Business Cycle and Stock Market Volatility: A Particle Filter Approach

Roberto Casarin[†‡¶] and Carmine Trecroci[‡§]

[†]CEREMADE and Dept. of Mathematics, University Paris Dauphine

[‡]Dept. of Economics, University of Brescia

[§]Dept. of Economics, University of Glasgow

This Version: February, 2007

Abstract

The recent observed decline of business cycle variability suggests that broad macroeconomic risk may have fallen as well. This may in turn have some impact on equity risk premia. We investigate the latent structures in the volatilities of the business cycle and stock market valuations by estimating a Markov switching stochastic volatility model. The novelty of the work relies upon the on-line analysis of the macroeconomic risk based on a new sequential Monte Carlo technique, which allow for a Bayesian sequential inference on both the parameters and the latent variables. The proposed multiscale particle filter is particularly useful when exploring sequentially the posterior distribution of the parameter and of the latent variable. In our univariate setting, we find that the switch to lower variability has occurred in both business cycle and stock market variables along similar patterns.

KEYWORDS: Markov Switching, Stochastic Volatility, Business Cycle, Equity Market, Particle Filters, Bayesian Inference.

JEL CLASSIFICATION: C11, C15, C22, C63, E32, E44.

[¶]Address: CEREMADE, University Paris Dauphine, 16 place du Maréchal de Lattre de Tassigny, 75775 Paris Cédex 16. E-mail: `casarin@ceremade.dauphine.fr`. We are grateful to Christian P. Robert for his useful suggestions and to the University of Brescia for financial support. We thank R. Shiller for providing his valuable dataset and participants to the *Computational Statistics and Data Analysis Meeting*, Limassol, Cyprus, 28-31 October 2005 and to *BSMF* seminars, Brescia, 2006, for helpful comments.

1 Introduction

The observed behaviour of the volatility of stock market returns is in strident clash with most of the implications of standard theoretical asset pricing models. For instance, standard consumption-based models imply a constant price-dividend ratio, which in turn requires that the volatility of returns should equal the volatility of dividend growth. The post-WWII average of US stock market monthly return volatility has been about 16% per year, compared with only 7% for dividends. Moreover, return volatility is not just high, but it is also time-varying. Historically, monthly market return volatility has fluctuated between 20 - 25 % in the 1930s to less that 2% in the middle of the 1960s. In the long run, the dynamics of dividend volatility of course depends on the variability of companies' earnings, and these likely co-vary with the business cycle. In principle, some relationship between the volatilities of market returns and aggregate economic activity should emerge.

Influential contributions on the business cycle argue that the early 1980s marked a structural shifts in business cycle volatility (McConnell and Perez-Quiros (2000); Stock and Watson (2002)). In the US, since 1984-86 expansions have lasted longer and slowdowns have been less frequent and shallower than in the previous several decades. Stock and Watson (2002) dubbed this reduction of volatility as the "Great Moderation". Can any similar shift in volatility be observed in stock market returns?

In theory, the run-up of stock prices in the late '90s might be related to a decline of the equity premium. This fall could be due to a significant decrease of broad macroeconomic risk, as represented by an exogenous decline in business cycle volatility. Indeed, some authors (Lettau et al. (2007)) try to offer explicit rationalizations of this linkage. Now, besides possible explanations for the widely noticed reduction in the volatility of economic activity (see for example Gordon (2005), Dynan et al. (2006)), there is a widespread perception that the relationship between business cycle and stock markets volatilities is far from straightforward to measure.

There are at least two intertwined reasons why asset prices might experience a boom. A permanent rise in total factor productivity could translate into a persistently higher level of earnings, which in turn could raise, for example, stock market valuations. On the other hand, non-fundamental shocks in the equity or housing markets, perhaps due to over-optimistic expectations about future productivity and returns, could boost prices in the short to medium term. In practice, stock prices represent a measure of the marginal value of firms' installed capital. Greater uncertainty about earnings prospects should immediately translate into more volatile stock prices, while equity and investment goods prices (returns) should therefore both be pro-cyclical (countercyclical) across the economy.

In this work we study the historical volatilities of business cycle and stock market fluctuations, trying to detect the existence of some common pattern. We focus on quarterly time series of U.S. stock indices and key macroeconomic variables over the past 40 years, and analyse their filtered volatilities using the methodology of particle filters. This evaluation exercise should also help to test the hypothesis that the documented reduction in business cycle volatility is associated to a fall in the equity premium, and then to the run-up in stock valuations at the end of 1990s. To this end one could, for instance, test the hypothesis that the decline in macroeconomic volatility affects the investors' perception of macroeconomic risk and the size of the equity premium by fitting a hidden Markov-switching model. This way one could pick up abrupt changes in volatility patterns, and identify common patterns across output (or its components) and stock market variables volatility. In this paper we perform a similar exercise, by employing the particle filter methodology to extract latent components in the time series of our interest. More in detail, we follow a Bayesian approach to time-series modelling (see Harrison and West (1997), Bauwens, Lubrano and Richard (1999), Kim and Nelson (1999)), and apply sequential Monte Carlo techniques to the joint estimation of the parameters and latent factors.

Particle filters are now widely employed in the evaluation of models for the financial market, in particular for stochastic volatility models (see for example Kim, Shephard and Chib (1998), Pitt and Shephard (1999) and Lopes and Marinho (2001)). Recently, some studies have proposed their application to macroeconometrics, for the analysis of general equilibrium models (Villaverde and Ramirez (2004a) and (2004b)), the extraction of information from the yield curve (Chopin and Pelgrin (2004)), and for the estimation of latent factors in business cycle analysis (Billio, Casarin and Sartore (2004)). The main advantage of these simulation-based techniques lies in their great flexibility when treating nonlinear dynamic models, which cannot be successfully handled through the traditional Kalman-Bucy or Hamilton-Kitagawa filters. Another advantage comes from the sequential nature of the particle filters, which allows the use of large datasets and to build on-line applications. In this work we extend the sequential Monte Carlo approach proposed by Liu and Chen (1998) by introducing a multiple-bandwidth kernel estimator of the posterior distribution of the parameters.

To sum up our findings, we detect a parallel decline of macroeconomic and stock market variability, with the switch to low volatility occurring first in the business cycle indicators.

More in detail, our business cycle indicators follow a low-volatility regime for the second part of our 1966-2003 sample. On the other hand, we identify a similar pattern for the volatility of all our stock market indicators: the market index return and its price-earnings and dividend-price ratios all switch to persistent low-volatility around 1991. Therefore, we cannot reject the view that the widely observed decline in US business cycle volatility prompted a similar persistent reduction in stock market volatility.

The paper is organized as follows. Section 2 describes some stylized facts about the variables we use and introduces the simple Bayesian dynamic model employed to identify the latent switching structures for stochastic volatility. Section 3 introduces particle filters for the estimation of the latent factors and discusses some parameter estimation issues. Section 3.2 discusses the multiple-bandwidth regularised filter and the convergence of the resulting algorithm. Section 4 comments on our estimation results on simulated and actual data. Section 5 concludes.

2 Are Economic Activity and Stock Market Volatilities Related?

2.1 Some Stylized Facts

Figure 1 shows in the top chart the quarterly growth rate of US real personal consumption expenditure per capita, and in the middle chart the change in real residential and non-residential fixed investment (RI and NRI, respectively, from 1966Q2 to 2003Q3[1].

In all our Figures, grey vertical bars denote NBER-dated contraction episodes. In the case of consumption one can easily notice a fall of volatility starting from mid-1980s. Aggregate investment data somehow yield the same visual impression. However, as the data span only a limited number of full economic cycles, some caution is in order. Volatilities appear to be markedly procyclical: inflation outbursts, oil price shocks and well-known phases of macroeconomic expansions and contractions all coincide with apparent volatility shifts.

This simple graphical evidence and more formal evidence gathered, for example, by Gordon (2005), show that the volatilities of residential investment and personal consumption expenditure do show a decline. However, NRI, which is perhaps the most interesting of our investment series, does not display any clear volatility shift. In fact,

[1] The definition and sources of all data are detailed in the Data Appendix.

recently its volatility appears to have been, if anything, slightly increasing. As to the causes, the absence of significant exogenous shocks and the generalized fall of inflation levels in the eighties and nineties might well account for the observed decrease in PCE and RI volatility. Also, sizeable portfolio reallocations triggered by the expansion of global liquidity and credit around and after the millennium certainly had some impact on investment choices, as claimed, inter alia, by Borio and Lowe (2004). All in all, however, these data confirm Stock and Watson (2002) conclusion that the volatility of economic activity in the US has markedly dropped after 1984. Given the observed decline in business cycle variability, the perception of macroeconomic risk since early 1980s might have decreased as well[2]. Let us therefore look at some stock market data.

The bottom chart of Figure 1 plots the quarterly real S&P500 price/earnings ratio, defined over a 10-year moving average of earnings, from 1946 to 2004. The sustained growth of equity valuations over the past two decades is apparent, as is their partial reversion at the turn of the millennium. Price-dividend ratio data (not shown) display a similarly trended pattern, with unprecedented low levels of the dividend yield in correspondence of the tech bubble. Do these data tell us anything about the link between macroeconomic and stock market volatilities? The bottom chart of Figure 1 also shows the output gap, computed by using the potential GDP measure of the Congressional Budget Office (see the Data Appendix). A cursory comparison of this series with the P/E data confirms an impressive common behaviour. However, we cannot accept this as evidence of a joint systematic relationship in the volatilities of stock returns and economic activity, as the plotted series' mean-reverting properties clearly differ.

According to theory, there are at least two ways to describe the problem at hand. First, the classical consumption-based asset pricing model[3] states that risk premia are proportional to the covariance of returns with consumption growth. Let us suppose that the preferences of the representative investor are time-separable

$$\mathbb{E}_t \left(\sum_{s=0}^{\infty} \beta^s u \left(c_{t+s} \right) \right) \tag{1}$$

with $\mathbb{E}_t(\cdot)$ representing the conditional expectation.

Optimal allocation of resources to consumption and investment implies that the marginal value of wealth and the marginal utility of consumption are equal. The first-

[2]Lettau et al. (2007) point to the same conclusion.

[3]Ferson (2003) and Cochrane (2005) are excellent surveys.

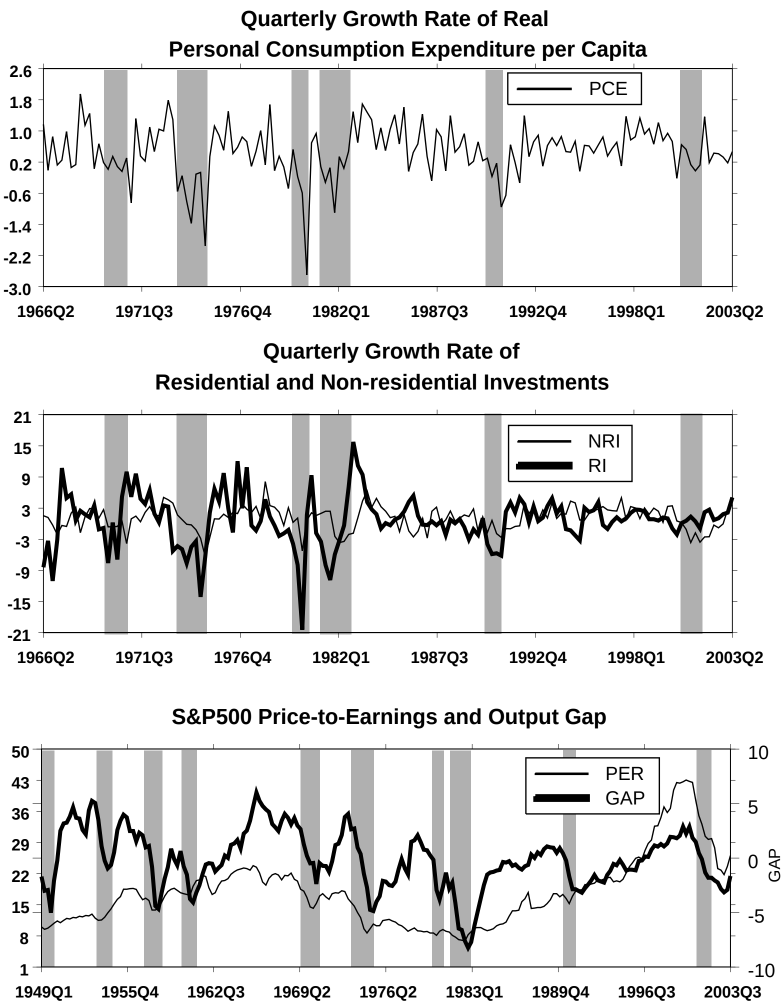

Figure 1: *Quarterly growth rate of real personal consumption expenditure per capita (top chart); quarterly growth rate of residential and non-residential investment expenditure (middle chart); S&P500-based price-to-earnings ratio (left-hand scale) and output gap (right-hand scale) (bottom chart). NBER recessions are denoted by shaded bars.*

order condition of the optimization problem therefore yields

$$\mathbb{E}_t\left(R^{ei}_{t+1}\right) = -\mathbb{C}ov_t\left(R^{ei}_{t+1}, \frac{u'\left(c_{t+1}\right)}{u'\left(c_t\right)}\right), \forall\, i \tag{2}$$

where $R^{ei}_t = R^i_t - R^0_t$ is the return of asset i in excess of a reference asset return R^0_t, and $\mathbb{C}ov_t\left(\cdot,\cdot\right)$ denotes conditional covariance. Using the conventional power utility function,

$u(c) = \frac{c^{1-\gamma}-1}{1-\gamma}$, we can rewrite equation (2) as

$$\mathbb{E}_t \left(R^{ei}_{t+1} \right) = \gamma \mathbb{C}ov_t \left(R^{ei}_{t+1}, \frac{c_{t+1}}{c_t} \right) \tag{3}$$

The last expression states that assets generate expected excess returns for their systematic risk, as summarized by the covariance of returns with the growth in the marginal utility of wealth (the so-called *stochastic discount factor*, *pricing kernel*, or *state-price density*). The covariance represents systematic risk because it measures the component of return that contributes to fluctuations in the marginal utility of wealth. By re-arranging (3) we can write

$$\mathbb{E}_t \left(R^{ei}_{t+1} \right) = \gamma \mathbb{V}ar_t \left(R^{ei}_{t+1} \right) \mathbb{V}ar_t \left(\Delta c_{t+1} \right) \rho_t \left(\Delta c, R^{ei}_{t+1} \right) \tag{4}$$

Risk premia vary over time. They move when the conditional variance of excess returns, $\mathbb{V}ar_t \left(R^{ei}_{t+1} \right)$, changes over time. But fluctuations in risk aversion (γ), the conditional correlation $\rho_t \left(\Delta c, R^{ei}_{t+1} \right)$, and consumption risk, $\mathbb{V}ar_t \left(\Delta c_{t+1} \right)$ as well explain changes in risk premia[4].

An alternative way of looking at the problem is by studying the present-value relations between stock prices, dividends and returns. As is well known, these relations become nonlinear when risk premia vary over time, and Campbell and Shiller (1988) provided an approximated loglinear framework that allows to derive tractable implications for prices and returns. For instance, one can write

$$r_{t+1} \approx k + \rho p_{t+1} + (1-\rho) d_{t+1} - p_t \tag{5}$$

where p_t and d_t denote the period t ex-dividend log price and log dividend payment, respectively, $r_t \equiv p_t + d_t - p_{t-1}$ is the one-period total return, and ρ and k are linearization parameters. Equation (5) says that returns are a weighted average of the log stock price and the log dividend. Solving forward, imposing $\lim_{j \to \infty} \rho^j p_{t+j} = 0$ (infinitely lived bubbles are barred) and $p_t = \mathbb{E}_t (p_t)$, one obtains

$$p_t = \frac{k}{1-\rho} + \sum_{j=0}^{\infty} \rho^j \left((1-\rho) d_{t+1+j} - r_{t+1+j} \right) \tag{6}$$

which says that if the stock price is high today, investors must be expecting some combination of high future dividends and low future returns. In terms of the log dividend-

[4]Note that equation (4) could well read backwards, with risk premia determining consumption volatility, conditional correlation, etc..

price ratio,

$$d_t - p_t = -\frac{k}{1-\rho} + \mathbb{E}_t \left(\sum_{j=0}^{\infty} \rho^j \left(-\Delta d_{t+1+j} + r_{t+1+j} \right) \right) \tag{7}$$

In this accounting framework it is straightforward to show (see Cochrane (2005) for details) that the log dividend-price ratio, stock returns and dividend growth are all stationary. This means that asymptotically the variance of returns must equal the variance of dividend growth:

$$\lim_{\tau\to\infty} \frac{1}{\tau} \mathbb{V}ar\left(p_{t+\tau} - p_t\right) = \lim_{\tau\to\infty} \frac{1}{\tau} \mathbb{V}ar\left(d_{t+\tau} - d_t\right) \tag{8}$$

More to our point, this implies that long-run price and dividend growth have an asymptotic correlation equal to one.

Including wealth and consumption in this picture makes for even clearer implications. Stock market valuations and aggregate consumption should be cointegrated: rising stock market wealth would lead to rising consumption levels, and vice versa. In fact, Lettau and Ludvigson (2001, 2004) find that the ratio of consumption to wealth forecasts future stock market returns. Moreover, while price/dividend and price/earnings ratios do forecast future share prices returns, they are not leading indicators of consumption growth. Therefore, most of short-run stock prices volatility is transitory. This also means that market data make the transmission of wealth effects from equity prices to consumption a quite complex process to disentangle.

More generally, aggregate market valuations respond to permanent and transitory changes in expected earnings, in turn affected by the current state of the economy. Expectations about future productivity and returns determine the equilibrium on the capital markets: share prices represent a measure of the marginal value of firms' installed capital. On the other hand, changes in share and house prices affect consumption via wealth effects, and alter the rental cost of capital for firms, through a variety of channels. Investment and consumption volatilities might therefore depend on past and expected changes in stock market values. All the above suggests that in studying whether the volatilities of the stock market and economic activity have some common component, a simpler, unrestricted approach, like a univariate approach, could yield vital information. We therefore choose to pursue this route in order to identify the volatility patterns of market valuations and business cycle variables, as motivated in the next subsection.

2.2 A Bayesian Dynamic Model

Our broad aim is to check for example whether common regime changes affect the stochastic volatility profile of stock market and macroeconomic variables. To this end we adopt the latent-variable modelling framework, which is now widely used in statistics (see for example Perpiñán (2001) for a review) and econometrics. In structural time series modelling (see Harvey (1989)), the evolution of the observed variable is described by means of a set of unobservable (latent) variables that allow capturing time-heterogenous behaviour of the observed series.

In the following, we consider a Bayesian representation of a dynamic model. The proposed model consists of the observed process $\{y_t, t \in \mathbb{N}_0\}$, with values in the measurable *observation space* $(E, \mathcal{E})$, $E \subset \mathbb{R}$, and of two latent processes: the volatility regime $\{s_t, t \in \mathbb{N}\}$ and the log-volatility process $\{x_t, t \in \mathbb{N}\}$, with values in the measurable *latent product space* $(F \times G, \mathcal{F} \otimes \mathcal{G})$, where $F \subset \mathbb{N}_0$ and $G \subset \mathbb{R}$.

Let us introduce $x_{s:t} \stackrel{\Delta}{=} (x_s, \dots, x_t)$, with $s \leq t$. Our Bayesian model is specified through the time-conditional distributions

$$p(x_t|x_{t-1}, \theta) \sim \mathcal{N}\left(\alpha_{s_t} + \phi x_{t-1}, \sigma^2\right) \tag{9}$$

$$p(y_t|x_t, \theta) \sim \mathcal{N}\left(\mu + \rho y_{t-1}, e^{x_t}\right) \tag{10}$$

where $\mathcal{N}(\cdot,\cdot)$ denotes the normal distribution, $\theta = (\alpha_1, \alpha_2, p_{11}, p_{22}, \phi, \sigma^2, \mu, \rho)$ is the parameter vector and the parameter α_{s_t} is indexed to the current volatility regime. The latent process $\{s_t, t \in \mathbb{N}\}$, which drives the stochastic log-volatility, is a time-homogeneous Markov chain with transition probabilities

$$p(s_t|s_{0:t-1}, \theta) \sim \mathbb{P}\left(s_t = k|s_{t-1} = l\right) = p_{lk} \quad \text{with} \quad k, l \in \{1, \dots, K\} \tag{11}$$

where K is the number of states (or regimes).

Our approach posits that two volatility regimes characterize the behaviour of stock market and business cycle variables: *low volatility* ($s_t = 1$) and *high volatility* ($s_t = 2$). This is a reasonable assumption, which has been already discussed in many empirical studies on the business cycle (see for example Watson (1994)) and financial variables (see for example So, Lam and Li (1998)). Despite its simplicity, the model and the proposed inference framework easily extend to account for a time-varying number of hidden states, as in Chopin (2001) and Chopin and Pelgrin (2004).

When making Bayesian inference, the parameters are random variables[5] and the initial values of the latent processes are random quantities as well. We assume the following diffuse priors

$$
\begin{aligned}
(\alpha_1, \alpha_2, p_{00}, p_{11}) &\sim \mathcal{N}(0.5, 100)\mathbb{I}_{(-\infty,\alpha_2]}(\alpha_1)\mathcal{N}(-0.5, 100)\mathcal{B}e(100, 2)\mathcal{B}e(100, 2) \quad (12) \\
(\sigma^2, \phi) &\sim \mathcal{IG}(10, 0.01)\mathcal{N}(0.5, 100)\mathbb{I}_{(0,1)}(|\phi|) \quad (13) \\
((s_0 - 1), x_0|\theta) &\sim \mathcal{B}(0.5)\mathcal{N}(\alpha_{s_0}, \phi) \quad (14)
\end{aligned}
$$

where $\mathcal{B}$ denotes the Bernoulli distribution, $\mathcal{IG}$ the Inverse Gamma and $\mathcal{B}e$ the Beta distribution.

This simple univariate *Markov-Switching Stochastic-Volatility* (MSSV) model nevertheless picks up important features of financial and macroeconomic time series, such as heteroscedasticity, volatility clustering and switches, and heavy-tails unconditional distributions. The estimation of the MSSV model poses also some challenging statistical problems, as described in the subsequent sections (see also Lopes and Marinho (2001) and Casarin (2004) for further details).

3 Particle Filters

The estimation of the latent-variable model presented in Section 2 configures a problem of nonlinear filtering (see Jazwinski (1970) and Arulampalam et al. (2001)) with unknown parameters. Parameters can be estimated jointly or separately w.r.t. the state filtering, as expressed by

$$
\begin{aligned}
p(\mathbf{x}_{t+1}, \theta|\mathbf{y}_{1:t+1}) &= \frac{p(\mathbf{y}_{t+1}|\mathbf{x}_{t+1}, \theta)p(\mathbf{x}_{t+1}|\mathbf{x}_t, \theta)p(\mathbf{x}_{0:t}|\mathbf{y}_{1:t})}{p(\mathbf{y}_{t+1}|\mathbf{y}_{1:t})}p(\theta|\mathbf{y}_{1:t}) \quad (15) \\
&= \frac{p(\mathbf{y}_{t+1}|\mathbf{x}_{t+1}, \theta)p(\mathbf{x}_{t+1}|\mathbf{x}_t, \theta)}{p(\mathbf{y}_{t+1}|\mathbf{y}_{1:t})}p(\mathbf{x}_{0:t}, \theta|\mathbf{y}_{1:t}). \quad (16)
\end{aligned}
$$

Equation (15) shows that the filtering problem can be treated as conditional on the parameters. It is possible for example to use the Kalman Filter or the HMM filtering algorithms to estimate the states and the particle filter to estimate the parameters (see Chopin (2001)). In our MSSV model neither the Kalman nor the HMM filters can be used, thus we employ a full Bayesian estimation approach. Unique among the existing approaches, the Bayesian methodology is general enough to allow the treatment of nonlinear

[5] Parameter estimation will be discussed in Section 3.1.

problems, but it also accounts for prior information on the parameters and permits the use of simulation methods in the inference process.

3.1 Joint Estimation of Parameters and Latent Variables

We include the parameters into the state vector of the system and then, following Liu and West (2001), we apply a *Regularised-Auxiliary Particle Filter* (R-APF) for estimating sequentially the hidden states and the unknown parameters of the Bayesian latent-variable model. Let $\mathcal{Y} \subset \mathbb{R}^{n_y}$, $\mathcal{X} \subset \mathbb{R}^{n_x}$ and $\Theta \subset \mathbb{R}^{n_\theta}$ be the observations, state and parameter spaces respectively. A probabilistic representation of a stochastic dynamic model (see Harrison and West (1997)) is

$$\mathbf{x}_t \sim p(\mathbf{x}_t|\mathbf{x}_{t-1},\theta) \qquad \text{(transition density)} \tag{17}$$

$$\mathbf{y}_t \sim p(\mathbf{y}_t|\mathbf{x}_{0:t},\theta) \qquad \text{(measurement density)} \tag{18}$$

$$(\mathbf{x}_0,\theta) \sim p(\mathbf{x}_0,\theta) \qquad \text{(prior density)} \tag{19}$$

with $t = 1, \ldots, T$, $\mathbf{x}_t \in \mathcal{X}$, $\mathbf{y}_t \in \mathcal{Y}$ and $\theta \in \Theta$. Boldface indicates that state and observation variables could possibly be vectors. In our MSSV model $\mathbf{x}_t = (x_t, s_t)'$ and $\mathbf{y}_t = y_t$.

Let us denote with $\delta_x(y)$ the Dirac's mass centered in x. The model given in Eq. (17)-(19) can be restated assuming the following dynamics for the parameter vector: $\theta_t \sim \delta_{\theta_{t-1}}(\theta_t)$, with initial condition $\theta_0 = \theta$ a.s.. Let us include the parameter θ_t into the hidden states and denote with $\mathbf{z}_t = (\mathbf{x}_t', \theta_t')'$ the augmented state vector and with $\mathcal{Z} = \mathcal{X}\times\Theta$ the corresponding augmented state space. The filtering, one-step-ahead prediction and smoothing densities associated to the model given in Eq. (17)-(19) are

$$p(\mathbf{z}_{t+1}|\mathbf{y}_{1:t}) = \int_{\mathcal{Z}} p(\mathbf{x}_{t+1}|\mathbf{x}_t,\theta_{t+1})\delta_{\theta_t}(\theta_{t+1})p(\mathbf{z}_t|\mathbf{y}_{1:t})d\mathbf{z}_t \tag{20}$$

$$p(\mathbf{y}_{t+1}|\mathbf{y}_{1:t}) = \int_{\mathcal{Z}} p(\mathbf{y}_{t+1}|\mathbf{x}_{t+1},\theta_{t+1})p(\mathbf{z}_t|\mathbf{y}_{1:t})d\mathbf{z}_{t+1} \tag{21}$$

$$p(\mathbf{z}_{t+1}|\mathbf{y}_{1:t+1}) \propto p(\mathbf{y}_{t+1}|\mathbf{x}_{t+1},\theta_{t+1})p(\mathbf{x}_{t+1}|\mathbf{x}_t,\theta_{t+1})\delta_{\theta_t}(\theta_{t+1}) \tag{22}$$

$$p(\mathbf{z}_s|\mathbf{y}_{1:t}) = p(\mathbf{z}_s|\mathbf{y}_{1:s})\int_{\mathcal{Z}} \frac{(\mathbf{z}_{s+1}|\mathbf{z}_s)p(\mathbf{z}_{s+1}|\mathbf{y}_{1:t})}{p(\mathbf{x}_{s+1}|\mathbf{y}_{1:t})}d\mathbf{z}_{s+1} \tag{23}$$

with $s < t$.

The analytical solution of the discrete-time filtering problem exists in a few cases: the linear and Gaussian dynamic model and the linear model with a countable finite-dimension latent space, where Kalman-Bucy (see Kalman (1960)) and Hamilton-Kitagawa (see Hamilton (1989)) filters respectively apply. Except for these cases simulation methods are usually applied.

In the simulation-based framework, many authors address the problem of the extraction of latent factors in the business cycle through the use of *Monte Carlo Markov Chain* (MCMC) techniques[6]. In Bayesian analysis MCMC techniques are considered as the most suitable tool for solving integration problems (see Casella and Robert (2004)), which arise in parameters and latent-variables estimation, and in hypothesis testing. Nevertheless, many applications reveal that MCMC methods have some drawbacks too. For example, the choice of the scale parameter of the random walk process in a Metropolis-Hastings algorithm can severely affect the convergence to the posterior distribution. Moreover, the generated Markov chain can get trapped in a local mode of the posterior distribution.

We follow a route alternative to MCMC, which relies upon sequential Monte Carlo (see Gordon, Salmond and Smith (1993), and Berzuini et al. (1997)). In particular we apply *Particle Filters* (PF)[7] also known as *bootstrap filters*, *interacting particle filters*, *condensation algorithms*, *Monte Carlo filters*.

Particles filters make use of weighted Monte Carlo samples to approximate the hidden-state posterior distribution. This distribution is then updated over time. The sequential nature of the particle filters makes them suitable for real-time applications, to deal with nonlinear models and to detect sudden changes in both the hidden states and parameters. As our MSSV model is concerned, the latter feature is particularly helpful.

As pointed out by Storvik (2002), the inclusion of the parameters into the state vector produces a negative effect on the estimation of the state and parameter joint posterior distribution, which degenerates into a Dirac mass. Different solutions to this kind of degeneracy problem have been proposed in the literature[8]. We address the degeneracy problem by employing a modified version of the *Regularised Auxiliary Particle Filter* (R-APF) due to Liu and West (2001) and apply a kernel-based regularisation technique similar to the one used by Musso, Oudjane and LeGland (2001).

Let $\{\mathbf{z}_0^i, w_0^i\}_{i=1}^N$ be a sequence, also called *particle set*, of N weighted Monte Carlo samples, simulated by importance sampling from the prior distribution of the dynamic model. Assume that a particle set $\{\mathbf{z}_t^i, w_t^i\}_{i=1}^N$, from the augmented-state posterior,

[6]See for example Kim and Nelson (1999), who provide an introduction to the use of MCMC methods in the estimation of Markov-switching business cycle models

[7]See Doucet, Freitas and Gordon (2001) for an updated review of the sequential Monte Carlo methods.

[8]See for example Kitagawa (1998), Storvik (2002) and Polson, Stroud and Müller (2002, 2003).

$p(\mathbf{z}_t|\mathbf{y}_{1:t})$, is available at time t. This distribution represents the empirical prior for the states at time $t+1$, and is written as

$$p_N(\mathbf{z}_t|\mathbf{y}_{1:t}) = \sum_{i=1}^{N} w_t^i \delta_{\{\mathbf{z}_t^i\}}(\mathbf{z}_t), \tag{24}$$

where $\delta_{\{y\}}(x)$ denotes the Dirac point-mass centered in y. Using the empirical prior, the one-step-ahead prediction density in Eq. (20) and the filtering density in Eq. (22) can be approximated by means of the *empirical prediction density* and the *empirical filtering density*, given by

$$p_N(\mathbf{z}_{t+1}|\mathbf{y}_{1:t}) = \int_{\mathcal{Z}} p(\mathbf{z}_{t+1}|\mathbf{z}_t) p_N(\mathbf{z}_t|\mathbf{y}_{1:t}) d\mathbf{z}_t = \sum_{i=1}^{N} w_t^i p(\mathbf{z}_{t+1}|\mathbf{z}_t^i) \tag{25}$$

$$p_N(\mathbf{z}_{t+1}|\mathbf{y}_{1:t+1}) \propto p(\mathbf{y}_{t+1}|\mathbf{z}_{t+1}) p_N(\mathbf{z}_{t+1}|\mathbf{y}_{1:t}) = \sum_{i=1}^{N} p(\mathbf{y}_{t+1}|\mathbf{z}_{t+1}) p(\mathbf{z}_{t+1}|\mathbf{z}_t^i) w_t^i \tag{26}$$

The empirical filtering density in Eq. (26) is then regularised through a multiple-bandwidth Gaussian-kernel estimator

$$\begin{aligned} p_N(\mathbf{z}_{t+1}|\mathbf{y}_{1:t+1}) &\propto \sum_{i=1}^{N} p(\mathbf{y}_{t+1}|\mathbf{x}_{t+1}, \theta_{t+1}) p(\mathbf{x}_{t+1}|\mathbf{x}_t^i, \theta_t^i) \delta_{\theta_t^i}(\theta_{t+1}) w_t^i \\ &\approx \frac{1}{N} \sum_{i=1}^{N} w_t^i p(\mathbf{y}_{t+1}|\mathbf{x}_{t+1}, \theta_t^i) p(\mathbf{x}_{t+1}|\mathbf{x}_t^i, \theta_t^i) \mathcal{N}_{n_\theta}(\theta_{t+1}|m_t^i, B_t^i V_t) \end{aligned} \tag{27}$$

where B_t^i is a scale matrix, which is positive-definite for all i and t, $m_t^i = a\theta_{t-1} + (1-a)\bar{\theta}$ with $a \in (0,1)$, $V_t = \sum_{i=1}^{N} (\theta_t^i - \bar{\theta}_t)(\theta_t^i - \bar{\theta}_t)' w_t^i$ and $\bar{\theta}_t = \sum_{i=1}^{N} \theta_t^i w_t^i$ (see Liu and West (2001)).

Due to the *regularisation step*, the original transition density, $\delta_{\{\theta_t^i\}}(\theta_{t+1})$, has been replaced by a Gaussian transition $\mathcal{N}_{n_\theta}(\theta_{t+1}|m_t^i, B_t^i V_t)$ with particle-specifics: scale matrix, B_t^i and location, m_t^i. The choice of B_i^t will be discussed in Section 3.2.

After the kernel reconstruction of the posterior density, a new set of particles $\{\mathbf{z}_{t+1}^i, w_{t+1}^i\}_{i=1}^N$ can be generated through a simulation procedure. This simulation step was introduced in the earlier particle filters, (e.g. the *Sequential Importance Sampling* (SIS)), to avoid the degeneracy of the particle weights (see Gordon, Salmond and Smith (1993), Gilks and Berzuini (2001), Doucet (2000) and Musso, Oudjane and LeGland (2001)).

In this work we apply the particle selection scheme due to Pitt and Shephard (1999). Note that the empirical filtering density given in Eq. (27) is a mixture of distributions. It can be demarginalized as follows

$$p_N(\mathbf{x}_{t+1}, \theta_{t+1}, i) = w_t^i p(\mathbf{y}_{t+1}|\mathbf{x}_{t+1}, \theta_t^i) p(\mathbf{x}_{t+1}|\mathbf{x}_t^i, \theta_t^i) \mathcal{N}_{n_\theta}(\theta_{t+1}|m_t^i, B_t^i V_t) \tag{28}$$

The *selection step* is performed by sampling the mixture index i (the auxiliary variable), and the *mutation step* by simulating conditionally on the sampled index, the states $\mathbf{x}_{t+1}$ and the parameters θ_{t+1}. A sample from the joint density (27) is obtained by importance sampling, with importance density

$$q(\mathbf{x}_{t+1}, \theta_{t+1}, i|\mathbf{y}_{1:t+1}) = p(\mathbf{x}_{t+1}|\mathbf{x}_t^i, \theta_t^i) \mathcal{N}_{n_\theta}(\theta_{t+1}|m_t^i, B_t^i V_t) q(i|\mathbf{y}_{1:t+1}) \tag{29}$$

where the importance density for the random index is $q(i|\mathbf{y}_{1:t+1}) = p(\mathbf{y}_{t+1}|\mu_{t+1}^i, m_t^i) w_t^i$. From previous assumptions we weights update as follows

$$\tilde{w}_{t+1}^j \propto \frac{p(\mathbf{y}_{t+1}|\mathbf{x}_{t+1}^j, \theta_t^j) p(\mathbf{x}_{t+1}^j|\mathbf{x}_t^{i^j}, \theta_t^j) \mathcal{N}_{n_\theta}(\theta_{t+1}^j|m_t^{i^j}, B_t^{i^j} V_t) w_t^{i^j}}{p(\mathbf{y}_{t+1}|\mu_{t+1}^{i^j}, m_t^{i^j}) p(\mathbf{x}_{t+1}^j|\mathbf{x}_t^{i^j}, \theta_t^{i^j}) \mathcal{N}_{n_\theta}(\theta_{t+1}^j|m_t^{i^j}, B_t^{i^j} V_t) w_t^{i^j}} \tag{30}$$

and $w_{t+1}^i = \tilde{w}_{t+1}^i \left(\sum_{j=1}^N \tilde{w}_{t+1}^j\right)^{-1}$ with $j = 1, \ldots, N$.

3.2 Multiple-Bandwidth Regularisation and Convergence Issues

The R-APF avoids the degeneracy problem by introducing variability in the particle set via the regularisation of the empirical densities. We extend the regularisation technique due to Liu and West (2001) and consider time-varying local-bandwidth. We will use the same scale factor in each direction of the parameter space: $B_t = b_t^2 I_{n_\theta}$, with I_{n_θ} denoting the $(n_\theta \times n_\theta)$ identity matrix and set $a_t = (3\lambda_t - 1)(2\lambda_t)^{-1}$ and $b_t^2 = 1 - a_t^2$ (see Liu and West (2001)). However our regularisation technique easily extends to the case of component-specific scale, that is $(B_t)^{1/2} = \text{diag}\{b_{1t}, \ldots, b_{n_\theta t}\}$.

Allowing the bandwidth to vary over the particle set, amounts performing a multiple-bandwidth (or sample-point) kernel estimation of $p(\theta_{t+1}|y_{1:t+1})$

$$\hat{p}_N(\theta_{t+1}|y_{1:t+1}) = \frac{1}{N} \sum_{i=1}^N \frac{1}{(b_t^i)^{n_\theta}} K((b_t^i)^{-n_\theta}(\theta_{t+1} - \theta_t^i)) \tag{31}$$

Sample-point estimators have locally adapted bandwidths and show to be more promising, than fixed bandwidth estimators. Fixed bandwidths deal badly with local scale variations of the empirical densities and produce undersmoothing of the tails and oversmoothing of the peaks.

Although more elaborate self-tuning iterative procedures could be proposed, in the following we use a bandwidth randomly distributed on the interval $[(b_N^*)^\alpha, (b_N^*)^\beta]$, with $b_N^* = (4/((2+n_\theta)N))^{\frac{1}{n_\theta+4}}$, $\alpha < 1$ and $\beta > 1$ and focus instead on the convergence properties of our multi-scales filter.

The convergence of standard PF has been studied in Crisan (2001) and Crisan and Doucet (2000, 2002)[9]. Let us focus on our multi-scale particle filter without selection step. Convergence results can be extended to account for the selection step (see Rossi (2004)). The L_2-convergence of the empirical filtering density w.r.t. the number of particles N, is given in the following.

Theorem 3.1. *(Quadratic-mean convergence)*
Let K be a Gaussian kernel on $\mathbb{R}^{n_\theta}$, $\{b_t^i\}_{i=1}^N$ a sequence of positive scale factors (bandwidths) with value in the finite interval $[b_{N,\min}, b_{N,\max}]$ and $\{\theta_t^i\}_{i=1}^N$ a sequence of $\mathbb{R}^{n_\theta}$-valued i.i.d. samples simulated from $\hat{p}(\theta_t|y_{1:t})$. If

$$\lim_{N\to\infty} b_{N,\max} = 0, \quad \textit{and} \quad \lim_{N\to\infty} (b_{N,\min})^d N = \infty \tag{32}$$

then the functional estimator, $\hat{p}_N(\theta_{t+1}|y_{1:t+1})$ defined in (31), converges in quadratic mean to the true density $\hat{p}(\theta|y_{1:t+1})$

$$\hat{p}_N(\theta_{t+1}|y_{1:t+1}) \xrightarrow[N\to\infty]{L^2} p(\theta_{t+1}|y_{1:t+1}). \tag{33}$$

Proof. See Appendix C. □

In Appendix C we also prove the a.c. and a.s. convergence of the kernel estimator. The two imply the L_1-convergence that is given in the following.

Theorem 3.2. *(L_1-convergence)*
Let K be a Gaussian kernel on $\mathbb{R}^{n_\theta}$, $p(\theta_t|y_{1:t}) \in L^1(\mathbb{R}^{n_\theta})$ a density, $\{b_t^i\}_{i=1}^N$ a sequence of positive scale factors (bandwidths) with value in the interval $[b_{N,\min}, b_{N,\max}]$ and $\{\theta_t^i\}_{i=1}^N$ a sequence of $\mathbb{R}^{n_\theta}$-valued i.i.d. samples simulated from $p(\theta_t|y_{1:t})$. If

$$\lim_{N\to\infty} b_{N,\max} = 0\,, \quad \lim_{N\to\infty} \frac{(b_{N,\min})^d N}{\log N} = \infty \tag{34}$$

[9]For a theoretical analysis of the particle methods see also Del Moral (2004) and Bartoli and Del Moral (2001). The convergence of the regularised particle filter of Liu and West(2001) is given in Stavropoulos and Titterington (2001).

then the sample-point estimators, $\hat{p}_N(\theta_{t+1}|y_{1:t+1})$ *defined in (31), converges in* L_1 *to the true density*

$$\hat{p}_N(\theta_{t+1}|y_{1:t+1}) \underset{N\to\infty}{\overset{L^1}{\longrightarrow}} p(\theta_{t+1}|y_{1:t+1}). \tag{35}$$

Proof. See Appendix C. □

4 Estimation Results

We apply our particle filter to the MSSV model given in Eq. (11), (9) and (10). In order to verify the efficiency of the algorithm in the estimation of parameters and hidden variables, and also to detect possible degeneracy problems, we first test the regularised APF on synthetic data. The sequential Monte Carlo filter is subsequently applied to actual US stock market and business cycle data.

4.1 Simulated Data

In the simulations we refer to the MSSV model given in Section 2 and apply our R-APF with multiple bandwidth. The resulting filter, given in Appendix A, allows us to sequentially estimate the parameter and to filter and predict the hidden state. We use a set of $N = 2,000$ particles to estimate the filtering and prediction densities[10].

Figure 2 plots on-line estimates (i.e. the average over the particle set) of the latent factors $\{x_t\}$ and $\{s_t\}$ obtained form a typical run of the R-APF algorithm on the synthetic dataset plotted in the first panel of the same figure. In order to detect the absence of degeneracy in the output of the APF algorithm we evaluate at each time step the *survival rate*

$$SR_t = 1 - \frac{1}{N}\sum_{i=1}^{N} \mathbb{I}_{\{J_{i,t}=\emptyset\}} \tag{36}$$

where $J_{i,t} = \{j \in \{1,\dots,N\}|i_t^j = i\}$ is the set of all random index values, which are selecting, at time t, the i-th particle. If at time t the particle k does not survive to the selection step, then the set $J_{k,t}$ becomes empty. Particles sets degenerate when persistently exhibiting a high number of dead particles from a generation to the next. In Figure 2 the survival rate does not decline, thus we conclude that the APF does not degenerate in our simulation study. Figure 3 shows sequential estimates of parameters α_1, α_2, p_{11}, p_{22}, ϕ

[10] All computations have been carried out on a Pentium IV 2.4 Ghz, and the APF algorithm has been implemented in GAUSS 7.0.

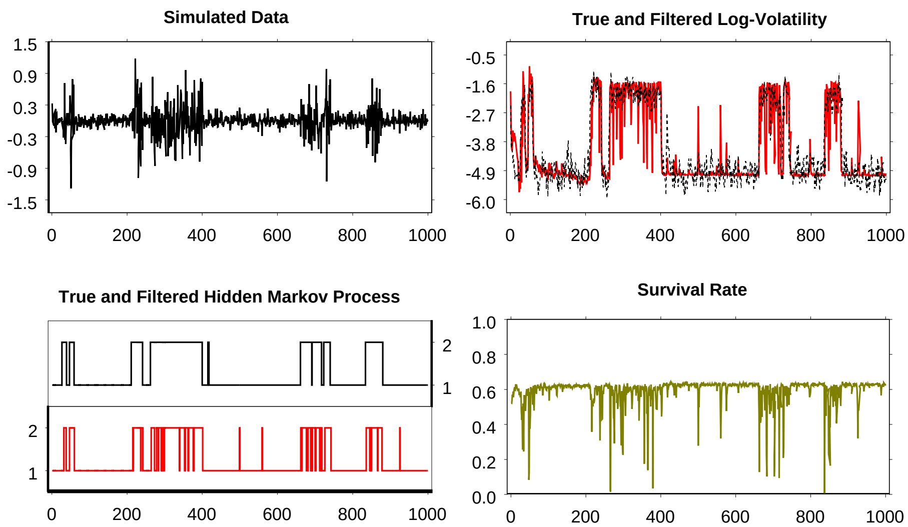

Figure 2: *The simulated (black line) and sequentially filtered (grey line) latent factors and the survival rate of the particle set* ($N = 2,000$)*, over a sample of* $T = 1,000$ *observations.*

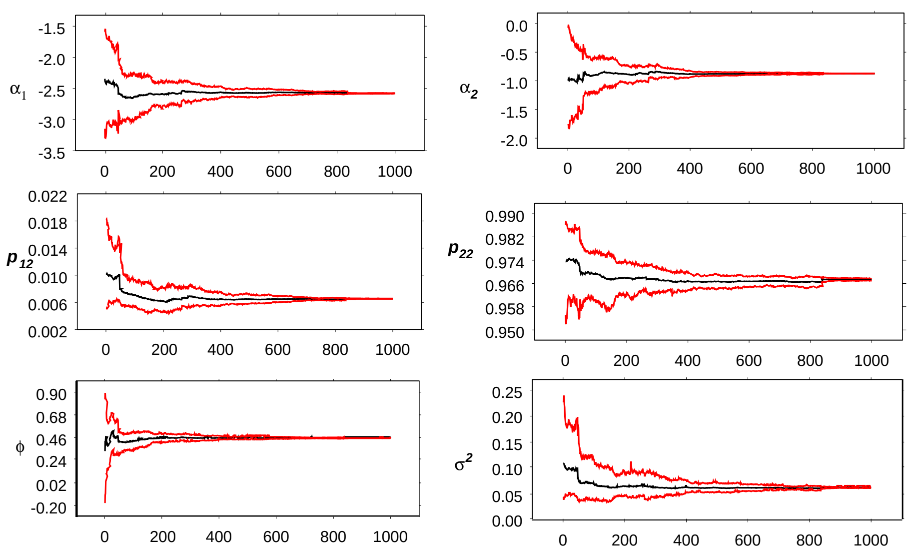

Figure 3: *On-line parameter estimates. Graphs show at each date the empirical mean and the quantiles at 0.025 and 0.975 for each parameter.*

and σ^2. We remind that these are the parameters of the conditional densities of the latent log-volatility $\{x_t\}$, the switching process $\{s_t\}$ and the observation $\{y_t\}$[11].

4.2 Volatility Patterns in Macroeconomic and Stock Market Data

The aim of our exercise is to detect similarities in the volatility profiles of macroeconomic and stock market variables. Our attention therefore focuses on the variability of stock market returns over the past three decades, during which the volatility of several macroeconomic variables has clearly declined.

We apply the sequential Monte Carlo filter to two groups of actual US data. The first group (the logarithmic return on the S&P500 stock market index and its corresponding dividend yield and price-to-earnings ratio) tracks the aggregate behaviour of US stock market. The second group of variables (the log-difference of industrial production, non-residential investment expenditure, real personal consumption expenditure per capita, and the level of the output gap) accounts for the evolution of the business cycle[12].

We start by studying the volatility patterns displayed by stock market returns. We therefore first estimate the MSSV model on the log-difference of the Standard & Poor's 500 index (S&P), and of the dividend yield (DP) and price-to-earnings ratio (PE) based on the market index.

The data are quarterly observations collected on the sample period 1966Q2-2003Q3 (see the Data Appendix for details on data sources and definitions). The filtered volatility regimes and log-volatility processes are shown in Fig. 4, 5 and 6. The top left-hand panel of each figure plots the time series accompanied by vertical shaded areas representing the contraction phases of US GDP as detected by Business Cycle Dating Committee of the NBER. The contraction starts at the peak of the cycle and ends at the through. Our sample covers the following GDP contraction periods (the corresponding quarter is in parenthesis) that are of interest in our study: December 1969 (IV)-November 1970 (IV), November 1973 (IV)-March 1975 (I), January 1980 (I)-July 1980 (III), July 1981 (III)-November 1982 (IV), July 1990 (III)-March 1991 (I) and March 2001 (I)-November 2001 (IV).

From the charts in Figures 4 to 6 one can easily notice that S&P, PE and DP all follow remarkably similar volatility patterns. The estimated filtered log-volatility and filtered Markov processes show that the regimes of high and low volatility alternate along very

[11]See Section 2.2 above.

[12]Gordon (2005) finds that government spending and net exports too might have played a role (albeit a relatively small one) in explaining the decline of output volatility.

Table 1: *Parameter estimates (posterior-mean estimates) with the 0.975 and 0.025 quantiles, for quarterly log-differences of the observations (except for the Output Gap) on the sample period 1966Q3-2003Q3. Give* h_t *as the log-volatility, our model* $h_t = \alpha_1 \mathbb{I}_{\{1\}}(s_t) + \alpha_2(1 - \mathbb{I}_{\{1\}}(s_t)) + \phi h_{t-1} + \sigma\varepsilon_t$ *with* ε_t *i.i.d. standard normals.*

	S&P Index			Price-to-Earnings Ratio			Dividend Yield		
θ	$\hat{\theta}$	$q_{0.025}$	$q_{0.975}$	$\hat{\theta}$	$q_{0.025}$	$q_{0.975}$	$\hat{\theta}$	$q_{0.025}$	$q_{0.975}$
α_1	-3.264	-3.378	-3.115	-3.183	-3.303	-3.057	-5.617	-5.953	-5.234
α_2	-2.747	-2.866	-2.634	-2.862	-2.952	-2.781	-5.117	-5.472	-4.793
p_{12}	0.014	0.012	0.017	0.018	0.017	0.020	0.011	0.005	0.019
p_{22}	0.990	0.989	0.991	0.974	0.972	0.977	0.989	0.983	0.994
σ^2	0.096	0.086	0.110	0.108	0.097	0.119	0.094	0.059	0.136
ϕ	0.465	0.421	0.515	0.468	0.440	0.491	0.457	0.417	0.497

	Industrial Production			Consumption Expenditure		
θ	$\hat{\theta}$	$q_{0.025}$	$q_{0.975}$	$\hat{\theta}$	$q_{0.025}$	$q_{0.975}$
α_1	-4.911	-5.025	-4.788	-0.898	-0.928	-0.871
α_2	-4.228	-4.400	-4.073	-0.307	-0.349	-0.260
p_{12}	0.020	0.017	0.023	0.009	0.008	0.010
p_{22}	0.999	0.988	0.991	0.979	0.975	0.983
σ^2	0.369	0.318	0.421	0.204	0.169	0.244
ϕ	0.498	0.471	0.525	0.531	0.427	0.627

	Non-Resid. Investment			Output Gap		
θ	$\hat{\theta}$	$q_{0.025}$	$q_{0.975}$	$\hat{\theta}$	$q_{0.025}$	$q_{0.975}$
α_1	0.626	0.535	0.736	-0.803	-0.999	-0.606
α_2	0.783	0.624	0.959	0.107	-0.080	0.304
p_{12}	0.019	0.154	0.023	0.020	0.016	0.024
p_{22}	0.980	0.976	0.983	0.989	0.987	0.991
σ^2	0.124	0.101	0.143	0.201	0.166	0.241
ϕ	0.579	0.469	0.687	0.495	0.310	0.677

similar sequences across the three variables. First, S&P, PE and (to a slightly smaller extent) DP all see periods of low volatility lasting much longer than those of high volatility. The last twenty years in particular seem to have witnessed only few and short reversions to high-volatility regimes. Second, the vertical bars point to volatility rising almost exclusively around periods of GDP contraction. The 1990-1991 and 2001 GDP recessions triggered a shift to high volatility for S&P, PE and DP, although for the latter the high-volatility regime predates the start of GDP contraction, and for S&P our algorithm detects a similar regime for the latter part of 1990s. Third, since the mid-eighties volatility switches do appear to take place less frequently. The latter two findings are *prima facie* support for some relationship between lower stock market volatility and decreased GDP volatility.

For all the stock market series the survival rate is well-behaved, therefore pointing to the absence of degeneracy phenomena in the posterior distribution. Figures 12 to 18 in Appendix D show the sequential estimates of the parameter α_1, α_2, p_{12}, p_{22}, ϕ and σ^2 with the 0.975 and 0.025 quantiles. The values of the parameter estimates at the last iteration of the filter are in Table 1.

Turning to business cycle variables, we study the volatility of the following series: Industrial Production (IP), Personal Consumption Expenditure per capita (PCE), Non-Residential Investments (NRI), all in their log-differences, and the level of output gap (YGAP)[13]. The filtered stochastic log-volatility and volatility-regime processes are shown in Figures 7, 8, 9 and 10. At the outset, one can notice some affinity in the behaviour of volatility in the cases of industrial production and non-residential investment. The volatilities of IP and NRI appear to have switched to a persistent low-volatility regime in the second half of the 1980s. IP reverts to high volatility only briefly after the 1990-91 contraction, and NRI does the same around the 2001 recession and the subsequent recovery. The volatility of PCE too enters a low-level regime, but only since 1993. The fact that in the preceding part of the sample high volatility dominates more than for the other macroeconomic series makes the switch to low volatility even more striking.

Finally, let us turn to the output gap. We remind that we study its level. Even a bird's eye view of the series (top left-hand chart of Figure 10) confirms that since the mid-eighties expansions have had much longer durations, while slowdowns have both occurred less often and been milder than in the earlier part of the sample. Indeed, the estimated MSSV model shows that YGAP behaved according to a persistent high-volatility regime essentially until

[13]We do not include here results for residential fixed investment expenditure.

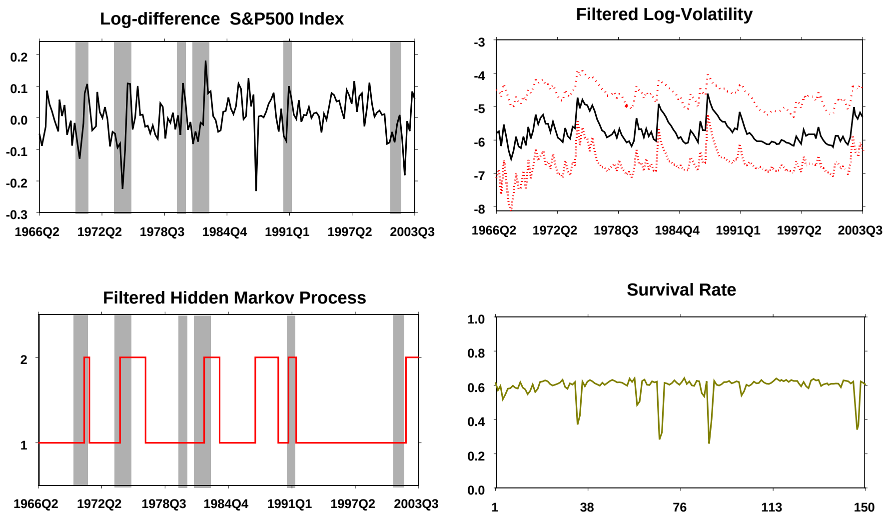

Figure 4: *Log-differences of the S&P500 price index; sequentially filtered log-volatility with 0.025 and 0.975 quantiles (dotted lines); filtered volatility regimes and survival rate of the particle set, over the sample of* $T = 150$ *observations.*

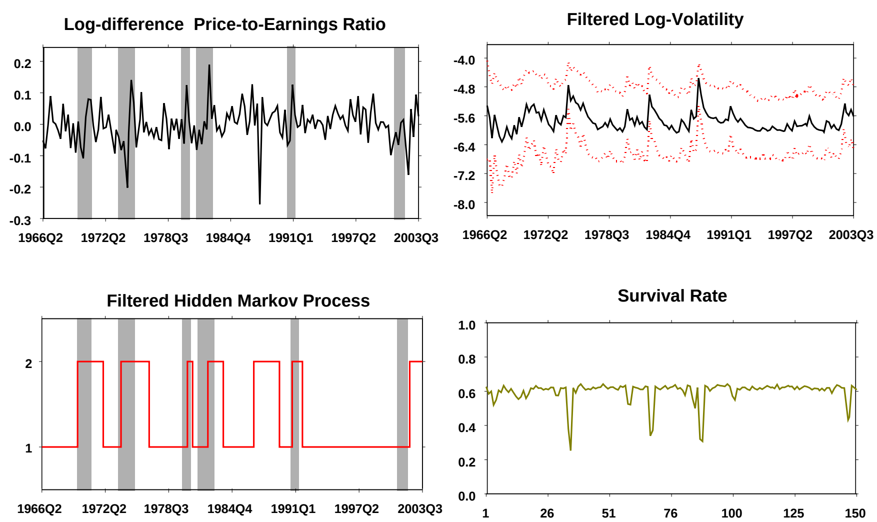

Figure 5: *Log-differences of the Price-to-Earnings series; sequentially filtered log-volatility with 0.025 and 0.975 quantiles (dotted lines); filtered volatility regime and survival rate of the particle set, over the sample of* $T = 150$ *observations.*

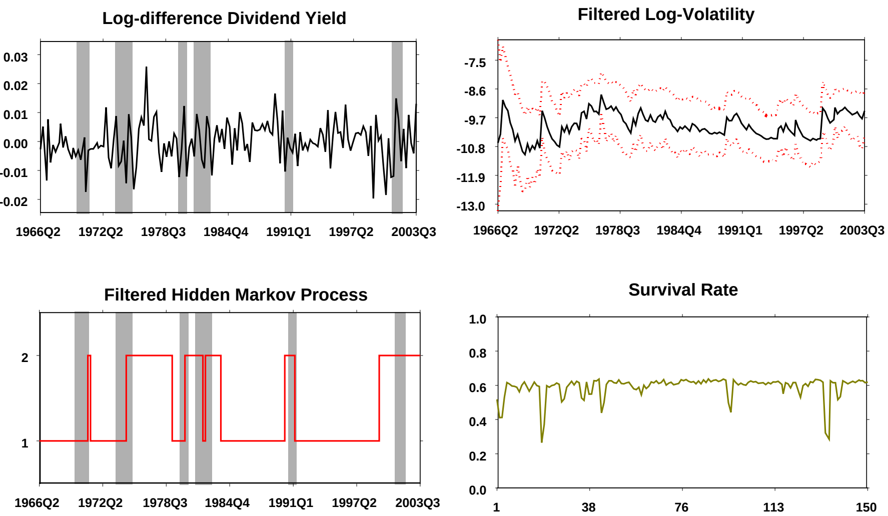

Figure 6: *Log-differences of the Dividend Yield series; sequentially filtered log-volatility with 0.025 and 0.975 quantiles (dotted lines); filtered volatility regime and survival rate of the particle set, over the sample of* $T = 150$ *observations.*

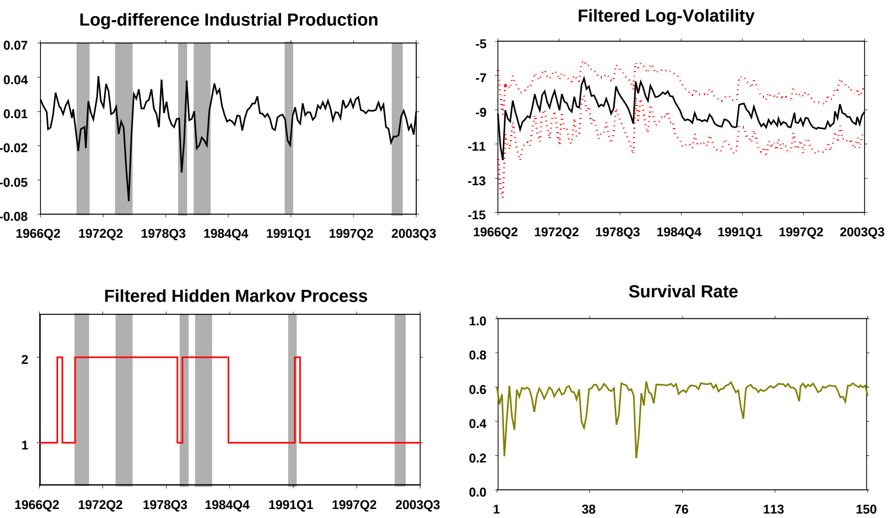

Figure 7: *Log-differences of the Industrial Production index; sequentially filtered log-volatility with 0.025 and 0.975 quantiles (dotted lines); filtered volatility regime and survival rate of the particle set, over the sample of* $T = 150$ *observations.*

1984. Afterwards, it switched to a persistent low-volatility regime, with brief spells of high volatility around the 1990-1991 and 2001 contractions. Thus, our evidence says that YGAP volatility has followed a pattern that is somewhat intermediate between those of the other business cycle indicators, but overall closer to that of IP and NRI.

What can we draw as to the detection of similarities in the filtered volatilities of macroeconomic and stock market variables? The economic activity indicators follow a low-volatility regime for most of the second part of the sample. In particular, persistently low volatility characterizes industrial production, non-residential investment and the output gap; personal consumption expenditure switches to low volatility too, but with a lag relative to the other series. Our MSSV estimates detect essentially the same pattern for the stock market indicators: the return on the market index and its price-earnings and dividend-price ratios all switch to low volatility around 1991. This switch appears to be permanent, as reversions to high volatility occurred only after the longest period of low volatility in our sample.

The estimates in Table 1 complement the graphical evidence we have just commented. In detail, the estimated values of α_1 and α_2 provide the mean level of log-volatility of the series under the regime of low and high volatility, respectively. These values confirm that PCE, NRI and YGAP, that is, three out of the four macroeconomic variables, have the highest mean volatility level. As to the values of the transition probabilities, we remind that they represent an indirect measure of the persistence of each volatility regime. We observe that p_{12}, the probability of the variable switching from a low- to a high-volatility regime, is marginally higher for IP and YGAP, and smaller for S&P and DP. On the other hand, p_{22}, the probability of the variable staying in a high-volatility regime, is highest for S&P and YGAP, and lowest for PE and PCE[14]. Finally, $(1-\phi)$ accounts for the speed of reversion of log-volatility to its mean value. We obtain the highest values for the ϕs of NRI and PCE, which therefore tend to revert to their mean volatility levels much more slowly than all other series. Stock market variables, as expected, are the quickest to revert to their mean values of log-volatility.

In our approach all estimated parameters are time-varying. This allows us to gauge the stability of our estimated models: graphs in Appendix D confirm that estimated parameters are overall stable.

Figure 11 makes easier to summarize the volatility profiles of two pairs of variables. The

[14]We remind that $p_{11} = (1 - p_{12})$ is the probability of log-volatility remaining in the low-volatility regime, whereas $p_{21} = (1 - p_{22})$ is the probability of log-volatility migrating from a high to a low-volatility regime.

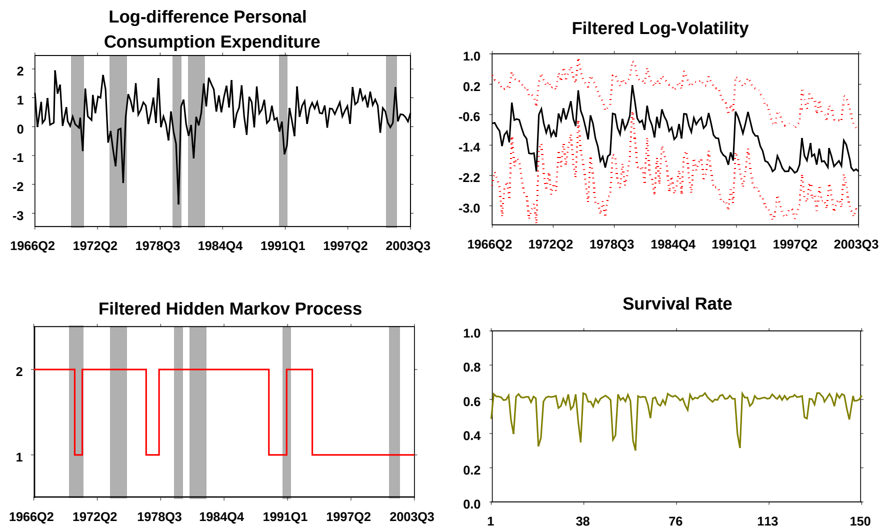

Figure 8: *Log-differences of Personal Consumption Expenditure per capita; sequentially filtered log-volatility with 0.025 and 0.975 quantiles (dotted lines); filtered volatility regime and survival rate of the particle set, over the sample of* $T = 150$ *observations.*

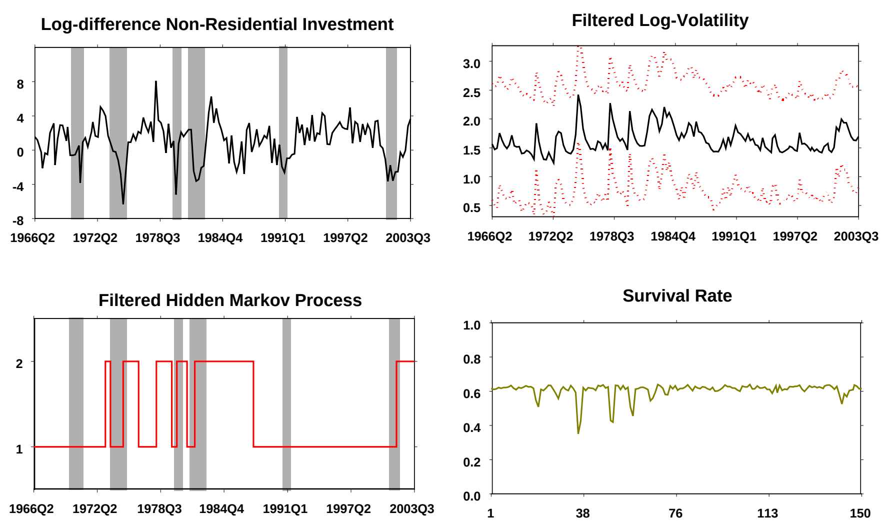

Figure 9: *Log-differences of Non-Residential Investments; sequentially filtered log-volatility with 0.025 and 0.975 quantiles (dotted lines); filtered volatility regime and survival rate of the particle set, over the sample of* $T = 150$ *observations.*

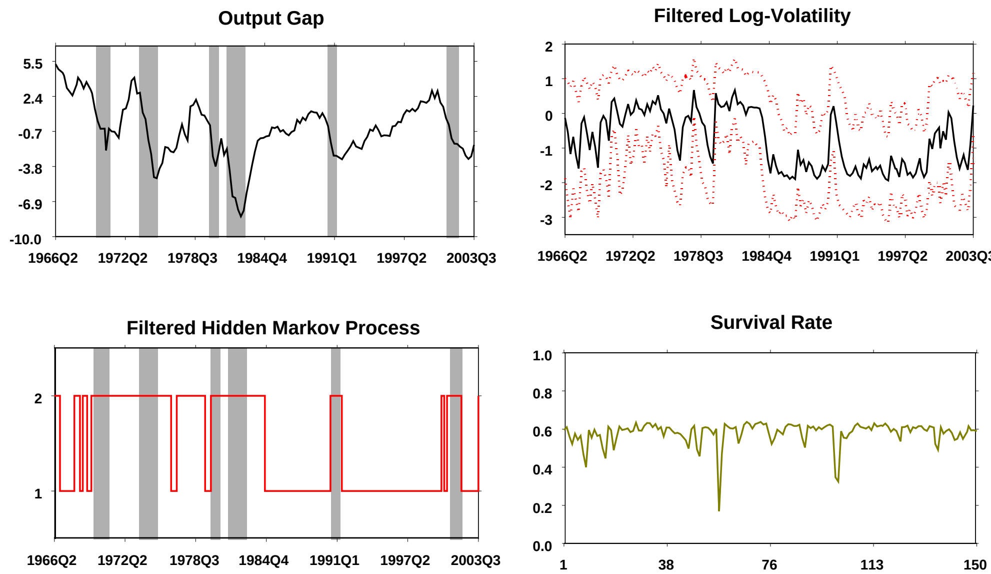

Figure 10: *Output Gap; sequentially filtered log-volatility with 0.025 and 0.975 quantiles (dotted lines); filtered volatility regime and survival rate of the particle set, over the sample of* $T = 150$ *observations.*

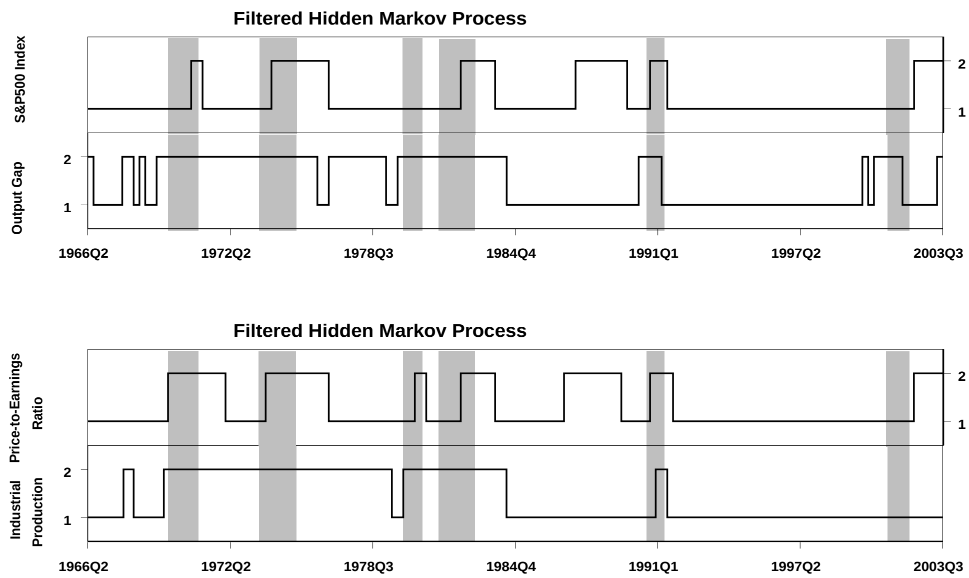

Figure 11: *Filtered volatility regimes for selected variables.*

top panel shows the volatility regimes of S&P and YGAP, the lower panel those of PE and IP. Both charts highlight that all series have switched to low volatility over the last half of the sample. In particular, the shift seems to have occurred in the business cycle variables first, and then in the stock market data. Firmer conclusions on the issue at hand would require the natural development of this work, e.g, the extension to a multivariate setting. Such framework would enable to deal with causality and make more robust inference on the timing of the switches between volatility regimes. In addition, one might want to allow for a finer definition of the consumption data, perhaps distinguishing between durable and nondurable spending.

5 Conclusions

In this work univariate Markov switching stochastic volatility models are estimated to extract the latent volatility regime processes for some macroeconomic and aggregate stock market variables. We make Bayesian inference and propose a particle-filter framework in order to jointly estimate the parameters and the hidden states of the dynamic model. The view we put to test is that the widely observed decline in US business cycle volatility translated into a similar long-term decrease in stock market volatility. According to this view, the relentless rise of stock prices in the late '90s might be due to a reduction of the equity premium, in turn generated by a decrease of broad macroeconomic risk, as represented by the noticed decline in business cycle volatility.

The evidence uncovered by our univariate estimates supports a decline of both macroeconomic and stock market variability, with the switch to low volatility occurring first in the business cycle indicators. Of course, *post hoc ergo propter hoc* (or *cum hoc ergo propter hoc*, for that matter) is a common fallacy and we certainly do not want to incur in it, but the findings of our empirical exercise are not inconsistent with the above view.

Appendix A - Data

Stock market data are from Robert Shiller's website. His data set consists of monthly stock price, dividends, and earnings data, and the consumer price index (to allow conversion to real values), all starting January 1871. Monthly dividend and earnings data are computed from the S&P four-quarter tools for the quarter since 1926, with linear interpolation to monthly figures. Stock price data are monthly averages of daily closing prices.

Personal Consumption Expenditures and Real Gross Domestic Product data are from the U.S. Department of Commerce, Bureau of Economic Analysis, and are quarterly, seasonally adjusted observations. Real Potential Gross Domestic Product is from the U.S. Congress, Congressional Budget Office, and is a quarterly, seasonally adjusted series, in billions of chained 2000 Dollars. The output gap is calculated as the (log) difference between Real Gross Domestic Product and Real Potential Gross Domestic Product. Personal Consumption Expenditure is converted to real values using the Personal Consumption Expenditures Chain-type Price Index, quarterly, seasonally adjusted, index 2000=100, taken from the U.S. Department of Commerce, Bureau of Economic Analysis. Real Private Non-Residential Fixed Investment and Real Residential Fixed Investment are from NIPA Tables, Bureau of Economic Analysis, and are quarterly, seasonally adjusted observations, in billions of chained 2000 dollars.

We converted all monthly data into quarterly averages. In many charts vertical shaded areas represent the contraction phases of the US GDP as detected by Business Cycle Dating Committee of the NBER. The contraction starts at the peak of the cycle and ends at the through. In particular, our sample covers the following GDP contraction periods (the corresponding quarter is in parenthesis) that are of interest in our study: December 1969 (IV)-November 1970 (IV), November 1973 (IV)-March 1975 (I), January 1980 (I)-July 1980 (III), July 1981 (III)-November 1982 (IV), July 1990 (III)-March 1991 (I) and March 2001 (I)-November 2001 (IV).

Appendix B - A Regularised Filtering Algorithm

In the following we give the multiple-bandwidth, kernel-regularised particle filter for the Markov Switching Stochastic Volatility model.

Algorithm 1. ***Kernel-Regularised Particle Filter***

Given an initial set of particles $\{x_t^i, s_t^i, \theta_t^i, \lambda_t^i, w_t^i\}_{i=1}^N$:

1. *Compute* $V_t = \sum_{i=1}^N (\theta_t^i - \bar{\theta}_t)(\theta_t^i - \bar{\theta}_t)' w_t^i$ *and* $\bar{\theta}_t = \sum_{i=1}^N \theta_t^i w_t^i$
2. *For* $i = 1, \ldots, N$ *and with* a *and* b *tuning parameters, calculate:*
   (a) $\tilde{S}_{t+1}^i = \underset{l \in \{1,2\}}{\arg\max} \mathbb{P}(s_{t+1} = l | s_t = s_t^i)$
   (b) $\tilde{X}_{t+1}^i = \alpha_t^i(\tilde{S}_{t+1}^i) + \phi_t^i x_t^i$
   (c) $\tilde{\theta}_t^i = a(\lambda_t^i)\theta_t^i + (1 - a(\lambda_t^i))\bar{\theta}_t$, *where* $\tilde{\theta} = (\tilde{\alpha}_{1,2}, \tilde{\phi}, \tilde{\sigma}, \tilde{\rho}, \tilde{\mu}, \tilde{p}_{11}, \tilde{p}_{22})$
3. *For* $i = 1, \ldots, N$:
   (a) *Simulate* k^i *from* $\sum_{k=1}^N \mathcal{N}(y_{t+1}; \tilde{\mu}_t^k + \tilde{\rho}_t^k y_t, e^{\tilde{X}_{t+1}^k}) w_t^k \delta_{\{k\}}(dk^i)$
   (b) *Simulate* s_{t+1}^i *from* $\mathbb{P}(s_{t+1}^i = l | s_t^{k^i})$ *with* $l \in \{1, 2\}$
   (c) *Simulate* x_{t+1}^i *from* $\mathcal{N}(x_{t+1}; \alpha_t^{k^i}(s_{t+1}^i) + \phi_t^{k^i} x_t^{k^i}, \sigma_t^{k^i})$
   (d) *Simulate* θ_{t+1}^i *from* $\mathcal{N}_{n_\theta}(\theta_{t+1}; \tilde{\theta}_t^{k^i}, b^2(\lambda_t^{k^i})V_t)$
4. *Update:* $\tilde{w}_{t+1}^i \propto \mathcal{N}(y_{t+1}; \tilde{\mu}_t^{k^i} + \tilde{\rho}_t^{k^i} y_t, e^{x_{t+1}^i}) / \mathcal{N}(y_{t+1}; \tilde{\mu}_t^{k^i} + \tilde{\rho}_t^{k^i} y_t, e^{\tilde{X}_{t+1}^{k^i}})$
5. *Normalize:* $w_{t+1}^i = \tilde{w}_{t+1}^i \left(\sum_{i=1}^N \tilde{w}_{t+1}^i\right)^{-1}$.

Appendix C - Convergence Issues

Let K be a Gaussian kernel and f a density function both defined on $\mathbb{R}^d$. Note that $\int K(y)dy = 1$. We denote by $K_h(y) = h^{-d}K(y/h)$ a Gaussian kernel with *bandwidth* h and by $(f * K_h)(x) = \int f(x-y)K_h(y)dy$ the convolution product. Let us define the following simple multiple-bandwidth kernel estimator.

Definition 5.1. *(Multiple-bandwidth estimator)*
Let $X_1, \ldots, X_N$ be a sequence of i.i.d. random vectors with density f. The Multiple-bandwidth estimator of the density f is

$$f_N(x) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{h_i^d}K(h_i^{-1}(x - X_i)). \tag{37}$$

where $\{h_i\}_{i=1}^N$ is a sequence of positive numbers (bandwidths).

Note that the proposed sample-point estimator is positive and integrates to one. Thus it is a density function and the Glick's theorem applies to its. The theorem gives a relation between a.s.-convergence and L_1-convergence for functional estimators.

Theorem 5.1. *(Glick's theorem)*
Let $\{f_N(x)\}$ be a sequence of probability density functions on $\mathbb{R}^d$, which are measurable functions of x and of $X_1, \ldots, X_N$ and such that $f_N \xrightarrow[N\to\infty]{a.s.} f$ almost everywhere in x. Then

$$\int_{\mathbb{R}^d} |f_N(x) - f(x)|dx \xrightarrow[N\to\infty]{L_1} 0. \tag{38}$$

Proof. See Glick (1974). □

The following lemma is a useful preliminary result before showing some convergence results for the proposed multi-scale density estimator.

Lemma 5.1. *Let K be a Gaussian kernel on $\mathbb{R}^d$, $f \in L^1(\mathbb{R}^d)$ a density function and $\{h_i\}_{i=1}^N$ a sequence of positive numbers (scale factors), which takes values in the interval $[h_{N,\min}, h_{N,\max}]$. Let $h_{N,\max}$ satisfy*

$$\lim_{N\to\infty} h_{N,\max} = 0 \tag{39}$$

then

$$\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}(f * K_{h_i})(x) = f(x), \quad \forall x \in \mathbb{R}^d \tag{40}$$

and

$$\lim_{N\to\infty} \sup_{x} \left| \frac{1}{N} \sum_{i=1}^{N} (f * K_{h_i})(x) - f(x) \right| = 0. \tag{41}$$

Proof. This result extends the Bochner's lemma to the case of a variable-bandwidth kernel. Thus to prove the theorem we follow the technique used in Bosq and Lecoutre (1987) (p. 61). Let $\delta > 0$, then

$$\begin{aligned}
&\left| \frac{1}{N} \sum_{i=1}^{N} (f * K_{h_i})(x) - f(x) \right| \leq \qquad (42)\\
&\leq \frac{1}{N} \sum_{i=1}^{N} \int_{\|x_i\| \leq \delta} |(f(x - x_i) - f(x)) K_{h_i}(x_i)| \, dx_i + \\
&\quad + \frac{1}{N} \sum_{i=1}^{N} \int_{\|x_i\| > \delta} |(f(x - x_i) - f(x)) K_{h_i}(x_i)| \, dx_i \\
&\leq \frac{1}{N} \sum_{i=1}^{N} \sup_{\|x_i\| \leq \delta} |(f(x - x_i) - f(x))| \int_{\|x_i\| \leq \delta} |K_{h_i}(x_i)| \, dx_i + \\
&\quad + \frac{1}{N} \sum_{i=1}^{N} \int_{\|x_i\| > \delta} \frac{|f(x - x_i)|}{\| x_i \|^d} \left\| \frac{x_i}{h_i} \right\|^d |K(x_i / h_i)| \, dx_i + \\
&\quad + \frac{1}{N} \sum_{i=1}^{N} |f(x)| \int_{\|x_i\| > \delta / h_i} |K(z_i)| \, dz_i \\
&\leq \frac{1}{N} \sum_{i=1}^{N} \sup_{\|x_i\| \leq \delta} |(f(x - x_i) - f(x))| \int_{\mathbb{R}^d} |K(z_i)| \, dz_i + \\
&\quad + \frac{1}{N} \sum_{i=1}^{N} \delta^{-d} \int_{\mathbb{R}^d} |f(x - z_i h_i)| \sup_{\|z_i\| > \delta / h_i} \| z_i \|^d |K(z_i)| \, dz_i + \\
&\quad + \frac{1}{N} \sum_{i=1}^{N} |f(x)| \int_{\|z_i\| > \delta / h_i} |K(z_i)| \, dz_i
\end{aligned}$$

Let δ be fixed and $N \to \infty$ then $h_{N,\max} \to 0 \Rightarrow h_i \to 0, \forall\, i$ and the last two summations tend to zero, because for Gaussian kernels $\| z \|^d |K(z)| \to 0$ as $\| z \|^d \to \infty$. Now let δ tend to zero, also the first summation equals to zero.

The second part of the lemma follows from the inequalities

$$\sup_x \left| \frac{1}{N} \sum_{i=1}^{N} (f * K_{h_i})(x) - f(x) \right| \leq \frac{1}{N} \sum_{i=1}^{N} \sup_x |(f * K_{h_i})(x) - f(x)| \tag{43}$$

$$\leq \frac{1}{N} \sum_{i=1}^{N} \left\{ \sup_x \sup_{\|x_i\| \leq \delta} |(f(x - x_i) - f(x))| \int |K(z_i)| \, dz_i \right\} +$$

$$+ \frac{1}{N} \sum_{i=1}^{N} \left\{ 2 \sup_x |f(x)| \int_{\|z_i\| > \delta / h_i} |K(z_i)| \, dz_i \right\}$$

which tends to zero due to result in the first part of the lemma. □

Theorem 5.2. *(Quadratic-mean convergence)*
Let K be a Gaussian kernel on $\mathbb{R}^d$, $f \in L^1(\mathbb{R}^d)$ a density, $\{x_i\}_{i=1}^N$ i.i.d. samples with common density f and $\{h_i\}_{i=1}^N$ a sequence of positive numbers (bandwidths) with values in the interval $[h_{N,\min}, h_{N,\max}]$. Let the bounds satisfy

$$\lim_{N \to \infty} h_{N,\max} = 0 \,, \quad \lim_{N \to \infty} (h_{N,\min})^d N = \infty \tag{44}$$

then the sample-point estimators, f_N, converges in L_2 to the true density

$$f_N(x) \xrightarrow[N \to \infty]{L_2} f(x), \quad a.e. \text{ in } x. \tag{45}$$

Proof. Take the expectation of the functional estimator with respect to the sequence of i.i.d. random vectors $\{X_i\}$

$$\begin{aligned} & \mathbb{E}\left(f_N(x)\right) = \\ = & \int_{\mathbb{R}^{Nd}} \left(\frac{1}{N} \sum_{i=1}^{N} \frac{1}{h_i^d} K(h_i^{-1}(x - x_i)) \right) f(x_1) \ldots f(x_N) dx_1 \ldots dx_N \\ = & \frac{1}{N} \sum_{i=1}^{N} \int_{\mathbb{R}^d} \frac{1}{h_i^d} K(h_i^{-1}(x - x_i)) f(x_i) dx_i \\ = & \frac{1}{N} \sum_{i=1}^{N} (f * K_{h_i})(x) \end{aligned} \tag{46}$$

which is the quantity studied in Lemma 5.1. Now the quadratic error decomposes as follows

$$
\begin{aligned}
& \mathbb{E}(f_N(x) - f(x))^2 = \qquad (47)\\
= \ & \mathbb{E}[f_N(x) - \mathbb{E}(f_N(x))]^2 + [\mathbb{E}(f_N(x)) - f(x)]^2 \\
= \ & \mathbb{E}\left(\frac{1}{N}\sum_{i=1}^{N}\frac{1}{h_i^d}K(h_i^{-1}(x-x_i)) - \frac{1}{N}\sum_{i=1}^{N}(f * K_{h_i})(x)\right)^2 + \\
& +[\mathbb{E}(f_N(x)) - f(x)]^2 \\
= \ & \mathbb{E}\left(\frac{1}{N}\sum_{i=1}^{N}\frac{1}{h_i^d}K(h_i^{-1}(x-x_i))\right)^2 - \left(\frac{1}{N}\sum_{i=1}^{N}(f * K_{h_i})(x)\right)^2 + \\
& +[\mathbb{E}(f_N(x)) - f(x)]^2
\end{aligned}
$$

$$
\begin{aligned}
& = \frac{1}{N^2}\sum_{i=1}^{N}\int\frac{1}{h_i^{2d}}K^2(h_i^{-1}(x-x_i))f(x_i)dx_i \\
& \quad + \frac{1}{N^2}\sum_{i\neq j}\int\frac{1}{h_i^d h_j^d}K(h_i^{-1}(x-x_i))K(h_j^{-1}(x-x_j))f(x_j)f(x_i)dx_j dx_i \\
& \quad - \left(\frac{1}{N}\sum_{i=1}^{N}(f * K_{h_i})(x)\right)^2 + [\mathbb{E}(f_N(x)) - f(x)]^2 \\
& = \frac{1}{N^2}\sum_{i=1}^{N}\frac{1}{h_i^d}(f * K_{h_i}^2)(x) - \frac{1}{N^2}\sum_{i=1}^{N}\frac{1}{h_i^{2d}}\left(\int K(h_i^{-1}(x-x_i))f(x_i)\right)^2 dx_i + \\
& \quad +[\mathbb{E}(f_N(x)) - f(x)]^2 \\
& \leq \frac{\int K^2}{N h_{N,\min}^d}\left(\frac{1}{N}\sum_{i=1}^{N}(f * \tilde{K}_{h_i}^2)(x)\right) - \frac{1}{N^2}\left(\sum_{i=1}^{N}\mathbb{E}^2\left(K_{h_i}(x-x_i)\right)\right) + \\
& \quad +[\mathbb{E}(f_N(x)) - f(x)]^2
\end{aligned}
$$

where $\tilde{K}_{h_i}^2(x) = K_{h_i}^2(x)/\int K^2$ is a normalized kernel, integrating to one.

Lemma 5.1 and the assumption: $h_{N,\max} \to 0$ as $N \to \infty$, insure that the first term in parenthesis converges to $f(x)$ and the last term, converges to zero. Each element of the sum in the middle is bounded. The two elements tend to zero providing that $\lim_{N\to\infty}(h_{N,\min})^d N = \infty$. □

The following theorem states the almost sure (a.s.) convergence of the proposed sample-point estimator.

Theorem 5.3. *(a.s. convergence)*
Let K be a Gaussian kernel on $\mathbb{R}^d$, $f \in L^1(\mathbb{R}^d)$ a density, $\{x_i\}_{i=1}^N$ i.i.d. samples from

f and $\{h_i\}_{i=1}^N$ a sequence of positive numbers (bandwidths) with values in the interval $[h_{N,\min}, h_{N,\max}]$. Let the bounds satisfy

$$\lim_{N\to\infty} h_{N,\max} = 0\,, \quad \lim_{N\to\infty} \frac{(h_{N,\min})^d N}{\log N} = \infty \tag{48}$$

then the sample-point estimators, f_N, converges a.s. to the true density

$$f_N(x) \xrightarrow[N\to\infty]{a.s.} f(x) \tag{49}$$

Proof. We follow Devroye and Wagner (1979) and apply the triangular inequality to decompose the estimation error into the variance and bias components.

$$|f_N(x) - f(x)| \leq |f_N(x) - \mathbb{E}(f_N(x))| + |\mathbb{E}(f_N(x)) - f(x)|. \tag{50}$$

We first consider the second term on the RHS of the inequality and show that by a straightforward application of the Lemma 5.1

$$\mathbb{E}(f_N(x)) \xrightarrow[N\to\infty]{a.s.} f(x).$$

The second result in Lemma 5.1 insures the pointwise convergence of $\mathbb{E}(f_N(x))$ to $f(x)$ uniformly in x and implies

$$|\mathbb{E}(f_N(x)) - f(x)| = \left| \frac{1}{N} \sum_{i=1}^N (f * K_{h_i})(x) - f(x) \right| \xrightarrow[N\to\infty]{a.s.} 0 \tag{51}$$

providing that $h_{N,\max} \to 0$ as $N \to \infty$ and $\parallel z \parallel^d |K(z)| \to 0$ as $\parallel z \parallel^d \to \infty$.

Consider now the first term on the RHS of the inequality,

$$|f_N(x) - \mathbb{E}(f_N(x))| = \left| \frac{1}{N} \sum_{i=1}^N \left(\frac{1}{h_i^d} K(h(x - x_i)) - (f * K_{h_i})(x) \right) \right|. \tag{52}$$

Let $M = \sup_z K(z)$ and note that

$$K_{h_i} \leq h_i^{-d} M \leq h_{N,\min}^{-d} M$$

and

$$\mathbb{E}(K_{h_i}^2) = h_i^{-d} (f * K_{h_i}^2)(x) \leq h_{N,\min}^{-d} h_i^{-d} M (f * K_{h_i})(x).$$

By Bernstein-Fréchet's inequality[15] ,

$$
\begin{aligned}
& P\left(|f_N(x) - \mathbb{E}(f_N(x))| \geq \varepsilon\right) \leq \\
\leq \quad & 2\exp\left(-\frac{\varepsilon^2 N^2}{4h_{N,\min}^{-d} M \sum_i \mathbb{V}(K_{h_i}) + 2\varepsilon N h_{N,\min}^{-d} M}\right) \\
< \quad & 2\exp\left(-\frac{\varepsilon^2 N^2 h_{N,\min}^{d}}{4M \sum_i \mathbb{E}(K_{h_i}) + 2\varepsilon N M}\right) \\
\leq \quad & 2\exp\left(-\frac{\varepsilon^2 N^2 h_{N,\min}^{d}}{4M \sum_i h_i^{-d}(f * K_{h_i})(x) + 2\varepsilon N M}\right) \\
= \quad & 2\exp\left(-\frac{\varepsilon^2 N h_{N,\min}^{d}}{4M\mathbb{E}(f_N(x)) + 2\varepsilon M}\right) \leq \exp\left(\alpha N h_{N,\min}^{d}\right)
\end{aligned}
\tag{53}
$$

which is bounded a.e. in x and for all N, due to the a.s. convergence of $\mathbb{E}(f_N(x))$ to $f(x)$.. The above inequality insures the almost complete (a.c.) convergence and consequently the a.s. convergence (see Bosq and Lecoutre (1987)) of the estimator

$$
f_N(x) \xrightarrow[N\to\infty]{a.c.} \mathbb{E}(f_N(x)),
$$

providing that $\sum_{N=1}^{\infty} e^{-\alpha N h_{N,\min}^{d}} < \infty$. This condition is clearly implied by the assumption: $\lim\limits_{N\to\infty} A_N = \infty$ where $A_N = N h_{N,\min}^{d} \log(N)^{-1}$. For an arbitrarily chosen $\eta > 0$ there exists M s.t. for all $N > M$, $A_N > \eta$ and

$$
\begin{aligned}
& \sum_{N=1}^{\infty} e^{-\alpha N h_{N,\min}^{d}} \sum_{N=1}^{\infty} e^{-\alpha \log(N) A_N} < \\
& < \sum_{N=1}^{M} e^{-\alpha \log(N) A_N} + \sum_{N=M+1}^{\infty} e^{-\alpha \log(N)\eta} \\
& < \sum_{N=1}^{M} e^{-\alpha \log(N) A_N} + \sum_{N=1}^{\infty} N^{-\alpha\eta} < \infty
\end{aligned}
\tag{54}
$$

choosing η such that the series is convergent. □

Theorem 5.4. *(L_1-convergence)*
Let K be a Gaussian kernel on $\mathbb{R}^d$, $f \in L^1(\mathbb{R}^d)$ a density, $\{x_i\}_{i=1}^N$ i.i.d. samples from f and $\{h_i\}_{i=1}^N$ a sequence of positive numbers (bandwidths) with values in the interval $[h_{N,\min}, h_{N,\max}]$. Let the bounds satisfy

$$
\lim_{N\to\infty} h_{N,\max} = 0\,, \qquad \lim_{N\to\infty} \frac{(h_{N,\min})^d N}{\log N} = \infty \tag{55}
$$

[15] See Bosq and Lecoutre (1987) (Th. I.2, p. 41), with $p = 3$, and $\mathbb{E}|K_{h_i} - \mathbb{E}(K_{h_i})|^3 \leq h_{N,\min}^{-d} M \mathbb{V}(K_{h_i})$.

then the sample-point estimators, f_N, converges in L_1 to the true density

$$f_N(x) \xrightarrow[N\to\infty]{L_1} f(x). \tag{56}$$

Proof. Note first that the assumptions of the theorem insure the a.s.-convergence (see Theorem (5.3)) of the multiple-bandwidth estimator. The a.s. convergence implies the L_1-convergence by the Glick's theorem. □

Appendix D - Recursive Parameter Estimates

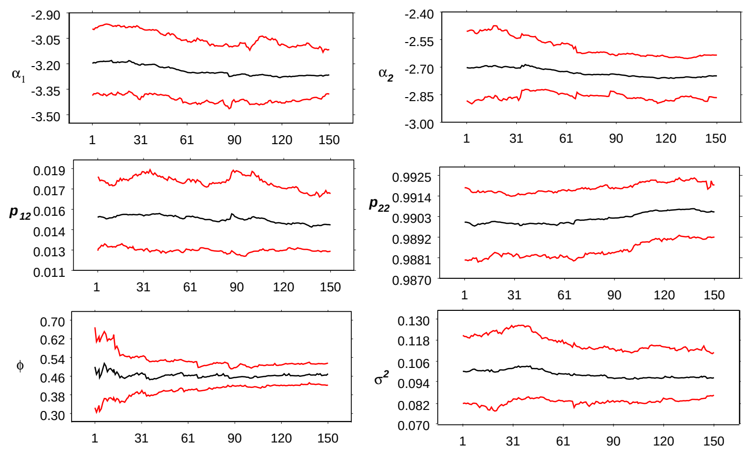

Figure 12: *On-line parameter estimates for the log-returns on the S&P price index. Graphs show at each date the empirical mean and the quantiles at 0.025 and 0.975 for each parameter.*

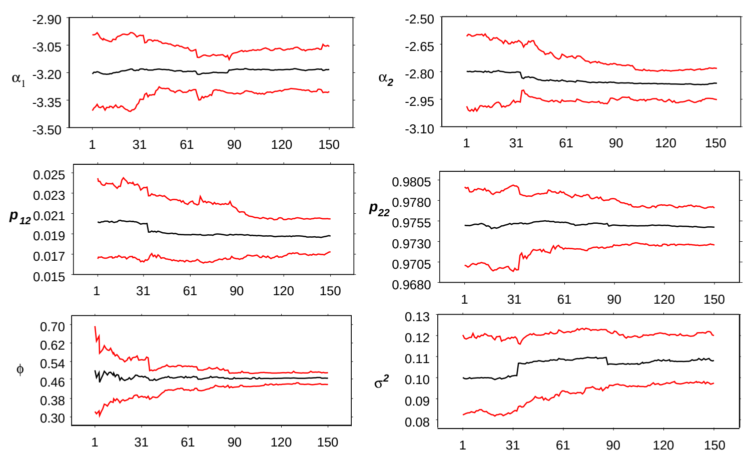

Figure 13: *On-line parameter estimates for the Price-to-Earnings. Graphs show at each date the empirical mean and the quantiles at 0.025 and 0.975 for each parameter.*

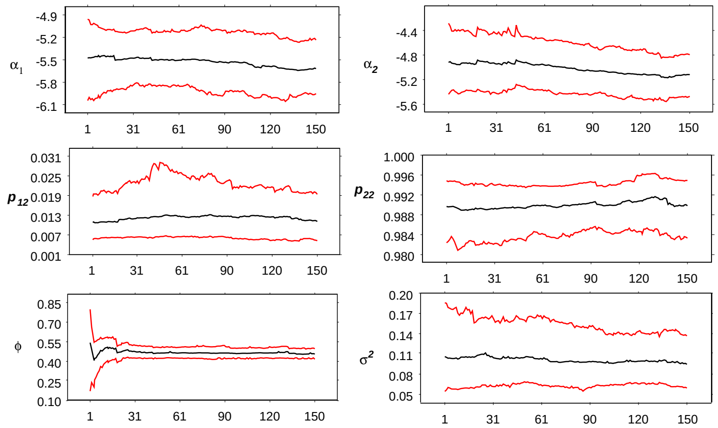

Figure 14: *On-line parameter estimates for the Dividend Yields. Graphs show at each date the empirical mean and the quantiles at 0.025 and 0.975 for each parameter.*

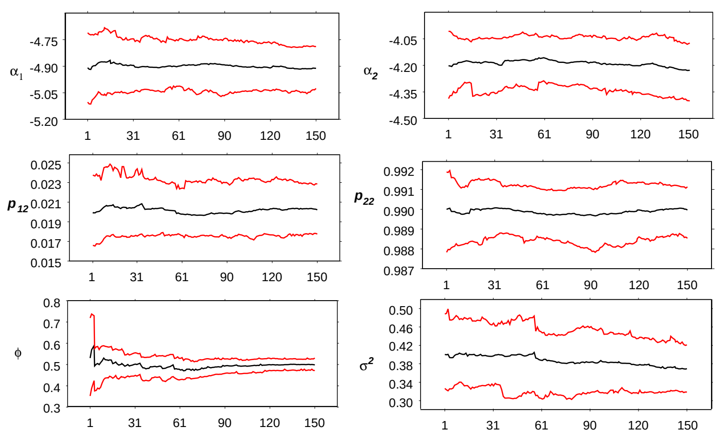

Figure 15: *On-line parameter estimates for the log-returns on the Industrial Production index. Graphs show at each date the empirical mean and the quantiles at 0.025 and 0.975 for each parameter.*

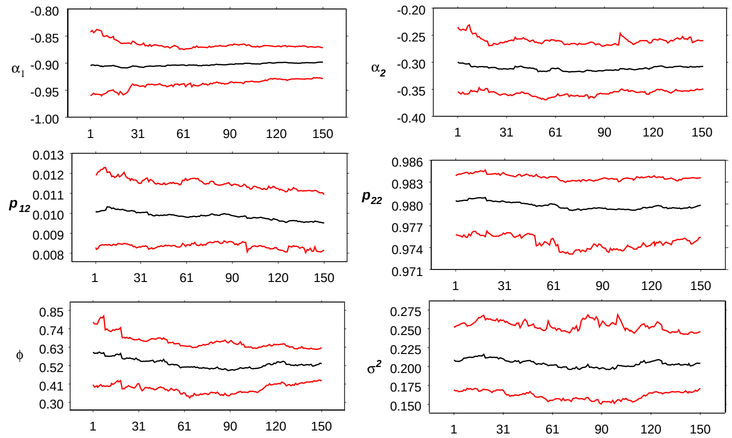

Figure 16: *On-line parameter estimates for the Personal Consumption Expenditure. Graphs show at each date the empirical mean and the quantiles at 0.025 and 0.975 for each parameter.*

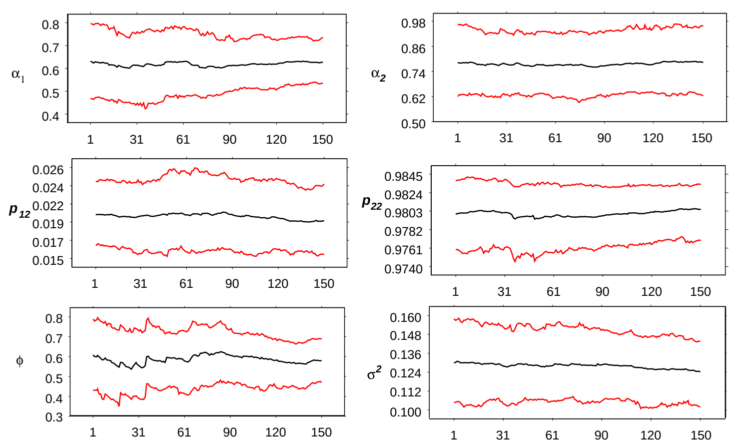

Figure 17: *On-line parameter estimates for the non-residential investments series. Graphs show at each date the empirical mean and the quantiles at 0.025 and 0.975 for each parameter.*

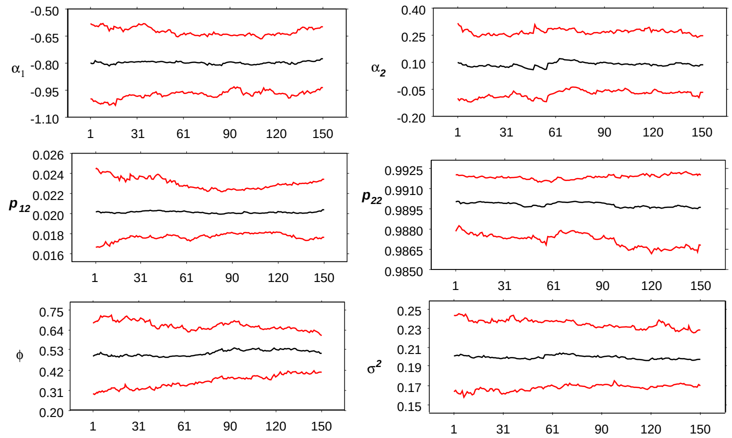

Figure 18: *On-line parameter estimates for the output gap series. Graphs show at each date the empirical mean and the quantiles at 0.025 and 0.975 for each parameter.*

References

ARULAMPALAM, S., S. MASKELL, N. GORDON, AND T. CLAPP (2001): "A Tutorial on Particle Filters for On-line Nonlinear/Non-Gaussian Bayesian Tracking," Discussion paper, Technical Report, QinetiQ Ltd., DSTO, Cambridge.

BARTOLI, N., AND P. DEL MORAL (2001): *Simulation et algorithmes stochastiques*. Cépaduès Éditions, Paris.

BAUWENS, L., M. LUBRANO, AND J. F. RICHARD (1999): *Bayesian Inference in Dynamic Econometric Models*. Oxford University Press, New York.

BERZUINI, C., N. G. BEST, W. R. GILKS, AND C. LARIZZA (1997): "Dynamic conditional independence models and Markov chain Monte Carlo methods," *Journal of the American Statistical Association*, 92(440), 1403–1441.

BERZUINI, C., AND W. R. GILKS (2001): "Following a moving average-Monte Carlo inference for dynamic Bayesian models," *Journal of Royal Statistical Society, B*, 63(1), 127–146.

BILLIO, M., R. CASARIN, AND D. SARTORE (2004): "Bayesian Inference on Dynamic Models with Latent Factors," Official Monography, Edited by EUROSTAT.

Borio, C., and P. Lowe (2004): "Securing sustainable price stability: should credit come back from the wilderness?," BIS Working Papers N. 157.

Bosq, D., and J. P. Lecoutre (1987): *Théorie de l'estimation fonctionelle*. Economica, Paris.

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

Casarin, R. (2004): "Bayesian Inference for Markov Switching Stochastic Volatility Models," Cahier du CEREMADE N. 0414, Université Paris Dauphine.

Chopin, N. (2001): "Sequential inference and state number determination for discrete state-space models through particle filtering," Technical Report N. 34, CREST.

Chopin, N., and F. Pelgrin (2004): "Bayesian inference and state number determination for hidden Markov models: an application to the information content of the yield curve about inflation," *Journal of Econometrics*, 123(2), 327–344.

Cochrane, J. H. (2005): *Asset Pricing (Revised edition)*. Princeton University Press, Princeton.

Crisan, D. (2001): "Particle filters, A theoretical perspective," in *Sequential Monte Carlo Methods in Practice*, ed. by A. Doucet, J. Freitas, and J. Gordon. Springer-Verlag, New York.

Crisan, D., and A. Doucet (2000): "Convergence of sequential Monte Carlo methods," Technical Report N. 381, CUED-F-INFENG.

——— (2002): "A survey of convergence results on particle filtering for practitioners," *IEEE Trans. Signal Processing*, 50, 736–746.

Del Moral, P. (2004): *Feynman-Kac Formulae*. Springer Verlag, New York.

Devroye, L. P., and T. J. Wagner (1979): "The L_1 Convergence of Kernel Density Estimates," *The Annals of Statistics*, 7(5), 1136–39.

Doucet, A., J. G. Freitas, and J. Gordon (2001): *Sequential Monte Carlo Methods in Practice*. Springer Verlag, New York.

DOUCET, A., S. GODSILL, AND C. ANDRIEU (2000): "On sequential Monte Carlo sampling methods for Bayesian filtering," *Statistics and Computing*, 10(3), 197–208.

DYNAN, K. E., D. W. ELMENDORF, AND D. E. SICHEL (2006): "Can financial innovation help to explain the reduced volatility of economic activity?," *Journal of Monetary Economics*, 53, 123–150.

FERSON, W. (2003): "Test of Multi-Factor Pricing Models, Volatility, and Portfolio Performance," in *Handbook of the Economics of Finance, 1B*, ed. by G. M. Constantinides, M. Harris, and R. M. Stulz. Elsevier, North-Holland.

GLICK, N. (1974): "Consistency conditions for probability estimators and integrals of density estimators," *Utilitas Mathematica*, 6, 61–74.

GORDON, N., D. SALMOND, AND A. F. M. SMITH (1993): "Novel Approach to Nonlinear and Non-Gaussian Bayesian State Estimation," *IEE Proceedings-F*, 140, 107–113.

GORDON, R. J. (2005): "What caused the decline in U.S. business cycle volatility?," NBER Working Paper N. 11777.

HAMILTON, J. D. (1989): "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica*, 57(2), 357–384.

HARRISON, J., AND M. WEST (1997): *Bayesian Forecasting and Dynamic Models, 2nd Ed.* Springer Verlag, New York.

HARVEY, A. C. (1989): *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press, Cambridge.

JAZWINSKI, A. H. (1970): *Stochastic Processes and Filtering Theory*. Academic Press, New York.

KALMAN, R. E., AND R. S. BUCY (1960): "New results in linear filtering and prediction problems," *Transaction of the ASME-Journal of Basic Engineering, Series D*, 83, 95–108.

KIM, C. J., AND C. R. NELSON (1999): *State-Space Models with Regime Switching*. MIT press, Cambridge.

KIM, S., N. SHEPHARD, AND S. CHIB (1998): "Stochastic volatility: likelihood inference and comparison with ARCH models," *Review of Economic Studies*, 65, 361–393.

Kitagawa, G. (1998): "A self-organizing state space model," *Journal of the American Statistical Association*, 93, 1203–1215.

Lettau, M., and S. C. Ludvigson (2001): "Consumption, aggregate wealth, and expected stock returns," *Journal of Finance*, 3, 815–849.

——— (2004): "Understanding Trend and Cycle in Asset Value: Reevaluating the Wealth Effect on Consumption," *American Economic Review*, 94, 276–299.

Lettau, M., S. C. Ludvigson, and J. A. Wachter (2007): "The declining equity premium: what role does macroeconomic risk play?," *Review of Financial Studies, forthcoming.*

Liu, J., and M. West (2001): "Combined Parameter and State Estimation in Simulation Based Filtering," in *Sequential Monte Carlo Methods in Practice*, ed. by A. Doucet, J. Freitas, and J. Gordon. Springer-Verlag, New York.

Liu, J. S., and R. Chen (1998): "Sequential Monte Carlo Methods for Dynamical System," *Journal of the American Statistical Association*, 93(443), 1032–1044.

Lopes, H. F., and C. Marigno (2001): "A particle filter algorithm for the Markov switching stochastic volatility model," Working paper, University of Rio de Janeiro.

McConnell, M. M., and G. Perez-Quiros (2000): "Output fluctuations in the United States: What has changed since the early 1980s?," *American Economic Review*, 90(5), 1464–1476.

Musso, C., N. Oudjane, and F. LeGland (2001): "Improving Regularised Particle Filters," in *Sequential Monte Carlo in Practice*, ed. by A. Doucet, J. Freitas, and J. Gordon. Springer Verlag, New York.

Perpiñán, M. A. C. (2001): "Continuous latent variable models for dimensionality reduction and sequential data reconstruction," PhD Thesis in Computer Science, University of Sheffield.

Pitt, M., and N. Shephard (1999): "Filtering via Simulation: Auxiliary Particle Filters," *Journal of the American Statistical Association*, 94(446), 590–599.

Polson, N. G., J. R. Stroud, and P. Müller (2002): "Practical Filtering with sequential parameter learning," Tech. report, Graduate School of Business, University of Chicago.

——— (2003): "Practical Filtering for Stochastic Volatility Models," Tech. report, Graduate School of Business, University of Chicago.

Robert, C. P., and G. Casella (2004): *Monte Carlo Statistical Methods, 2nd ed.* Springer Verlag, New York.

Rossi, V. (2004): "Filtrage non linéaire par noyaux de convolution. Application à un procédé de dépollution biologique," Thèse du Doctorat en Sciénce, Ecole Nationale Superiore Agronomique de Montpellier.

So, M. K. P., K. Lam, and W. K. Li (1998): "A stochastic volatility model with Markow switching," *Journal of Business & Economic Statistics*, 16, 244–253.

Stavropoulos, P., and D. M. Titterington (2001): "Improved Particle filters and smoothing," in *Sequential Monte Carlo Methods in Practice*, ed. by A. Doucet, J. Freitas, and J. Gordon. Springer-Verlag, New York.

Stock, J. H., and M. W. Watson (2002): "Has the Business Cycle Changed and Why?," in *NBER Macroeconomics Annual*, ed. by M. Gertler, and K. Rogoff. MIT Press, Cambridge.

Storvik, G. (2002): "Particle filters for state space models with the presence of unknown static parameters," *IEEE Trans. on Signal Processing*, 50, 281–289.

Villaverde, J. F., and J. F. Ramirez (2004a): "Estimating Nonlinear Dynamic Equilibrium Economies: A Likelihood Approach," Atlanta FED Working Paper N. 03.

——— (2004b): "Estimating Nonlinear Dynamic Equilibrium Economies: Linear versus Nonlinear Likelihood," Atlanta FED Working Paper N. 01.

Watson, J. (1994): "Business cycle durations and postwar stabilization of the U.S. economy," *American Economic Review*, 84, 24–46.