Board of Governors of the Federal Reserve System

International Finance Discussion Papers

Number 461

December 1993

FLUCTUATING CONFIDENCE AND STOCK-MARKET RETURNS

Alexander David

Abstract

The drift of two different diffusion processes (asset returns) is determined by a state variable which can take on two values. It jumps between the two according to Poisson increments (this is called a 'regime-switch'). For any given position of the state variable the drift of one process is high and the other is low. I find that the posterior probability that the 1st asset has higher average returns, conditional on observing the path (returns) of each process, follows a diffusion process and calculate its infinitesimal parameters. I also derive analytical expresssions for its stationary density and for some of its path properties. I compare the filtering problem to the Kalman Filtering problem and find that even though the dynamics of the mean of the distribution are very similar, the dynamics of the variance are subject to stochastic fluctutations. The model is parsimonious in that the conditional mean and variance are functions of a single variable.

I characterize the interest-rate and total-returns processes in a Cox-Ingersoll-Ross[1985] style model where the productivities of assets are unobserved, but inferred as above. I find that this model is capable of reproducing three stylized facts of stock-market returns and interest-rates. These are the skewness and kurtosis of returns and the 'Predictive-Asymmetry' of returns: excess-returns and future changes in volatility are negatively correlated. Further negative returns cause reactions of larger magnitude. The success of the model in generating these features depends on the speed of learning about the regime switches. Parameter values which lead to faster learning, are consistent with large negative skewness of returns and the Predictive Asymmetry property. The slower learning version leads to greater kurtosis of returns. I show that a model based on the same fundamentals but with observed 'regime-shifts' is not reconcilable with these features. My analysis suggests that learning about the productivities of assets of the kind introduced here may be an important determinant of portfolio choices and observed asset returns.

Fluctuating Confidence and Stock Market Returns

Alexander David [1]

This paper has three purposes. Firstly, the presentation of a Filter in continuous time which characterizes the dynamics of Bayesian learning about *recurrent* 'regime-switches' to be defined shortly. Secondly, to show how 'fluctuating confidence' which arises due to this updating process, is reflected in the statistical properties of interest rate and stock-return processes in a Cox-Ingersoll-Ross[1985] (henceforth CIR) stochastic production economy. I also discuss the nature of the risk associated with these fluctuations and the form of the optimally chosen portfolios to hedge this risk. Finally I draw some relationships between the speed of learning and the ability of the model to replicate three stylized facts about stock-market returns.

A 'regime-switch' is said to happen when the *average productivities* of two different sets of assets in the economy are reversed. The switching occurs due to a Poisson process. We assume that the switching is unobserved and that the total output from each asset which is the sum of the average productivity and 'noise' is observed. In contrast to learning models based on Gaussian distributions of the the underlying state variables and 'noise', the updating process here exhibits stochastically fluctuating conditional variance. The regime-switch defined here is to be distuingshed from that of Hamilton[1991], where a switch is a change in the *rate of growth of ouptput* in an

[1]The author is a staff economist in the Division of International Finance. This paper represents the views of the author and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or other members of its staff. I thank the members of the International Finance Division and specially those in Trade and Financial Studies for helping me make a smooth transition. A large part of this paper is from my PhD dissertation written at UCLA. The financial support of the National Science Foundation is gratefully acknowledged. I am grateful to Michael Brennan, Bryan Ellickson, Jon Faust, Robert Jones, David Levine, Thomas Mountford, Simon Potter, Keunkwan Ryu, Kwanho Shin and Bill Zame for helpful conversations. I thank David Senouf and Ronald Fedkiw for clarifying some issues in numerical analysis and Yun Sun for research assistance. I am specially grateful to Joe Ostroy for his guidance, patience and support.

economy. Both here and in Hamilton's model the regime switching is described by a 2-state Markov chain. However, the implications of a regime-switch on output growth in this model depend on the speed of learning about these switches, the expected-time to retain certain levels of confidence and the resulting portfolio choices. Further, due to unobservability of these switches there are effectively a continuum of states for the decision maker, a state being identified by the belief of the agent regarding the current regime.

The inference process is described as follows. The agent receives signals continuously from a source which *jumps* 'infrequently' between two values. I emphasise that this is different from the assumption used in the Kalman Filtering problem, Jazwinski[1970], and its variants, that the source moves according to a diffusion process. I find that the updating process follows a diffusion process and calculate its infinitesimal parameters. This implies that the updating process has continuous sample paths even though the source follows a jump process. Furthermore the updating process satisfies certain regularity conditions which enable us to characterize its stationary density and several of its path properties. I show that the dynamics of the conditional mean of the agent's estimate are very similar to that of the Kalman filtering problem. However as opposed to the Kalman Filtering problem, the dynamics of the conditional variance are stochastic and subject to cyclical fluctuations. This makes my model particularly suitable for studying the effects of fluctuating confidence.

It is somewhat surprising that the updating process has continuous sample paths even though the drift may jump by a large amount, and the process is observed continuously. The increments of a diffusion process are the sum of a term which is proportional to the drift and the the length of the observation period, and a noise term whose increments are distributed like a Brownian motion. The standard deviation of the increments in the noise term are proportional to the square-root of the observation interval. Over intervals of small length, the drift has negligible effect while the amount of noise is of a larger magnitude. So over very small intervals the increments of the observed process are not informative and do not lead to large changes in estimates.

A model of Non-Gaussian learning has been worked out by Detemple[1991]. In that model it is assumed that the prior distribution is Non-Gaussian. In that economy too the conditional variance process is stochastic. The mean and a set of sufficient of statistics for the conditional variance characterize the updated distribution. In the model presented the mean and variance of the distribution are completely characterized by a single parameter. This is completely analagous to the static binomial distribution. This parsimony allows us to make precise several properties of the updating process including the stationary distribution of the updated priors, the boundary behaviour of the process and the amount of time spent in various regions of the state space characterized by hitting times. Further we are able to solve for optimal decision rules and equilibrium rates of return in a one-factor CIR model.

I point to three stylized facts about stock-market returns. These are the observed kurtosis and skewness of excess-returns and the asymmetric feed-back effect of excess-returns on stock-market volatility. A number of time-series models have been written to replicate these features and to estimate the strengths of these effects. Without attempting an exhaustive survey of the literature I refer readers for a vivid description of the facts to Black[1976] and to the the models of Nelson[1991] and Campbell and Hentschel[1992]. These models are generalizations of ARCH and ARCH-M models (Bollerslev, Engle and Woolridge[1985] and Engle, Lilien and Robins[1987]) which documented the autoregressive fluctuating volatility of stock-returns. In the model economy presented here fluctuations in returns and their feedback to conditional variance arise due to unobservability of the regime shifts. The inherent inertia in Bayesian learning and changing portfolio choices account for the changing means and the autocorrelation in the volatility of return processes.

I show that in an economy with the same fundamentals and *observable* regime shifts the interest rate and volatility of returns is constant and the three features of stock returns cannot be replicated. The ability of the model with unobservable regime-shifts to replicate these features depends on the parameters of the model chosen. In models where the difference in drifts is large, the level of noise is low and regime switches are less frequent, the agent spends a large proportion of time in states with high confidence regarding his knowledge of the current regime. I call this the 'Fast Learning' model and the one with the opposite set of parameters the 'Slow Learning' model. I find that the Slow Learning model generates returns which exhibit excess-kurtosis (fatter tails than the Normal Distribution) and that the Fast Learning generates negative- skewness. A model with an intermediate speed of learning is needed to replicate both features simultaneously. I also find that the Fast Learning model generates a negative relationship between realized excess-returns and future increases in volatility (as found in the data) and that the relationship is not so clear in the Slow Learning model.

Before carrying out the analysis I refer to some empirical evidence for 'reallocative' shocks, i.e. shocks which create a desire to move resources between 'sectors'. The literature has developed since the *Sectoral Shifts Hypothesis* of Lilien[1982], who argued that employment fluctuations in the U.S. economy can be explained due to shocks which unevenly affect the productivities of different sectors in the U.S. economy. The exact definition of a 'sector' has been a subject of debate. I do not attempt an exhaustive survey of the literature . Loungani, Rush and Tave[1990] created a dispersion measure and found evidence of reallocative shocks between 60 industrial indeces constructed by Standard and Poor. Davis and Haltiwanger[1992] argue that the effects of these shocks is experienced at the plant rather than the industry level. The literature is still evolving yet supportive of the Reallocative Shock Hypotheis at different levels of disaggreation. I find it satisfactory to assume for now that the effects of asymmetric shocks to different industries may be a useful paradigm for understanding various features of stock-market returns.

The plan for the rest of this paper is as follows. In Section 2 the structure of the two models is presented. In Section 3 I solve the model under the assumption that regime-switches are observed. In Section 4 the filtering problem associated with unobserved regime-switches is solved. In Section 5 the model with unobserved regime-switches is analyzed and the nature of optimally chosen portfolios is discussed. In Section 6 some results from the numerical evaluation of the model and statistics from simulations are presented. The success of the model in its ability to replicate stylized facts about stock-market returns is evaluated in Section 7. The conclusion is in Section 8.

2 Structure of the Models

In this section we introduce the main features of the model. These are closely related to those in CIR[1985].

Feature 1. Single Good There is a single physical good which may be allocated to consumption or investment. All values are expressed in terms of units of this good.

Feature 2. Production Technology Production possibilities consist of two linear activities. The transformation of an investment of β_i amount of the good in the ith production process, is governed by a stochastic differential equation of the form

$$\frac{d\beta_{it}}{\beta_{it}} = \alpha_i(z_t) \cdot dt + \sigma \cdot d\zeta_{it} \tag{1}$$

for $i = 1, 2$. ζ_i are independent S.B.M.'s (Simple Brownian Motions),[1] where

$$\alpha_1(z_t) = z_t$$

$$\alpha_2(z_t) = a + b - z_t$$

$a > b$

(1) specifies the growth of an initial investment when the output of each process is continually reinvested in that same process. The production processes have stochastic constant returns to scale in the sense that the distribution of the rate of return on an investment in any process is independent of the scale of the investment. The drift rates of the processes are determined by a state variable z_t, which accounts for random productivity switches between technology 1 and 2.

[1] A real valued process ζ_i is a S.B.M. if (i) ζ_i has continous sample paths with probability one (ii) ζ_i has independent increments and (iii) $\zeta_{is} - \zeta_{it}$ has a normal distribution with mean zero and variance $s - t$.

Feature 3. Productivity Switches The dynamics of the state variable z_t which affects the drift rate of the production technologies is given by the Poisson stochastic differential equation

$$dz_t = (b-a)\cdot[1-\frac{2\cdot(z_t-a)}{b-a}]\cdot dq_t \tag{2}$$

$a > b$, q_t is a Poisson Process i.e.,

$q_{t+\Delta t} - q_t = 0$ with probability $1-\lambda\Delta t + o(\Delta t)^2$.

$q_{t+\Delta t} - q_t = 1$ with probability $\lambda\Delta t + o(\Delta t)$

$q_{t+\Delta t} - q_t = n, n \geq 2$ with probability $o(\Delta t)$

We call (2) *The Transition Equation* The Transition Equation implies that the unobserved state variable z_t can take only the two values a and b. It switches between the two with Poisson increments. The probability of no switches in a 'small' interval of time is $1-\lambda dt$, of one switch λdt, and of two or more switches is 'negligible'.

Feature 4. Representative Consumer There is a representative consumer in the economy. He seeks to maximize an objective function of the form:

$$E_t[\int_t^\infty \exp(-\rho\cdot(s-t))U[C_s]ds] \tag{3}$$

In (3), E_t is an expectations operator conditional on the current state of the economy. C_t is the consumption flow at time t. Throughout this paper we shall assume that

$$U[C_t] = \frac{C_t^\gamma}{\gamma} \tag{4}$$

(4) implies that the consumer's utility function exhibits constant relative risk-aversion with a risk-aversion coefficient of $1-\gamma$.

[2] We shall use the standard notation, $x(\Delta t) = o(\Delta t)$ if $\lim_{\Delta t\to 0}\frac{x(\Delta t)}{\Delta t} = 0$

Feature 5. Two Industries Investment is done through competitive value-maximizing firms in two different industries. Firms in each industry have access to one of the technologies mentioned above. There is free entry of firms within each industry.

Feature 6. Financial Assets In Zero Net Supply There is a market for instantaneous borrowing and lending at an interest rate r. The market clearing rate is the rate at which lending is at zero net supply. The market clearing rate is determined as part of the competitive equilibrium of the economy.

Feature 7. Continous Decision Making Physical investment and trading in real and financial claims take place continously in time. Trading takes place only at equilibrium prices.

3 Model 0. Observable Regime Switches

I first solve the stochastic control problem associated with the social-planning problem. I then define equilibrium in the economy, decentralize the decision-making and provide explicit expressions for the risk-free rate and the excess-returns.

In this model, I shall assume that the state variable z_t is observed by the representative agent. With these assumptions the firm faces, in the terminology introduced by Merton[1973], a *constant opportunity set* . At any moment of time, there exist two assets with average rates of return of a and b respectively. Since the chance of z_t switching from a to b in a small interval of time Δt is $\lambda \cdot \Delta t$ and asset choices can be revised after this small span of time, the firms expected payoffs are unaffected up to an order $o(\Delta t)$ by the potential switch. In this situation the only risk in the firm's return is due to the noise in the payoff of each technology. The aggregate wealth dynamics are given by

$$W_{t+\Delta t} = [W_t - C_t \Delta t] \cdot [\sum_{i=1}^{2} w_{it}(1 + \frac{d\beta_i}{\beta_{it}})] + o(\Delta t)$$

Proposition 1 *When regime-switches are observed the value function of the social-planner's problem is of the form*

$$J[W_t, t] = \exp(-\rho t) \cdot A \cdot \frac{W^\gamma}{\gamma} \tag{5}$$

where A is a constant determined by the parameters of the problem. Optimal portfolio choices are constant and satisfy

$$\begin{aligned} w_1 &= \frac{1}{2} + \frac{1}{2} \cdot \frac{(a-b)}{\sigma^2(1-\gamma)} \\ w_2 &= \frac{1}{2} + \frac{1}{2} \cdot \frac{(b-a)}{\sigma^2(1-\gamma)} \end{aligned} \tag{6}$$

The Proof is in Appendix 0. The analysis is completely straightforward and can be found in for example Ingersoll[1988].

Comments on Proposition 1 The value function does not depend on the state z_t. Both states offer the same opportunities to invest and grow. The portfolio choices reflect the desire to hedge noise, the only source of risk in this model. The portfolio is more diversified when the gap in average productivities $a - b$ is smaller, the level of noise is higher and the coefficient of risk-aversion is higher.

Decentralization Investment is done through competitve value maximizing firms. We recall there are three types of firms, each with access to one production process. With free-entry within each industry and stochastic constant returns to scale, there is no incentive for firms to enter or leave the industry if and only if the returns on the shares of each firm (the rate at which it can acquire capital) are identical to the technologically determined physical returns on that process. The equilibrium scale of each industry would then be determined by the supply of investment to that industry. Let w_0 be the share of his wealth allocated to riskless borrowing /lending.

The agent's wealth dynamics at t are, $W_{t+\Delta t} =$

$$[W_t - \hat{C}_t\Delta t] \cdot [\sum_{i=1}^{2} \hat{w}_{it}(1 + \alpha_i(z_t)\Delta t + \sigma\Delta\zeta_{it}) + \hat{w}_{0t}(1 + r_t\Delta t)] + o(\Delta t) \tag{7}$$

where $\hat{w}_{it} \geq 0$, $\hat{w}_{0t} + [\sum_{i=1}^{2} \hat{w}_{it}] = 1$ and $\hat{w}_{0t}$ can take either sign.

Definition 1 *An equilibrium is defined as a set of stochastic processes* $(\hat{C}_t, \hat{w}_{it}, r_t)$ *satisfying the first order conditions (38), (39) of the social-planning problem in Appendix 0, and the market clearing conditions* $\hat{w}_{it} \geq 0$ *and* $[\sum_{i=1}^{2} \hat{w}_{it}] = 1$ *and* $\hat{w}_0 = 0$

The *excess returns* in the economy are the difference in the rates of return on optimally invested wealth and the riskless security. I.e. the excess returns equal $\sum_{i=1}^{2} \alpha_i(z_t) \cdot w_{it} - r_t$.

Corollary 1 *The excess returns in the observable economy are constant and equal* $(1 - \gamma) \cdot (\sum_{i=1}^{2} w_i^2) \cdot \sigma^2$, *where* w_i *are the choices in Proposition 1.*

The proof is in Appendix 0.

4 The Filtering Problem

The filtering problem is described by two equations. The first describes the dynamics of an unobserved state variable and is called the Transition Equation. The second, describes the composition of a variable which is observed. It consists of the sum of a drift term which is determined by the state variable and a noise term, which has increments distributed like a Simple Brownian Motion. It is called the Observation Equation. Both equations fall under the class of Ito stochastic differential equations.

THE TRANSITION EQUATION

$$dz_t = (b-a)\cdot[1 - \frac{2\cdot(z_t - a)}{b-a}]\cdot dq_t \tag{8}$$

$a > b$, q_t is a Poisson Process i.e.,

$q_{t+\Delta t} - q_t = 0$ with probability $1 - \lambda\Delta t + o(\Delta t)^3$.

$q_{t+\Delta t} - q_t = 1$ with probability $\lambda\Delta t + o(\Delta t)$

$q_{t+\Delta t} - q_t = n, n \geq 2$ with probability $o(\Delta t)$

The Transition Equation implies that the unobserved state variable z_t can take only the two values a and b. It switches between the two with Poisson increments. The probability of no switches in a 'small' interval of time is $1 - \lambda dt$, of 1 switch λdt, and of two or more switches is 'negligible'.

THE OBSERVATION EQUATION

$$d\beta_t = (z_t)\cdot dt + \sigma\cdot d\eta_t \tag{9}$$

$\sigma > 0$, η_t is a Simple Brownian Motion i.e., if $t_1, t_2, t_3, ...$ is any sequence of times, then

[3] We shall use the standard notation, $x(\Delta t) = o(\Delta t)$ if $\lim_{\Delta t \to 0} \frac{x(\Delta t)}{\Delta t} = 0$

(i) the increments $\eta_{t_{i+1}} - \eta_{t_i}$ are independent.

(ii) $\eta_{t_{i+1}} - \eta_{t_i}$ is distributed Normal with mean zero and variance $t_{i+1} - t_i$.

The Observation Equation falls under the class of Ito Processes which we define below.

Definition 2 *(Ito Process)* [4] *Let $(\Omega, \mathcal{F}, P)$ be a complete probability space, with $(\mathcal{F}_t, t \in T)$ being a right continuous filtration defined on it. Let $(\eta_t, \mathcal{F}_t, \in T)$ be a Brownian motion process. The continuous random process $(X_t, \mathcal{F}_t, t \in T)$ is called an Ito Process (relative to the Brownian motion process $(W_t, \mathcal{F}_t, t \in T)$ if there exist two nonanticipative $\mathcal{F}_t$-measurable random processes $a_t(\omega)$ and $b_t(\omega)$ satisfying for each $t \in T$*

$$\int_0^t |a_s(\omega)| ds < \infty \qquad (a.s.) \tag{10}$$

$$\int_0^t |b_s(\omega)|^2 ds < \infty \qquad (a.s.) \tag{11}$$

with $b_t(\omega)$ being left continuous, and if, with probability 1, $X_t(\omega)$ satisfies the integral equation

$$X_t(\omega) = X_0(\omega) + \int_0^t a_s(\omega) ds + \int_0^t b_s(\omega) d\eta_s \qquad t \in T$$

or equivalently its stochastic differential equation (S.D.E)representation as

$$dX_t(\omega) = a_t(\omega) dt + b_t(\omega) d\eta_t, \qquad t \in T \tag{12}$$

Since z_t switches by Poisson increments, its paths are piece- wise continuous and 10 is satisfied. Also, for the Observation Equation 11 is trivially satisfied (b_t is a non-deterministic and constant process) and hence the integral form of the Observation Equation is an Ito Process.

THE PROBLEM Let $\mathcal{F}_t$ be the σ-field generated by the sample path $(\beta_\tau)_{0 \leq \tau \leq t}$. We are interested in characterizing the infinitesimal dynamics of $\pi_t = Prob[z_t = a \ \mathcal{F}_t]$.

[4] Definition 6.2.1 in Krishnan[1984]

Theorem 1 *π_t is the solution of the Ito stochastic differential equation*

$$\begin{aligned} d\pi_t &= (1-2\cdot\pi_t)\cdot\lambda dt \\ &+\frac{\pi_t\cdot(1-\pi_t)\cdot(a-b)}{\sigma^2}\cdot[d\beta_t-(a\cdot\pi_t+b\cdot(1-\pi_t))\cdot dt] \end{aligned} \tag{13}$$

To prove the Theorem we shall find $\lim_{\Delta t\to 0}[\pi_{t+\Delta t}-\pi_t]$. The analytical expression for $[\pi_{t+\Delta t}-\pi_t]$ is found in lemmas 1 to 3. We state the lemmas below but leave the proofs for the appendix. Since both the Normal and Poisson distributions satisfy a *temporal homogeneity* property (Chung and Williams[1990]) , the analytical expression does not depend on the length Δt. The limiting expression gives the dynamics of the inference process when *no information* is received between t and $t+\Delta t$, and Δt tends to zero.

Lemma 1 *Let L_a, L_b be the densities of observable increments $\Delta\beta_t=\beta_{t+\Delta t}-\beta_t$ conditional on the drift of the output being a, b respectively. Then ,*

$$L_a = L_b\cdot[1+\frac{(a-b)\cdot(\Delta\beta_t-b\cdot\Delta t)}{\sigma^2}]+o(\Delta t)$$

Lemma 2 $\pi_{t+\Delta t}=$

$$\frac{\pi_t\cdot(1-\lambda\Delta t)\cdot L_a+(1-\pi_t)\cdot\lambda\Delta t\cdot L_b}{\pi_t\cdot(1-\lambda\Delta t)\cdot L_a+(1-\pi_t)\cdot\lambda\Delta t\cdot L_b+\pi_t\cdot\lambda\Delta t\cdot L_a+(1-\pi_t)\cdot(1-\lambda\Delta t)\cdot L_b}$$
$$+o(\Delta t)$$

Lemma 3

$$\begin{aligned} \pi_{t+\Delta t}-\pi_t &= (1-2\cdot\pi_t)\cdot\lambda\Delta t \quad + \\ \frac{\pi_t\cdot(1-\pi_t)\cdot(a-b)}{\sigma^2} &\cdot[\Delta\beta_t-(a\cdot\pi_t+b\cdot(1-\pi_t))\cdot\Delta t]+o(\Delta t) \end{aligned}$$

Lemma 1 calculates the difference of the densities of the Observation Process arising from the two possible regimes. Lemma 2 is mostly an expression of Bayes Law, with a tiny twist explained in its proof. Lemma 3 simply completes the algebra.

REMARKS ON THEOREM 1 The Ito S.D.E. is the sum of two components. The first $(1 - 2 \cdot \pi) \cdot \lambda dt$ we call a *mean-reverting* component. It is an adjustment to accommodate for the constant unobserved Poisson switching probability. The second is a product of two terms. $\frac{\pi_t \cdot (1-\pi_t) \cdot (a-b)}{\sigma^2}$ which we call the *information weighting* term and $[d\beta_t - (a \cdot \pi_t + b \cdot (1 - \pi_t)) \cdot dt] \equiv d\nu_t$ the *innovations* component, because it contains new information. We notice that $(a \cdot \pi_t + b \cdot (1 - \pi_t)) = E^{\mathcal{F}_t} z_t(\omega)$. Result 1, below, from filtering theory is used to show that the process ν_t is a Brownian Motion with respect to $(\mathcal{F}_t)$ and that the S.D.E. 13 can be written in the form

$$d\pi_t(\omega) = \mu_t(\pi_t(\omega))dt + \sigma_t(\pi_t(\omega))d\nu_t(\omega) \tag{14}$$

In this form the infinitesimal coefficients of the S.D.E. are functions of π_t only. Theorem 1 only states that the updating process satisfies the S.D.E. 13. It will be useful in defining the updating process as a diffusion process which arises as a solution to a S.D.E.. Result 2 below provides conditions on the coefficients of the S.D.E. (14) to have a unique solution. Result 3 states that under the same conditions of the coefficients the coefficients of the S.D.E. 14 are the *infinitesimal mean and variance* of the process. We will then use results from theory for diffusion processes to provide several interesting properties of the updating process in Subsection 3.

Result 1 [5] *Let $\{X_t, \mathcal{B}_t, t \in T\}$ be an Ito process defined on a complete probability space $(\Omega, \mathcal{F}, P)$ represented by the S.D.E.*

$$dX_t = a_t(\omega)dt + dW_t(\omega) \qquad t \in T$$

where

$$\int_{t\in T} E|a_t(\omega)|dt < \infty$$

with $\mathcal{B}_t$ being the σ-field generated by $(a_s, W_s, s \leq t, t \in T)$. Let $\{\mathcal{F}_t, t \in T\}$ be the right continuous σ-field generated by $\{X_s, s \leq t, t \in T\}$ with $\mathcal{F}_t \subset \mathcal{B}_t \subset \mathcal{F}$, and define

[5] Theorem 8.1.1 Krishnan[1984]

a functional

$$\alpha_t(X_t(\omega)) = E^{\mathcal{F}_t} a_t(\omega)$$

Then the innovations process ζ_t *given by*

$$d\zeta_t = dX_t - \alpha_t(X_t)dt \qquad t \in T \tag{15}$$

is an $\mathcal{F}_t$*-measurable Brownian motion process, and the Ito process* (X_t) *has a S.D.E. representaion,*

$$dX_t = \alpha_t(X_t)dt + d\zeta_t \qquad t \in T$$

As discussed earlier, the process (9) is an Ito process of the form (12) and that $(a \cdot \pi_t + b \cdot (1 - \pi_t)) = E^{\mathcal{F}_t} z_t(\omega)$. Hence, $[d\beta_t - (a \cdot \pi_t + b \cdot (1 - \pi_t)) \cdot dt] \equiv d\nu_t$ is an "innovations" process in the sense of Result 1 and the process (13) has the representation

$$d\pi_t = (1 - 2 \cdot \pi_t) \cdot \lambda dt + \frac{\pi_t \cdot (1 - \pi_t) \cdot (a - b)}{\sigma^2} \cdot d\nu_t \tag{16}$$

To ascertain that a solution to the S.D.E (16) exists and that the solution is a diffusion process (defined below) we shall use a result from Karlin and Taylor[1982].

Definition 3 [6] *A continuous time parameter stochastic process which possesses the (strong) Markov property and for which the sample paths* X_t *are (a.s.) continuous functions of* t *is called a diffusion process.*

Definition 4 *(Growth Condition) The coefficients* $\mu(\pi, t)$ *and* $\sigma(\pi, t)$ *satisfy the growth condition if there exists a constant* K *independent of* t *and* π *such that*

$$\mu^2(\pi, t) + \sigma^2(\pi, t) \leq K(1 + \pi^2)$$

[6] from Karlin and Taylor[1982], Chapter 15

Definition 5 *(Lipschitz Condition) The coefficients $\mu(x,t)$ and $\sigma(x,t)$ satisfy the Lipschitz Condition if there exists a constant L independent of t and x such that*

$$|\mu(x,t) - \mu(y,t)| + |\sigma(x,t) - \sigma(y,t)| \leq L|x-y|$$

Result 2 [7] *Let $\mu(x,t)$ and $\sigma(x,t)$ satisfy the growth and Lipshitz conditions defined above. Then there exists a unique solution to the S.D.E. (14) as a continuous process.*

It is easy to check that the coefficients of (16) satisfy the growth and Lipshitz conditions and therefore the conclusions of Result 2 hold.

Result 3 [8] *Let π_t be the diffusion that arises as a solution of the S.D.E. (14), where the coefficients $\mu(x,t)$ and $\sigma(x,t)$ satisfy the growth and smoothness conditions. Then,*

$$\lim_{h \to 0} \frac{1}{h} E[\pi_{t+h} - \pi_t | \pi_t = x] = \mu(x,t)$$

$$\lim_{h \to 0} \frac{1}{h} E[\{\pi_{t+h} - \pi_t\}^2 | \pi_t = x] = \sigma^2(x,t)$$

Result 3 suggests the names 'infinitesimal mean' and 'infinitesimal variance' for the coefficients $\mu(\pi,t)$ and $\sigma^2(\pi,t)$ of the S.D.E. (16)

4.1 Properties Of The Updating Process

We shall briefly introduce some notation and terminology standard to the literature of diffusion processes[9]. We then focus our attention on three properties of the updating process. Finally we look at a numerical example which has been of interest in formulating a model of Business Cycles in David[1992b].

[7] Theorem 16.5, Chapter 15 Karlin and Taylor[1982]

[8] Theorem 16.6, Chapter 15, Karlin and Taylor[1982]

[9] This subsection relies heavily on the analysis in Chapter 15, Karlin and Taylor[1982]

4.1.1 Terminology

Let $\mu(x,t)$ and $\sigma(x,t)$ be the coefficients of a S.D.E. as in (14). Define,

$$s(x) = exp\{-\int^{x}[2\frac{\mu(\xi)}{\sigma(\xi)}]d\xi\}$$

$$S(x) = \int^{x} s(\eta)d\eta = \int^{x} exp\{-\int^{\eta}[2\frac{\mu(\xi)}{\sigma(\xi)}]d\xi\}d\eta$$

$$m(x) = \frac{1}{s(x)\cdot\sigma^2(x)}$$

HEURISTICS: From a classical viewpoint $1/s(x)$ is an integrating factor for the differential operator L defined by

$$Lf(x) = \mu(x)f'(x) + \frac{1}{2}\sigma^2(x)f''(x)$$

A modern view of the function $s(x)$ is the following. Let l and r be the left and right boundaries of the diffusion process X_t. Let $u(x) = Prob(T_l < T_r|X_0 = x)$, where T_a is the hitting time to a. $u(x)$ is the probability that the process hits l before hitting r, starting at x. It can be shown that $u(x) = \frac{S(x)-S(a)}{S(b)-S(a)}$. So the function $S(x)$ can be used to rescale the state space (l,r) in terms of probabilities of achieving various levels and is hence named the *scale function*. We note that the process $Y_t = S(X_t)$ has a linear scale function in that

$$Prob\{T_a(Y) < T_b(Y)|Y_0 = y\} = \frac{b-y}{b-a}$$

i.e. its hitting probabilities are proportional to actual distances. The modern and classical views are reconciled when we realize that $u(x)$ satisfies the differential equation $Lu(x) = 0, u(a) = 0, u(b) = 1$.

The function $m(x)$ is called the *speed-density* of the process. The name is motivated by the fact that

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon^2}E[T_{x-\epsilon,x+\epsilon}|X_0 = x] = m(x)$$

where $T_{a,b} = min\{T_a, T_b\}$. The function $m(x)$ can also be thought of as a measure of 'volatility' of the process at x. The functions introduced here will be used to classify the diffusion process.

4.1.2 Boundary Classification

An entrance boundary is one that cannot be reached from the interior of the state space. It is possible to consider processes that start there. Such processes quickly move to the interior never to return to the entrance boundary.

Let $S[a,b]$ and $M[a,b]$ be the Stieltjes measures induced on the state space by the functions $s(x)$ and $m(x)$ respectively, for example $S[a,b] = \int_a^b s(x)dx$. Let $N(l) = \int_l^x S[\eta, x]dM(\eta) = \int_l^x M(l,\xi]dS(\xi)$. $N(l)$ roughly measures the time it takes to reach an interior point x in (l,r) starting at the boundary l. To show that a boundary l is entrance it suffices to establish [10] that $S(l,x] = \infty$ while $N(l) < \infty$ (please note that it is sufficient to establish this for any point in the interior of the state space and so the argument x is suppressed in the definition of $N(l)$.

Property 1 *(Entrance Boundary) 0 and 1 are entrance boundaries of the S.D.E.* π_t *as defined in (14)*

We recall for the updating process π_t, $\mu(\pi_t, t) = (1 - 2\pi_t) \cdot \lambda$ and that $\sigma(\pi_t, t) = \frac{(\pi_t)\cdot(1-\pi_t)\cdot(a-b)}{\sigma}$. Before proving the statement we argue that these parameters suggest that 0 and 1 are entrance boundaries. The infinitesimal mean of the process is of non-negligible size and pulls the process towards the center as the process moves close to its boundaries, while the infinitsimal variance declines to zero as the process approaches either boundary. So the process never hits either boundary.

Proof We shall prove that $S(0,x] = \infty$ and $N(0) < \infty$ The proof for the other boundary i.e. 1 is similar,because all the functions considered are symmetric about

[10] Page 234 Karlin and Taylor[1982]

$\frac{1}{2}$.

$$
\begin{aligned}
s(x) &= exp\{-\int^{x}[2\frac{\mu(\xi)}{\sigma^2(\xi)}]d\xi\} \\
&= exp\{\frac{\lambda\cdot\sigma^2}{a-b}\cdot\frac{1}{x(1-x)}\}
\end{aligned} \tag{17}
$$

Since, $\int_0^x \frac{1}{\xi(1-\xi)} = \infty$, $S(0,x] = \infty$.

$$
\begin{aligned}
N(0) &= \int_0^x\left(\int_\xi^x exp[\frac{K}{\eta(1-\eta)}d\eta\right)\frac{1}{J\cdot\xi^2(1-\xi)^2\cdot exp[\frac{K}{\xi(1-\xi)}]}d\xi \\
&\le \int_0^x\left(\int_\xi^x \frac{exp[\frac{K}{\eta(1-\eta)}]}{exp[\frac{K}{x(1-x)}]}d\eta\right)\frac{1}{J\cdot\xi^2(1-\xi)^2}d\xi \\
&\le \int_0^x \frac{x-\xi}{J\cdot\xi^2(1-\xi)^2} < \infty\Box
\end{aligned}
$$

4.1.3 Stationary Distribution

If it exists, a stationary density $\psi(x)$ of a diffusion process X_t necessarily satisfies

$$
\psi(y) = \int \psi(x)\cdot p(t,x,y)dx \qquad \forall t > 0
$$

where $p(t,x,y)$ is the transition density function i.e., $P(t,x,y) = Prob(X_t \le y|X_0 = x)$ and $\frac{dP(t,x,y)}{dy} = p(t,x,y)$. It can be shown that the stationary density $\psi(x)$ satisfies,

$$
\frac{1}{2}\frac{\partial^2}{\partial y^2}[\sigma^2(y)\psi(y)] - \frac{\partial}{\partial y}[\mu(y)\psi(y)]
$$

Solving the differential equation yields,

$$
\psi(x) = m(x)[C_1S(x) + C_2]
$$

where C_1 and C_2 are constants that guarantee that $\psi(x) \ge 0$ on (l,r) and $\int_l^r \psi(\xi)d\xi = 1$

When l and r are both entrance boundaries (0 and 1 are both entrance boundaries for π_t), then $S(x) \to \infty$ as $x \to l$ or $x \to r$, as discussed in the previous sub-subsection. In this case, C_2 is chosen equal to zero and the stationary density is

$$\psi(x) = \frac{m(x)}{\int_l^r m(\xi)d\xi} \tag{18}$$

where $m(x) = \frac{1}{[\sigma^2(x)s(x)]}$ as defined earlier.

Property 2 *(Stationary Distribution) For the updating process π_t the stationary distribution $\psi(x)$ is*

$$\psi(x) = C_1 \cdot exp\{\frac{-2\lambda\sigma^2}{(a-b)^2} \cdot \frac{1}{x(1-x)}\} \cdot \frac{\sigma^2}{(a-b)^2 x^2 (1-x)^2} \tag{19}$$

Proof Substitute $s(x)$ from (17) into (18). $\Box$

The shape of the stationary distribution depends on the characteristics of the learning process which depends on the parameters λ, σ^2 and $(a-b)^2$. When λ is large a 'lot' of switching occurs and so a relatively large amount of time is spent with π close to a $\frac{1}{2}$. When σ^2 is large the signal are noiser and again π spends a larger time around a $\frac{1}{2}$. When $(a-b)^2$ is large the diferrence in the drifts from the two regimes is large, so learning occurs faster and a relatively large amount of time is spent near the boundaries.

4.1.4 Path Properties

So far we have characterized the infinitesimal dynamics of the learning process in Theorem 1 and the very long-term behaviour of the process as measured by the stationary distribution in the previous subsubsection. For some problems in decision theory it is interesting to have estimates of the 'intermediate-term'dynamics of the sample paths of the process. In this sub-subsection, we estimate the expected time to be spent in different 'regions' of the state space conditional on being in the region.

For example it might be important to calculate how long the agent expects to be in a region of 'low-confidence' which may be defined as $\pi_t \in (.4, .6)$ or the length of time the agent expects to be 'confident' in a regime as measured by the time spent in $\pi_t \in (.8, 1)$.

Define $v(x) = E[T_{a,b}|X_0 = x]$ $x \in (a, b)$. $v(x)$ is the length of time the agent expects to be in (a, b) conditional on being at x. It can be shown [11] that $v(x)$ is the solution of the differential equation

$$Lv(x) = -1 \qquad v(a) = v(b) = 0$$

The solution to this is

$v(x) =$

$$2\{u(x, a, b) \cdot \int_x^b [S(b) - S(\xi)]m(\xi)d\xi + [1 - u(x, a, b)] \int_a^x [S(\xi) - S(a)]m(\xi)d\xi\} \quad (20)$$

where,

$$u(x, a, b) = Prob\{T_b < T_a | X_0 = x\}$$

We give a numerical example to illustrate these properties next.

4.1.5 Example

Time is measured in 'years'. $\lambda = 2$ which implies that there are approximately two switches in a year and a 9 percent chance of having a switch. $a = .07$ and $b = -.045$ and $\sigma = .02$ The stationary distribution is almost U-shaped and is drawn in Figure 1. With these parameter values the process spends a relatively large proportion of times at the ends rather than in the middle. When $\pi_t = .5$, the expected time to be spent in the interval $(.3, .7) = .09$ years, roughly a month. Once the process hits the level .9, the expected time spent before it goes back to a level of .6 is .47 years. The process moves out of regions faster when in the middle. This represents fairly

[11] page 192 Karlin and Taylor[1982]

quick learning when the level of confidence reaches a low level. An agent in such an environment knows that when he reaches a level of middling beliefs, he expects to receive news relatively soon and move fairly quickly to a level of higher confidence. Once he reaches a level of high confidence in one of the regimes, he expects to retain approximately the same beliefs for a relatively long period. In David[1992b], we find that the environment described here, causes cyclical investment patterns, if there is some inflexibility regarding the agents investment choices. When all parameters are kept at these values, except that the level of noise is raised to .07, learning is slow and the stationary distribution has most of its mass around .5. This is shown in Figure 2.

5 Comparison With The Kalman Filter

In this section we briefly discuss some similarities and differences in the assumptions and results between the Kalman Filtering problem, Jazwinski[1970] and the filter introduced in this paper. We will not write the Kalman filtering problem in its most general form, but instead concentrate on a few broad features which are apparent from the simplified version discussed here.

The Kalman Filtering Problem As in the problem considered in Section 1, the problem is described by the Transition and Observation equations.

THE TRANSITION EQUATION

$$dz_t = fz_t + qd\zeta_t \tag{21}$$

$f > 0$, $q > 0$ and ζ_t is a Standard Brownian Motion. The unobserved state variable z_t follows a diffusion process. This is different from the transition described in (1), where the state variable could take on only two values and switched between the two with Poisson increments. The distribution of z at time zero is assumed Gaussian with mean $\hat{z}_0$ and variance v_0.

THE OBSERVATION EQUATION

$$d\beta_t = z_t dt + \sigma d\eta_t \tag{22}$$

$\sigma > 0$ and η_t is a Simple Brownian Motion. This is identical to (2), the Observation Equation for the filtering problem considered in this paper.

Let $\hat{z}_t = E^{\mathcal{F}_t}[z_t]$, where $\mathcal{F}_t$ is the σ-field generated by the sample path $(\beta_\tau)_{0\leq\tau\leq t}$. Let $v_t = Var^{\mathcal{F}_t}[z_t]$.

Result 4 *(Results of the Kalman Filtering Problem)*

$$
\begin{aligned}
d\hat{z}_t &= f \cdot \hat{z}_t dt + \frac{b\sigma + v_t}{\sigma^2}(d\beta_t - \hat{z}dt) && (23)\\
\frac{dv_t}{dt} &= (2fv_t + q^2) - (\frac{b\sigma + v_t}{\sigma})^2 && (24)
\end{aligned}
$$

Remarks UPDATING THE CONDITIONAL MEAN The dynamics (23), of the conditional mean are analagous to the dynamics of $d\pi$ described by (13). Notice that (23) may be decomposed into the sum of two parts. The first is a deterministic drift component and it equals $E^{\mathcal{F}_t}[dz_t]$. This is the counterpart of the 'mean-reverting' component of (13) which equalled $E^{\mathcal{F}_t}[dz_t]$ in that context.

The second component is the product of two terms. $\frac{1}{\sigma}(d\beta_t - \hat{z}dt)$ is an *innovations* process as defined in (15), and is a $\mathcal{F}_t$-measurable Brownian Motion by Result 1. We recall that the updating rule (13) had a similar 'innovations' term.

$\frac{b\sigma+v_t}{\sigma}$ is the weight give to the 'innovation' term in updating. This matches the 'information weighting' term in (13) which was $\frac{\pi_t\cdot(1-\pi_t)\cdot(a-b)}{\sigma}$. It is easy to show in both cases that the information weighting term equals $\frac{Cov^{\mathcal{F}_t}[d\beta_t, dz_t]}{Var^{\mathcal{F}_t}[d\beta_t]}$. The correspondence of the conditional mean dynamics is complete. This identification of the information weighting term also shows that the estimator π, which we obtained from Bayesian Updating, coincides with the optimal *Least Squares* estimator.

UPDATING THE CONDITIONAL VARIANCE The dynamics of the conditional variance are deterministic and this is reflected in the fact that (24) is an Ordinary Differential Equation. In particular, the path taken by the conditional variance is completely determined by the parameters of the problem and can be pre-computed. The solution of (24) is[12]

$$
\begin{aligned}
v_t &= \frac{v_0\sigma^2}{v_0 t + \sigma^2} && if \quad w = 0\\
v_t &= 2w[1 - \frac{1}{1 + \frac{1}{2w - v_0} v_0 \exp(\frac{2wt}{\sigma^2})}] && if \quad w \neq 0 \qquad (25)
\end{aligned}
$$

[12]See for example Gennotte[1986].

where $w = f\sigma^2 - g\sigma$. If $w < 0$, $\lim_{t\to\infty} v_t = 0$. If $w > 0$, $\lim_{t\to\infty} v_t = 2w$. Intuitively, if the ratio of the speed of the drift to the strength of the signal as in the case $w > 0$, then the agent is always a step behind in his evaluation. Even asymptotically he does not get an exact estimation of the drift. The tracking problem in this paper also has this property, since the drift can always jump.

The deterministic path of the conditional variance is a property particular to the Kalman Filtering problem. The deterministic dynamics of the conditional variance for the Kalman Filtering problem is based on a well known result on Normal Distributions in statistics. From (22) $d\beta_t$ and z_t have a jointly Normal distribution, conditional on all information until until t. More precisely[13] if (X,Y) has a nondegenerate $\mathcal{N}[\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho]$ distribution, then the conditional distribution of X given $Y = y$ is $\mathcal{N}[\mu_1 + \rho\frac{\sigma_1}{\sigma_2}(y - \mu_2), \sigma_1^2(1 - \rho^2)]$, i.e. the conditional variance of X given $Y = y$ is a constant. Our model lacks the property that makes the Kalman Filter tractable. However, since both the conditional mean and conditional variance depend only on π_t, we were able to write the updating as a diffusion process in one variable and analyze its properties.

[13]Theorem 1.4.2, Bickel and Doksum[1976]

6 Model 1 Unobservable Regime Switches

In this model I shall assume that the state variable z_t is unobserved. The agent observes the rates of return on both physical assets which are as in (1). The agent has beliefs about the underlying value of the state variable z_t. I assume that he uses Bayes Law in updating his beliefs. The agent updates his belief by observing the difference in the rates of returns from the two investments. If the difference in the rates is positive his beliefs shift in the direction of asset 1 having the higher drift and vice versa. This filtering problem has been solved in Section 4. Unobservability of the regime switches lead to various degrees in 'confidence' for the representative agent. These can also be viewed as wealth or 'prospect' changes. When the agent is confident about the current regime, he can invest in two assets with similar expected returns, approximately the average productivity of the two assets in the economy. On the other hand if he is confident of the current regime, then one asset has high expected productivity and the agent can receive higher returns by investing mostly in that asset. Furthermore the expected time to be spent in different regions of the state space changes with changes in beliefs. This implies that these changes in confidence have payoff consequences for periods larger than an 'instant'. So unlike the situation in Model 0 where the agent faced a constant opportunity set, the agent here faces an entire *spectrum* of opportunity sets. The constantly changing 'confidence' / 'prospects' introduces a class of risk, which we outline next.

Recall that the belief updating process is a diffusion process. It contains a disturbance term which has a standard deviation of the order $o((\Delta t)^{\frac{1}{2}})$. In other words uncertainty about the future position of belief is 'large' relative to the length of the interval considered. Lets consider the case where the probability of asset 1 having drift a is greater than .5. In this case asset 1 promises a higher average return. When the total return of asset 1, which is the sum of the drift and noise , is high it increases the updated probability of the regime being in favor of asset 1 When the return of asset 2 is high it lowers this probability. Effectively, asset 1 pays off in states of *greater wealth* (or lower marginal utility) and asset 2 in states of *lower wealth* (higher marginal utility) Therefore holdings in asset 2 help hedge the risk of wealth changes at this infinitesimal horizon. In particular if the the size of the infinitesimal movements is large, the agent can hedge the risk at the cost of lower expected returns. This risk-hedging behaviour is illustrated in the portfolio choices in Corollary 2.

The Control Problem Let $J[W_t, \pi_t, t]$ be the value function of the social-planning problem at time t, given the wealth and beliefs of the agent. The dynamics of the beliefs are given by (13) in Section 4. The updating depends on $d\beta_t$, which depends on z_t the underlying productivity variable. (13) is not a Markov process in π. Therefore the consumer's decision problem with (13) defining the state variable dynamics is not a Markov decision problem. However the identification of the 'innovations' process as a Simple Brownian Motion on the filteration of the agent's information, in Result 1 rectifies this. A similar result has been proved for models based on the Kalman Filter such as Dothan and Feldman[1986], Gennotte[1986] and Sundaresan[1984].

Proposition 2 *(Separation of Updating and Optimization) The consumer's consumption and portfolio decision in the economy with unobservable regime-shifts would be the same as in an economy with an exogenous state variable which follows a diffusion process (in particular Markov).*

Proof By Result 1, the statisical distribution of (16) is identical to that of (13) on the filteration of the agent's information. □.

Henceforth (16) shall be used to define the dynamics of beliefs. As for the previous model the value function is separable in wealth, beliefs and other state variables.

$$J[W_t, \pi_t, t] = \exp(-\rho t)\frac{W_t^\gamma}{\gamma} \cdot I[\pi_t] \tag{26}$$

Furthermore the optimal policy is independent of time and wealth and depend only on beliefs.

Proposition 3 *When regime-switches are unobserved $I[\pi]$ satisfies the differential equation*

$$0 = \max_{\sum_{i=1}^{2} w_i = 1, w_i \geq 0} \left(\frac{1}{\gamma} - 1\right)(I[\pi])^{\frac{\gamma}{\gamma-1}} - \rho \cdot I[\pi] +$$

$$I[\pi]\left(\sum_{i=1}^{2} \alpha_i(\pi) w_i + \frac{1}{2}(\gamma - 1)\left(\sum_{i=1}^{2} w_i^2\right)\sigma^2 \right) +$$

$$I_\pi[\pi]\left(\frac{\mu(\pi)}{\gamma} + (w_1 - w_2) \cdot (\pi) \cdot (1 - \pi) \cdot (a - b) \right) +$$

$$\frac{1}{2}\sigma^2(\pi) I_{\pi\pi}(\pi) \tag{27}$$

The proof is in Appendix 2. The transversality condition for this model has not been explicitly discussed. If parameter values are chosen so that the condition is satisfied for Model 0, then it will be satisfied for this model. This is because for any strategy, the value for the unobservable case cannot be larger than for the observable case.

Boundary Conditions. Its evident that the value function is symmetric about .5. Therefore it reaches a local minimum at .5 and we impose $I'[.5] = 0$. In Section 4 we showed in Property 1 that the belief process did not hit the boundaries 0 and 1 although it got arbitraily close to them. As π approaches either boundary it is pulled inward with probability 1. I therefore use the Reflecting Boundary condition $I'[1] = 0$.

As in Model 0 the portfolio choices are partly determined by the desire to hedge noise. In this model the agent also attempts to hedge the risk due to wealth changes. This lowers his demand for asset 1 and raises it for assets 2.

Corollary 2 *When regime-switches are unobserved portfolio choices are made to hedge the risk associated with 'fluctuating-confidence'. In the case of interior solutions the portfolio choices are determined by*

$$w_1(\pi) = \frac{1}{2} + \frac{1}{2} \cdot \frac{\alpha_1(\pi) - \alpha_2(\pi)}{\sigma^2(1-\gamma)} + \frac{(\pi) \cdot (1-\pi) \cdot (a-b)}{2 \cdot (1-\gamma) \cdot \sigma^2} \cdot \frac{I_\pi[\pi]}{I[\pi]}$$

$$w_2(\pi) = \frac{1}{2} + \frac{1}{2} \cdot \frac{\alpha_2(\pi) - \alpha_1(\pi)}{\sigma^2(1-\gamma)} - \frac{(\pi) \cdot (1-\pi) \cdot (a-b)}{2 \cdot (1-\gamma) \cdot \sigma^2} \cdot \frac{I_\pi[\pi]}{I[\pi]} \tag{28}$$

The excess-returns vary with the 'level-of-confidence' of the agent(s). and are *given by*

$$er[\pi] = -(w_1 - w_2)(\pi)(1-\pi)(a-b)\frac{I'[\pi]}{I[\pi]} + (1-\gamma)\sigma^2 \sum_{i=1}^{2} w_i^2 \tag{29}$$

The proof is in Appendix 2.

Just as in the observable case the agent diversifies the noise in the payoff from each asset. However, in this model the opportunity cost of diversification, the difference in the drifts depends on the beliefs of the agents varies with the agent's beliefs. The 2nd term in the portfolio choices reflects the noise diversification. Further, the agent hedges the risk of the changing opportunity set as described in the introduction of this section. The last term in the first portfolio choice equation measures the reduction in holdings of this asset, because its payoff has positive comovements with growth 'prospects'.

The characterization of the risk-free rate in this model, the expected returns minus the excess-returns from Corollary 2, is formally identical to that in the Consumption-Based-Asset Pricing literature, for example Lucas[1978]. This follows from Theorem 1 CIR[1985], which only depends on the state variable following a diffusion process. I provide a proof which is similar to theirs, but is specialized to this model. It extends the CIR result to the case of unobservable productivity parameters. Let $MU[\pi]$ denote the marginal-utility of consumption, which by the envelope condtion equals the marginal-utility of optimally invested wealth. Let $rf[\pi]$ be the risk-free rate when the agents beliefs are π. This is as calculated in Corollary 2. Then

Proposition 4

$$rf[\pi] = -\frac{1}{\Delta t} \cdot \frac{\Delta MU[\pi]}{MU[\pi]}.$$

The proof is in Appendix 2. From Corollary 2, the risk-free rate equals the best net-returns in value terms from any asset in the economy. The CCAPM literature prices a risk-free asset in zero net-supply so that its equilibrium rate equals minus the expected rate of change of the marginal rate of substitution. Propostion 4 shows that at an optimum these two quantities are the same.

A VERY ROUGH UPPER BOUND FOR THE EXCESS RETURNS IN THIS MODEL. A very rough bound on the size of $(\pi)(1-\pi)(a-b)\frac{I'[\pi]}{I[\pi]}$ can be obtained from the solution of Model 0. Let $\overline{A}$ be the value when π is always .5 and $\underline{A}$ be the value when π is always 1. These numbers can be calculated from (44). Then the above quantity is less than

$$.25 \cdot (a-b) \cdot \frac{\overline{A} - \underline{A}}{2 \cdot \underline{A}} \tag{30}$$

This puts a bound on the excess returns arising out of unobservability. For example if $a = .07$, $b = -.05$, $\gamma = -1$, $\sigma = .03$, then the bound equals 0.015 or 1.5 %.

7 Fluctuating Confidence and Properties of Stock Market Returns

I state three stylized facts about stock-market data. The facts are well documented in the literature. Here I do not cite all the papers but direct the reader to Black[1976] for a description of the facts and to Campbell and Hentschel[1992] and Nelson[1991] for formal time-series models which measure these effects. To evaluate the capability of Model 1 these features, I solve the model numerically and then simulate sample paths of real and financial variables. Equation (27), the value function for the consumer's problem is solved using an implicit finite-difference method[14].

I show that with appropriate parameter choices, the stock-market return process of Model 1 is capable of replicating these features, while Model 0 the case with observable regime-switches is not. Illustrative pictures and informal explanations for the success of the model are provided. I also discuss to what extent models based on Kalman Filter learning models might replicate these facts. I do not argue that the model presented here is the only one capable of replicating these features. The idea is to examine the extent to which learning about the relative productivities of different assets in the economy affects the qualitative properties of equlibrium returns.

THREE STYLIZED FACTS

(i) Kurtosis or Fat-Tails of Excess-Returns. Large realizations of returns happen more often than consistent with normality. Using stock-returns monthly data from CRSP (1/26 - 12/88) and T.Bill Data from Ibbotson Associates, Campbell and Hentschel[1992] report Excess-kurtosis of 6.82 in excess-returns.

(ii) Skewness of Excess-Returns. Large negative returns are more common than large positive ones. Campbell and Hentschel report a skewness parameter of -0.443 for the above series.

[14]The program is written in Gauss. It is available upon request from the author.

(*iii*) The Predictive Asymmetry of Stock-Market Volatility. Black[1976], Nelson[1991] and Campbell and Hentschel[1992] find a negative correlation between current returns and future returns volatility. Further, reactions to unfavorable news tend to be larger than reactions to favorable events.

I start by showing that Model 0, the case of observable regime switches is inconsistent with these stylized facts.

Proposition 5 *The conditional distribution of returns is unchanging in Model 0. The unconditional distribution of returns over any horizon (may be non-infinitesimal) is Gaussian.*

Proof. In Proposition 1 in Section 3 I showed that the fractions of resources invested in the high and low productivity is constant over time and independent of z the productivity variable and wealth. Therefore the statistical distribution of rate of return to optimally invested wealth is unchanging. The rate of return over horizon Δt therefore has a normal distribution, with a mean $(w_1 \cdot a + w_2 \cdot b) \cdot \Delta t$ and variance $(w_1^2 + w_2^2) \cdot \sigma \cdot \Delta t$, where w_1 and w_2 are given by (6). $\square$.

The success of Model 1 in generating these stylized facts depends on the parameter chosen. I distinguish between parameters which lead to a U-shaped stationary distribution of beliefs from one which lead to an inverse U-shape. The former is called the Fast and the latter Slow learning model. Numerical results from the two models are presented in Figures 3.1 - 3.8 and 4.1 - 4.8 respectively which I discuss now.

Figures 3.1 and 4.1 show the empirical densities of beliefs for the two models respectively. The shapes are the same as the theoretical shapes discussed in Section 4. Figures 3.2 and 4.2 show the conditional variance of beliefs. The conditional variance is given by $\frac{\pi^2 \cdot (1-\pi)^2 \cdot (a-b)^2}{\sigma^2}$. The shapes of the two curves are the same, though the absolute values differ. Figures 3.3 and 4.3 show the percentage of the portfolio allocated to the high productivity asset. The portfolio choices are as given in (28) in Section 6. I discussed the motives to diversify in that section. The figures show that the agent plunges less readily in the Slow Learning model. This reflects both the greater level of noise and the larger variance of the belief process. Figures 3.4 and 4.4 give the risk-free rate in the two situtions as calculated in Section 6. The interest rate in both models is lowest when confidence is lowest, i.e. $\pi = .5$, This reflects because in this state beliefs are most volatile. This leads to a large volatility of the consumption process and by Proposition 4 a low risk-free rate. Another way of interpreting this is that the risk-free rate is a measure of the opportunity cost of lending. Notice that the agent completely diversifies under these conditions and effectively faces a market with poor prospects.

Figures 3.5 and 4.5 show the conditional excess-returns in the two models. Figures 3.6 and 4.6 show the expected market returns. The market returns reflect the fact that when the agent is confident of the current regime and therefore allocates his asset in the high-productivity asset, he recieves high expected returns. In the Slow learning case the agent waits to get more confident before he plunges and correspondingly the expected rate of returns rises at a slower rate than in the Fast learning case. The excess returns reflect that the risk in the market portfolio does essentially increases as the agent becomes more confident. The relationship is not monotonic however. As beliefs approach 0 or 1, the agent has already plunged and taken the maximum risk with respect to noise. However, because the volatility of the belief process declines, the risk due to 'prospect' changes as discussed in Section 4 also declines. So the excess-returns due to this effect taper off as π approaches either 0 or 1.

Figures 3.7 and 4.7 show the relationship between the speed of learning and the negative skewness and excess-kurtosis of unconditional realized excess-returns. The pictures reveal that the Fast Learning model leads to negative skweness while the Slow Learning model leads to excess-kurtosis. Further results on these statistics are given in Table 1. In the Table I consider various sets of parameters and watch the effects of changing only the level of noise. A large level of noise gives slow learning. The same finding as the pictures are also found for other parameter choices. Besides the Z-values show that the statistics are statistically significant.

The relationship between the speed of learning and the potential negative skewness or excess-kurtosis of the unconditional distribution is as follows. The returns from the two assets over any given time horizon are given by normal distributions with different means. In the slow learning case, the agent spends a large amount of times in regimes with low confidence. Because his beliefs are rational, this means that essentially he is getting returns from the high and low mean asset with about equal probability. Therefore the distribution of returns is obtained from normal distributions with different means[15] The resulting distribution has fatter tails than a Normal density. In the Fast Learning case, the agent spends a large proportion of time with high confidence regarding his knowledge of the current regime. In this case he allocates his investment to the high productivity asset. Because his beliefs are rational, he gets returns from a Normal density with a high mean most of the time and an occasional small return from the smaller mean density. This makes his distribution negatively skewed. In Figure 5, the approximate modal positions in the Fast and Slow Learning cases is depicted.

Figures 3.8 and 4.8 show the conditional variance of realized returns for different levels of the belief variable. Again the pattern for the Slow and Fast learning cases are different. For the Fast learning model the conditional variance of returns is highest when the agent is least confident or $\pi = .5$. For the Slow learning model the opposite is true. The reason this happens is due to the relative importance of noise and different drifts as well as the portfolio choices in the two models. Over any horizon, the variation in returns is caused by two different sources of fluctuations. (1) The difference in the drifts of the assets and (2) the amount of unhedged noise. In both models when confidence is low, the agent choses diversified portfolios. Therefore the noise in the assets is hedged against. Returns are generated with close to equal

[15] If the means are very far apart, the unconditional distribution will be bi-modal. For the parameter choices which I made the distribution has a single mode.

probabilites with either drift a or drift b. The variance due to (2) is relatively large. When confidence is high, the returns are generated by the same drift with a high probability. Therefore the variation due to (2) is low. Most of the variation in returns is caused by (1), because noise is unhedged in high confidence states. In the Slow learning model, the level of noise is large relative to the difference in drifts. Therefore variance of the realized returns is smallest at $\pi = .5$ and largest when π is near 0 or 1. In the Fast learning model the opposite holds.

The connection to stylized fact (*iii*) is as follows. In the Fast learning model, the agent is mostly in a state of high confidence and low conditional variance of returns. A large negative realized returns leads to a a loss of confidence and his beliefs move closer to .5 where realized returns exhibit higher conditional variance. There is therefore a negative correlation between realzied returns and future changes in volatility. Further a large positive returns, merely confirms the agents' belief regarding the underlying regime and leads to small revisions. Therefore negative returns have lager effects in absolute value. Things are not so clear for the Slow learning case. Here the conditional variance of realized returns is high when the agent is confident. Therefore the relationship between realized returns and future changes in volatility is in contradiction with the stylized fact for the Slow learning case.

In passing I compare the success of my model to the explanation put forward by Clark[1973] regarding the kurtosis of returns. Clark showed that a process of returns which was 'Subordinated' to the Normal distribution could explain the kurtosis of returns. The subordinated process has a constant conditional mean but a varying condtional variance. The fact that large conditonal variance is followed by large conditional variance is enough to generate fat-tails. He motivated the fluctuating conditional variance by differing amount of trading volume in different periods. Large trades are motivated by differing 'news' recieved by traders as well as other idiosyncratic factors. However, the conditional mean of returns was constant. My explanation, has a clearer expostion of the 'news arrival' process, with the updating explicitly described. The results are driven by changing conditional means as well as changing conditional variance of returns both of which leads to differential portfolio choices. Further my model is also consistent with the other stylized facts discussed.

8 Conclusion

I have constructed a Cox-Ingersoll-Ross model which is capable of replicating three stylzed facts about the U.S. stock-market. The model is built around the dynamics of the beliefs of agents in the economy who are tracking unobserved regime-switches. I find that parameter values which permit faster learning are better able to replicate the observed stylized facts. The paper complements the Conditional Heteroskedasticity literature which argues that several U.S. economic time series are best described by ARCH and GARCH models and their variants. The analysis here has the added advantage of being in a General Equilibrium framework. The required persistence properties of stock-returns are inherited from the inertia in Bayesian Updating.

Appendix 0

Proof of Proposition 1.

The aggregate wealth dynamics are given by

$$W_{t+\Delta t} = [W_t - C_t\Delta t] \cdot [\sum_{i=1}^{2} w_{it}(1 + \frac{d\beta_i}{\beta_{it}})] + o(\Delta t) \tag{31}$$

$$\begin{aligned}\Delta W_t = &-C_t\Delta t + [W - C_t\Delta t] \\ &\cdot (1 - \lambda \cdot \Delta t) \cdot \sum_{i=1}^{2} w_{it}(\alpha_i(z_t)\Delta t + \sigma\Delta\zeta_{it}) \\ &+(\lambda \cdot \Delta t) \cdot \sum_{i=1}^{2} w_{it}(\alpha_i(a + b - z_t)\Delta t + \sigma\Delta\zeta_{it})] + o(\Delta t)\end{aligned} \tag{32}$$

Therefore,

$$\begin{aligned}\Delta W_t = &-C_t\Delta t \\ &+ [W - C_t\Delta t] \cdot [\sum_{i=1}^{2} w_{it}(\alpha_i(z_t)\Delta t + \sigma\Delta\zeta_{it})] + o(\Delta t)\end{aligned} \tag{33}$$

therefore.

$$\mathrm{E}_t[\Delta W_t] = -C_t\Delta t + W_t \cdot [\sum_{i=1}^{2} w_{it}\alpha_i(z_t)\Delta t] + o(\Delta t) \tag{34}$$

$$\mathrm{Var}_t[\Delta W_t] = W_t^2 \cdot [\sum_{i=1}^{2} w_{it}^2] \cdot \sigma^2\Delta t + o(\Delta t) \tag{35}$$

$$\mathrm{E}_t[(\Delta W_t)^k] = o(\Delta t) \tag{36}$$

for k $>$ 2.

$J[W, z, t]$ the derived utility of wealth function satisfies

$$0 = U[C_t] +$$

,

$$0 = \max_{s.t[\sum_{i=1}^{2} wit]=1} [U[C_t] + J_t + +J_W \cdot [-C_t + W_t[\sum_{i=1}^{n} w_{it}\alpha_i]$$

$$+\frac{W^2}{2}J_{WW}\cdot[\sum_{i=1}^{2}w_{it}^2]\sigma^2+o(\Delta t)] \tag{37}$$

The first order conditions are

$$U_C = J_W \tag{38}$$

$$J_W W\alpha_i(z_t) + J_{WW}W^2 w_{it}\sigma^2 \leq \lambda^0 \tag{39}$$

for $i = 1, 2, 3$. λ^0 is the Lagrange multiplier associated with the constraint $\sum_{i=1}^{2} w_{it} = 1$. The complementary slackness condition is

$$J_W W[\sum_{i=1}^{2}\alpha_i(z_t)w_{it}] + J_{WW}W^2[\sum_{i=1}^{2}w_{it}^2]\sigma^2 = [\sum_{i=1}^{2}w_{it}]\lambda^0 = \lambda^0 \tag{40}$$

As noted at the beginning of Section 3, the model has a *constant opportunity set* i.e. even though the productivities of activity 1 and 2 change over time, there is always one with an average productivity of a and one with an average productivity of b. This implies that the production decision may depend on the wealth at time t, but not on the value of t. Also with a power utility fuction and opportunities to scale up or down wealth proportionately that optimal decision rules are independent of wealth.

We now guess and verify that the value function is of the form

$$J[W_t, t] = \exp(-\rho t)\cdot A\cdot\frac{W^\gamma}{\gamma} \tag{41}$$

With this guess $J_t = -\rho J$, $J_W = \frac{J}{W}$ and $J_{WW} = \frac{(\gamma-1)J}{W^2}$. Subcripts of J denote partial derivatives. Substituting these into the first order conditions (39) gives us the conditions

$$\alpha_i(z_t) + (\gamma-1)w_{it}\sigma^2 \leq \frac{\lambda^0}{W^\gamma A} \tag{42}$$

and equality holding whenever $w_{it} > 0$ for $i = 1, 2, 3$ which imply that the portfolio choices w_{it} are independent of time. From now on we shall avoid the time subscript t. Let asset 1 have productivity a and asset 2 have productivity b with the understanding that the choices given below are for the case $z = a$ and that w_1 and w_2 would be reversed when $z = b$. We explicitly write the portfolio choices in the case of interior solutions.

$$\begin{aligned} w_1 &= \frac{1}{2} + \frac{1}{2} \cdot \frac{(a-b)}{\sigma^2(1-\gamma)} \\ w_2 &= \frac{1}{2} + \frac{1}{2} \cdot \frac{(b-a)}{\sigma^2(1-\gamma)} \end{aligned} \tag{43}$$

From (38) we get $C = A^{\frac{1}{\gamma-1}} W$. Substituting the optimal choices of consumption and portfolio shares and the expressions for the partial derivatives into (37) yields

$$0 = AW^{\gamma} \cdot [\frac{1-\gamma}{\gamma} A^{\frac{1}{\gamma-1}} - \frac{\rho}{\gamma} + \sum_{i=1}^{2}[w_i \alpha_i(a)] + (\gamma - 1)\sigma^2 \sum_{i=1}^{2}[w_i^2]]$$

which implies that

$$A = [\frac{\gamma}{1-\gamma}(\frac{\rho}{\gamma} + \sum_{i=1}^{2}[w_i \alpha_i(a)] + (\gamma - 1)\sigma^2 \sum_{i=1}^{2}[w_i^2])]^{\gamma-1} \tag{44}$$

The transversality condition[16] for this problem is that

$$\rho > \max\{0, \gamma[\sum_{i=1}^{2}[w_i \alpha_i(a)] + (\gamma - 1)\sigma^2 \sum_{i=1}^{2}[w_i^2]\} \qquad \square$$

Proof of Corollary 1. The first order conditions for the C, w_1 and w_2 are same as before. The condtion for w_0 is

$$J_W W r = \lambda^0 \tag{45}$$

λ^0 is the Lagrange muliplier associated with the constraint $\sum_{i=1}^{2} w_{it} = 1$. Let $J^*[W_t, t]$ be the value to the social planning problem, solved in the previous subsection λ^{*0} be the Lagrange multiplier associated with the constraint $[\sum_{i=1}^{2} w_{it}] = 1$, w_{it}^* the portfolio shares in the two industries, and C_t^* the consumption flow rate chosen by

[16]The transversality condition rules out strategies which make the value function unbounded.

the social-planner. Then we claim that $\hat{C}_t = C_t^*$, $\hat{w}_{it} = w_{it}^*$ and $\hat{r}_t = \frac{\lambda^{*0}}{W_t J_W^*}$ constitute a competitive equilibrium. Inspection of (45) reveals that $\hat{w}_{0t} = 0$ is optimal, so there is no borrowing / lending. Confirmation of the other conditions is straightforward.

Substituting the value of $\hat{r}_t$ into the complementary slackness condition and collecting terms implies

$$[\sum_{i=1}^{2} \alpha_i(z_t)\hat{w}_{it}] - \hat{r}_t] = W_t^2 J_{WW}(\sum_{i=1}^{2} \hat{w}_{it}^2)\sigma^2 \tag{46}$$

For the Power Utility function and the resulting value function the right-hand side of (46) is $(1-\gamma)(\sum_{i=1}^{2} w_{it}^2\sigma^2$. □

Appendix 1

Lemma 1 *Let L_a, L_b be the densities of observable increments $\Delta\beta_t = \beta_{t+\Delta t} - \beta_t$ conditional on the drift of the output being a, b respectively. Then ,*

$$L_a = L_b \cdot [1 + \frac{(a-b)\cdot(\Delta\beta_t - b\cdot\Delta t)}{\sigma^2}] + \quad o(\Delta t)$$

Proof

$$L_b = \frac{1}{(2\cdot\pi)^{\frac{1}{2}}} \cdot \frac{1}{(\sigma\cdot\Delta t)^{\frac{1}{2}}} \cdot exp[\frac{-1}{2\cdot\sigma^2\Delta t}\cdot(\Delta\beta_t - b\cdot\Delta t)^2]$$

$L_a =$

$$\frac{1}{(2\cdot\pi)^{\frac{1}{2}}} \cdot \frac{1}{(\sigma\cdot\Delta t)^{\frac{1}{2}}} \cdot exp[\frac{-1}{2\cdot\sigma^2\Delta t}\cdot(\Delta\beta_t - b\cdot\Delta t - (a-b)\cdot\Delta t)^2]$$

$+ \quad o(\Delta t)$

$$= L_b \cdot \exp[\frac{-1}{2\cdot\sigma^2\cdot\Delta t}\cdot[(a-b)^2\cdot(\Delta t)^2 - 2\cdot(a-b)\cdot\Delta t\cdot(\Delta\beta_t - b\cdot\Delta t)]$$

$+ \quad o(\Delta t)$

since, $e^x = 1 + x + \frac{x^2}{2!} + \ldots$

$L_a =$

$$L_b \cdot [1 + \frac{(a-b)\cdot(\Delta\beta_t - b\Delta t)}{\sigma^2} - \frac{(a-b)^2}{2\cdot\sigma^2}\Delta t + \frac{(a-b)^2}{2\cdot\sigma^4}\cdot(\Delta\beta_t)^2]$$

$+ \quad o(\Delta t)$

but $(\Delta\beta_t)^2 = \sigma^2\Delta t$ a.s., since the Quadratic variation process (Chung and Williams[1990]) of a diffusion process is indistinguishable from $\{\sigma^2\cdot t, t\in\mathcal{R}_+\}$. Therefore,

$$L_a = L_b \cdot [1 + \frac{(a-b)\cdot(\Delta\beta_t - b\Delta t)}{\sigma^2}] + \quad o(\Delta t)\square$$

Lemma 2

$$\pi_{t+\Delta t} = \frac{\pi_t \cdot (1-\lambda\Delta t)\cdot L_a + (1-\pi_t)\cdot \lambda\Delta t \cdot L_b}{\pi_t \cdot (1-\lambda\Delta t)\cdot L_a + (1-\pi_t)\cdot \lambda\Delta t \cdot L_b + \pi_t \cdot \lambda\Delta t \cdot L_a + (1-\pi_t)\cdot (1-\lambda\Delta t)\cdot L_b}$$
$$+o(\Delta t)$$

Proof By Bayes Law,

$$\pi_{t+\Delta t} = \frac{\pi_t \cdot (1-\lambda\Delta t)\cdot L_a + (1-\pi_t)\cdot \lambda\Delta t \cdot L_a}{\pi_t \cdot (1-\lambda\Delta t)\cdot L_a + (1-\pi_t)\cdot \lambda\Delta t \cdot L_a + \pi_t \cdot \lambda\Delta t \cdot L_b + (1-\pi_t)\cdot (1-\lambda\Delta t)\cdot L_b} \tag{47}$$
$$+ \quad o(\Delta t)$$

We have used the property of the Poisson process that one switch occurs with probability $\lambda\Delta t + \quad o(\Delta t)$,no switches with probability $(1-\lambda\Delta t) + \quad o(\Delta t)$ and $n \geq 2$ switches with probability $o(\Delta t)$. Please recall from lemma 1 that L_a and L_b are conditional densities.

Now in (47) interchange L_a and L_b, whenever they are multiplied by the $\lambda\Delta t$ term, but not when they are multiplied by the $(1-\lambda\Delta t)$ term. This gives an error of size $o(\Delta t)$, since $L_a - L_b = o((\Delta t)^{\frac{1}{2}})$ by lemma 1. □

Lemma 3

$$\pi_{t+\Delta t} - \pi_t =$$
$$(1 - 2\cdot\pi_t)\cdot\lambda\Delta t \quad +$$
$$\frac{\pi_t \cdot (1-\pi_t)\cdot(a-b)}{\sigma^2}\cdot[\Delta\beta_t - (a\cdot\pi_t + b\cdot(1-\pi_t))\cdot\Delta t]$$
$$+ \quad o(\Delta t)$$

Proof By lemma 2,

$\pi_{t+\Delta t} =$

$$\frac{\pi_t \cdot (1-\lambda\Delta t)\cdot \frac{L_a}{L_b} + (1-\pi_t)\cdot\lambda\Delta t}{\pi_t \cdot (1-\lambda\Delta t)\cdot \frac{L_a}{L_b} + (1-\pi_t)\cdot\lambda\Delta t + \pi_t\cdot\lambda\Delta t\cdot\frac{L_a}{L_b} + (1-\pi_t)\cdot(1-\lambda\Delta t)}$$

$+o(\Delta t)$ $\pi_{t+\Delta t} =$

$$\frac{\pi_t \cdot (1-\lambda\Delta t)\frac{L_a}{L_b} + (1-\pi_t)\cdot\lambda\Delta t}{\pi_t\cdot\frac{L_a}{L_b} + (1-\pi_t)} + \quad o(\Delta t)$$

substituting for $\frac{L_a}{L_b}$ from Lemma 1,

$$= \frac{\pi_t\cdot(1-\lambda\Delta t)\cdot[1+\frac{(a-b)\cdot(\Delta\beta_t - b\Delta t)}{\sigma^2}] + (1-\pi_t)\lambda\Delta t}{1-\pi_t\cdot[\frac{-(a-b)\cdot(\Delta\beta_t - b\Delta t)}{\sigma^2}]}$$

$+ \quad o(\Delta t)$

since $\frac{1}{1-x} = 1 + x + x^2 + ...$

$\pi_{t+\Delta t} =$

$$(Numerator)\cdot[1 - \frac{\pi_t\cdot(a-b)\cdot(\beta_t - b\Delta t)}{\sigma^2} + \frac{\pi_t^2\cdot(a-b)^2\cdot\sigma^2\cdot\Delta t}{\sigma^4}]$$

$+ \quad o(\Delta t)$

$= (Numerator)\cdot$

$$[1 - \frac{\pi_t\cdot(a-b)\cdot\Delta\beta_t}{\sigma^2} + \frac{(\pi_t^2\cdot(a-b)^2 - \pi_t\cdot(b-a)\cdot b)\Delta t}{\sigma^2}] \tag{48}$$

$+ \quad o(\Delta t)$

$Numerator =$

$$\lambda\Delta t - 2\cdot\pi_t\cdot\lambda\Delta t + \pi_t + \pi_t\cdot[\frac{(a-b)\cdot(\Delta\beta_t - b\Delta t)}{\sigma^2}]$$

$+ \quad o(\Delta t)$

$Numerator =$

$$\pi_t + \frac{\pi_t \cdot (a-b)}{\sigma^2}\Delta\beta_t + [\frac{\pi_t \cdot (b-a) \cdot b}{\sigma^2} + \lambda - 2 \cdot \pi_t \cdot \lambda]\Delta t \tag{49}$$

$+ \quad o(\Delta t)$

substituting (49) into (48),

$\pi_{t+\Delta t} =$

$$[1 + \frac{\pi_t \cdot (b-a) \cdot \Delta\beta_t}{\sigma^2} + \frac{\pi_t^2 \cdot (b-a)^2 - \pi_t \cdot (b-a) \cdot b)}{\sigma^2}\Delta t] \quad \times$$
$$[\pi_t - \frac{\pi_t \cdot (a-b)}{\sigma^2}\Delta\beta_t + \frac{\pi_t \cdot (b-a) \cdot b}{\sigma^2}\Delta t + \lambda \cdot (1 - 2 \cdot \pi_t)\Delta t]$$

$+ \quad o(\Delta t)$

$$= \pi_t + \frac{\pi_t \cdot (a-b)}{\sigma^2} \cdot \Delta\beta_t + \frac{\pi_t \cdot (b-a) \cdot b}{\sigma^2}\Delta t$$
$$+\lambda \cdot (1 - 2 \cdot \pi_t)\Delta t + \frac{\pi_t^2 \cdot (b-a)}{\sigma^2} \cdot \Delta\beta_t - \frac{\pi_t^2 \cdot (b-a)^2}{\sigma^2} \cdot \Delta t$$
$$+\pi_t^3 \cdot (a-b)^2 \Delta t - \pi_t^2 \cdot (b-a) \cdot b\Delta t$$

$+ \quad o(\Delta t)$

$$= \pi_t + \frac{\pi_t \cdot (1-\pi_t) \cdot (a-b)}{\sigma^2} \cdot \Delta\beta_t + (1 - 2 \cdot \pi_t) \cdot \lambda\Delta t$$
$$-\frac{\pi_t^2 \cdot (1-\pi_t) \cdot (a-b^2)}{\sigma^2} + \frac{\pi_t \cdot (b-a) \cdot b \cdot (1-\pi_t)}{\sigma^2}\Delta t$$

$+o(\Delta t)$ collecting terms,

$\pi_{t+\Delta t} - \pi_t =$

$$(1 - 2 \cdot \pi_t) \cdot \lambda\Delta t + \frac{\pi_t \cdot (1-\pi_t) \cdot (a-b)}{\sigma^2} \cdot [\Delta\beta_t - (a \cdot \pi_t + b \cdot (1-\pi_t)) \cdot \Delta t]$$

$+ o(\Delta t)\square$

Appendix 2

Proof of Propostion 3 The 1st moment of ΔW_t is analagous to (34) with z_t replaced by π_t, its expected value. The 2nd moment and kth moments are the same as (35) and (36). In addition, from (16) and because the agent takes the difference in output from the two industries to update

$$\text{Cov}[\Delta W_t \Delta\pi] = W\cdot(w_1 - w_2)\cdot(\pi)\cdot(1-\pi)\cdot(a-b) \tag{50}$$

Using Bellman's Principle of Optimality,

$$\begin{aligned} 0 = \max_{s.tC_t,\sum_{i=1}^{2} wit=1, w_i\geq 0} & [U[C_t] + J_t + +J_W\cdot[-C_t + W_t[\sum_{i=1}^{2} w_{it}\alpha_i] \\ & +J_\pi\cdot(1-2\pi)\cdot\lambda + J_{W\pi}\cdot W(w_1-w_2)\cdot\pi\cdot(1-\pi)\cdot(a-b) \\ & +\frac{W^2}{2}J_{WW}\cdot[\sum_{i=1}^{3} w_{it}^2]\sigma^2 + J_{\pi\pi}\pi^2\cdot(1-\pi)^2(a-b)^2 + o(\Delta t)] \end{aligned} \tag{51}$$

The first order conditions are

$$U_C = J_W \tag{52}$$

$$\begin{aligned} J_W\cdot W\cdot\alpha_1(z_t) + J_{W\pi}\cdot W\cdot\pi(1-\pi)(a-b) + J_{WW}W^2 w_{1t}\sigma^2 &\leq \lambda^{(1)} \\ J_W\cdot W\cdot\alpha_2(z_t) - J_{W\pi}\cdot W\cdot\pi(1-\pi)(a-b) + J_{WW}W^2 w_{2t}\sigma^2 &\leq \lambda^{(1)} \end{aligned} \tag{53}$$

$\lambda^{(1)}$ is the Lagrange muliplier associated with the constraint $\sum_{i=1}^{2} w_{it} = 1$. Equality holds for i only when $w_i > 0$. Summing the complementary slackness conditions for the assets implies

$$\begin{aligned} & J_W W[\sum_{i=1}^{2}\alpha_i(z_t)w_{it}] + J_{WW}W^2[\sum_{i=1}^{2} w_{it}^2]\sigma^2 \\ + \quad & J_{W\pi}\cdot(w_{1t}-w_{2t})\pi\cdot(1-\pi)\cdot(a-b) = \lambda^{(1)} \end{aligned} \tag{54}$$

Now guessing that the value function is separable as in (26) and substituting completes the proof. $\square$

Proof of Corollary 2 The decentralization for this model is the same as in Model 0 (and that in CIR) and we shall be brief. The risk-free rate r_t, the rate at which borrowing and lending are in zero net supply equals $\frac{\lambda^1}{W_t \cdot J_W}$. This is the same characterization as in Model 0 and for the general case considered in CIR. Using the complementary slackness condition (54) yields the result. $\square$

Proof of Proposition 4 The value function is of the form (26). Suppose the statement is true.

$$MU_{t+\Delta t}[\pi + \Delta\pi] = exp(-\rho \cdot (t + \Delta t)) \cdot I[\pi + \Delta\pi] \cdot (W + \Delta W)^{\gamma-1} \tag{55}$$

Using Taylor's theorem,

$$(W + \Delta W)^{\gamma-1} = W^{\gamma-1} + (\gamma - 1) \cdot W^{\gamma-1}$$

$$\cdot[\ \sum_{i=1}^{2} w_i \alpha_i(\pi)\Delta t + \sigma \cdot (\sum_{i=1}^{2} w_i d\zeta_i - (I[\pi])^{\frac{1}{\gamma-1}}\Delta t + (\gamma - 2) \cdot \frac{\sigma^2}{2}(\sum_{i=1}^{2} w_i^2)\ ]$$

$$+o(\Delta t)$$

and using the dynamics of π from (13)

$$I[\pi + \Delta\pi] = I[\pi]$$

$$+I'[\pi] \cdot [\ (1 - 2\pi) \cdot \lambda\Delta t + \frac{\pi(1-\pi)(a-b)}{\sigma^2} \cdot (\sigma(d\zeta_1 - d\zeta_2) - (2\pi - 1)(a-b)\Delta t\ ]$$

$$+\frac{1}{2} \cdot \frac{(a-b)^2}{\sigma^2} \cdot \pi^2(1-\pi)^2 I''[\pi]\Delta t + o(\Delta t)$$

and $\exp(-\rho(t + \Delta t)) = \exp(-\rho t) \cdot (1 - \rho\Delta t)$

Substituting these into (55) and taking expectations conditonal on π implies that

$$rf[\pi] = \rho - (\gamma - 1) \cdot [\ \sum_{i=1}^{2} w_i\alpha_i(\pi) + (\gamma - 2)\frac{\sigma^2 \sum_{i=1}^{2} w_i^2}{2}$$

$$-I^{\frac{1}{\gamma-1}}[\pi] + \frac{I'[\pi]}{I[\pi]}\pi(1-\pi)(a-b)(w_1 - w_2) + \frac{1}{2} \cdot \frac{I''[\pi]}{I[\pi]} \cdot \frac{\pi^2(1-\pi)^2}{(a-b)^2}\sigma^2\ ]$$

From Proposition 3,

$$rf[\pi] = \sum_{i=1}^{2} w_i\alpha_i(\pi) + (w_1 - w_2)\pi(1-\pi)(a-b)\frac{I'[\pi]}{I[\pi]} + \frac{1}{2}\sigma^2(\sum_{i=1}^{2} w_i^2)$$

These two characterizations of the risk-free rate are the same only if the differential equation (27) is satisfied. □

Table 1

Model 1 Skewness and Kurtosis

Unobservable Regime Switches

#	a	b	λ	σ	kr	Z_k	sk	Z_s
1.	.09	-.04	2	.02	2.94	-.6	-.32	-9.23
2.	.09	-.04	2	.04	2.99	-.008	-.05	-1.5
3.	.09	-.04	2	.07	3.19	2.11	.-.06	-1.9
4.	.09	-.04	2	.09	3.16	1.79	.04	1.24
5.	.09	-.04	1	.02	2.99	-.004	-.36	-10.37
6.	.09	-.04	1	.04	3.005	.05	-.10	-2.9
7.	.09	-.04	1	.07	3.08	.93	-.02	-.58
8.	.09	-.04	1	.09	3.23	2.488	-.014	-.39

a: High Productivity Drift Value

b: Low Productivity Drift Value

λ: Poisson Regime Switching Parameter

σ: Noise Level

kr: Kurtosis

Z_k: Z value for Kurtosis. Null Hypothesis of Normality

sk: Skewnesss

Z_s: Z value for Skewness.Null Hypothesis of Normality

ρ: Impatience Parameter. 0.04 for all simulations

γ: Risk-Aversion Parameter. CRRA 4 for all simulations

Figure 1
Implementation of Filter

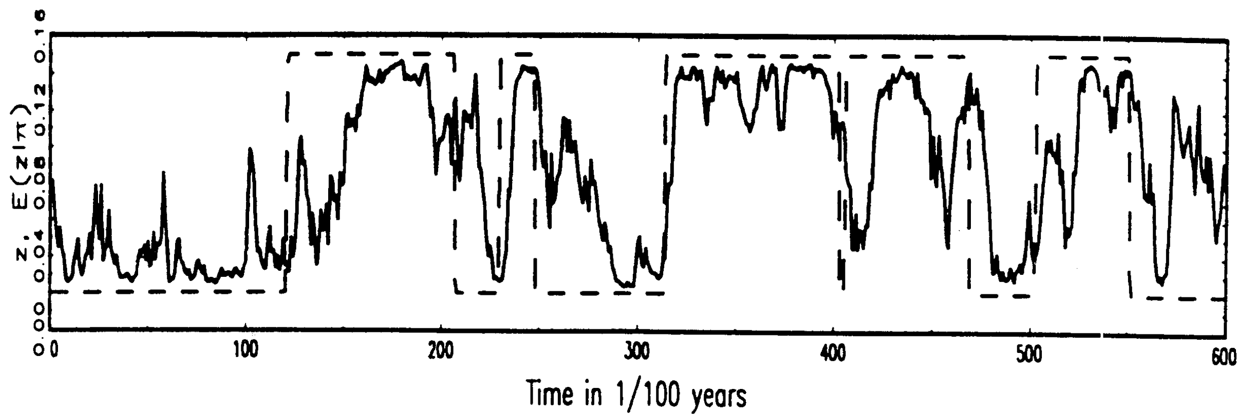


Figure 2.1
Fast Learning

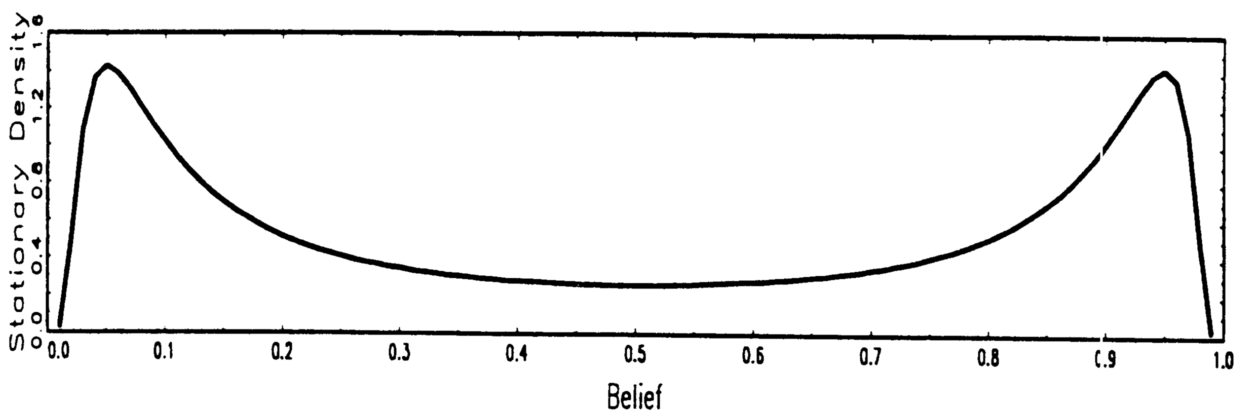


Figure 2.2
Slow Learning

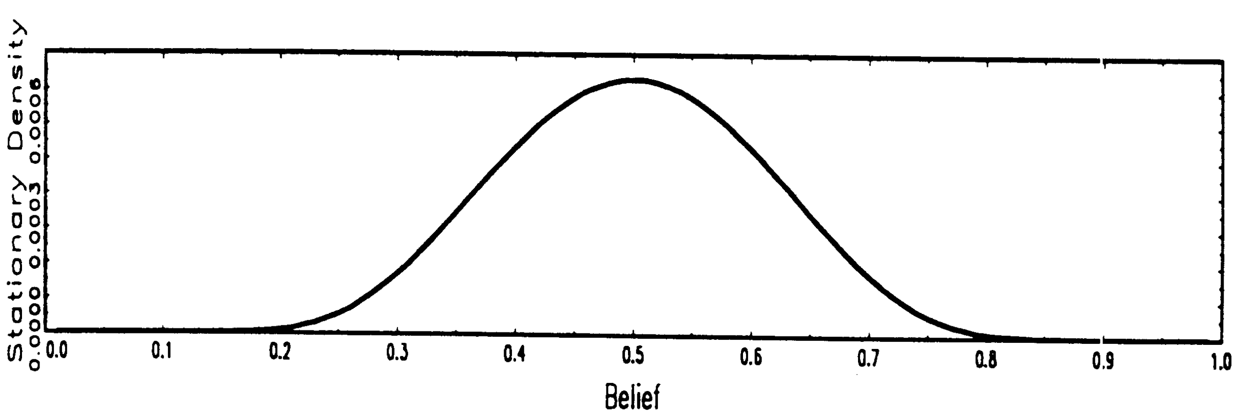


Figure 3.1

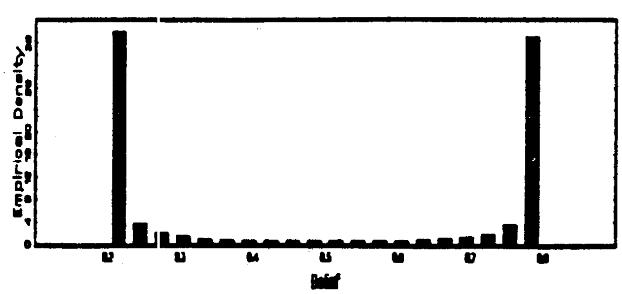


Figure 3.2

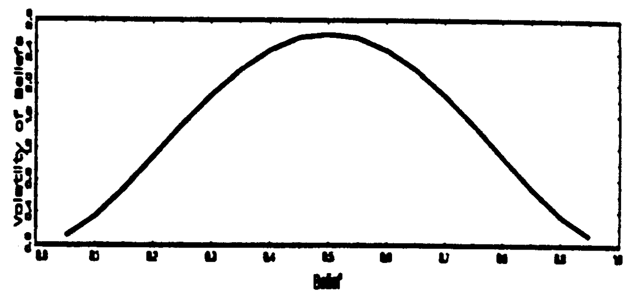


Figure 3.3

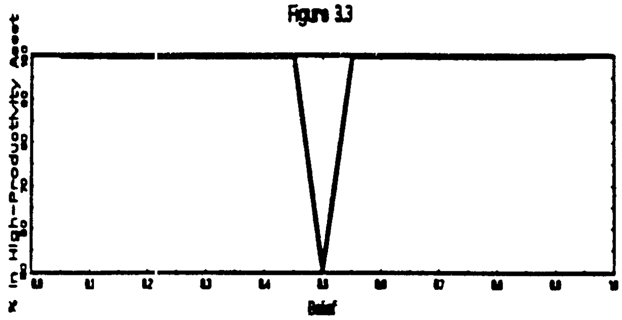


Figure 3.4

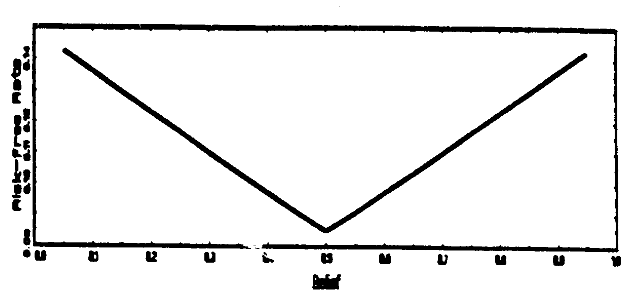


Figure 3.5

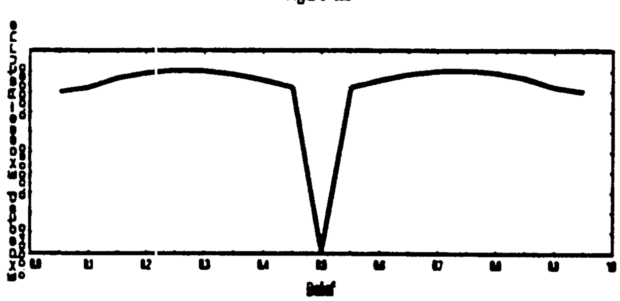


Figure 3.6

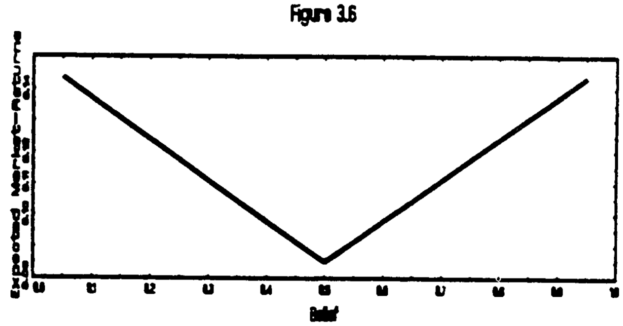


Figure 3.7

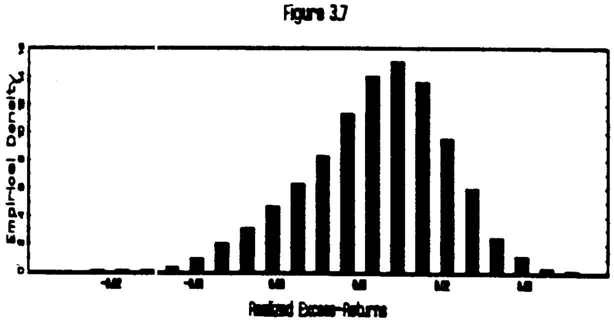
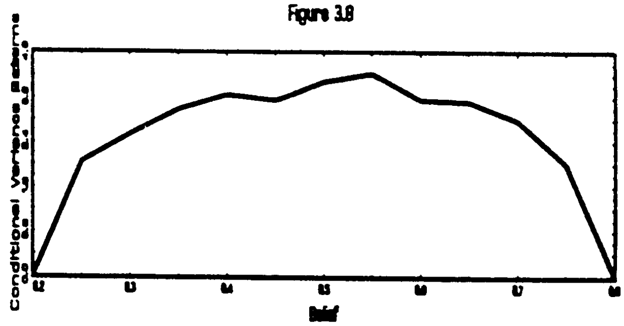


Figure 3.8



Fast Learning $\quad a=.15 \; b=.02 \; \lambda= 2 \; \sigma =.02$

Figure 4.1

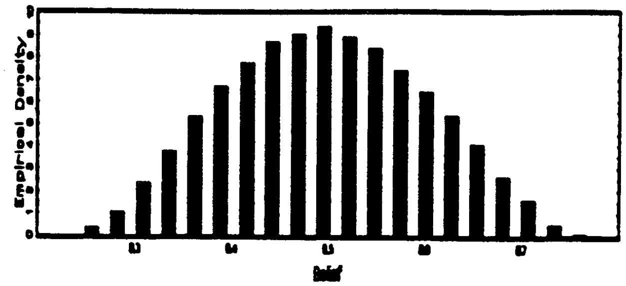


Figure 4.2

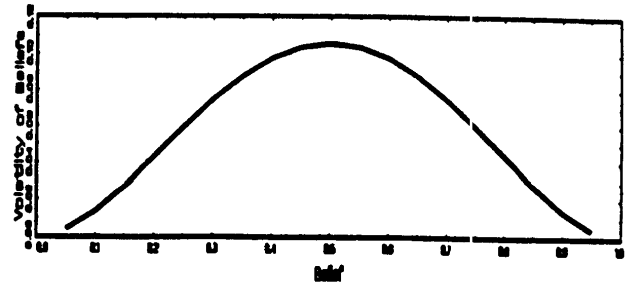


Figure 4.3

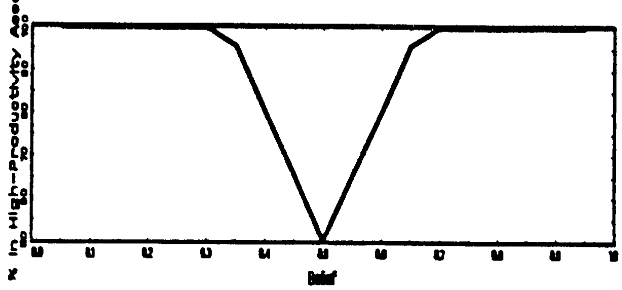


Figure 4.4

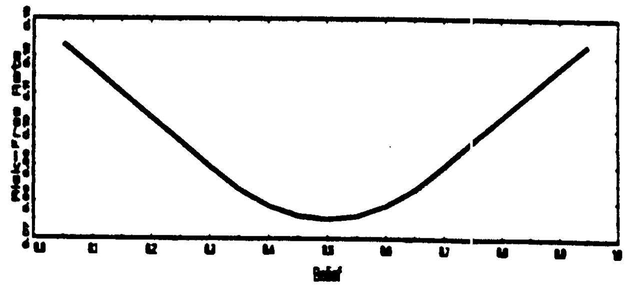


Figure 4.5

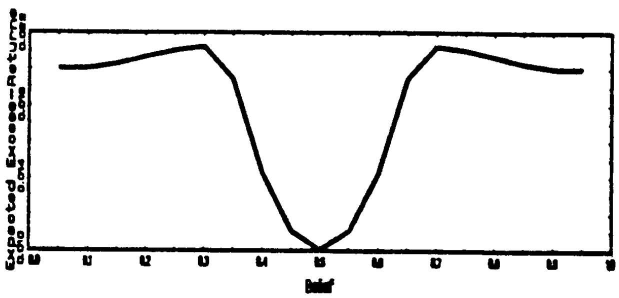


Figure 4.6

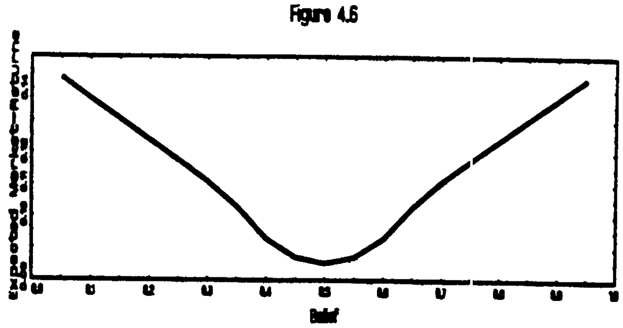


Figure 4.7

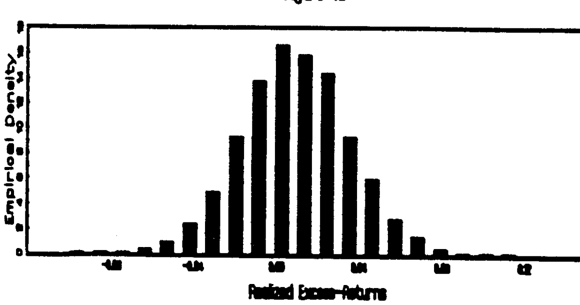
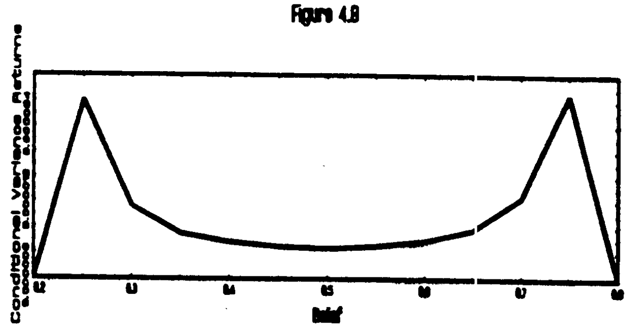


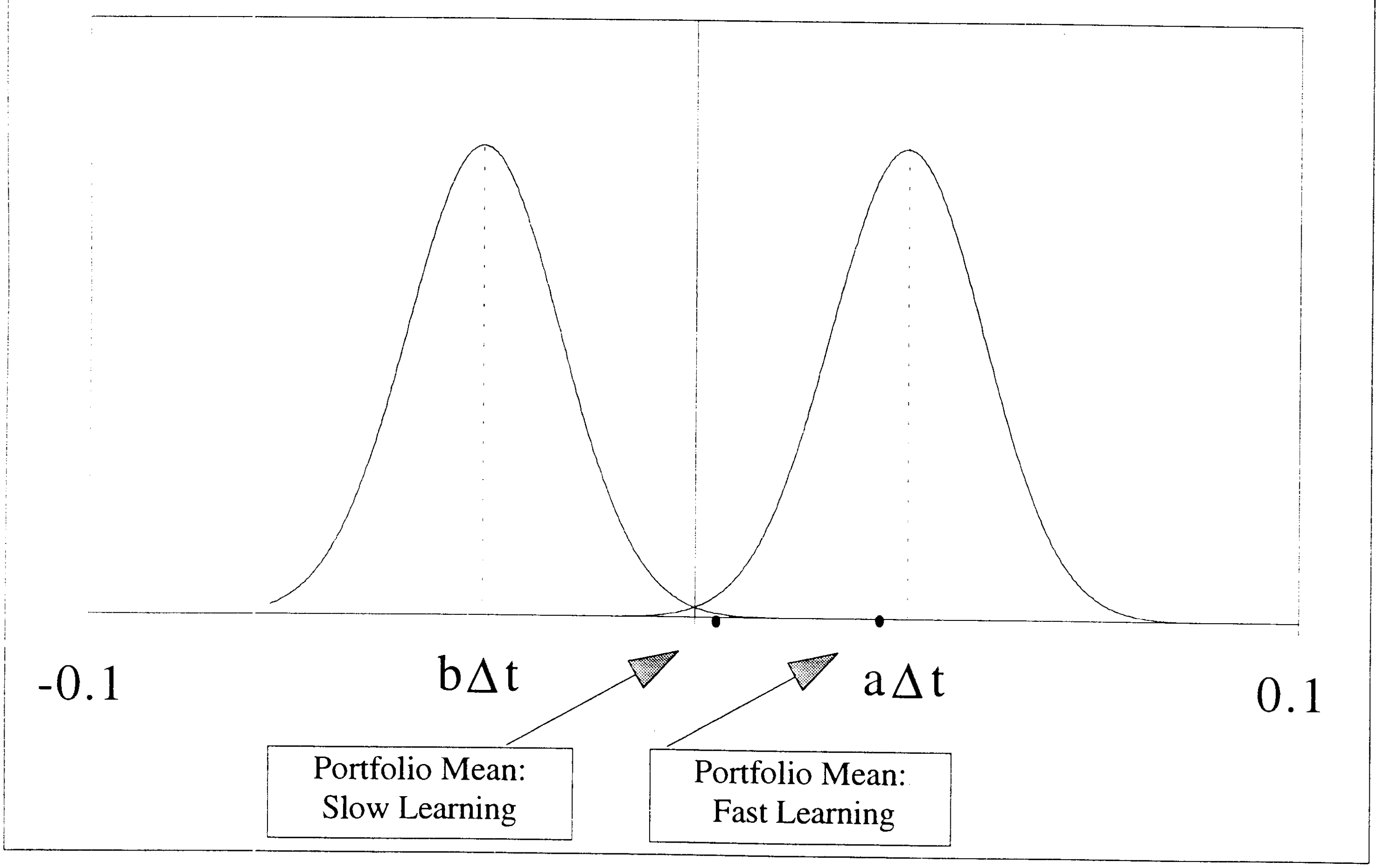
Figure 4.8



Slow Learning $\quad a=.15 \; b=.02 \; \lambda=2 \; \sigma=.1$

Figure 5

Understanding the Unconditional Distribution of Stock-Market Returns



References

P.J. Bickel and K.A. Doksum[1976]. "Mathematical Statistics. Basic Ideas and Selected Topics." *Holden-Day, Inc.*

F. Black[1976]. "Studies of Stock Price Volatility Changes." *Proceeding of the Business and Economics Statistics Section, American Statistical Association.*

T. Bollerslev, R.F. Engle, and J.M. Woolridge[1988]. "A Capital Asset Pricing Model with Time Varying Covariances." *Journal of Political Economy, 10.*

J.Y. Campbell and L. Hentschel[1992]. "No News is Good News: An Asymmetric Model of Changing Volatility in Stock Returns." *Journal of Financial Economics, 31..*

K. Chung and R. Williams[1990]. "Introduction to Stochastic Integration." *Birkhauser.*

P. Clark[1973]. "A Subordinated Stochastic Process Model With Finite Variance for Speculative Prices." *Econometrica, Vol. 41, No. 1.*

J. Cox, J. Ingersoll and S. Ross[1985]. "An Intertemporal General Equilibrium Model of Asset Prices." *Econometrica, 53.*

A. David[1992]. "Business Cycle Risk and the Equity Premium." *Working Paper, Department of Economics, UCLA.*

S. Davis and J. Haltiwanger. "Gross Job Creation, Gross Job Destruction and Employment Reallocation." *Quarterly Journal of Economics*, 1992.

J.B. Detemple[1991]. "Further Results on Asset Pricing with Incomplete Information." *Journal of Economic Dynamics and Control, 15.*

U. Dothan and D. Feldman[1986]. "A Theory of Asset Prices and the Term Structure of Interest Rates in a Partially Observable Economy." *The Journal of Finance, VOL XLI, No.3.*

R.F. Engle, D.M. Lilien and R.P. Robbins[1987]. "Estimating Time Varying

Risk Premia in the Term Structure." *Econometrica, 55.*

G. Gennotte[1986]. " Optimal Portfolio Choice Under Incomplete Information." *The Journal of Finance, VOL XLI, No. 3.*

J.D. Hamilton[1989]. "A New Approach to the Economic Analysis of Non-Stationary Time Series and the Business Cycle." *Econometrica, Vol 57 # 2.*

A. Jazwinski[1970]. "Stochastic Processes and Filtering Theory." *Academic Press.*

S. Karlin and H. Taylor[1982]. "A Second Course in Stochastic Processes." *Academic Press.*

V. Krishnan[1982]. "Non-Linear Filtering and Smoothing: An Introduction to Martingales, Stochastic Integrals and Estimation." *Wiley.*

D. Lilien. "Sectoral Shifts and Cyclical Unemployment." *Journal of Political Economy, 90* 1982.

P. Loungani, M. Rush and W. Tave. "Stock Market Dispersion and Unemployment." *Journal of Monetary Economics, 25*, 1990.

R.E. Lucas[1978]. "Asset Prices in an Exchange Economy. *Econometrica, 46.*

R. Merton[1973]. "An Intertemporal Capital Asset Pricing Model." *Econometrica, 41.*

D. Nelson[1991]. "Conditional Heteroskedasticity in Asset Returns: A New Approach." *Econometrica, Vol. 59 # 2.*

M. Sundaresan[1984]. "Consumption and Equilibrium Interest Rates in Stochastic Production Economies." *The Journal of Finance, VOL XXXIX.*

International Finance Discussion Papers

IFDP Number	Titles	Author(s)
	1993	
461	Fluctuatinng Confidence and Stock-Market Returns	Alexander David
460	Dollarization in Argentina	Steven B. Kamin Neil R. Ericsson
459	Union Behavior, Industry Rents, and Optimal Policies	Phillip Swagel
458	A Comparison of Some Basic Monetary Policy Regimes: Implications of Different Degrees of Instrument Adjustment and Wage Persistence	Dale W. Henderson Warwick J. McKibbin
457	Cointegration, Seasonality, Encompassing, and the Demand for Money in the United Kingdom	Neil R. Ericsson David F. Hendry Hong-Anh Tran
456	Exchange Rates, Prices, and External Adjustment in the United States and Japan	Peter Hooper Jaime Marquez
455	Political and Economic Consequences of Alternative Privatization Strategies	Catherine L. Mann Stefanie Lenway Derek Utter
454	Is There a World Real Interest Rate?	Joseph E. Gagnon Mark D. Unferth
453	Macroeconomic Stabilization Through Monetary and Fiscal Policy Coordination Implications for Monetary Union	Jay H. Bryson
452	Long-term Banking Relationships in General Equilibrium	Michael S. Gibson
451	The Role of Fiscal Policy in an Incomplete Markets Framework	Charles P. Thomas
450	Internal Funds and the Investment Function	Guy V.G. Stevens
449	Measuring International Economic Linkage with Stock Data	John Ammer Jianping Mei
448	Macroeconomic Risk and Asset Pricing: Estimating the APT with Observable Factors	John Ammer

Please address requests for copies to International Finance Discussion Papers, Division of International Finance, Stop 24, Board of Governors of the Federal Reserve System, Washington, D.C. 20551.

International Finance Discussion Papers

IFDP Number	Titles	Author(s)
	1993	
447	Near observational equivalence and unit root processes: formal concepts and implications	Jon Faust
446	Market Share and Exchange Rate Pass-Through in World Automobile Trade	Robert C. Feenstra Joseph E. Gagnon Michael M. Knetter
445	Industry Restructuring and Export Performance: Evidence on the Transition in Hungary	Valerie J. Chang Catherine L. Mann
444	Exchange Rates and Foreign Direct Investment: A Note	Guy V.G. Stevens
443	Global versus Country-Specific Productivity Shocks and the Current Account	Reuven Glick Kenneth Rogoff
442	The GATT's Contribution to Economic Recovery in Post-War Western Europe	Douglas A. Irwin
441	A Utility Based Comparison of Some Models of Exchange Rate Volatility	Kenneth D. West Hali J. Edison Dongchul Cho
440	Cointegration Tests in the Presence of Structural Breaks	Julia Campos Neil R. Ericsson David F. Hendry
	1992	
439	Life Expectancy of International Cartels: An Empirical Analysis	Jaime Marquez
438	Daily Bundesbank and Federal Reserve Intervention and the Conditional Variance Tale in DM/$-Returns	Geert J. Almekinders Sylvester C.W. Eijffinger
437	War and Peace: Recovering the Market's Probability Distribution of Crude Oil Futures Prices During the Gulf Crisis	William R. Melick Charles P. Thomas
436	Growth, Political Instability, and the Defense Burden	Stephen Brock Blomberg
435	Foreign Exchange Policy, Monetary Policy, and Capital Market Liberalization in Korea	Deborah J. Lindner
434	The Political Economy of the Won: U.S.-Korean Bilateral Negotiations on Exchange Rates	Deborah J. Lindner