Rachev G. Samorodnitsky

22	2005	The Role of Contracting Schemes for the Welfare Costs of Nominal Rigidities over the Business Cycle	Matthias Paustian
23	2005	The cross-sectional dynamics of German business cycles: a bird's eye view	J. Döpke, M. Funke S. Holly, S. Weber
24	2005	Forecasting German GDP using alternative factor models based on large datasets	Christian Schumacher
25	2005	Time-dependent or state-dependent price setting? – micro-evidence from German metal-working industries –	Harald Stahl
26	2005	Money demand and macroeconomic uncertainty	Claus Greiber Wolfgang Lemke
27	2005	In search of distress risk	J. Y. Campbell, J. Hilscher, J. Szilagyi
28	2005	Recursive robust estimation and control without commitment	Lars Peter Hansen Thomas J. Sargent
29	2005	Asset pricing implications of Pareto optimality with private information	N. R. Kocherlakota Luigi Pistaferri
30	2005	Ultra high frequency volatility estimation with dependent microstructure noise	Y. Aït-Sahalia, P. A. Mykland, L. Zhang
31	2005	Umstellung der deutschen VGR auf Vorjahrespreisbasis – Konzept und Konsequenzen für die aktuelle Wirtschaftsanalyse sowie die ökonomische Modellierung	Karl-Heinz Tödter



32	2005	Determinants of current account developments in the central and east European EU member states – consequences for the enlargement of the euro area	Sabine Herrmann Axel Jochem
33	2005	An estimated DSGE model for the German economy within the euro area	Ernest Pytlarczyk
34	2005	Rational inattention: a research agenda	Christopher A. Sims
35	2005	Monetary policy with model uncertainty: distribution forecast targeting	Lars E.O. Svensson Noah Williams
36	2005	Comparing the value relevance of R&D reporting in Germany: standard and selection effects	Fred Ramb Markus Reitzig
37	2005	European inflation expectations dynamics	J. Döpke, J. Dovern U. Fritsche, J. Slacalek
38	2005	Dynamic factor models	Sandra Eickmeier Jörg Breitung
39	2005	Short-run and long-run comovement of GDP and some expenditure aggregates in Germany, France and Italy	Thomas A. Knetsch
40	2005	A “wreckers theory” of financial distress	Ulf von Kalckreuth
41	2005	Trade balances of the central and east European EU member states and the role of foreign direct investment	Sabine Herrmann Axel Jochem
42	2005	Unit roots and cointegration in panels	Jörg Breitung M. Hashem Pesaran
43	2005	Price setting in German manufacturing: new evidence from new survey data	Harald Stahl

1	2006	The dynamic relationship between the Euro overnight rate, the ECB's policy rate and the term spread	Dieter Nautz Christian J. Offermanns
2	2006	Sticky prices in the euro area: a summary of new micro evidence	Álvarez, Dhyne, Hoeberichts Kwapil, Le Bihan, Lünnemann Martins, Sabbatini, Stahl Vermeulen, Vilmunen
3	2006	Going multinational: What are the effects on home market performance?	Robert Jäckle
4	2006	Exports versus FDI in German manufacturing: firm performance and participation in international markets	Jens Matthias Arnold Katrin Hussinger
5	2006	A disaggregated framework for the analysis of structural developments in public finances	Kremer, Braz, Brosens Langenus, Momigliano Spolander
6	2006	Bond pricing when the short term interest rate follows a threshold process	Wolfgang Lemke Theofanis Archontakis
7	2006	Has the impact of key determinants of German exports changed? Results from estimations of Germany's intra euro-area and extra euro-area exports	Kerstin Stahn
8	2006	The coordination channel of foreign exchange intervention: a nonlinear microstructural analysis	Stefan Reitz Mark P. Taylor
9	2006	Capital, labour and productivity: What role do they play in the potential GDP weakness of France, Germany and Italy?	Antonio Bassanetti Jörg Döpke, Roberto Torrini Roberta Zizza
10	2006	Real-time macroeconomic data and ex ante predictability of stock returns	J. Döpke, D. Hartmann C. Pierdzioch

11	2006	The role of real wage rigidity and labor market frictions for unemployment and inflation dynamics	Kai Christoffel Tobias Linzert
12	2006	Forecasting the price of crude oil via convenience yield predictions	Thomas A. Knetsch
13	2006	Foreign direct investment in the enlarged EU: do taxes matter and to what extent?	Guntram B. Wolff
14	2006	Inflation and relative price variability in the euro area: evidence from a panel threshold model	Dieter Nautz Juliane Scharff
15	2006	Internalization and internationalization under competing real options	Jan Hendrik Fisch
16	2006	Consumer price adjustment under the microscope: Germany in a period of low inflation	Johannes Hoffmann Jeong-Ryeol Kurz-Kim
17	2006	Identifying the role of labor markets for monetary policy in an estimated DSGE model	Kai Christoffel Keith Küster Tobias Linzert
18	2006	Do monetary indicators (still) predict euro area inflation?	Boris Hofmann
19	2006	Fool the markets? Creative accounting, fiscal transparency and sovereign risk premia	Kerstin Bernoth Guntram B. Wolff
20	2006	How would formula apportionment in the EU affect the distribution and the size of the corporate tax base? An analysis based on German multinationals	Clemens Fuest Thomas Hemmelgarn Fred Ramb

21	2006	Monetary and fiscal policy interactions in a New Keynesian model with capital accumulation and non-Ricardian consumers	Campbell Leith Leopold von Thadden
22	2006	Real-time forecasting and political stock market anomalies: evidence for the U.S.	Martin Bohl, Jörg Döpke Christian Pierdzioch
23	2006	A reappraisal of the evidence on PPP: a systematic investigation into MA roots in panel unit root tests and their implications	Christoph Fischer Daniel Porath
24	2006	Margins of multinational labor substitution	Sascha O. Becker Marc-Andreas Müндler
25	2006	Forecasting with panel data	Badi H. Baltagi
26	2006	Do actions speak louder than words? Household expectations of inflation based on micro consumption data	Atsushi Inoue Lutz Kilian Fatma Burcu Kiraz
27	2006	Learning, structural instability and present value calculations	H. Pesaran, D. Pettenuzzo A. Timmermann

## Series 2: Banking and Financial Studies

1	2005	Measurement matters – Input price proxies and bank efficiency in Germany	Michael Koetter
2	2005	The supervisor's portfolio: the market price risk of German banks from 2001 to 2003 – Analysis and models for risk aggregation	Christoph Memmel Carsten Wehn
3	2005	Do banks diversify loan portfolios? A tentative answer based on individual bank loan portfolios	Andreas Kamp Andreas Pfingsten Daniel Porath
4	2005	Banks, markets, and efficiency	F. Fecht, A. Martin
5	2005	The forecast ability of risk-neutral densities of foreign exchange	Ben Craig Joachim Keller
6	2005	Cyclical implications of minimum capital requirements	Frank Heid
7	2005	Banks' regulatory capital buffer and the business cycle: evidence for German savings and cooperative banks	Stéphanie Stolz Michael Wedow
8	2005	German bank lending to industrial and non-industrial countries: driven by fundamentals or different treatment?	Thorsten Nestmann
9	2005	Accounting for distress in bank mergers	M. Koetter, J. Bos, F. Heid C. Kool, J. Kolari, D. Porath
10	2005	The eurosystem money market auctions: a banking perspective	Nikolaus Bartzsch Ben Craig, Falko Fecht
11	2005	Financial integration and systemic risk	Falko Fecht Hans Peter Grüner

12	2005	Evaluating the German bank merger wave	Michael Koetter
13	2005	Incorporating prediction and estimation risk in point-in-time credit portfolio models	A. Hamerle, M. Knapp, T. Liebig, N. Wildenauer
14	2005	Time series properties of a rating system based on financial ratios	U. Krüger, M. Stötzel, S. Trück
15	2005	Inefficient or just different? Effects of heterogeneity on bank efficiency scores	J. Bos, F. Heid, M. Koetter, J. Kolatri, C. Kool
01	2006	Forecasting stock market volatility with macroeconomic variables in real time	J. Döpke, D. Hartmann C. Pierdzioch
02	2006	Finance and growth in a bank-based economy: is it quantity or quality that matters?	Michael Koetter Michael Wedow
03	2006	Measuring business sector concentration by an infection model	Klaus Düllmann
04	2006	Heterogeneity in lending and sectoral growth: evidence from German bank-level data	Claudia M. Buch Andrea Schertler Natalja von Westernhagen
05	2006	Does diversification improve the performance of German banks? Evidence from individual bank loan portfolios	Evelyn Hayden Daniel Porath Natalja von Westernhagen

## **Visiting researcher at the Deutsche Bundesbank**

The Deutsche Bundesbank in Frankfurt is looking for a visiting researcher. Among others under certain conditions visiting researchers have access to a wide range of data in the Bundesbank. They include micro data on firms and banks not available in the public. Visitors should prepare a research project during their stay at the Bundesbank. Candidates must hold a Ph D and be engaged in the field of either macroeconomics and monetary economics, financial markets or international economics. Proposed research projects should be from these fields. The visiting term will be from 3 to 6 months. Salary is commensurate with experience.

Applicants are requested to send a CV, copies of recent papers, letters of reference and a proposal for a research project to:

Deutsche Bundesbank  
Personalabteilung  
Wilhelm-Epstein-Str. 14

D - 60431 Frankfurt  
GERMANY

