CONVERGENCE FROM DISCRETE TO CONTINUOUS TIME FINANCIAL MODELS*

Hua He

Haas School of Business
University of California, Berkeley
Berkeley, CA 94720

October 12, 1989

*This paper is based on the second chapter of my Ph.D. dissertation at MIT. I thank Ayman Hindy and Mark Rubinstein for comments, Chi-fu Huang and Robert Merton for numerous helpful conversations and comments, and Neil Pearson for comments and editorial assistance. Financial support from the Berkeley Program in Finance is gratefully acknowledged. Address reprint requests to Hua He, Haas School of Business, University of California at Berkeley, Berkeley, CA 94720.
Abstract

This paper generalizes the Cox, Ross and Rubinstein (1979) binomial option pricing model, and develops a theory of convergence from discrete time multivariate multinomial models to continuous time multi-dimensional diffusion models in the context of contingent claim pricing and consumption-portfolio choice. The key to our approach is to approximate the \( N \)-dimensional diffusion price process by a sequence of \( N \)-variate, \( N + 1 \)-nomial processes. It is shown that contingent claim prices and optimal consumption-portfolio policies derived from the discrete time models converge to their corresponding continuous time limits. In contrast to solving a partial differential equation (PDE) with possibly more than one state variable, this approach provides a simpler numerical procedure for computing contingent claim prices and optimal consumption-portfolio policies.
1 Introduction

To model securities markets and trading behaviors, continuous and discrete time models represent two alternative approaches. While the continuous time approach has been prevalent in a large part of financial economic theory, it is widely understood that continuous and discrete time models are approximations of each other [Kreps (1982) and Merton (1982)]. In their seminal paper, Cox, Ross and Rubinstein (CRR) (1979) establish the convergence of certain binomial processes to geometric Brownian motions and show that the Black-Scholes (1973) option pricing formula is a limit of the discrete time binomial option pricing formula.\(^1\) Indeed, the binomial option pricing technique provides an extremely powerful tool for evaluating various types of derivative securities.

Recent papers by Evnine (1983), Boyle (1988), Cheyette (1988), Hull and White (1988), Madan, Milne and Shafrin (1988), and Boyle, Evnine and Gibbs (1989) have attempted to generalize the CRR binomial model in order to approximate a multi-dimensional geometric Brownian motion. With the exception of the paper by Cheyette, none of these have provided satisfactory answers.\(^2\) While convergence results have been established for option prices or more generally contingent claim prices, much less is known about the convergence of optimal consumption and portfolio policies.

In this paper we develop a theory of the convergence of discrete time multivariate multinomial models to general continuous time multi-dimensional diffusion models for both contingent claim prices and optimal consumption-portfolio policies. Similar to CRR we use weak convergence as our convergence tool. The diffusion model considered in this paper consists of \(N\) risky stocks and one riskless bond where the stocks and the bond form a dynamically complete securities market. We approximate the \(N\)-dimensional diffusion process for stock prices by a sequence of \(N\)-variate, \(N + 1\)-nominal processes. Thus the stocks and the bond in the discrete time models also form a dynamically complete securities market. Dynamic completeness allows us to price contingent claims whose payoffs depend upon the values of the assets in general ways.

We show that contingent claim prices and replicating portfolio strategies derived from the discrete time models converge weakly to the corresponding contingent claim price and replicating portfolio strategy of the limiting continuous time model. In particular, the discrete time contingent claim prices and replicating portfolio strategies calculated at time 0 converge (numerically) to the corresponding continuous time limits. We also show that optimal consumption policies, optimal

\(^1\)The binomial or two state variable approach was also developed independently in Sharp (1978) and Rendleman and Bartter (1979).

\(^2\)A number of trinomial and multinomial models have been proposed by these authors. An important distinction between our approximations and those of other authors is that others usually let the discrete time price processes converge to the corresponding continuous time price process under the risk neutral probability measure, i.e., the probability measure under which the expected return for the risky asset is equal to the riskless rate. This allows them to price options by taking expectations.
portfolio policies and optimal expected utilities derived from the discrete time models converge weakly to their corresponding continuous time limits.

There are two reasons for establishing such a convergence theory. The first one is the intuitive appeal of discrete time models. While the economic intuition that continuous trading in stocks and bond can span infinitely many states of nature is difficult to convey without using advanced mathematics, the simple discrete time models provide an easy way of explaining how uncertainties are resolved in the continuous time model. The second reason is that the discrete time models provide an elegant numerical alternative to the partial differential equations (PDE) obtained in continuous time models. In contrast to solving a PDE with possibly more than one state variable, the multinomial approximation approach provides a simple numerical procedure for computing contingent claim prices and optimal consumption-portfolio policies.

The question of approximating a price system with two stocks and one bond where the stock prices follow two correlated geometric Brownian motions has been of interest for a long time. Intuitively one would think that if one geometric Brownian motion can be approximated by one binomial process, then two geometric Brownian motions should be approximated by two binomial processes. This leads to a multinomial process with four uncertain states following each trading date. Since there are four uncertain states and only two stocks and one bond are available for trading, markets cannot be completed by dynamic trading and options cannot be priced by arbitrage. This phenomenon is counter to the intuition of the familiar continuous time model in which markets can be completed by continuous trading in two stocks and one bond.

To overcome this problem, Evnine (1983) proposes a "multiple" binomial model which approximates the increments of two geometric Brownian motions by three sequential moves. His idea is first to let the value of the first stock move stochastically while letting the value of the other grow at the riskless rate. He then lets the value of the second stock move stochastically while letting the value of the first grow at the riskless rate. Finally, he lets the values of both stocks move together in order to capture the correlation. While Evnine manages to show that the discrete price process matches the continuous price process in distribution in the limit, the dynamic portfolio trading strategy implied by his model is always indeterminant because the return on one of the two stocks is correlated perfectly with the return on the riskless bond. Madan, Milne and Shefrin (1988) construct an \( N + 1 \)-nomial process for \( N \) stocks such that the discrete time price process for each individual stock converges weakly to a one-dimensional geometric Brownian motion. Since they fail to specify the correlations among different assets and establish joint convergence for \( N \) stock prices, their model does not imply convergence for general contingent claims prices, such as an option on the maximum of two stocks.
In this paper we resolve this controversy by showing that an $N$-dimensional diffusion process for stock prices can be approximated by an $N$-variate, $N + 1$-nomial process.\textsuperscript{3} We utilize the fact that the increments of $N$ independent Brownian motions can be approximated by $N$ uncorrelated, not necessarily independent, random variables. For example, we can use a trinomial model to approximate two geometric Brownian motions. A crucial distinction between our approximation and that of Madan, Milne and Shefrin is that the implicit Arrow-Debreu state price processes derived from the discrete time models converge to the corresponding continuous time limit.

While we establish convergence results for contingent claim prices as many others do, we also establish convergence results for optimal consumption and portfolio policies. Convergence theory as such is of theoretical interest, because it enhances our understanding of the relationship between discrete and continuous time models. In addition, it provides a simpler numerical procedure for finding optimal consumption and portfolio strategies.

Computationally, our method is also likely to perform better than those proposed by others. For example, in the case of two-dimensional geometric Brownian motions, the number of nodes in our model grows at a rate of $n^2$ ($n$ is the number of time steps), which is significantly slower than the rate of growth in the models of Hull and White (1988) and Boyle, Evnine and Gibbs (1989).

Other related work has been done by Cheyette (1988) in which he approximates the returns of $N$ assets by an $N + 1$-nomial process, where the asset prices follow an $N$-dimensional geometric Brownian motion. Similar to our model, the complete market property of the continuous time model is also preserved in the discrete time models and the convergence of contingent claim prices is achieved. While Cheyette's approach is similar to that of this paper, we emphasize that his method only applies to the geometric Brownian motion case, and his approximation procedure as well as his proof of convergence are less general and less elegant than the ones we provide in this paper. In particular, our results are based upon the convergence of Arrow-Debreu state prices, which he does not explore.

The rest of this paper is organized as follows. Section 2 sets out the definition of weak convergence and derives some of the basic properties of weak convergence to be used in later sections. Section 3 presents a multivariate multinomial approximation to the multi-dimensional diffusion process for stock prices and the bond price. Section 4 deals with the convergence of contingent claim prices and replicating portfolio strategies, while Section 5 establishes the convergence of optimal consumption-portfolio policies and optimal expected utilities. We conclude this paper in Section 6 with some comments and suggestions for future research.

\textsuperscript{3}The number $N$ is the martingale multiplicity discussed in Duffie and Huang (1985). When an $N + 1$-nomial process is used the martingale multiplicities in both discrete and continuous time models are exactly the same.
2 Preliminaries on Weak Convergence

This section sets out the definition of weak convergence and derives some of the basic properties used in later sections. In particular, we present a martingale central limit theorem to be used as a tool for establishing weak convergence from a sequence of Markov chains to a diffusion process.

We take the sample space to be $D^M[0,1]$ which is the space of functions from $[0,1]$ to $\mathbb{R}^M$ that are right-continuous with left limits (RCCLL). Since diffusion processes have continuous sample paths and the multinomial processes to be considered in later sections are piecewise constant, the choice of $D^M[0,1]$ as our sample space is appropriate. The topology to be used for $D^M[0,1]$ is the Skorohod topology discussed by Billingsley (1968). The Borel field generated by the Skorohod topology is denoted by $\mathcal{D}$. Let $X^n$ be a sequence of stochastic processes in $D^M[0,1]$, and let $P_n$ be the probability measure on $D^M[0,1]$ associated with $X^n$, i.e., for any $A \in \mathcal{D}$, $P_n(A)$ is the probability of event $\{X^n \in A\}$. A sequence of stochastic processes $X^n$ in $D^M[0,1]$ is said to converge to $X$ in $D^M[0,1]$ weakly or in distribution, denoted by $X^n \Rightarrow X$, if for any bounded continuous function $h$ mapping from $D^M[0,1]$ to $\mathbb{R}$, we have $E_n[h(X^n)] \to E[h(X)]$ where $E_n$ denotes the expectation under $P_n$ and $E$ denotes the expectation under the probability $P$ associated with $X$.

Since the multinomial processes to be used in later sections are Markov chains we would like to have a mathematical tool for establishing the weak convergence of a sequence of Markov chains to a diffusion process. The martingale central limit theorem in Ethier and Kurtz (1986) is exactly what we need. Let $X$ be determined by the stochastic differential equation

$$dX_t = b(X_t) \, dt + \sigma(X_t) \, dw_t,$$

(1)

where $b(x)$ is an $M \times 1$ vector, $\sigma(x)$ is an $M \times N$ matrix for all $x \in \mathbb{R}^M$, and $w$ is an $N$-dimensional Brownian motion. For the balance of the paper, we use $A^\top$ to denote the transpose of matrix $A$ and $|A| \equiv \text{trace}(A^\top A)^{\frac{1}{2}}$ to denote the norm of matrix $A$. The following proposition gives the martingale central limit theorem, the proof of which can be found in Ch. 7, Ethier and Kurtz (1986).

Proposition 1 (Martingale Central Limit Theorem) Let $b$ and $\sigma$ be continuous such that (1) admits a unique weak solution.\(^4\) Let $X$ be the weak solution to (1) with $X_0 = x_0$. Suppose

\(^4\)Equation (1) is said to have a weak solution $X$ with initial distribution $\mu$ if there exists $(\Omega, \{\mathcal{F}_t\}, P)$ such that $\omega$ is an $(\mathcal{F}_t)$-Brownian motion, $X_0$ has law $\mu$, and $X$ satisfies

$$X_t = X_0 + \int_0^t b(X_s) \, ds + \int_0^t \sigma(X_s) \, dw_s,$$

where all of the stochastic integrals are well-defined. The weak solution to (1) is said to be unique if whenever $\{X_t, t \geq 0\}$ and $\{\tilde{X}_t, t \geq 0\}$ are two solutions such that the distributions of $X_0$ and $\tilde{X}_0$ are the same, then the distributions of $X_t$ and $\tilde{X}_t$ are the same [Rogers and Williams (1987)].
$X^n$ is a sequence of Markov processes with sample paths in $D^M[0,1]$. Let $L^n$ and $A^n$ be $N \times 1$ and (symmetric) $N \times N$ matrix-valued processes respectively, such that each of their elements has sample path in $D^1[0,1]$ and $A_t^n - A_s^n$ is non-negative definite for $t > s \geq 0$. Define $\tau^q_n = \inf\{t \leq T : |X_t^n| \geq q \text{ or } |X^n_t - X^n_s| \geq q\}$, and suppose further that

a) $X^n_0 \rightarrow x_0$ in distribution;

b) $M^n \equiv X^n - L^n$ and $M^nM^n^\top - A^n$ are martingales;

c) for all $q > 0$,

$$\lim_{n \to \infty} E_n \left[ \sup_{t \leq \tau^q_n} |X^n_t - X^n_{t-}|^2 \right] = 0,$$

$$\lim_{n \to \infty} E_n \left[ \sup_{t \leq \tau^q_n} |L^n_t - L^n_{t-}|^2 \right] = 0,$$

$$\lim_{n \to \infty} E_n \left[ \sup_{t \leq \tau^q_n} |A^n_t - A^n_{t-}| \right] = 0;$$

d) $\sup_{t \leq \tau^q_n} |L^n_t - \int_0^t b(X^n_s) ds| \rightarrow 0$ in probability for all $q > 0$, as $n \to \infty$,

$$\sup_{t \leq \tau^q_n} |A^n_t - \int_0^t a(X^n_s) ds| \rightarrow 0$$

in probability for all $q > 0$, as $n \to \infty$.

Then $X^n \Rightarrow X$.

Either of the following two conditions are sufficient for (1) to have a unique weak solution [Rogers and Williams (1987)]:

i) $b$ and $\sigma$ satisfy a uniform Lipschitz condition, i.e., there exists a constant $L > 0$ such that for all $x, y \in \mathbb{R}^M$,

$$|b(x) - b(y)| + |\sigma(x) - \sigma(y)| \leq L|x - y|;$$

ii) $a(x) = \sigma(x)\sigma(x)^\top$ is strictly positive definite for each $x \in \mathbb{R}^M$ and there exists a constant $K$ such that for all $x \in \mathbb{R}^M$,

$$|a(x)| \leq K(1 + |x|^2), \quad |b(x)| \leq K(1 + |x|).$$

Before leaving this section we record two lemmas to be used below. The proofs of these can be found in Billingsley (1968).

**Lemma 1 (Continuous Mapping Theorem)** Assume that $X^n \Rightarrow X$.

i) Suppose that $h(x)$ is a continuous function mapping from $\mathbb{R}^N$ to $\mathbb{R}^M$, except for at most countably many points, and let $Y^n_t = h(X^n_t)$ and $Y_t = h(X_t)$. Then $Y^n \Rightarrow Y$. 

5
ii) Suppose that $H$ is a continuous mapping from $D^M[0, 1]$ to $\mathbb{R}$. Then $H(X^n)$ converges to $H(X)$ in distribution.

Lemma 2 Let $X_n$ and $X \in \mathbb{R}^1$ be random variables and suppose that $X_n \rightarrow X$ in distribution as $n \rightarrow \infty$. If $\{X_n\}$ is uniformly integrable, then $E_n(X_n) \rightarrow E(X)$ as $n \rightarrow \infty$.

3 Multinomial Approximation

In this section we construct a sequence of $N$-variate, $N+1$-nomial processes for stock prices that converge weakly to the $N$-dimensional diffusion price process of interest. Moreover, we show that the sequence of implicit Arrow-Debreu state price processes derived from the discrete time model converges weakly to the corresponding continuous time limit. This result plays an important role in establishing the weak convergence of contingent claim prices and optimal consumption-portfolio policies in later sections.

We consider a securities market consisting of $N$ risky stocks and one locally riskless bond. The price processes for the stocks and the bond are described by the stochastic differential equations

$$dS_t = b(S_t) \ dt + \sigma(S_t) \ d\omega_t,$$

$$dB_t = B_t \ r(S_t) \ dt, \quad B_0 = 1,$$

where $\omega_t$ is an $N$-dimensional standard Brownian motion defined on a probability space $(\Omega, \mathcal{F}, P)$, $S_t$ is an $N$-dimensional vector of the stock prices, and $B_t$ is the price of the bond. We assume that the functions $b : \mathbb{R}^N \rightarrow \mathbb{R}^{N \times 1}$, $\sigma : \mathbb{R}^N \rightarrow \mathbb{R}^{N \times N}$ and $r : \mathbb{R}^N \rightarrow \mathbb{R}$ are continuous and that $\sigma$ is non-singular. We assume further that $b$ and $\sigma$ satisfy the uniform Lipschitz condition defined in the previous section. This implies that $b$ and $\sigma$ also satisfy a linear growth condition, i.e., there exists a constant such that for all $z \in \mathbb{R}^N$

$$|b(z)| + |\sigma(z)| \leq K(1 + |z|).$$

We assume that $r$ is non-negative and that the time span is $[0, 1]$. The regularity conditions imposed on the drift and the diffusion terms are satisfied for the Black-Scholes price system.

As in the Black-Scholes economy, we assume that markets are dynamically complete. That is, contingent claims such as an option written on the stocks can be created by dynamic trading in the stocks and the bond.\footnote{Rigorously speaking, dynamic completeness requires the existence of a unique equivalent martingale measure and a proper choice of the space of feasible trading strategies [see Cox and Huang (1987) for details].} To rule out arbitrage opportunities, we assume that there exists a unique equivalent martingale measure or a risk neutral probability measure for the price system.
defined by (2) and (3). The martingale measure, denoted by \( Q \), has the following form. Let
\[
\kappa(S_t) = -\sigma(S_t)^{-1}(b(S_t) - r_t S_t)
\]
and \( \xi_t = \exp\left( \int_t^T \kappa_s dw_s - \frac{1}{2} \int_t^T |\kappa_s|^2 ds \right) \). Then the probability \( Q \) is defined by
\[
Q(A) = \int_A \xi_t(\omega) dP.
\]
One can easily verify that the stochastic process \( \{\xi_t\} \) satisfies the stochastic differential equation
\[
d\xi_t = \kappa(S_t) \xi_t \, dw_t, \quad \xi_0 = 1.
\]
(4)

In the literature \( \xi_t(\omega) \) is usually interpreted as the implicit Arrow-Debreu state price per unit of probability for a security that pays off one unit of the consumption good at time \( t \) and state \( \omega \) and nothing otherwise [see Cox and Huang (1987)].

We now proceed with the construction of a sequence of multivariate \( N + 1 \)-nomial processes. The basic idea is to approximate the increments of \( N \) independent Brownian motions by \( N \) uncorrelated random variables. We first construct the random variables.

Let \( A \) be an \((N + 1) \times (N + 1)\), real, orthogonal matrix such that the last column of \( A \) is \((1/\sqrt{N + 1}, \ldots, 1/\sqrt{N + 1})^T \) (such a matrix always exists!). Define a sequence of \( N \)-dimensional, independent and identically distributed random vectors, \( \tilde{\xi}_k = (\tilde{\xi}_k^1, \ldots, \tilde{\xi}_k^N)^T \), \( k = 1, 2, \ldots, n \), as follows. Let \( \Omega_k = \{\omega_k^1, \ldots, \omega_k^{N+1}\} \) be the sample space on which \( \tilde{\xi}_k \) is defined. For \( j = 1, \ldots, N \), let
\[
P[\tilde{\xi}_k(\omega_k) = a_{s,j} \sqrt{N + 1}] = \frac{1}{N + 1}, \quad s = 1, \ldots, N + 1,
\]
where \( a_{s,j} \) is the \( s \)-th element in \( j \)-th column of \( A \). That is, we assign equal probability to each state \( \omega_k^j \). It is easy to verify that, for fix \( k \), \( \tilde{\xi}_k^1, \ldots, \tilde{\xi}_k^N \) are uncorrelated and have mean zero and variance 1.

Next, we divide \([0, 1]\) into \( n \) equally spaced subintervals, each with length \( \frac{1}{n} \). Define a sequence of discrete time stock and bond prices \((\tilde{S}_t^n, \tilde{B}_t^n)\) as follows. Let \( S_t^n \) and \( B_t^n \) be determined by the following stochastic difference equations:
\[
S_{k+1}^n = S_k^n + b(S_k^n) \frac{1}{n} + \sigma^1(S_k^n) \frac{\tilde{\xi}_k^1}{\sqrt{n}} + \cdots + \sigma^N(S_k^n) \frac{\tilde{\xi}_k^N}{\sqrt{n}},
\]
(5)
\[
B_{k+1}^n = B_k^n \left( 1 + \frac{r(S_k^n)}{n} \right),
\]
(6)
\[
S_0^n = S_0, \quad B_0^n = 1,
\]
where \( \sigma^j \) is the \( j \)-th column of \( \sigma \). Letting \( \tilde{B}_t^n = B_{[nt]}^n \) and \( \tilde{S}_t^n = S_{[nt]}^n \), the processes \( \tilde{B}_t^n \) and \( \tilde{S}_t^n \) are well-defined for all \( t \in [0, 1] \) and their sample paths are right-continuous with left limits. Moreover, the sample paths are piecewise constant and jump only at \( t = \frac{k}{n} \). Equations (5) and (6) can be viewed as finite difference approximations to the stochastic differential equations (2) and (3).
Since there are $N + 1$ assets traded and $N + 1$ possible uncertain states following each trading date, markets are dynamically complete. The unique Arrow-Debreu state price $\pi s^n_k(\omega^k_s)$ at time $\frac{k}{n}$ and stock price $S^n_k$ for a security that pays off one unit of the consumption good at time $\frac{k+1}{n}$ and state $\omega^k_s$ but nothing otherwise must satisfy the relation

$$\sum_{s=1}^{N+1} \pi s^n_k(\omega^k_s) S^n_{k+1}(\omega^k_s) = S^n_k. \tag{7}$$

Solving this equation by substituting (5) into (7), we obtain

$$\pi s^n_k(\omega^k_s) = \frac{1}{N+1} \left( 1 + \frac{\kappa(S^n_k)^T}{\sqrt{n}} \tilde{c}_k(\omega^k_s) \right) \left( 1 + \frac{r(S^n_k)}{n} \right)^{-1}. \tag{8}$$

Note that $\pi s^n_k$ is the one period Arrow-Debreu state price (from time $\frac{k}{n}$ to time $\frac{k+1}{n}$). To obtain the Arrow-Debreu state price at time 0 for a security that pays off one unit of the consumption good at time $\frac{k}{n}$, we need to multiply together all of the one period Arrow-Debreu state prices from period 1 to period $k$, where period $i$ is from time $\frac{i-1}{n}$ to time $\frac{i}{n}$. Letting $\pi_k$ denote this state price, then $\pi^n_k = \pi s^n_{k-1} \pi s^n_{k-2} \cdots \pi s^n_0$ for $k \geq 1$, and $\pi^n_0 = 1$. Moreover, $\pi^n_k$ satisfies the stochastic difference equation

$$\pi^n_{k+1} = \frac{\pi^n_k}{N+1} \left( 1 + \frac{\kappa(S^n_k)^T}{\sqrt{n}} \tilde{c}_k(\omega^k_s) \right) \left( 1 + \frac{r(S^n_k)}{n} \right)^{-1}. \tag{9}$$

We assume that all of the $\pi$'s are non-negative for sufficiently large $n$. A sufficient condition for this to be true is that $\kappa$ is bounded.6

To relate $\{\pi^n_k\}$ to the implicit Arrow-Debreu state price process $\{\xi_t\}$ defined in (4), we introduce a new variable $\xi^n_k$,

$$\xi^n_k = \pi^n_k B^n_k(N + 1)^k. \tag{10}$$

One can verify that $E_n[\xi^n_n] = 1$ using the fact that $\sum_{s=1}^{N+1} \pi s^n_k(\omega^k_s) = (1 + \frac{r(S^n_k)}{n})^{-1}$. Now, defining

$$Q^n(A) = \int_A \xi^n_n P_n(d\omega), \tag{11}$$

then (7) implies that $Q^n$ is an equivalent martingale measure, i.e., the discounted stock price processes become martingales under this measure. We therefore call $\xi^n_k$ the implicit Arrow-Debreu state price (per unit of the probability) as we did $\xi_t$ in the continuous time case. The implicit state price also can be represented by a stochastic difference equation

$$\xi^n_{k+1} = \xi^n_k \left( 1 + \kappa(S^n_k)^T \tilde{c}_k \right)^{-1}. \tag{12}$$

---

6When this condition is not satisfied, we need to add a higher order term to the right hand side of (5) so that the resulting price system admits no arbitrage opportunities while still permitting weak convergence for prices and Arrow-Debreu state prices.
For \( t \in [0, 1] \), set \( \tilde{\xi}_t^n = \xi_{\lfloor t \rfloor} \). Our first claim for this \( N+1 \)-nomial approximation is the following convergence theorem on asset prices and the implicit Arrow-Debreu state prices.

**Theorem 1** Let \( \tilde{X}^n = (\tilde{S}^n, \tilde{B}^n, \tilde{\xi}^n) \) and \( X = (S, B, \xi) \). Then \( \tilde{X}^n \Rightarrow X \).

**Proof.** First, we argue that the uniform Lipschitz condition and the linear growth condition for \( b \) and \( \sigma \) guarantee that the weak solution of (2) and (3) is unique. As a result, the weak solution of (4) is also unique. Next, letting

\[
L_t^n = \begin{pmatrix}
\int_0^t b(\tilde{S}_s^n)ds \\
\int_0^t r(\tilde{S}_s^n)\tilde{B}_s^n ds \\
0
\end{pmatrix},
\]

\[
A_t^n = \begin{pmatrix}
\int_0^t \sigma(\tilde{S}_s^n)\sigma^\top(\tilde{S}_s^n)ds & 0 & -\int_0^t \sigma(\tilde{S}_s^n)\kappa(\tilde{S}_s^n)\tilde{\xi}_s^n ds \\
0 & 0 & 0 \\
-\int_0^t \kappa(\tilde{S}_s^n)^\top\sigma(\tilde{S}_s^n)^\top\tilde{\xi}_s^n ds & 0 & \int_0^t |\kappa(\tilde{S}_s^n)\tilde{\xi}_s^n|^2 ds
\end{pmatrix},
\]

then the matrix \( A_t^n - A_s^n \) is non-negative definite for \( t \geq s \), and \( M^n \equiv X^n - L^n \). \( M^nM^n^\top - A^n \) are martingales. If we can verify that conditions c) and d) of Proposition 1 are satisfied, we can conclude that \( \tilde{X}^n \) converges to \( X \) weakly.

To verify c), we observe that with the control of stopping time and the continuity of \( r, b, \sigma \) and \( \kappa \), \( |X_t^n - X_t^n| \) is of order \( n^{-1/2} \), \( |L_t^n - L_t^n| \) is of order \( n^{-1} \) and \( |A_t^n - A_t^n| \) is of order \( n^{-1} \). Thus, c) is satisfied. A similar argument applies to d). This completes the proof. \( \blacksquare \)

**Remark 1** a) This proposition implies that the local movements of \( N \) one-dimensional and independent Brownian motions can be approximated by \( N \) uncorrelated but possibly dependent random variables \( \{ \tilde{\xi}_k, j = 1, \ldots, N \} \).

b) \( N + 1 \) is the minimum number of branches one can allow in order to keep \( \{ \tilde{\xi}_k \}, j = 1, \ldots, N \) mutually uncorrelated.

As our first example we consider the Black-Scholes price system with one stock and one bond \( (N = 1) \). The Black-Scholes price system is

\[
B_t = e^{rt}, \quad S_t = S_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma \sigma(t)}.
\]

If we choose \( A \) to be

\[
A = \begin{pmatrix}
\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\
-\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}
\end{pmatrix}
\]

9
we obtain the binomial approximation

\[ S_{k+1}^n = \begin{cases} S_k^n + \mu S_k^n \frac{n}{\sqrt{n}} + \frac{\sigma S_k^n}{\sqrt{n}} \\ S_k^n + \mu S_k^n \frac{n}{\sqrt{n}} - \frac{\sigma S_k^n}{\sqrt{n}} \end{cases} \]

More generally, consider the following price system with two correlated lognormal processes,

\[
\begin{align*}
    dS_{t,1} & = \mu_1 S_{t,1} dt + \sigma_1 S_{t,1} dw_1, \\
    dS_{t,2} & = \mu_2 S_{t,2} dt + \rho \sigma_2 S_{t,2} dw_1 + \sigma_2 \sqrt{1 - \rho^2} S_{t,2} dw_2,
\end{align*}
\]

where \( \sigma_i^2 \) is the volatility of the return on the \( i \)-th asset, and \( \rho \) is the correlation coefficient of the returns on these two assets. If we choose \( A \) to be

\[
    A = \begin{pmatrix}
        \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\
        0 & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\
        -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}}
    \end{pmatrix},
\]

we obtain the trinomial approximation

\[
\begin{align*}
    S_{k+1,1}^n & = \begin{cases} S_{k,1}^n + \frac{\mu_1 S_{k,1}^n}{n} + \sigma_1 S_{k,1}^n \frac{\sqrt{3}}{2n} \\
                     S_{k,1}^n + \frac{\mu_1 S_{k,1}^n}{n} \\
                     S_{k,1}^n + \frac{\mu_1 S_{k,1}^n}{n} - \sigma_1 S_{k,1}^n \frac{\sqrt{3}}{2n} \end{cases} \\
    S_{k+1,2}^n & = \begin{cases} S_{k,2}^n + \frac{\mu_2 S_{k,2}^n}{n} + \sigma_1 S_{k,2}^n \frac{\sqrt{3}}{2n} + \sigma_2 \sqrt{1 - \rho^2} S_{k,2}^n \frac{1}{2n} \\
                     S_{k,2}^n + \frac{\mu_2 S_{k,2}^n}{n} - \sigma_2 \sqrt{1 - \rho^2} S_{k,2}^n \frac{1}{2n} \\
                     S_{k,2}^n + \frac{\mu_2 S_{k,2}^n}{n} - \sigma_1 S_{k,2}^n \frac{\sqrt{3}}{2n} + \sigma_2 \sqrt{1 - \rho^2} S_{k,2}^n \frac{1}{2n} \end{cases}
\end{align*}
\]

The choice of such an orthogonal matrix \( A \) is not unique. In fact, there exist infinitely many such matrices satisfying the requirements stated earlier. Also, it is important to point out that by collapsing those states that have same levels of stock prices, the number of nodes in this case grows at a rate \((n + 1)(n + 2)/2\), which is significantly slower than the rate of growth in the models proposed by Hull and White (1988) and Boyle, Evnine, Gibbs (1989). This suggests that our approximation scheme could be more efficient for computation.
4 Contingent Claim Pricing

An immediate application of Theorem 1 is the convergence of contingent claim prices. We demonstrate in this section that a contingent claim price process obtained from the discrete time model based on a no arbitrage argument converges weakly to its continuous time counterpart. Moreover, the dynamic portfolio strategy that replicates the payoff of the underlying contingent claim also converges weakly. It then follows that the contingent claim price and the replicating portfolio strategy computed at time 0 converge (numerically) to the corresponding continuous time limits.

We begin with a definition of a contingent claim. Let $g$ be a measurable function mapping from $\mathbb{R}^N$ to $\mathbb{R}$. A contingent claim on the stock price process $S$ is defined to be a security which pays $g(S_t)$ units of consumption good at the final date (i.e. $t = 1$). This formulation subsumes all of the usual examples. For example, it includes as special cases an option on one stock, $\max(S^{(1)}_t - K, 0)$, and an option on the maximum of the two stocks, $\max(\max(S^{(1)}_t, S^{(2)}_t) - K, 0)$, where $S^{(i)}_t$ denotes the price of the $i$-th security at time $t$.

Following Harrison and Kreps (1979), the price of a contingent claim at time $t$ can be evaluated by taking the conditional expectation of the discounted final payoff under the equivalent martingale measure $Q$, i.e.,

$$V(S_t, t) = E_Q \left[ e^{-\int_t^1 r(s, r)dr} g(S_1) | S_t \right] = E \left[ \xi_t e^{-\int_t^1 r(s, r)dr} g(S_1) | S_t \right].$$

(9)

Alternatively, if $V$ is continuously differentiable with respect to $t$ and twice continuously differentiable with respect to $S$, then $V$ can be determined as a solution to the following partial differential equation [see Cox, Ingersoll and Ross (1985)]:

$$\frac{1}{2} \text{trace}[\sigma^T V_{SS} \sigma] + r S^T V_S + V_t - rV = 0,$$

(10)

$$V(S, 1) = g(S),$$

where

$$V_S(S, t) = \left[ \frac{\partial V(S, t)}{\partial S_1}, \cdots, \frac{\partial V(S, t)}{\partial S_N} \right]^T,$$

$$V_{SS}(S, t) = \left[ \frac{\partial^2 V(S, t)}{\partial S_i \partial S_j} \right]_{N \times N}.$$
Equation (10) is usually called the fundamental partial differential equation in the option pricing literature. The dynamic portfolio strategy which replicates the final payoff of this claim is given by

$$\theta(S_t, t) = V(S_t, t),$$

$$\alpha(S_t, t) = (V(S_t, t) - \theta(S_t, t)^T S_t) / B_t,$$

where $\theta_i$ and $\alpha$ denote the number of shares held in the $i$-th stock and bond respectively.

This valuation technique can be easily applied to the discrete time model. The price of the contingent claim at any time $\frac{k}{n}$, denoted by $V^n$, can be evaluated by taking the conditional expectation under the equivalent martingale measure $Q^n$, i.e.,

$$V^n(S^n_k, \frac{k}{n}) = E^n \left[ g(S^n_k) B^n_k / B^n_n \mid S^n_k \right]$$

$$= E^n \left[ \frac{\xi^n_k}{\xi^n_n} g(S^n_k) B^n_k / B^n_n \mid S^n_n \right].$$

(11)

Alternatively, $V^n$ can be determined as a solution to the recurrent equation

$$V^n(S^n_k, \frac{k}{n}) = \sum_{i=1}^{N+1} \pi S^n_k(\omega_s^k) V^n(S^n_{k+1}(\omega_s^k), \frac{k+1}{n}),$$

$$V^n(S^n_n, 1) = g(S^n_n).$$

(12)

The dynamic portfolio strategy that replicates the final payoff of this claim is determined by the following system of linear equations:

$$\alpha^n_k B^n_{k+1}(\omega^k_s) + \theta^n_{k,1} S^n_{k+1,1}(\omega^k_s) + \cdots + \theta^n_{k,N} S^n_{k+1,N}(\omega^k_s) = V(S^n_{k+1}(\omega^k_s), \frac{k+1}{n}),$$

for $s = 1, \cdots, N+1$. Since there are $N+1$ equations and $N+1$ unknowns, the solution for $(\alpha^n_k, \theta^n_k)$ is uniquely determined. Moreover, $\alpha^n_k$ and $\theta^n_k$ are functions of the stock prices $S^n_k$ and time $\frac{k}{n}$. The following theorem establishes the convergence of contingent claim prices from discrete time models to the continuous time diffusion model. We need a definition to begin.

A function $f : \mathbb{R}^N \rightarrow \mathbb{R}^M$ is said to satisfy a polynomial growth condition if there exist constants $K, \beta > 0$ such that

$$|f(x)| \leq K(1 + |x|^\beta), \quad \forall x \in \mathbb{R}^N.$$

**Theorem 2 (Contingent Claim Pricing)** Suppose the following:

(a) $r(x)$ and $r(x)x$ satisfy a uniform Lipschitz condition for $x \in \mathbb{R}^N$, and $r(x), \sigma(x)$ are continuously differentiable w.r.t. $x$ up to the sixth order with all of these derivatives satisfying a polynomial growth condition; and
(b) \( g(x) \) is piecewise and continuously differentiable up to the sixth order with all of these derivatives satisfying a polynomial growth condition.

Then,

1. Letting \( V_t = V(S_t, t) \), \( V^n_k = V^n(S^n_k, t^n_k) \) and \( \tilde{V}^n_k = V^n_{[n]} \), we have \( \tilde{V}^n_k \to V_t \);

2. Letting \( \alpha_i = \alpha(S_t, t), \theta_i = \theta(S_t, t), \alpha^n_i = \alpha^n_{[n]} \) and \( \tilde{\theta}^n_i = \theta^n_{[n]} \), we have \( (\alpha^n_i, \tilde{\theta}^n_i) \Rightarrow (\alpha_i, \theta_i) \).

In particular, \( \tilde{V}^n_0 \to V_0 \), and \( (\tilde{\alpha}^n_0, \tilde{\theta}^n_0) \to (\alpha_0, \theta_0) \) as \( n \to \infty \).

**Proof.** The proof of this theorem is lengthy and we leave it to the Appendix.

**Remark 2.** a) The conditions imposed on the drift and the diffusion terms are readily satisfied for a price system in which the stock prices follow a multi-dimensional geometric Brownian motion and bond price grows at a constant rate of interest.

b) The piecewise differentiability assumption on \( g(x) \) allows contingent claims such as call option with truncated final payoff.

In contrast to solving a PDE such as (10) with possibly more than one state variable, this theorem provides us with a simple numerical procedure for computing the price of a contingent claim and the replicating portfolio strategy. Specifically, at \( t = 0 \), \( \tilde{V}^n_0 \), \( \tilde{\alpha}^n_0 \) and \( \tilde{\theta}^n_0 \) can be used as the numerical approximations to the corresponding continuous time limits.

## 5 Optimal Consumption-Portfolio Policies

In this section we study the convergence of optimal consumption-portfolio policies. We focus on the following three economic variables: the optimal consumption policy, the optimal portfolio policy, and the optimal expected utility. Our objective is to show that all of these variables converge weakly to their continuous time counterparts.

We first consider the convergence of optimal consumption policies. According to Cox and Huang (1987), the dynamic consumption and portfolio problem for a single individual can be formulated in terms of the following static expected utility maximization problem:

\[
\sup_{c, w \geq 0} \mathbb{E} \left[ \int_0^1 u(c_t, t) dt + V(W) \right]
\]

s.t. \( \mathbb{E} \left[ \int_0^1 \xi_t c_t / B_t dt + \xi_1 W / B_1 \right] \leq W_0 \),

\[ (13) \]

Traditionally, the dynamic consumption and portfolio problem is solved using stochastic dynamic programming [see Merton (1971)]. However, for our purposes the static approach of Cox and Huang (1987) turns out to be much easier to work with.
where \( u \) is the utility function for the consumption rate process \( \{c_t; t \in [0, 1]\} \), \( V \) is the utility function for the final wealth \( W \) which represents a bequest to the next generation, and \( u \) and \( V \) are continuous, increasing and concave. We assume that the individual’s consumption decisions \((c, W)\) depend only upon the information available at that date.\(^9\) Equation (13) is called the static budget constraint, which says that the value at time 0 of the consumption process \( \{c_t\} \) and the final wealth \( W \) is equal to the initial wealth \( W_0 \).

Under the static setup, one first solves for the optimal consumption policy and then implements it by some dynamic portfolio strategy. A nice feature of this formulation is that the optimal consumption policy can be expressed explicitly as a function of the implicit Arrow-Debreu state prices. Specifically, if we define

\[
\hat{f}(y; t) = \inf \{ x \geq 0 : u'_x(x; t) \leq y \},
\]

\[
\hat{g}(y) = \inf \{ x \geq 0 : V'_x(x) \leq y \},
\]

then applying the Lagrangian theory we can find a \( \lambda_0 > 0 \) such that

\[
c_t = \hat{f}(\lambda_0 \xi_t / B_t; t), \quad W = \hat{g}(\lambda_0 \xi_1 / B_1),
\]

where \( \lambda_0 \) is determined through the budget constraint

\[
E \left[ \int_0^1 \hat{f}(\lambda_0 \xi_t / B_t; t) \xi_t / B_t dt + \hat{g}(\lambda_0 \xi_1 / B_1) \xi_1 / B_1 \right] = W_0. \tag{14}
\]

We assume that \( \lambda_0 \) is uniquely determined.\(^{10}\) Moreover, if we introduce a new state variable \( Z_t = B_t / \lambda_0 \xi_t \), then \((S, Z)\) forms a Markov process and the optimal wealth process is a function of \( Z_t, S_t \), and \( t \). In fact, letting \( F \) denote this function, we have that

\[
F(S_t, Z_t, t) = B_t E \left[ \int_t^1 \hat{f}(Z_s^{-1}, s; B_s ds + \hat{g}(Z_1^{-1}) / B_1 | S_t, Z_t) \right]
= Z_t E \left[ \int_t^1 Z_s^{-1} \hat{f}(Z_s^{-1}, s; B_s ds + Z_1^{-1} \hat{g}(Z_1^{-1}) / B_1 | S_t, Z_t) \right]
\]

and satisfies the linear partial differential equation

\[
\frac{1}{2} \text{trace}[\sigma^T F SS \sigma] + \frac{1}{2} |Z\sigma|^2 Z FS + Z \kappa^T \sigma^T F S + r S^T F + r Z F + F_t - r F + \hat{f}(Z^{-1}, t) = 0, \tag{15}
\]

\[
F(S, Z, 1) = \hat{g}(Z^{-1}). \tag{16}
\]

\(^9\)Let \( \mathcal{F}_t \) be the smallest sigma-field containing all of the \( P \)-measure zero set with respect to which \( \{\omega_s, 0 \leq s \leq t\} \) is measurable. We require that \( c_t \) be measurable with respect to \( \mathcal{F}_t \) and \( W \) be measurable with respect to \( \mathcal{F}_t \).

\(^{10}\)A sufficient condition for this to be true is that \( u \) and \( V \) are strictly concave.
Furthermore, the dynamic portfolio strategy that produces the optimal consumption policy is determined by

$$
\theta_t = F_S + (\sigma_t^T \sigma_t)^{-1}(b_t - rS)ZF_t,
$$

$$
\alpha_t = (F - \theta_t^T S_t)/B_t.
$$

We refer the reader to Cox and Huang (1987) for details.

This treatment can be easily applied to the discrete time model. Analogously, we formulate a discrete time static consumption problem as follows:

$$
\max_{c_{k}^{n} \geq 0} E_n \left[ \sum_{k=0}^{n-1} u(c_{k}^{n}, t_{k}^{n}) \Delta t_{k}^{n} + V(W^{n}) \right]
$$

subject to

$$
E_n \left[ \sum_{k=0}^{n-1} \xi_{k}^{n} c_{k}^{n}/B_{k}^{n} \Delta t_{k}^{n} + \xi_{n}^{n} W^{n}/B_{n}^{n} \right] = W_{0},
$$

where \( t_{k}^{n} = \frac{k}{n} \) and \( \Delta t_{k}^{n} = \frac{1}{n} \). Applying the Lagrangian theory, we can find a \( \lambda_n > 0 \) such that

$$
c_{k}^{n} = \tilde{f}(\lambda_n \xi_{k}^{n}/B_{k}^{n}, t_{k}^{n}), \quad W^{n} = \hat{f}(\lambda_n \xi_{n}^{n}/B_{n}^{n}),
$$

for \( k = 0, 1, \ldots, n-1 \), where \( \lambda_n \) is determined by the budget constraint (18). Therefore, the optimal consumption policy for the discrete time model also can be expressed explicitly as a function of the implicit Arrow-Debreu state prices. Now, we define \( \tilde{\xi}_{t}^{n} = c_{[nt]}^{n} \) for \( t \in [0, 1] \). Since \( \tilde{f} \) is continuous except for at most countably many points, following Lemma 1 and Theorem 1, \( (\tilde{\xi}_{t}^{n}, W^{n}) \) converges to \( (\xi, W) \) weakly if and only if \( \lambda_n \rightarrow \lambda_0 \). The fact that \( \lambda_n \rightarrow \lambda_0 \) is proved in the following theorem.

Without loss of generality, we assume that \( V = 0 \) and \( r(t) = 0 \).

**Theorem 3 (Consumption Policies)** If \( \int_{0}^{1} \tilde{\xi}_{t}^{n} \tilde{f}(\gamma \tilde{\xi}_{t}^{n}, [nt]/n)\,dt \) is uniformly integrable for any constant \( \gamma > 0 \), then \( \tilde{\xi}_{t}^{n} \Rightarrow \xi \) as \( n \rightarrow \infty \).

**Proof.** We demonstrate \( \lambda_n \rightarrow \lambda_0 \) by contradiction. Suppose it is not true that \( \lambda_n \rightarrow \lambda_0 \), then there exists a subsequence of \( \lambda_n \), denoted \( \lambda_{n_j} \), which converges to some number \( \tilde{\lambda} \neq \lambda_0 \) (\( \tilde{\lambda} \) could be zero or infinity, but is always non-negative). Let us first assume that \( \tilde{\lambda} \) is finite and positive. According to Lemma 1, \( \tilde{\xi}_{t}^{n} \Rightarrow \xi_{t} \), where \( \xi_{t} = \hat{f}(\tilde{\lambda} \xi_{t}, t) \). Hence, \( \int_{0}^{1} \tilde{\xi}_{t}^{n} \tilde{f}_{t}^{n} \,dt \) converges to \( \int_{0}^{1} \xi_{t} \tilde{f}(\tilde{\lambda} \xi_{t}, t) \,dt \) in distribution.

If \( \tilde{\lambda} > \lambda_0 \), then there exists sufficiently small number \( \epsilon \) such that \( \lambda_{n_j} > \tilde{\lambda} - \epsilon > \lambda_0 \) for \( j > j_0 \). Hence, \( \tilde{f}(\lambda_{n_j} \xi_{t}^{n_j}, [nt]/n) \leq \tilde{f}((\tilde{\lambda} - \epsilon) \xi_{t}^{n_j}, [nt]/n) \) for \( j > j_0 \), where we have used the fact that \( \tilde{f}(x, t) \) is decreasing in \( x \). Since \( \int_{0}^{1} \tilde{\xi}_{t}^{n} \tilde{f}(\gamma \tilde{\xi}_{t}^{n}, [nt]/n)\,dt \) is uniformly integrable, we have

$$
E_n \left[ \int_{0}^{1} \tilde{f}(\lambda_{n_j} \xi_{t}^{n_j}, [nt]/n)\,dt \right] \leq E_n \left[ \int_{0}^{1} \tilde{f}((\tilde{\lambda} - \epsilon) \xi_{t}^{n_j}, [nt]/n)\,dt \right] \rightarrow E \left[ \int_{0}^{1} \xi_{t} \tilde{f}((\tilde{\lambda} - \epsilon) \xi_{t}, t)\,dt \right], \quad (19)
$$
where the convergence on the right-hand-side follows from Lemma 2. Since the left-hand side of (19) is equal to $W_0$ by (18), we obtain that $E \left[ \int_0^t \xi_t \hat{f}(\lambda - \epsilon) \xi_t \, dt \right] \geq W_0$. But this inequality contradicts (14).

If $\tilde{\lambda} < \lambda_0$, a similar technique leads to a contradiction. We omit the proof for the cases when $\tilde{\lambda} = 0$ and $\tilde{\lambda} = \infty$. \[ \Box \]

We remind the readers that the uniform integrability condition required by Theorem 3 is not difficult to check. For example, if $\kappa$ is bounded and $\hat{f}(x^{-1}, t)$ satisfies a polynomial growth condition, i.e., $\exists K, \beta > 0$ such that

$$\hat{f}(x, t) \leq K(1 + x^{-\beta}) \quad \forall x > 0,$$

then the uniform integrability condition is satisfied. To demonstrate this, we need only to show that for some $\eta > 0$,

$$\sup_n E_n \left[ \frac{1}{n} \sum_{k=0}^{n-1} \left( \xi_k^n + (\xi_k^n)^{1-\beta} \right)^{1+\eta} \right] < \infty$$

[see Stroock (1987)]. In fact, by convexity we have

$$E_n \left[ \frac{1}{n} \sum_{k=0}^{n-1} \left( \xi_k^n + (\xi_k^n)^{1-\beta} \right)^{1+\eta} \right] \leq E_n \left[ \frac{1}{n} \sum_{k=0}^{n-1} \left( \xi_k^n + (\xi_k^n)^{1-\beta} \right)^{1+\eta} \right]$$

$$\leq 2^{1+\eta} E_n \left[ \frac{1}{n} \sum_{k=0}^{n-1} \left( (\xi_k^n)^{1+\eta} + (\xi_k^n)^{(1-\beta)(1+\eta)} \right) \right].$$

Applying the technique used in the proof of Lemma 3 (in Appendix), one can show that for all $\alpha \in \mathcal{R}$,

$$E_n[\xi_k^n]^{\alpha} \leq C,$$

where $C$ depends only on $\alpha$ and the bound of $|\kappa|$, but not on $k$ or $n$. We conclude that

$$\sup_n E_n \left[ \frac{1}{n} \sum_{k=0}^{n-1} \xi_k^n + (\xi_k^n)^{1-\beta} \right]^{1+\eta} < \infty$$

for any $\eta > 0$. Therefore, when $\kappa$ is bounded and $\hat{f}(x^{-1}, t)$ satisfies a polynomial growth condition, we obtain $\mathcal{R} \Rightarrow c$.

Theorem 3 basically says that if the implicit Arrow-Debreu state price process converges to the continuous time limit, then subject to some regularity conditions on the utility functions and the parameters of the price processes, the optimal consumption process always converges to the

\footnote{A specific example is $u(x) = \frac{x^{1-b}}{1-b}$ with $b > 0$.}
corresponding continuous time limit. However, this theorem does not tell us whether the optimal portfolio policy converges or not.

It turns out that the convergence of optimal portfolio policies can be established in the same way we established the convergence of the replicating portfolio strategies in contingent claim pricing. In the previous case, we established the convergence of the replicating portfolio strategies by comparing the solution of the PDE to the solution of a recurrent equation that determines the price of the claim. This motivates us to study the convergence of the discrete time optimal portfolio policy to the continuous time optimal portfolio policy by studying the optimal wealth processes and their partial derivatives. For notational convenience, we assume \( N = 1 \) (\( N > 1 \) is analogous).

Let \( F \) be the optimal wealth process that satisfies (15) and (16) and let \( \theta \) be the optimal portfolio policy defined by (17). Similarly, we define a discrete time optimal wealth process \( F^n \) and a discrete time optimal portfolio policy \( \theta^n \) that replicates the optimal consumption process. As in the continuous time case, we introduce an additional state variable \( \tilde{Z}^n \) by setting \( Z^n_k = \lambda^n B^n_k / \xi^n_k \) and \( \tilde{Z}^n_t = Z^n_{[nt]} \). It should be obvious that the value process \( F^n \) is a function of \( (S^n_k, Z^n_k, t^n_k) \) (see (20) below). With these definitions, \( F^n \) must satisfy the recurrent equation

\[
F^n(S^n_k, Z^n_k, t^n_k) = \frac{1}{n} f(Z^n_{k-1}, t_k) + \alpha n S^n_k (\cdot) F^n(S^n_{k+1}, Z^n_{k+1}, t^n_{k+1}) + \alpha n S^n_k (\cdot) F^n(S^n_{k-1}, Z^n_{k+1}, t^n_{k+1}),
\]

\[
F^n(S^n_k, Z^n_k, t^n_k) = \bar{g}(Z^n_{n-1}),
\]

(20)

where \(+\) and \(-\) denote the states \( \bar{e}_k = 1 \) and \( \bar{e}_k = -1 \) respectively, and

\[
S^n_{k+1} = S^n_k + \frac{b(S^n_k)}{n} + \sigma(S^n_k) \sqrt{n},
\]

\[
S^n_{k+1} = S^n_k + \frac{b(S^n_k)}{n} - \sigma(S^n_k) \sqrt{n},
\]

\[
Z^n_{k+1} = Z^n_k (1 + \frac{r(S^n_k)}{n})(1 + \kappa(S^n_k) \sqrt{1/n})^{-1},
\]

\[
Z^n_{k-1} = Z^n_k (1 + \frac{r(S^n_k)}{n})(1 - \kappa(S^n_k) \sqrt{1/n})^{-1}.
\]

Equation (20) says that the sum of the values of current consumption and future wealth is equal to the current wealth. Given the wealth process, the optimal portfolio policy that replicates the optimal consumption policy is determined by

\[
\theta^n_k = \frac{F^n(S^n_{k+1}, Z^n_{k+1}, t^n_{k+1}) - F^n(S^n_{k-1}, Z^n_{k+1}, t^n_{k+1})}{S^n_{k+1} - S^n_{k-1}},
\]

\[
\alpha^n_k = (F^n(S^n_k, Z^n_k, t^n_k) - \theta^n_k^T S^n_k) / B^n_k.
\]

(21)

We want to show that \( (\bar{\alpha}^n, \bar{\theta}^n) \) converges to \( (\alpha, \theta) \). To do so, we an approximation technique similar to that used in the proof of Theorem 2. Let \( e^n_k \) be the truncation errors of the optimal
wealth process defined by

$$e_k^n = F(S^n_k, Z^n_k, t^n_k) - F(S^n_k, Z^n_k, t^n_k).$$

Substituting this into (21), we get

$$\theta^n_k = \frac{\sqrt{n}}{2\sigma(S^n_k)} \left( F(S^n_{k+1}, Z^n_{k+1}, t^n_{k+1}) - F(S^n_{k+1}, Z^n_{k+1}, t^n_{k+1}) \right) + \frac{\sqrt{n}}{2\sigma(S^n_k)} (e^n_{k+1} - e^n_{k+1}).$$

The following theorem shows that

(a) \( \frac{\sqrt{n}}{2\sigma(S^n_k)} (F(S^n_{k+1}, Z^n_{k+1}, t^n_{k+1}) - F(S^n_{k+1}, Z^n_{k+1}, t^n_{k+1})) \Rightarrow \theta, \)

(b) \( \frac{\sqrt{n}}{2\sigma(S^n_k)} (e^n_{k+1} - e^n_{k+1}) \Rightarrow 0. \)

**Theorem 4 (Portfolio Policies)** Suppose the following

a) \( \tilde{Z}^n \) converges to \( Z \) weakly (or equivalently \( \lambda_n \rightarrow \lambda_0 \));

b) \( \sigma, r \) and \( \kappa \) are continuously differentiable up to the sixth order with all these derivatives satisfying a polynomial growth condition;

c) \( r(x)y, \kappa(x)y \) and \( \kappa(x)y \) satisfy a uniform Lipschitz condition and a linear growth condition;

d) \( \hat{f}(x^{-1}, t) \) and \( \hat{g}(x^{-1}) \) satisfy a polynomial growth condition.

Then \((\tilde{\alpha}_n, \tilde{\theta}_n) \Rightarrow (\alpha, \theta)\). In particular, \( \tilde{\alpha}_n^0 \rightarrow \alpha_0 \) and \( \tilde{\theta}_n^0 \rightarrow \theta_0 \) as \( n \rightarrow \infty \).

**Proof.** The proof for this theorem is similar to that of Theorem 2 and we delegate it to the Appendix.

**Remark 3** The uniform Lipschitz condition imposed on \( r(x)y \) and \( \kappa(x)y \) implies that both \( r \) and \( \kappa \) are bounded. This condition can be weakened and replaced by a uniform Lipschitz condition on \( r, \kappa \) and \( |\kappa|^2 \), provided that \( \hat{f}(e^{-x}, t) \) and \( \hat{g}(e^{-x}) \) satisfy a polynomial growth condition in \( x \).

Similar to Theorem 2, we obtain a simple numerical procedure for computing the optimal consumption and the optimal portfolio policy as compared to the standard method of solving a PDE such as (15) with at least two state variables, namely \( S \) and \( Z \).

Our last theorem is on convergence of optimal expected utilities, the proof of which follows immediately from Lemma 2.

**Theorem 5 (Expected Utilities)** Suppose that the sufficient conditions for Theorem 3 are satisfied, and suppose further that \( \{\int_0^1 |u(c_t, nt^n)|dt\} \) is uniformly integrable. Then

$$E_n \left[ \int_0^1 u(c^n_t, nt^n)dt \right] \Rightarrow E \left[ \int_0^1 u(c_t, t)dt \right]. \quad (22)$$

18
Remark 4 A sufficient condition for \( \int_0^1 |u(x, t)| \, dt \) to be uniformly integrable is that there exist constants \( A, B, K, \) and \( 0 < b < 1, \beta > 0 \) such that for all \( x > 0 \)

\[
|u(x, t)| \leq A + B \frac{x^{1-b}}{1-b}, \quad \hat{f}(x, t) \leq K(1 + x^{-\beta})
\]

and that \( \kappa(x) \) is uniformly bounded. This statement can be proved using the approach we used before.

A stronger convergence result is established in He (1989) in which it is shown that the optimal expected utilities as well as the optimal conditional expected utilities converge to the corresponding continuous time limits.\(^{12}\) While it is natural to require convergence of optimal expected utilities in the context optimal consumption-portfolio choice, convergence of optimal conditional expected utilities is also useful to require, since at each point in time the objective of an individual is always to maximize the conditional expected utility of consumption conditioning upon the past resolution of asset prices. We refer the reader to He (1989) for the proof of the convergence of conditional expected utilities.

This concludes our demonstration of the convergence of optimal consumption-portfolio policies and optimal expected utilities. To summarize our results, the convergence theorems established in Sections 3, 4 and 5 confirm our economic intuition that discrete time models and continuous time models are approximations of each other. In cases where closed form solutions exist, the continuous time approach provides a convenient analytic tool for studying contingent claim prices and optimal consumption-portfolio policies. However, in cases where closed form solutions do not exist, the discrete time approach provides an elegant numerical solution to the continuous time model.

6 Concluding Remarks

The leading assumption for the convergence results established in this paper is that markets in continuous time models are dynamically complete. Dynamic completeness is necessary for contingent claims to be priced by arbitrage, but it is not necessary for solving dynamic consumption and portfolio problem. We conjecture that similar convergence results would also be obtained for optimal consumption-portfolio policies in which markets are dynamically incomplete. In this case, it is reasonable to expect that markets in discrete time models would also be dynamically incomplete. The exact implementation is subject to future research.

\(^{12}\) The proof of convergence of conditional expected utilities uses a notion of extended weak convergence developed in Aldous (1981).
References


Rubinstein, M., 1976, "The strong case for the generalized logarithmic utility model as the premier model of financial markets," *Journal of Finance*.


Appendix: Proofs

Proof for Theorem 2. We need two lemmas to proceed with the proof.

Lemma 3 For any integers $m, l, k \geq 0$ where $l \leq k \leq n$, there exists a constant $A > 0$, depending only upon $m$ and $K$, such that

$$\hat{E}_n|S^n_k|^{2m} \leq A(1 + \hat{E}_n|S^n_l|^{2m}),$$

where $K$ is a constant such that the linear growth condition is satisfied for $b(x)$, $\sigma(x)$ and $x\sigma(x)$, and $\hat{E}_n$ denotes the expectation under $Q_n$.

Proof. We demonstrate this inequality for $N = 1$ (the proof for $N > 1$ is analogous to the case with $N = 1$). The proof follows closely the proof of Theorem 2.3 of Friedman (1975, pp. 107), where he obtained this result for the diffusion process.

We can find some $K' > 0$, depending only upon $K$, such that for any $x \in \mathbb{R}$,

$$|b(x)| \leq K'(1 + |x|), \quad |b(x)|^2 \leq K'(1 + |x|^2), \quad |\sigma(x)| \leq K'(1 + |x|),$$

$$|\sigma(x)|^2 \leq K'(1 + |x|^2), \quad |xb(x)| \leq K'(1 + |x|^2), \quad |b(x)\sigma(x)| \leq K'(1 + |x|^2),$$

$$|x^2\sigma(x)| \leq K'(1 + |x|^2).$$

Applying a Taylor expansion to the function $x^{2m}$, we obtain

$$|S^n_{k+1}|^{2m} = |S^n_k|^{2m} + 2m|S^n_k|^{2m-1}(S^n_{k+1} - S^n_k) + m(2m - 1)|S^n_k|^{2m-2}(S^n_{k+1} - S^n_k)^2$$

$$= |S^n_k|^{2m} + 2m|S^n_k|^{2m-1}\left(\frac{b}{n} + \frac{\sigma}{\sqrt{n}}\tilde{e}_k\right) + m(2m - 1)|S^n_k|^{2m-2}\left(\frac{b^2}{n^2} + \frac{2b\sigma}{n\sqrt{n}}\tilde{e}_k + \frac{\sigma^2}{n}\tilde{e}_k^2\right),$$

where $\tilde{S}^n_k = S^n_k + \beta\left(\frac{b}{\sqrt{n}} + \frac{\sigma}{\sqrt{n}}\tilde{e}_k\right)$ for some $\beta \in [0, 1]$. Taking expectation $\hat{E}_n$ on both sides and noticing that $|\tilde{S}^n_k| \leq |S^n_k| + |b| + |\sigma|$, we obtain

$$\hat{E}_n|S^n_{k+1}|^{2m} \leq \hat{E}_n|S^n_k|^{2m} + \frac{2mK'}{n}\hat{E}_n(\hat{E}_n|S^n_k|^{2m-2} + |S^n_k|^{2m})$$

$$+ \frac{m(2m - 1)}{n}\hat{E}_n\left(|S^n_k| + |b| + |\sigma|)^{2m-2}(b^2 + 2|b\sigma| + \sigma^2\right),$$

where we have used the fact that $x^{2m} \leq 1 + x^{2m}$ and $\hat{E}_n[\tilde{e}_k] = \frac{\sigma(S^n_k)}{\sqrt{n}}$. Noticing further that $x^{2m-2} \leq 1 + x^{2m}$ and $(x + y)^m \leq 2^m(x^m + y^m)$ when $x, y > 0$, we obtain

$$\hat{E}_n|S^n_{k+1}|^{2m} \leq \hat{E}_n|S^n_k|^{2m} + \frac{2mK'}{n}\hat{E}_n[1 + 2|S^n_k|^{2m}]$$

23
\[ + \frac{4K'm(2m - 1)}{n} \hat{E}_n \left(2K' + (1 + 2K')|S^n_k|^2 + 2(1 + |S^n_k|^2) \right) \]
\leq \hat{E}_n |S^n_k|^{2m} + \frac{2mK'}{n} \hat{E}_n [1 + 2|S^n_k|^{2m}]
\[ + \frac{4K'm(2m - 1)}{n} 2^{2m-2} \left((2K')^{2m-2}(1 + \hat{E}_n |S^n_k|^2) + (1 + 2K')^{2m-2} \hat{E}_n ||S^n_k||^{2m-2} + |S^n_k|^{2m} \right) \]

Since \(\hat{E}_n |S^n_k|^2 \leq 1 + \hat{E}_n |S^n_k|^{2m}\) for \(m \geq 1\) and \(\hat{E}_n |S^n_k|^{2m-2} \leq 1 + \hat{E}_n |S^n_k|^{2m}\), we can find a constant \(C > 0\), depending only upon \(m\) and \(K'\), such that

\[ \hat{E}_n |S^n_{k+1}|^{2m} \leq \frac{C}{n} + (1 + \frac{C}{n}) \hat{E}_n |S^n_k|^{2m}. \]

This implies that

\[ \hat{E}_n |S^n_k| \leq (1 + \frac{C}{n})^{k-1} (1 + \hat{E}_n |S^n_k|^{2m}). \]

Thus \(A = \sup_n \left(1 + \frac{C}{n}\right)^n\) is our choice. \(\blacksquare\)

**Lemma 4.** Under assumptions (a) and (b), all of the partial derivatives of \(V\) up to the third order satisfy a polynomial growth condition.

**Proof.** First, we rewrite \(V\) as

\[ V(x, t) = E_Q \left[ e^{-\Psi_t g(\bar{S}_1)} \right], \]

where \((\bar{Y}_r, \bar{S}_r)\) satisfies the stochastic differential equation

\[
\begin{align*}
    d\bar{Y}_r &= r(\bar{S}_r)dr \\
    d\bar{S}_r &= r(\bar{S}_r)\bar{S}_rdr + \sigma(\bar{S}_r)dw^*_r \\
    \bar{Y}_t &= 0, \quad \bar{S}_t = x,
\end{align*}
\]

where \(\tau \geq t, \bar{Y} \geq 0\) and \(w^*_r = w_r - \int_0^r \kappa(s)ds\) is a Brownian motion under \(Q\) by Girsanov's theorem [Friedman (1975)].

Following the argument used in Theorem 5.5 of Friedman (1975, pp. 122) and Theorem 1 of Gihman and Skorohod (1972, pp. 61), we conclude that \(V\) is continuously differentiable with respect to \(x\) up to the sixth order and that all of these derivatives satisfy a polynomial growth condition. Moreover, since \(V\) satisfies the PDE and \(V_{t}, V_{tt}\) and \(V_{ttt}\) can be expressed as functions of the partial derivatives of \(V\) w.r.t. \(x\) up to at most the sixth order, we deduce that \(V\) is also continuously differentiable with respect to \(t\) up to the third order. This plus the fact that \(b, \sigma, r\) and all of their derivatives satisfy a polynomial growth condition leads us to further deduce that all of the derivatives with respect to \(x\) and \(t\) up to the third order satisfy a polynomial growth condition. \(\blacksquare\)
PROOF OF THEOREM 2. We prove this theorem for $N = 1$ (the proof for $N > 1$ is analogous to the case with $N = 1$). The basic idea is to substitute the true value function $V$ into the recurrent equation that defines $V^n$ in order to get better estimates for the truncation errors defined as 

$$
\epsilon^n_k = V(S^n_k, \frac{k}{n}) - V^n(S^n_k, \frac{k}{n}).
$$

We use $+$ and $-$ to denote the states $\bar{\epsilon}_k = 1$ and $\bar{\epsilon}_k = -1$ respectively, and define 

$$
S^{n+}_{k+1} = S^n_k + \frac{b(S^n_k)}{n} + \frac{\sigma(S^n_k)}{\sqrt{n}},
$$

$$
S^{n-}_{k+1} = S^n_k + \frac{b(S^n_k)}{n} - \frac{\sigma(S^n_k)}{\sqrt{n}},
$$

where $S_{k+1}^{n+}$ and $S_{k+1}^{n-}$ denote the prices at time $t_{k+1}$ when $\bar{\epsilon}_k = 1$ and $\bar{\epsilon}_k = -1$ respectively. Similarly, we define two functions $f_{+}^{k,n}$ and $f_{-}^{k,n}$ as follows: 

$$
f_{+}^{k,n}(r) = V(S^n_k + r(S_{k+1}^{n+} - S^n_k), t^n_k + r(t^n_{k+1} - t^n_k)),
$$

$$
f_{-}^{k,n}(r) = V(S^n_k + r(S_{k+1}^{n-} - S^n_k), t^n_k + r(t^n_{k+1} - t^n_k)).
$$

where $t^n_k = \frac{k}{n}$. Since $V$ is differentiable up to the third order, $f_{+}^{k,n}$ and $f_{-}^{k,n}$ admit the following Taylor expansions (we omit the superscripts for $f$):

$$
f_{+}(1) = f_{+}(0) + f'_{+}(0) + \frac{1}{2} f''_{+}(0) + R^n_k,
$$

$$
f_{-}(1) = f_{-}(0) + f'_{-}(0) + \frac{1}{2} f''_{-}(0) + Q^n_k,
$$

where 

$$
R^n_k = \frac{1}{6} \int_0^1 (1-s)^3 f^{(3)}_{+}(s) ds,
$$

$$
Q^n_k = \frac{1}{6} \int_0^1 (1-s)^3 f^{(3)}_{-}(s) ds.
$$

We have 

$$
f'_{+}(0) = V_S(S_{k+1}^{n+} - S^n_k) + \frac{1}{n} V_t,
$$

$$
f'_{-}(0) = V_S(S_{k+1}^{n-} - S^n_k) + \frac{1}{n} V_t,
$$

$$
f''_{+}(0) = V_{SS}(S_{k+1}^{n+} - S^n_k)^2 + V_{tt} \frac{1}{n^2} + 2 V_{St} \frac{1}{n} (S_{k+1}^{n+} - S^n_k),
$$

$$
f''_{-}(0) = V_{SS}(S_{k+1}^{n-} - S^n_k)^2 + V_{tt} \frac{1}{n^2} + 2 V_{St} \frac{1}{n} (S_{k+1}^{n-} - S^n_k),
$$

25
where $V$ and the partial derivatives of $V$ are evaluated at $(S^n_k, t^n_k)$. Now, to obtain an estimate for $e^n_k$, we substitute the expressions for $f^n_+(1)$ and $f^n_-(-1)$ into the recurrent equation. Using the fact that $V$ satisfies (10), we obtain,

$$\pi S^n_k (+) f_+ (1) + \pi S^n_k (-) f_- (1) = V(S^n_k) + \frac{1}{2} \left( V_{SS} \left( \frac{b^2}{n^2} + \frac{2b_0 \kappa}{n^2} \right) + 2 V_{St} \frac{r S^n_k}{n^2} \right) \left( 1 + \frac{r}{n} \right)^{-1}$$

$$+ \frac{1}{2} V_{tt} \left( 1 + \frac{r}{n} \right)^{-1} - \gamma^n_k \left( 1 + \frac{r}{n} \right)^{-1},$$

where $\gamma^n_k = -(\pi S^n_k (+) R^n_k + \pi S^n_k (-) Q^n_k)$. Letting $h = -\frac{1}{2} V_{SS} (b^2 + 2b_0 \kappa) - V_{St} r S - \frac{1}{2} V_{tt}$, we obtain the following recurrent equation for $e^n_k$:

$$e^n_k = \pi S^n_k (+) e^n_{k+1} + \pi S^n_k (-) e^n_{k-1} + \frac{1}{n^2} h(S^n_k, \frac{k}{n})(1 + \frac{r}{n})^{-1} + \gamma^n_k \left( 1 + \frac{r}{n} \right)^{-1}. $$

This equation allows us to argue that

$$|e^n_k| \leq \frac{C}{\sqrt{n}} (1 + |S^n_k|^{2q}), \quad (23)$$

for some constant $C > 0$ and integer $q > 0$ (which are independent of $k$ and $n$). To see that, we express $e^n_k$ as follows:

$$e^n_k = \beta_k \left[ \frac{1}{n^2} h(S^n_m, \frac{m}{n}) B^n_m / B^n_{m+1} + \gamma^n_m B^n_m / B^n_{m+1} | S^n_k \right].$$

According to Lemma 2, $h$ satisfies a polynomial growth condition. Therefore, we can find a constant $C' > 0$ and an integer $q > 0$ such that

$$|h(x, t)| \leq C'(1 + |x|^{2q}).$$

Applying Lemma 3, we obtain

$$\beta_k \leq C' \left( 1 + A(1 + |S^n_k|^{2q}) \right),$$

for $m \geq k$. This takes care of the first summation on the right hand side of (24). For the second summation, we argue that it is of order $\frac{1}{\sqrt{n}}$. To see this, one can express $R^n_k$ and $Q^n_k$ explicitly by writing down $f^n_+(s)$ and $f^n_-(s)$, allowing us to argue that they are the sum of terms, of order $n^{-3/2}$ or higher. For example, a typical term of $R^n_k$ has the following form:

$$\int_0^1 (1 - s)^3 V_{SS} \left( ((1 - s) S^n_k + s S^n_{k+1}, \ldots) (S^n_{k+1} - S^n_k)^3 ds$$

$$= \int_0^1 (1 - s)^3 V_{SS} \left( ((1 - s) S^n_k + s S^n_{k+1}, \ldots) \frac{\sigma^2(S^n_k)}{n^{3/2}} ds + \ldots$$

26
Since $V_{SSS}$ satisfies a polynomial growth condition, so does the integrand. We can choose $q$ to be larger enough so that all of the polynomial growth conditions have the same power $q$. Now applying the same procedure, we can get an inequality for $\gamma_m^n$ which is of order $n^{-3/2}$. Therefore the summation is of order $n^{-1/2}$. Altogether, we conclude that (23) is true. Thus $\tilde{V}^n \Rightarrow V$ weakly.

We now substitute $f_-$ and $f_+$ into the equation that defines $\theta_k^n$. This yields

$$
\theta_k^n = \frac{f_+^k(n)(1) - f_-^k(n)(1)}{S_{k+1}^{n+} - S_{k+1}^{n-}}
= V_S + \frac{2b}{n} V_{SS} + \frac{2V_{St}}{n} + \frac{2\sqrt{n}}{\sigma} \left( R_k^n - Q_k^n \right).
$$

(25)

Since $R_k^n$ and $Q_k^n$ are of order $n^{-3/2}$ or higher, we conclude that

$$
\frac{2b}{n} V_{SS} + \frac{2V_{St}}{n} + \frac{2\sqrt{n}}{\sigma} \left( R_k^n - Q_k^n \right) \Rightarrow 0.
$$

Hence, $\theta_k^n \Rightarrow V_S$. Finally, since $\tilde{V}^n \Rightarrow V$ and $\tilde{\theta}_i^n = \tilde{\alpha}_i^n \tilde{R}_i^n + \tilde{\theta}_i^n \tilde{S}_i^n$, we conclude that $\tilde{\alpha}^n \Rightarrow \alpha$. This proves part 2). The claim that $\tilde{\theta}_0^n \to V_0$, $\tilde{\alpha}_0^n \to \alpha_0$ and $\tilde{\theta}_0^n \to \theta_0$ follows directly from the fact that they are non-stochastic. 

**Proof for Theorem 4.** We rewrite $F$ as follows:

$$
F(x, y, t) = \mathbb{E}Q \left[ \int_t^1 e^{-7_i} f(Z_{-1}^{-1}, r) dr + e^{-7_i} 7 \left( Z_{-1}^{-1} \right) \right],
$$

where $(\tilde{Y}_r, \tilde{Z}_r, \tilde{S}_r)$ satisfies the stochastic differential equation

$$
\begin{align*}
\frac{d\tilde{Y}_r}{d\tilde{S}_r} &= r(\tilde{S}_r) dr, \\
\frac{d\tilde{S}_r}{d\tilde{S}_r} &= r(\tilde{S}_r) \tilde{S}_r dr + \sigma(\tilde{S}_r) dw^*_r, \\
\frac{d\tilde{Z}_r}{d\tilde{S}_r} &= r(\tilde{S}_r) \tilde{Z}_r dr - \kappa(\tilde{S}_r) \tilde{Z}_r dw^*_r, \\
\tilde{Y}_t &= 0, \quad \tilde{S}_t = x, \quad \tilde{Z}_t = y, \quad r \in [t, 1],
\end{align*}
$$

and $w^*_r = w_r - \int_0^r \kappa_s ds$. An argument similar to that used in the proof of Theorem 2 allows us to conclude that conditions a), b) and c) imply that $F$ is continuously differentiable up to the third order, and that all of these derivatives satisfy a polynomial growth condition.

The basic idea of this proof is exactly the same as the proof of Theorem 2. We substitute the true function $F$ into the finite difference equation that defines $F^n$ in order to get better estimates for the truncation errors. We define two functions, $f_+^{kn}$ and $f_-^{kn}$ as follows:
\[ f_{+,n}^k(r) = F(S_k^r + r(S_{k+1}^r - S_k^r), Z_k^r + r(Z_{k+1}^r - Z_k^r), t_k^r + r(t_{k+1}^r - t_k^r)), \]
\[ f_{-,n}^k(r) = F(S_k^r + r(S_{k+1}^r - S_k^r), Z_k^r + r(Z_{k+1}^r - Z_k^r), t_k^r + r(t_{k+1}^r - t_k^r)). \]
Then functions \( f_{+,n}^k \) and \( f_{-,n}^k \) admit the following Taylor expansions (we omit the superscripts):
\[ f_{+}(1) = f_{+}(0) + f'_{+}(0) + \frac{1}{2} f''_{+}(0) + R_k^n, \]
\[ f_{-}(1) = f_{-}(0) + f'_{-}(0) + \frac{1}{2} f''_{-}(0) + Q_k^n, \]
where
\[ R_k^n = \frac{1}{6} \int_0^1 (1-s)^3 f^{(3)}_{+}(s) ds, \]
\[ Q_k^n = \frac{1}{6} \int_0^1 (1-s)^3 f^{(3)}_{-}(s) ds. \]
We need to show that \( \lim_{n \to \infty} (f_{+,n}^k(1) - f_{-,n}^k(1)) = 0 \). If we express \( f_{+,n}^k \) and \( f_{-,n}^k \) explicitly, and substitute them into the right hand side of (a), we obtain
\[ (a) = F_{s_k}(S_k^r, Z_k^r, t_k^r) - \frac{1}{\sigma(S_k^r)} F_{Z_k}(S_k^r, Z_k^r, t_k^r) \kappa(S_k^r) Z_k^r + \epsilon_k^n, \]
where \( \epsilon_k^n \) is the sum of several terms, all of which are of order \( n^{-1} \) or higher, and therefore go to zero.

To show (b), we need to use the fact that \( F \) satisfies (15). Expressing \( \pi_S^k(+) f_{+}(1) + \pi_S^k(-) f_{-}(1) \) in terms of \( f_{+}(0) \) and \( f_{-}(0) \), we obtain
\[ \pi_S^k(+) F(S_{k+1}^r, Z_{k+1}^r, t_{k+1}^r) + \pi_S^k(-) F(S_{k+1}^r, Z_{k+1}^r, t_{k+1}^r) + \frac{1}{n(1 + \frac{r}{n})} \tilde{f}(Z_k^r) = F(S_k^r, Z_k^r, t_k^r) + O\left(\frac{1}{n^{3/2}}\right), \]
where \( O\left(\frac{1}{n^{3/2}}\right) \) denotes the sum of those terms which are of order \( \frac{1}{n^{3/2}} \) or higher. This plus the budget equation (20) gives us the following relation for the truncation errors:
\[ \epsilon_k^n = \pi_S^k(+) \epsilon_{k+1}^n + \pi_S^k(-) \epsilon_{k+1}^n + O\left(\frac{1}{n^{3/2}}\right), \]
and \( \epsilon_0^n = 0. \) As in Theorem 2, this allows us to conclude that
\[ |\epsilon_k^n| = O\left(\frac{1}{n^{1/2}}\right). \]
Thus, \( \tilde{\beta}^n = F \) and (b) holds. Finally, since \( F^n \) and \( \theta^n \) converge to \( F \) and \( \theta \) respectively, we conclude that \( \tilde{\alpha}^n \rightarrow \alpha. \)