Optimal Portfolio and Consumption in a Switching Diffusion Market

Daniel Oliveira Cajueiro**
Takashi Yoneyama***

Abstract

This paper addresses the problem of finding the optimal portfolio and consumption of a small agent in an economy. The novelty of this work is in considering that the financial market, in contrast to the celebrated Black-Scholes model, is composed of two sources of uncertainties: a Brownian motion and a continuous time Markov chain. While the Brownian motion intends to model the normal oscillations of the asset prices, the continuous time Markov chain aims at taking into account the abrupt variations that can occur in the parameters of the asset dynamics due to changes that take place in the state of the economy. The problem is formulated in terms of classical optimal stochastic control and the Hamilton-Jacobi-Bellman equation is solved to yield the solution.

Keywords: Continuous Time Markov Chain, Financial Market, Jump Processes, Optimal Consumption, Optimal Portfolio, Optimal Stochastic Control.

JEL Code: G11.
1. Introduction

This paper is concerned with the problem of finding the optimal portfolio and consumption of a small risk-averse investor. This problem was introduced by Merton (1969, 1971) and it has already been studied in a variety of settings by many authors (Aase, 1984, Framstad et al., 1998, Barbachan, 2000, Hanson and Westman, 2002). The novelty of this paper is in studying the optimal portfolio and consumption problem of a small investor in a switching diffusion market.

The switching diffusion market, in contrast to the celebrated Black-Scholes model, is a market model composed of two sources of uncertainties: a Brownian motion and a continuous time Markov chain. While the Brownian motion describes the normal oscillations of the asset prices, the continuous time Markov chain aims at modeling the abrupt variations that can occur in the parameters of the dynamic model governing the evolution of the assets due to changes that can take place in an economy.

In fact, the idea of modeling by using switching models is not new in the financial literature. The seminal paper (Hamilton, 1989), which is set in the context of discrete time models, studies discrete shifts that may occur in the growth rate of nonstationary series. On the other hand, in the continuous time set, switching diffusion models have been used to value European options (Naik, 1993, Chourdakis, 2000, Guo, 2001), to choose the correct time to buy and sell stocks (Zhang, 2001) and to deal with portfolio selection in a mean variance sense (Yin et al., 2002). Additionally, three simpler results in the same field may be found in Cajueiro and Yoneyama (2002b,a) and David (1997).

In this setting, this paper intends to contribute to the understanding of the portfolio and consumption allocations of a small agent in a dynamic economy with jumping parameters. Moreover, one will see how switches in regimes may affect these optimal allocations and that, contrary to the single regime structures (Merton, 1969, 1971), there are feedbacks of the system which are relevant to the agent's decision.

This paper is organized as follows. In section 2, the switching diffusion market is introduced. In section 3, the problem described above is formally stated. In section 4, some preliminary results, which will be useful later in solving the problem, are presented. In section 4, the problem is solved. Finally, section 5 presents some conclusions.

2. The Switching Diffusion Market

This section intends to define the financial market that constitutes the object of study in this paper. In order to make the formulation precise, one should consider the following mutually statistically independent processes on a suitable probability space \((\Omega, \mathcal{F}, P)\):

a) A standard Brownian motion \(B = \{B(t), \mathcal{F}_t; s \leq t \leq T\}; \)
b) A homogeneous continuous time Markov chain \( \theta = \{\theta(t), F_t; s \leq t \leq T\} \), with right continuous trajectories, and taking values in the finite set \( S = \{1, 2, \ldots, n\} \), characterized by \( p(t) = \{p_1(t), p_2(t), \ldots, p_n(t)\} \), with \( p_i(t) = P(\theta(t) = i) \), for \( i \in S \). Moreover, \( p(t) \) satisfies the following Kolmogorov's forward equation \( dp/dt = p(t)\Lambda \) where \( \Lambda = [\lambda_{ij}] \) is the stationary \( n \times n \) transition rate matrix of \( \theta \) (the infinitesimal generator of a continuous time Markov chain) with \( \lambda_{ij} \geq 0 \) for \( i \neq j \), and \( \lambda_{ii} = -\sum_{i \neq j} \lambda_{ij} \). Thus, the process is supposed to be conservative.

**Remark 2.1** The filtration \( F_t \) represents the information available at time \( t \) and any decision should be made based on it.

In a manner analogous to the Black-Scholes approach (Black and Scholes, 1973), one may suppose that the prices \( X_0 \) of the risk-free asset and \( X_1 \) of the stock are given by

\[
dX_0(t) = \rho X_0(t)dt
\]

and

\[
dX_1(t) = \mu(\theta)X_1(t)dt + \sigma(\theta)X_1(t)dB(t)
\]

respectively, where the interest rate \( \rho \) is a constant, the appreciation rate \( \mu(\theta) \triangleq \sum_{i=1}^n \delta(i, \theta)\mu_i \), the volatility \( \sigma(\theta) \triangleq \sum_{i=1}^n \delta(i, \theta)\sigma_i \) and \( \delta(i, j) \) symbol. The statewise processes \( \mu_i \) and \( \sigma_i \), for \( i = 1, \ldots, n \), are constant.

Model (2) simply states that \( \mu \) and \( \sigma \) are modulated by a Markov chain and can present \( n \) distinct values. Between jumps, if the Markov chain is in a state \( i \in S \), the variable \( \ln X_1(t) \) is normally distributed with instantaneous mean \( (\mu_i - \frac{1}{2}\sigma_i^2)dt \) and variance \( \sigma_i^2 dt \). Moreover, with regard to the selection of parameters, the variable \( \ln X_1(t) \) may have a probability distribution similar to the normal one or a distribution that may present multimodality, asymmetry and kurtosis. Thus, this model is much more general than that which is considered by Merton (1969, 1971).

**Remark 2.2** When Lipschitz continuity and linear growth conditions are satisfied, equation (2) has a unique strong solution. The proof can be carried out along the same lines of the argument used in theorem 4.6 on page 128 in Liptser and Shiryayev (1977).
Remark 2.3 The matrix $\Lambda$ (the infinitesimal generator of the underline Markov chain) can be estimated, Yin et al. (2002) using a constrained optimization procedure and stochastic approximation methods. On the other hand, Chourdakis (2000) shows how all the parameters of model (2) may be estimated by a procedure which extends Hamilton's (1989) filter.

Remark 2.4. It is natural to require $\min\{\mu_i \in S\} > \rho$, a.s., since otherwise there would be no sense in investing in the risky stock.

Model $(X_1(t), \theta(t))$ tries to take into account a large class of significant changes in the financial structure. Examples of such changes are arrival of important new information, aggregate supply shocks, abrupt shifts in technology or sudden swings in economic policies and political regimes. More specifically, while the appreciation rate $\mu(\theta)$ is related to the state of economy or the situation of a specific company in this economy, the volatility $\sigma(\theta)$ is related to the information flow. In fact, there is evidence that the volatility changes over time because of many factors (Figlewski, 1989, Schwert, 1989, Stein, 1989). The volatility increases during recessions and for brief periods during and immediately following panic situations. In general, a rise in volatility in these situations is due to lack of reliable information.

Two situations that can be modeled by using model $(X_1(t), \theta(t))$ are presented below. One can see that these two situations are very common in emergent markets.

Example 2.1 Consider a company that has a client belonging to the international market as the biggest buyer of its products. It is clear that the profit of this company depends highly on the magnitude of the ratio of the national currency to the international currency. In general, this exchange ratio is stabilized within a band. However, a monetary shock can cause a large deviation in this ratio as well as in the appreciation rate of this company. Therefore, this appreciation rate remains (approximately) constant until a monetary shock occurs. Thus, it can be modeled by a continuous time Markov chain where each value $\mu_i$, for $i \in S$, represents the value of the appreciation rate in each cycle.

Example 2.2 Consider the situation of an economy that is likely to have a new president that is in opposition to the present government. It is clear that some months before the elections the market will be more volatile due to the uncertainty about the new monetary policy. After the elections, the volatility will decrease, since the uncertainty returns to the standard level. Model $((X_1(t), \theta(t))$ allows taking into account these cycles of volatility.\footnote{In Almeida and Pereira (2000) and Bohl et al. (2000), one may find interesting applications of discrete counterparts of the continuous time model presented here that take volatility cycles into account.}
3. Problem Statement

This section aims at formally stating the problem that is solved in this paper, i.e., the problem of finding the optimal portfolio and consumption of a small agent in an economy. To do this, it is first necessary to define the wealth process (or the budget equation).

So, one may denote the wealth of an agent $W(t)$ defined on the suitable complete probability space $(\Omega, \mathcal{F}, P)$ as

$$dW(t) = W(t) \left\{ (1 - \pi) \frac{dX_0(t)}{X_0(t)} + \pi \frac{dX_1(t)}{X_1(t)} \right\} - cd t$$

where $\pi$ is the fraction of the agent’s wealth that is invested in the risky asset, thereby investing the fraction $(1 - \pi)$ in the safe one and where $c$ is the consumption of the agent. Then, the evolution of the process $(W(t), \theta(t))$ is described by

$$dW(t) = (\pi(\mu(\theta) - \rho)W(t) + \rho W(t) - c)dt + \pi \sigma(\theta)W(t)dB(t)$$

and

$$P(\theta(t + \Delta t) = j / \theta(t) = i) = \begin{cases} \lambda_{ij} \Delta t + o(\Delta t) & \text{for } i \neq j \\ 1 + \lambda_{ij} \Delta t + o(\Delta t) & \text{for } i = j \end{cases}$$

where $\lim_{\Delta t \to 0} \frac{o(\Delta t)}{\Delta t} = 0$.

**Assumption 3.1** Let $\Pi$ be the set of admissible portfolio policies $\pi$. A process $\pi: \mathbb{R}^+ \times \mathbb{R} \times \mathcal{S} \to \Pi$ is called an admissible portfolio policy if $\pi \in [0,1]$ and $\pi$ satisfies the following conditions:

a) Restriction on growth condition

$$\pi^2(t, x(t), \varepsilon(t)) \leq L_1 \int_0^t (1 + x^2(s))dK(s) + L_2(1 + \varepsilon^2(t) + x^2(t))$$

b) Lipschitz condition

$$|\pi(t, x(t), \varepsilon(t)) - \pi(t, y(t), \varepsilon(t))|^2 \leq L_1 \int_0^t (x(t) - y(t))^2dK(s) + L_2(x(t) - y(t))^2$$

where $L_1$ and $L_2$ are positive constants, $K(\cdot)$ is a non-decreasing right continuous function, $0 \leq K(\cdot) \leq 1$, $x(\cdot)$ and $y(\cdot)$ are continuous measurable functions, $\varepsilon(t)$ is
a function that assumes values in \( S \) and \( s \leq t \leq T \).

**Assumption 3.2.** Let \( C \) be the set of admissible consumption policies \( c \). A process \( c : \mathcal{R}^+ \times \mathcal{R} \times S \rightarrow C \) is called an admissible consumption policy if \( c \) satisfies the following conditions:

a) **Restriction on growth condition**

\[
c^2(t, x(t), \varepsilon(t)) \leq L_1 \int_0^t (1 + x^2(s))dK(s) + L_2(1 + \varepsilon^2(t) + x^2(t)) \tag{8}
\]

b) **Lipschitz condition**

\[
|c(t, \cdot)| \leq \Phi(t) \tag{9}
\]

where \( L_1 \) and \( L_2 \) are positive constants, \( K(\cdot) \) is a non-decreasing right continuous function, \( 0 \leq K(\cdot) \leq 1 \), \( x(\cdot) \) and \( y(\cdot) \) are continuous measurable functions, \( \varepsilon(t) \) is a function that assumes values in \( S \) and \( s \leq t \leq T \).

**Remark 3.1** With the assumptions 3.1 and 3.2, the equation (4) has one unique strong solution. The details of the proof is along the same lines of theorem 4.6 on page 128 and theorem 4.9 on page 142 in Liptser and Shiryaev (1977).

**Assumption 3.3** In this work, the jump sizes are considered predictable in \( \theta = \{\theta(t), \mathcal{F}_t; s \leq t \leq T\} \), i.e., one does not know if a jump will occur, but if it does, its magnitude is known.

Now, one may define the so-called value function \(^2\Phi : [s, T] \times \mathcal{R} \times S \rightarrow \mathcal{R} \). So, the problem that will be solved in this paper is formulated as follows:

**Problem 3.1** Suppose that, by starting with the initial wealth \( W(s) = w \) at time \( t = s \), a small agent\(^3\) wants to maximize the performance index given by

\[
E[\int_s^T \exp(-\gamma t) U(c)dt + \exp(-\gamma T)U(W(T))/W(s) = w, \theta(s) = i],
\]

where \( T \) is a known fixed date in the future. If the utility function \( U \) of the investor is increasing and concave (the investor is risk-averse), then the problem is to find the value

\(^2\)The value function in a continuous time optimization problem plays the same role of the optimized function in a statical optimization problem, i.e., the function that attains its optimum when the optimal policy is used.

\(^3\)The assumption that the agent is small is a very significant assumption which says that the behavior of the agent by himself/herself does not affect demands and prices in the economy.
Optimal Portfolio and Consumption in a Switching Diffusion Market

function $\Phi(s, w, i)$ and a Markov policy $\{\pi, c\}$ such that $\Phi(s, w, i)$, for $i = 1, \cdots, n$, is given by

$$\Phi(s, w, i) = \sup_{\pi \in \Pi, c \in C} J(s, w, i, \pi, c) = J(s, w, i, \pi^*(s, w, i), c^*(s, w, i))$$

where

$$J(s, w, i, \pi, c) = E \left[ \int_s^T \exp(-\gamma t) U(c) dt + \exp(-\gamma T) U(W(T))/W(s) \right]$$

Assumption 3.4 In this paper, the utility functions of the investor are assumed to belong to the class of Constant Relative Risk Aversion (CRRA) utility functions, i.e., $U(W) = \frac{W^r}{r}$ and $U(c) = \frac{c^r}{r}$, where $0 < r < 1$.

The problem 3.1 with assumption 3.4 is exactly the same problem considered by Merton (1969), but now, with jumping parameters. In this problem, the small agent chooses his/her consumption and portfolio in a way to provide himself/herself the maximum utility of consumption over time, weighing it with a constant discount rate $\gamma$. In particular, assumption 3.4 plays a very important role here. Without it, this problem may not have an explicit solution or any solution at all.

4. Preliminary Results

This section deals with the presentation of some results that will be necessary to apply the dynamic programming methodology of optimal stochastic control theory. It follows the same lines of Fragoso and Hemerly (1991), who studied the LQG problem in this setting.

Proposition 4.1 The process $\{(W(t), \theta(t))\}$ is a Markov process.\(^4,5\)

Proof See appendix.

\(^4\)Markov processes are those which present the Markov property which says that if the state of a system at a particular time $s$ is known, additional information regarding the behavior of the system at times $t < s$ has no effect on our knowledge of the probable development of the system at $t > s$.

\(^5\)Although Markov processes have been widely used in economics, it is a restricted class of processes. In fact, this class does not include several types of processes, for instance, those which present long memory.
Proposition 4.1 presents an essential result here. If the joint process \((W(t), \theta(t))\) were not a Markov process, the dynamic programming methodology of optimal stochastic control theory should not be used here.

**Proposition 4.2** The process \((W(t), \theta(t))\) has sample paths that are continuous from the right.

**Proof** It follows directly from equations (4) and (5).

The following definition is necessary to determine the infinitesimal generator of the stochastic process \((W(t), \theta(t))\).

**Definition 4.1** (Semigroup Operator) One may define the operator \(T_h\) on the space \(B(\mathcal{R}^+ \times \mathcal{R} \times \mathcal{S})\) of bounded Borel measurable scalar functions \(\Phi\) defined on \(\mathcal{R}^+ \times \mathcal{R} \times \mathcal{S} = \mathcal{X}\) and equipped with the norm

\[
\|\Phi\| = \sup_{x \in \mathcal{X}} |\Phi(x)|
\]

as follows:

\[
T_h\Phi(s, w, i) = E[\Phi(s + h, W(s + h), \theta(s + h))/W(s) = w, \theta(s) = i]
\]

for \(h \geq 0\).

**Remark 4.1** One may see that \(T_h\Phi(s, w, i)\) is the mean value of \(\Phi(s + h, W(s + h), \theta(s + h))\) under the condition \(W(s) = w\) and \(\theta(s) = i\).

**Proposition 4.3** If \(\Phi : \mathcal{R}^+ \times \mathcal{R} \times \mathcal{S} \rightarrow \mathcal{R}\) is a continuous bounded function, then

\[
\lim_{h \to 0} T_h\Phi(s, W(s), \theta(s)) = \Phi(s, W(s) = w, \theta(s) = i)
\]

**Proof** It follows directly from definition 4.1.

Propositions 4.1, 4.2 and 4.3 allow the unique characterization of the process \((W(t), \theta(t))\) by means of its infinitesimal generator:

**Definition 4.2** (Infinitesimal Generator) The infinitesimal generator\(^6\) \(L\) of the Markov process \((W(t), \theta(t))\) defined on the space of bounded Borel measurable functions \(\Phi \in B(\mathcal{R}^+ \times \mathcal{R} \times \mathcal{S})\) is defined as

\(^6\)The role of the infinitesimal generator of a Markov process may be elucidated by means of the Dynkin's formula which is given by \(E[f(X_t)] = f(x) + E[\int_0^t Lf(X_s)ds/X_0 = x]\) where \(X\) is a Markov process and \(f\) is a measurable smooth function.
Optimal Portfolio and Consumption in a Switching Diffusion Market

\[ \lim_{h \to 0^+} \frac{T_h \Phi(s, W(s), \theta(s)) - T_0 \Phi(s, W(s), \theta(s))}{h} \]  

(12)

**Proposition 4.4** Let \( \Phi \) be a function that belongs to the class of functions with continuous first-order derivatives in \( t \) on \([s, T]\) and first and second order in \( W \in \mathcal{R} \), almost everywhere. The infinitesimal generator of \( \{(W(t), \theta(t)); s \leq t \leq T\} \), with \( W(t) \) that satisfies equation (4), \( \{\theta(t)); s \leq t \leq T\} \) that satisfies equation (5) and assumption 3.3, \( \pi \) that satisfies assumption 3.1 and \( c \) that satisfies assumption 3.2 is given by

\[
L^{\pi, c} \Phi(t, W, i) = \frac{\partial \Phi(t, W, i)}{\partial t} + ((\pi_i - \rho) + \rho)W - c \frac{\partial \Phi(t, W, i)}{\partial W} + \frac{1}{2} (\pi \sigma^2 W^2) \frac{\partial^2 \Phi(t, W, i)}{\partial W^2} + \sum_{j=1}^n \lambda_{ij} \Phi(t, W, j) 
\]

(13)

**Proof** See appendix.

The two theorems below respectively present the necessary and sufficient conditions for the existence of an optimal policy. Theorem 4.1 states that if there exists an optimal control policy associated with the problem at issue, it is given by the solution of the Hamilton-Jacobi-Bellman equation. Thus, using this theorem, the original control problem is transformed into the simpler problem of finding the maximum of a real function. On the other hand, theorem 4.2 is the converse of Hamilton-Jacobi-Bellman equation, i.e., it states that if the Hamilton-Jacobi-Bellman equation has a solution, this is actually the optimal control policy. Both are fundamental results in control theory and they provide a very nice solution to the optimal stochastic control problem.

**Theorem 4.1** The Hamilton-Jacobi-Bellman equation associated with this problem is given by

\[
\sup_{\pi, c} \{L^{\pi, c} \Phi(t, W, i) + \exp(-\gamma t)U(c)\} = 0 
\]

(14)

with boundary conditions \( \Phi(T, W, i) = \exp(-\gamma T)U(W) \), for \( i = 1, \ldots, n \), i.e., if there is an optimal policy \( \{\pi^*, c^*\} \), it is given by the solution of equation (14).

**Proof** It follows from Dynkin's formula and Bellman's optimality principle (Fleming and Rishel, 1975).
Theorem 4.2 (Dynamic Programming Verification Theorem) Let \( \Phi \) be the solution of the dynamic programming equation \( \sup_{\pi,c} \{ L_{\pi,c} \Phi(t,W,i) + \exp(-\gamma t)U(c) \} = 0 \) with boundary conditions \( \Phi(T,W,i) = \exp(-\gamma T)U(W) \), for \( i = 1, \ldots, n \). Then:

a) \( J(t,W,i,\pi,c) \leq \Phi(t,W,i) \) for any admissible policy \( \{ \pi,c \} \) and any initial data;

b) If \( \{ \pi^*,c^* \} \) is an admissible policy such that

\[
L_{\pi^*,c^*} \Phi(t,W,i) + \exp(-\gamma t)U(c^*) = \sup_{\pi,c} \{ L_{\pi,c} \Phi(t,W,i) + \exp(-\gamma t)U(c) \} = 0
\]

Thus \( \{ \pi^*,c^* \} \) is optimal;

Proof This proof is analogous to the proof of theorem 4.1 on page 159 in Fleming and Rishel (1975).

5. Problem Solution

This section aims at solving problem 3.1.

Theorem 5.1 The optimal policy \( \{ \pi^*,c^* \} \) that solves the problem 3.1 is given by

\[
\pi^*(W,i) = \min \left( \frac{\mu_i - \rho}{\sigma_i^2(1-r)}, 1 \right) \quad \text{for} \quad \theta(t) = i
\]

and

\[
c^*(t,W,i) = \frac{W}{g(t,i)} \quad \text{for} \quad \theta(t) = i
\]

and the value function is given by \( \Phi(t,W,i) = \exp(-\gamma t)g(t,i)^{1-r} \frac{W^r}{r} \) with boundary conditions \( \Phi(t,0,i) = 0 \), for \( i = 1, \ldots, n \) and where \( g(t,i) \), for \( i = 1, \ldots, n \) is the solution of the real positive system of

\[
\frac{dg(t,i)}{dt} + \frac{1}{1-r} \left( \rho r + \frac{(\mu_i - \rho)^2 r}{2 \sigma_i^2(1-r)} - \gamma \right) g(t,i) + \frac{1}{1-r} \sum_{j=1}^{n} \lambda_{ij} g^{1-r}(t,j) g^r(t,i) + 1 = 0
\]

with boundary conditions \( g(T,i) = 1 \), for \( i = 1, \ldots, n \).

Proof See Appendix.
With regard to theorem 5.1, several conclusions may be drawn about the optimal allocations of a small agent in a dynamic economy with jumping parameters.

One should first notice that the optimal portfolio given by equation (15) has the same form as the optimal portfolio for a single regime financial market (see Merton (1969, 1971), i.e., the agent will choose his/her portfolio based on the difference \((\mu(\theta) - \rho(\theta))\) and the size of the volatility \(\sigma(\theta)\). However, there is a fundamental difference. The state of the economy (given here by a continuous time Markov chain), which is one of the feedbacks of the system, is crucial for the portfolio decision. This situation differs from the situation presented by Merton (1969, 1971) where no feedback of the system is relevant for the portfolio decision and, therefore, the optimal portfolio is constant over time.

The influence of the state of the economy on consumption is conditional to the value of the discount rate \(\gamma\) which measures the agent’s “appetite for liquidity”. Notice from equation (18) that if \(\gamma\) is sufficiently small, then the state of the economy will strongly influence the solutions \(g(t, i)\) for \(i = 1, \ldots, n\) and, therefore, the optimal consumption policy [see equation (16)] through the term \(\frac{(\mu(\theta) - \rho)^2}{2\sigma(\theta)^2(1-\gamma)}\) in equation (18). On the other hand, if \(\gamma\) is sufficiently large, the state of the economy has almost no influence on the optimal consumption policy.

Actually, as explained above, with regard to the value of \(\gamma\), the optimal consumption may have a volatile behavior. This is not surprising. It is clear that the optimal consumption of the small agent depends on the level of wealth of the agent (this constraint is given by the budget equation 4) but also on the state of the economy. If the agent wants to maximize the optimal performance index given by equation (12) then he/she should allocate more wealth to consumption in good times for consumption and more wealth to investment in good times for investment.

In order to more precisely take into account the consumer behavior in an economy with jumping parameters, an example is provided below:

Example 5.1 Consider an economy which can be modelled by a two-state continuous time Markov chain:

\[\text{a) State 1 models a financial market that is down with parameters } \mu_1 = 0.0010 \text{ and } \sigma_1 = 0.1. \text{ Therefore, } \pi_1 = 0.4.\]

\[\text{7Actually, the consumption is smooth between jumps, but it may jump when the state of the economy changes.}\]

\[\text{8It is worth mentioning that in this work we are not thinking about aggregate consumption (and this is not an equilibrium model), we are dealing with the classical (microeconomic) problem where a small agent chooses between consumption and investment. To consider this agent a representative household may be a mistake. This hypothesis may go against the general macroeconomic assertion that agents try to smooth consumption over time. If the agent had this focus here (to smooth consumption over time), he/she would not be maximizing the performance index given by equation (12).}\]
b) State 2 models a financial market that is up with parameters $\mu_2 = 0.0022$ and $\sigma_2 = 0.07$. Therefore, $\pi_2 = 1$.

Consider also that $\rho = 0.0002$, $\gamma = 0.0001$ and $r = 0.8$. Three situations will be presented:

a) When the probability of state 1 and the probability of state 2 are the same, i.e., for example, $\lambda_{11} = -0.005$, $\lambda_{12} = 0.005$, $\lambda_{21} = 0.005$ and $\lambda_{22} = -0.005$.

b) When the probability of state 1 is much greater than the probability of state 2 (In this situation, one can say that in this economy the financial market is almost always not good to invest in), i.e., for example, $\lambda_{11} = -0.005$, $\lambda_{12} = 0.005$, $\lambda_{21} = 0.5$ and $\lambda_{22} = -0.5$.

c) When the probability of state 2 is much greater than the probability of state 1 (In this situation, one can say that in this economy the financial market is almost always good to invest in), i.e., for example, $\lambda_{11} = -0.5$, $\lambda_{12} = 0.5$, $\lambda_{21} = 0.005$ and $\lambda_{22} = -0.005$.

As was already introduced above and as one can see in figures 1, 2 and 3, if $\gamma$ is sufficiently small, then the infinitesimal generator of the CTMC has a strong influence on the optimal consumption in an economy. According to figure 1, when the financial market is up (state 2), the optimal consumption of an agent in an economy will be smaller than when the financial market is down (state 1). In this case, the agent prefers to invest. On the other hand, if an economy is basically dominated by state 2 which is good to invest in, as one can see in situation 2 (figure 2), then the agent allocates a bigger parcel of his/her wealth to investment. This occurs due to the low probability of state 1 occurring. Moreover, if an economy is basically dominated by state 1 which is bad to invest in, as one can see in situation 3 (figure 3), then the agent allocates a big parcel of his/her wealth to consumption. This happens due to the low probability of state 2 occurring. In these two latter situations, since the change of the state of the economy is unlikely, the agent will behave as he/she was in a single state economy.
Figure 1
Plot of $g(t, i)$ of an economy which is modelled by a CTMC with two states which have the same probability of happening

Figure 2
Plot of $g(t, i)$ of an economy whose financial market is up with high probability
6. Conclusions

This paper has studied the optimal investment and consumption problem in the switching diffusion market. The proposed problem has been completely solved by means of dynamic programming arguments. The Hamilton-Jacobi-Bellman equation was solved to provide a closed solution to the problem. As one can see, the optimal investment and consumption policies in this market depend strongly on the state of an economy which is modeled by a continuous time Markov chain.

References


Bohl, M. T., Siklos, P. L., & Werner, T. (2000). Do central banks react to the


Appendix

Proof of proposition 4.1

Proof Since, according to remark 3.1, equation (4) has a unique strong solution, then there exists a function $f$ (that can be determined by $\mu$ and $\sigma$) such that $W(t) = f(X(t_0), B(s), \theta(s); t_0 \leq s \leq t)$. Thus,

a) $W(t)$ depends only on $W(t_0)$ and $\mathcal{F}_{[t_0,t]} = \mathcal{F}^B_{[t_0,t]} \times \mathcal{F}^\theta_{[t_0,t]} = \sigma\{B(v) \mid B(t_0); t_0 \leq v \leq t\} \times \sigma\{\theta(v)\};$

b) It is obvious that $(\theta(v), B(v) - B(t_0))$, for $v \geq t_0$ is independent of $(\theta(l), B(l))$ for $l \leq t_0$, given $\theta(t_0);$

c) Since $\{B(t); s \leq t \leq T\}$ and $\{\theta(t); s \leq t \leq T\}$ are supposed to be mutually independent, it follows that $\theta(t)$ is independent of $\mathcal{F}_{[s,t_0]}$, given $\theta(t_0)$.

Therefore, from (a) and (b), it follows that $W(t)$ is independent of $\{W(v); v \leq t_0\}$ given $\theta(t_0)$. From (c) one can conclude that $(W(t), \theta(t))$ is independent of $\{(W(v), \theta(v)); v \leq t_0\}$ given $(W(t_0, \theta(t_0))$. Thus $(W(t), \theta(t))$ is a Markov process.

Proof of proposition 4.4

From equation (4.1), one may write

$$L^\pi \Phi(s, W, i) = \lim_{h \to 0} \frac{T_h \Phi(s, W(s), i) - T_0 \Phi(s, W(s), i)}{h}$$

$$= \lim_{h \to 0} \frac{1}{h} \{E_{W \times \theta}[\Phi(s + h, W(s + h), \theta(s + h))/W(s) - w, \theta(s) = i] - T_0 \Phi(s, W(s), i) \}$$

Additionally, noticing $E[\Phi(s + h, W(s + h), \theta(s + h))/W(s) = w, \theta(s) = i] = \sum_{j=1}^{n} E[\Phi(s + h, W(s + h), j)/W(s) = w] \cdot P[\theta(s + h) = j/\theta(s) = i]$ and from equation (5), one may get
\[ L^{\pi,e}\Phi(s, W, i) = \lim_{\theta \to 0} \frac{1}{\theta} \left\{ \sum_{j=1}^{n} E[\Phi(s + \theta, W(s + \theta), j)/W(s) = w] \cdot \theta \right\} - T_0 \Phi(s, W(s), i) \]

\[ = \lim_{\theta \to 0} \frac{1}{\theta} \left\{ \sum_{j=1}^{n} E[\Phi(s + \theta, W(s + \theta), j)/W(s) = w] \cdot (\lambda_{ij} \theta + o(\theta)) \right\} \]

\[ + E[\Phi(s + \theta, W(s + \theta), i)/W(s) = w] \cdot (1 + \lambda_{ii} \theta + o(\theta)) - T_0 \Phi(s, W(s), i) \]

\[ = \lim_{\theta \to 0} \frac{1}{\theta} \left\{ \sum_{j=1}^{n} E[\Phi(s + \theta, W(s + \theta), i)/W(s) = w] \right\} \]

And, from definition 4.1 and proposition 4.3

\[ L^{\pi,c}\Phi(s, W, i) = \lim_{\theta \to 0} \frac{1}{\theta} \left\{ \sum_{j=1}^{n} \lambda_{ij} \Phi(s, W(s), j) \right\} \]

Finally, from equation (4), one can notice that

\[ \lim_{\theta \to 0} \frac{1}{\theta} \left\{ \Phi(s + \theta, W(s + \theta), i)/W(s) = w \right\} \]

Thus, the proof is complete.

**Proof of theorem 5.1**

Before considering the main theorem, the following lemmas are required.

**Lemma A.1** The system

\[ \frac{dg(t, j)}{dt} + M_{ii}g(t, j) + \sum_{k=1}^{n} M_{jk}g_{-r}(t, k)g^{r}(t, j) + C = 0 \quad \text{for} \quad j = 1, \cdots, n \quad (A.1) \]
where $M$ is a real matrix and $C$ is a constant, has a unique solution.

**Proof** One should note that the system (A.4) can be rewritten as

$$
\frac{1}{1-r} \frac{dg^{1-r}(t,i)}{dt} + M_{ii}g^{1-r}(t,i) + \sum_{j=1, j \neq i}^{n} M_{ij}g^{1-r}(t,j) + Cg^{-r}(t,i) = 0
$$

Let $\bar{g}(t,i) = g(t,i)^{1-r}$, then

$$
\frac{d\bar{g}(t,i)}{dt} + \overline{M}_{ii}\bar{g}(t,i) + \sum_{j=1, i \neq j}^{n} \overline{M}_{ij}\bar{g}(t,j) + C\bar{g}^{-r}(t,i) = 0
$$

where $\overline{M} = (1-r)M$ and $\overline{C} = (1-r)C$.

The above equation satisfies the Lipschitz condition and, therefore, has a unique solution (for details, see the theorem that deals with the existence of nonlinear ordinary differential equations on page 205 in Hochstadt (1963).

**Lemma A.2** Suppose that $M_{jk} \geq 0$ for $j \neq k$ and $C$ is a positive constant. Then the system described by

$$
\frac{dg(t,j)}{dt} + M_{jj}g(t,j) + \sum_{k=1, j \neq k}^{n} M_{jk}g^{1-r}(t,k)g^{r}(t,j) + C = 0 \text{ for } j = 1, \ldots, n \quad (A.2)
$$

with final condition $g(T,j) = g_T > 0$ is a real positive system.

**Proof** Let $\tau = T - t$, then

$$
\frac{dg(T-\tau,i)}{d\tau} = M_{ii}g(T-\tau,i) + \sum_{j=1, i \neq j}^{n} M_{ij}g^{1-r}(T-\tau,j)g^{r}(T-\tau,i) + C
$$

Define $\overline{g}(\tau, i) = g(T-\tau,i)$ with initial condition $\overline{g}(0, i) = g_T > 0$, then

$$
\frac{d\overline{g}(\tau, i)}{d\tau} = M_{ii}\overline{g}(\tau, i) + \sum_{j=1, i \neq j}^{n} M_{ij}g^{1-r}(\tau,j)\overline{g}^{r}(\tau,i) + C \quad (A.3)
$$

One should note that the variable $\overline{g}(\tau, i)$ starts at $g_T > 0$ and if for any reason $\overline{g}(\tau, i) \to 0$ then its derivative $\frac{d\overline{g}(\tau, i)}{d\tau}$ becomes positive and, therefore, $\overline{g}(\tau, i)$ will become more positive. One should also note that system (A.3) is the forward form of the system (A.2) which is in the backward form.
Proof of the theorem

From equations (13) and (14), the Hamilton-Jacobi-Bellman equation can be written as

\[
\frac{\partial \Phi(t, W, i)}{\partial t} + \sup_{\pi, c} \left\{ (\pi (\mu_i - \rho) + \rho) W - c \frac{\partial \Phi(t, W, i)}{\partial W} + \frac{1}{2} \left( \sigma_i^2 W^2 \frac{\partial^2 \Phi(t, W, i)}{\partial W^2} \right) \right. \\
\left. + \exp(-\gamma t)U(c) \right\} + \sum_{j=1}^{n} \lambda_{ij} \Phi(t, W, j) = 0
\]

with boundary conditions \( \Phi(T, W, i) = \exp(-\gamma T)U(W) \), for \( i = 1, \cdots, n \) and \( \Phi(t, 0, i) = 0 \), for \( i = 1, \cdots, n \).

The policy \( \{ \pi^*, c^* \} \) that maximizes the above equation is given by

\[
\pi^* = -\frac{(\mu_i - \rho) \frac{\partial \Phi(t, W, i)}{\partial W}}{\sigma_i^2 W \frac{\partial^2 \Phi(t, W, i)}{\partial W^2}}
\]

and

\[
c^* = \left( \exp(\gamma t) \frac{\partial \Phi(t, W, i)}{\partial W} \right)^\frac{1}{\gamma - 1}
\]

Thus, the Hamilton-Jacobi-Bellman becomes

\[
\frac{\partial \Phi(t, W, i)}{\partial t} + \rho W \frac{\partial \Phi(t, W, i)}{\partial W} - \frac{1}{2} \frac{(\mu_i - \rho)^2 \left( \frac{\partial \Phi(t, W, i)}{\partial W} \right)^2}{\sigma_i^2 W \frac{\partial^2 \Phi(t, W, i)}{\partial W^2}} \\
- \left( \exp(\gamma t) \frac{\partial \Phi(t, W, i)}{\partial W} \right)^\frac{1}{\gamma - 1} \frac{\partial \Phi(t, W, i)}{\partial W} \\
+ \exp(-\gamma t)U \left( \left( \exp(\gamma t) \frac{\partial \Phi(t, W, i)}{\partial W} \right)^\frac{1}{\gamma - 1} \right) \\
+ \sum_{j=1}^{n} \lambda_{ij} \Phi(t, W, j) = 0
\]

Let \( U(c) = \frac{c^r}{r} \), then

246

Brazilian Review of Econometrics 24(2) November 2004
\[
\begin{align*}
\frac{\partial \Phi(t, W, i)}{\partial t} & + \rho W \frac{\partial \Phi(t, W, i)}{\partial W} - \frac{1}{2} \left( \frac{\partial^2 \Phi(t, W, i)}{\partial W^2} \right) \\
& - \left( \exp(\gamma t) \frac{\partial \Phi(t, W, i)}{\partial W} \right)^{r-1} \frac{\partial \Phi(t, W, i)}{\partial W} \\
& + \frac{1}{r} \exp(-\gamma t) \left( \exp(\gamma t) \frac{\partial \Phi(t, W, i)}{\partial W} \right)^{\frac{r}{r-1}} \\
& + \sum_{j=1}^{n} \lambda_{ij} \Phi(t, W, i) 
\end{align*}
\]

If \( \Phi(t, W, i) \neq \frac{W^r}{r} \), then one gets

\[
\begin{align*}
\frac{df(t, i)}{dt} & = \left( \rho r + \frac{1}{2} \frac{(\mu_i - \rho)^2 r}{\sigma_i^2 (1 - r)} \right) f(t, i) \\
& + \sum_{j=1}^{n} \lambda_{ij} f(t, i),
\end{align*}
\]

Let also \( f(t, i) \),

\[
\begin{align*}
\frac{dg(t, i)}{dt} & = \frac{1}{1 - r} \left( \rho r + \frac{1}{2} \frac{(\mu_i - \rho)^2 r}{\sigma_i^2 (1 - r)} - \gamma \right) g(t, i) \\
& + \frac{1}{1 - r} \sum_{j=1}^{n} \lambda_{ij} g^{1-r}(t, i),
\end{align*}
\]

with final condition \( g(T, i) = 1 \), for \( i = 1, \ldots, n \), which satisfies lemmas A.1 and A.2.