Risk Premium and Pricing of Derivatives in Complete Markets

Jin-Chuan Duan*

First Draft: March 2001

Abstract

Risk-neutral pricing of derivative assets (i.e., independent of the risk premium for the underlying asset) has typically been misconstrued as an inherent feature of complete markets since the seminar work of Black and Scholes (1973). We first construct a semi-recombined binomial lattice, which constitutes a complete-market model, and then use it to show that risk-neutral pricing is not an inherent property of complete markets. The paper shows that risk-neutral pricing is a result that has more to do with the assumption typically adopted to describe the price dynamic of the underlying asset. The limiting model of our semi-recombined binomial lattice is the continuous-time version of the GARCH option pricing model in Kallsen and Taqqu (1998). Over a set of discrete time points, the limiting model coincides with the GARCH option pricing model first derived by Duan (1995).

Key words: Risk-neutral, risk premium, complete market, GARCH, weak convergence.

*Rotman School of Management, University of Toronto and Department of Finance, Hong Kong University of Science & Technology; Tel: 416-946-5653, Fax: 416-971-3048, E-mail: jcduan@rotman.utoronto.ca; Web: http://www.rotman.utoronto.ca/~jcduan. Duan acknowledges the support received as the Manulife Chair in Financial Services at University of Toronto. This research was in part completed during Duan's visit (sponsored by Polaris Securities Group) to Department of Quantitative Finance, National Tsing Hua University, Taiwan.

1 Introduction

Black and Scholes (1973) and Merton (1973) established the celebrated option pricing theory, which has profoundly altered the course of academic research and industrial practices since. One of the most cherished results of the option pricing theory is its implication on risk-neutral valuation of derivative contracts, which was further developed in Cox and Ross (1976), Harrison and Kreps (1978) and Harrison and Pliska (1981). Risk-neutral valuation is typically attributed to complete markets, because replication (static or dynamic) of contingent claims is possible. In order to avoid arbitrage, contingent claims must therefore be priced identically to the values of their replicating portfolios. As a result, one can simply consider a hypothetical market in which agents have risk-neutral preferences. Under this hypothetical risk-neutrality, contingent claims can be priced by taking discounted expected value as long as the underlying asset also behaves according to the implication of risk-neutrality. Specifically in the diffusion modeling framework, one only need to change the drift term to the risk-free rate of return and keep the diffusion term intact. In other words, the expected return is changed to the risk-free rate but the volatility remains unchanged. This result naturally suggests that the risk premium of the underlying asset (or the risk-adjusted expected return) plays no direct role in determining the values of derivative contracts written on this underlying asset. For pedagogical purposes, the same risk-neutral valuation conclusion is often demonstrated using the binomial lattice, owing its popularity to Cox, Ross and Rubinstein (1979). This result is typically interpreted as the implication of complete markets. Since the contingent claim can be exactly replicated by the existing assets, the physical probabilities of states are irrelevant and there is no unhedgeable risk that requires the consideration of risk premium (see, for example, Wilmott (1998, p.74)).

This interpretation prevailing in the literature is, however, misconstrued. The risk-neutral valuation result turns out to be valid for diffusion models and others that bear similar distributional characteristics, but it is invalid in general. In this paper, we construct a semi-recombined binomial lattice to ascertain this claim. The semi-recombined binomial lattice constitutes a complete market in which all contingent claims can be perfectly replicated. However, the arbitrage-free values of contingent claims still depend on the risk premium of the underlying asset whose price dynamic is governed by the semi-recombined binomial lattice. In other words, the risk-neutral valuation principle is invalid in this complete market. In fact, this model serves to show that risk-neutral valuation has much to do with specific distributional features of a given model. To say the least, it does not follow logically from market completeness.

The semi-recombined binomial lattice behaves like the conventional binomial lattice up to some pre-specified time point and starts to branch out in an overall non-recombined fashion but with recombined sub-lattices until reaching another pre-specified time point. The process of branching out with recombined sub-lattices continues until reaching a final time point. The overall non-recombinedness of the lattice is caused by stochastic volatility, which evolves in a GARCH-like manner over the pre-specified time points. Although we show that stochastic volatility is crucial in obtaining our result, it alone is insufficient to establish the relevancy of risk premium. In fact, we need to have stochastic volatility driven by unanticipated return shocks to reach such a conclusion. Loosely speaking, a change of the probability measure for the original economy to one for the hypothetical risk-neutral economy alters the conditional mean as well as what is regarded as an unanticipated return shock. Risk premium of the underlying asset becomes important because the unanticipated return shock in the original economy is no longer an unanticipated shock. It requires an adjustment with risk premium to remain as an unanticipated return shock in the risk-neutral economy. Interestingly, the empirical success of the GARCH process in modeling volatility has a great deal to do with using unanticipated return shocks to drive the conditional volatility.

If one increases the number of binomial steps, the semi-recombined lattice converges to a continuous-time version of the GARCH model first proposed by Kallsen and Taqqu (1998). This continuous-time version of the GARCH model uses geometric Brownian motions to patch together the in-between points of a discrete-time GARCH model. Kallsen and Taqqu (1998) showed that this continuous-time GARCH model constitutes a complete market and derivatives can be priced by arbitrage and the pricing result is in agreement with that of Duan (1995) for which a discrete-time GARCH model was used. Duan's (1995) result suggests that the risk premium of the underlying asset will affect the price of the derivative contract. Since the discrete-time GARCH model does not constitute a complete market, such a result may not be surprising. However, the work of Kallsen and Taqqu (1998) suggests that the conclusion is not simply due to market incompleteness. The limiting model of our semi-recombined binomial lattice under the complete-market pricing probability turns out to be a particular continuous-time GARCH model same as the pricing conclusion derived by Kallsen and Taqqu (1998), which in turn agrees with Duan (1995) over the pre-specified set of time points for which the discrete-time GARCH model is defined. The relationship between this paper and Kallsen and Taqqu (1998) is, in a way, analogous to that between Cox, Ross and Rubinstein (1979) and Black and Scholes (1973), except for their opposite conclusions on the relevancy of risk premium. Rubinstein (1976) and Brennan (1979) have shown that risk-neutral valuation can also be established in some incomplete markets with the help of some combinations of restrictions on distributions and on the family of utility functions. The relationship between Duan (1995) and Kallsen and Taqqu (1998) can therefore be likened to that between Rubinstein (1976) and Brennan (1979) on one hand and Black and Scholes (1973) and Merton (1973) on the other, albeit their opposite conclusions on the relevancy of risk premium. The relationship among these papers is summarized in Figure 1.

Inclusion of risk premium in the pricing of derivatives is not a mere theoretical interest. We show that the volatility level under the complete-market pricing probability is typically higher than that under the physical probability. This suggests that the price assessed according to the complete-market pricing probability will be higher in comparison to that of using the physical probability. Given the general empirical evidence of underpricing by the

Black-Scholes model, it is interesting to observe that risk premium can serve to explain this underpricing pricing phenomenon and to potentially close the gap between the option pricing theory and the empirical regularity.

2 A semi-recombined binomial lattice

In this section, we present a binomial lattice for asset prices that is overall non-recombined but many of its sub-branches are recombined. Let $(\Omega, \mathcal{F}, \pi)$ denote the underlying probability space. Denote by $\{\epsilon_i; i = 1, 2, \cdots\}$ a sequence of independent standard normal random variables defined on $(\Omega, \mathcal{F}, \pi)$. Consider a finite time interval $[0, T]$, where T is an integer representing the number of discrete time periods of unit length. Divide this time interval into nT sub-intervals of length $s = 1/n$. Let $\phi_{k,n} = \left[\frac{k-1}{n}\right]$ denote the largest integer such that it is smaller than or equal to $\frac{k-1}{n}$. For $k = 1, 2, \cdots, nT$, we define the following stochastic process with the known initial value (S_0, h_1):

$$S_{ks}^{(n)} = S_{\phi_{k,n}}^{(n)} \prod_{i=n\phi_{k,n}+1}^{k} \left[u_{\phi_{k,n}}^{(n)} 1_{\{\epsilon_i \geq c(\phi_{k,n})\}} + d_{\phi_{k,n}}^{(n)} 1_{\{\epsilon_i < c(\phi_{k,n})\}} \right] \quad (1)$$

$$u_{\phi_{k,n}}^{(n)} = \exp\left(\frac{r}{n} + \sqrt{\frac{h_{\phi_{k,n}+1}^{(n)}}{n}} \right) \quad (2)$$

$$d_{\phi_{k,n}}^{(n)} = \exp\left(\frac{r}{n} - \sqrt{\frac{h_{\phi_{k,n}+1}^{(n)}}{n}} \right) \quad (3)$$

$$p_{\phi_{k,n}}^{(n)} = \max\left\{ v, \min\left[1 - v, \frac{\exp\left(\frac{r+\lambda\sqrt{h_{\phi_{k,n}+1}^{(n)}}}{n} \right) - d_{\phi_{k,n}}^{(n)}}{u_{\phi_{k,n}}^{(n)} - d_{\phi_{k,n}}^{(n)}} \right] \right\} \quad \text{for } 0 < v < \frac{1}{2} \quad (4)$$

$$h_{\phi_{k,n}+1}^{(n)} = \beta_0 + \beta_1 h_{\phi_{k,n}}^{(n)} + \beta_2 \left\{ \ln\left(\frac{S_{\phi_{k,n}}^{(n)}}{S_{\phi_{k,n}-1}^{(n)}} \right) - \left[r + \left(2p_{\phi_{k,n}-1}^{(n)} - 1 \right) \sqrt{n h_{\phi_{k,n}}^{(n)}} \right] - \theta \sqrt{h_{\phi_{k,n}}^{(n)}} \right\}^2 \quad (5)$$

where $c(\phi_{k,n})$ is the unique solution to $\Pr^\pi \{\epsilon_1 \geq c(\phi_{k,n})\} = p_{\phi_{k,n}}^{(n)}$; $\mathcal{F}_{\phi_{k,n}}$ be the σ-field generated by $\{\epsilon_i; i = 1, 2, \cdots, n\phi_{k,n}\}$ with $\mathcal{F}_0$ being defined as the trivial σ-field; $\beta_0 > 0, \beta_1 \geq 0$ and $\beta_2 \geq 0$; r is a constant interest rate (continuously compounded) over a period of unit length; and λ can be interpreted as the unit risk premium, which will become clear later. Note that the above definition ensures that $h_{\phi_{k,n}+1}^{(n)}$ is $\mathcal{F}_{\phi_{k,n}}$-measurable. Thus, $u_{\phi_{k,n}}^{(n)}$, $d_{\phi_{k,n}}^{(n)}$ and

$p^{(n)}_{\phi_{k,n}}$ are $\mathcal{F}_{\phi_{k,n}}$-measurable. The max and min operators are used to ensure $v < p^{(n)}_{\phi_{k,n}} < 1-v$ so that the model will not degenerate. From time 0 to 1, the above stochastic process is a standard recombined binomial lattice with the up (down) move multiplier $u^{(n)}_0$ ($d^{(n)}_0$) and the probability $p^{(n)}_0$ ($1-p^{(n)}_0$) which are determined solely by the known constant h_1. As the stochastic process progresses into the interval between time 1 and 2, the lattice begins to branch out depending on the value of $h^{(n)}_2$, which is in turn determined by the values of h_1, S_0 and S_1. Each branch, however, forms a recombined sub-lattice until reaching time 3. The process of branching out recombined sub-lattices continues. The up (down) probability is fixed within a recombined sub-lattice and differs for a different recombined sub-lattice. This semi-recombined binomial lattice over $[0,2]$ is illustrated in Figure 2.

Although we use the max and min operator to ensure $p^{(n)}_{\phi_{k,n}}$ being a legitimate probability bounded between v and $1-v$, it can be shown that for large n, $p^{(n)}_{\phi_{k,n}}$ naturally falls between v and $1-v$ in π-probability. Formally, we let

$$p^{*(n)}_{\phi_{k,n}} = \frac{\exp\left(\frac{r+\lambda\sqrt{h^{(n)}_{\phi_{k,n}+1}}}{n}\right) - d^{(n)}_{\phi_{k,n}}}{u^{(n)}_{\phi_{k,n}} - d^{(n)}_{\phi_{k,n}}}. \tag{6}$$

We prove in Appendix that $p^{*(n)}_{\phi_{k,n}} \to \frac{1}{2}$ in π-probability as $n \to \infty$. Since the particular combination of max and min operators in defining the relationship between $p^{*(n)}_{\phi_{k,n}}$ and $p^{(n)}_{\phi_{k,n}}$ is a continuous function, we conclude that $p^{(n)}_{\phi_{k,n}} \to \frac{1}{2}$ in π-probability as $n \to \infty$.

The conditional mean and variance of the continuously compounded return over one discrete time period of unit length can be derived. They are

$$\begin{aligned} & E\left[\ln\left(\frac{S^{(n)}_{\phi_{k,n}+1}}{S^{(n)}_{\phi_{k,n}}}\right)\middle|\mathcal{F}_{\phi_{k,n}}\right] \\ = \; & r + \left(2p^{(n)}_{\phi_{k,n}} - 1\right)\sqrt{nh^{(n)}_{\phi_{k,n}+1}} \\ = \; & \begin{cases} r + \lambda\sqrt{h^{(n)}_{\phi_{k,n}+1}} - \frac{1}{2}h^{(n)}_{\phi_{k,n}+1} + O_1\left(\frac{1}{\sqrt{n}}\right) & \text{if } v < p^{(n)}_{\phi_{k,n}} < 1-v \\ r + (1-2v)\sqrt{nh^{(n)}_{\phi_{k,n}+1}} & \text{if } p^{(n)}_{\phi_{k,n}} = 1-v \\ r - (1-2v)\sqrt{nh^{(n)}_{\phi_{k,n}+1}} & \text{if } p^{(n)}_{\phi_{k,n}} = v \end{cases} \end{aligned} \tag{7}$$

and

$$Var\left[\ln\left(\frac{S^{(n)}_{\phi_{k,n}+1}}{S^{(n)}_{\phi_{k,n}}}\right)\middle|\mathcal{F}_{\phi_{k,n}}\right] = \begin{cases} h^{(n)}_{\phi_{k,n}+1} + O_1\left(\frac{1}{n}\right) & \text{if } v < p^{(n)}_{\phi_{k,n}} < 1-v \\ 4v(1-v)h^{(n)}_{\phi_{k,n}+1} & \text{if } p^{(n)}_{\phi_{k,n}} = v \text{ or } 1-v \end{cases}, \tag{8}$$

where $O_k\left(\frac{1}{n^q}\right)$ denotes the quantity whose L^k-norm is in the order of $\frac{1}{n^q}$. The technical details are given in Appendix. For discussion purposes, we will view the conditional mean and variance as $r+\lambda\sqrt{h^{(n)}_{\phi_{k,n}+1}}-\frac{1}{2}h^{(n)}_{\phi_{k,n}+1}$ and $h^{(n)}_{\phi_{k,n}+1}$, respectively. In other words, we restrict our discussion to the case of $v<p^{(n)}_{\phi_{k,n}}<1-v$ and ignore the $O_1\left(\frac{1}{n}\right)$ and $O_1\left(\frac{1}{\sqrt{n}}\right)$ terms.

The above result implies that $h^{(n)}_{\phi_{k,n}+1}$ is indeed the conditional variance. The stochastic process $\left\{h_1,h^{(n)}_2,h^{(n)}_3,\cdots\right\}$ is thus the same as Engle and Ng's (1993) NGARCH(1,1) dynamic for the conditional variance of the continuously compounded asset return over integer time points. The conditional variance is shocked by the unanticipated return innovation over two adjacent integer time points. It is an unanticipated shock because $\left(2p^{(n)}_{\phi_{k,n}}-1\right)\sqrt{nh^{(n)}_{\phi_{k,n}+1}}$ can be interpreted as the total risk premium, which for all practical purposes equals $\lambda\sqrt{h^{(n)}_{\phi_{k,n}+1}}-\frac{1}{2}h^{(n)}_{\phi_{k,n}+1}$. Parameter λ can be interpreted as the unit risk premium (per unit of conditional standard deviation). As usual, the term $-\frac{1}{2}h^{(n)}_{\phi_{k,n}+1}$ arises naturally when one deals with continuously compounded returns. Parameter θ captures asymmetric responses to positive and negative unanticipated return shocks. If $\theta=0$, the volatility dynamic reduces to the linear GARCH(1,1) model of Bollerslev (1986) and Taylor (1986).

3 Risk premium dependent valuation in complete markets

The semi-recombined binomial lattice constitutes a complete market. This is obvious because any contract contingent on $S^{(n)}_{ks}$ can be replicated by locally forming a portfolio of the underlying asset and the risk-free asset. The standard arbitrage argument easily gives rise to the so-called risk-neutral probability of the up move for all binomial steps within a recombined sub-lattice in the interval $[\phi_{k,n},\phi_{k,n}+1]$ as $\left[\exp\left(\frac{r}{n}\right)-d^{(n)}_{\phi_{k,n}}\right]/\left(u^{(n)}_{\phi_{k,n}}-d^{(n)}_{\phi_{k,n}}\right)$. Putting together the whole complete-market pricing system, we have

$$S^{(n)}_{ks} = S^{(n)}_{\phi_{k,n}}\prod_{i=n\phi_{k,n}+1}^{k}\left[u^{(n)}_{\phi_{k,n}}1_{\{\epsilon_i\geq c^*(\phi_{k,n})\}}+d^{(n)}_{\phi_{k,n}}1_{\{\epsilon_i<c^*(\phi_{k,n})\}}\right] \tag{9}$$

$$u^{(n)}_{\phi_{k,n}} = \exp\left(r+\sqrt{\frac{h^{(n)}_{\phi_{k,n}+1}}{n}}\right) \tag{10}$$

$$d^{(n)}_{\phi_{k,n}} = \exp\left(r-\sqrt{\frac{h^{(n)}_{\phi_{k,n}+1}}{n}}\right) \tag{11}$$

$$q^{(n)}_{\phi_{k,n}} = \frac{\exp\left(\frac{r}{n}\right) - d^{(n)}_{\phi_{k,n}}}{u^{(n)}_{\phi_{k,n}} - d^{(n)}_{\phi_{k,n}}} \tag{12}$$

$$h^{(n)}_{\phi_{k,n}+1} = \beta_0 + \beta_1 h^{(n)}_{\phi_{k,n}} + \beta_2 \left\{ \ln\left(\frac{S^{(n)}_{\phi_{k,n}}}{S^{(n)}_{\phi_{k,n}-1}}\right) - \left[r + \left(2p^{(n)}_{\phi_{k,n}-1} - 1\right)\sqrt{nh^{(n)}_{\phi_{k,n}}}\right] - \theta\sqrt{h^{(n)}_{\phi_{k,n}}}\right\}^2 \tag{13}$$

where $c^*(\phi_{k,n})$ is the unique solution to $\Pr^\pi\{\epsilon_1 \geq c^*(\phi_{k,n})\} = q^{(n)}_{\phi_{k,n}}$. Note that $1_{\{\epsilon_i \geq c^*(\phi_{k,n})\}}$ is simply a Bernoulli random variable with a probability of obtaining 1 at $q^{(n)}_{\phi_{k,n}}$. One can always construct the sequence differently as opposed to using $\{\epsilon_i; i = 1, 2, \cdots\}$ as is in here. As long as $q^{(n)}_{\phi_{k,n}}$ is maintained, all pricing systems will be distributionally equivalent. Because $h^{(n)}_{\phi_{k,n}+1}$ is bounded below by $\beta_0 > 0$, we have $0 < q^{(n)}_{\phi_{k,n}} < 1$ always. Unlikely $p^{(n)}_{\phi_{k,n}}$ in the preceding section, there is no need to use the max and min operators to ensure a legitimate probability. In fact, this is the reason for having $u^{(n)}_{\phi_{k,n}}$ and $d^{(n)}_{\phi_{k,n}}$ in this particular form because for any n, the risk-neutral probability $q^{(n)}_{\phi_{k,n}}$ always remains strictly between 0 and 1.[1]

Since the up and down moves depend on the realization of the process up to the beginning point of a discrete period of unit length, all binomial steps within a recombined sub-lattice must share the same up and down risk-neutral probability. Note that $q^{(n)}_{\phi_{k,n}}$ is globally stochastic but locally deterministic due to the fact that both $u^{(n)}_{\phi_{k,n}}$ and $d^{(n)}_{\phi_{k,n}}$ are $\mathcal{F}_{\phi_{k,n}}$-measurable. In this market characterized by the semi-recombined binomial lattice, $q^{(n)}_{\phi_{k,n}}$ is stochastic and is a function of the total risk premium $\left(2p^{(n)}_{\phi_{k,n}-1} - 1\right)\sqrt{nh^{(n)}_{\phi_{k,n}}}$ (or the unit risk premium λ) unless $\beta_2 = 0$. In other words, the so-called risk-neutral probability is no longer independent of the expected return of the underlying asset. To avoid confusion, we will refer to $q^{(n)}_{\phi_{k,n}}$ as the complete-market pricing probability.

If $h^{(n)}_{\phi_{k,n}+1}$ is a constant, i.e., $\beta_1 = \beta_2 = 0$, the complete-market pricing probability becomes a constant and does not depend on the underlying asset's risk premium (or expected return). The lattice becomes a standard globally recombined binomial tree. Under such a scenario, the standard risk-neutral valuation result prevails; that is, the price of a derivative contract is independent of the expected return of the underlying asset. Our result has a profound implication for the general option pricing theory. Specifically, complete markets *per se* do not give rise to risk-neutral valuation of derivatives. The standard result has much to do

[1] In the Cox, Ross and Rubinstein (1979) constant-volatility construction, $u = \exp(\sigma/\sqrt{n})$ and $d = 1/u$. The corresponding risk-neutral probability is $q = \frac{\exp(r/n)-d}{u-d}$. Although one can ensure $0 < q < 1$ for large enough n, q may be greater than 1 for smaller values of n. This is not a serious constraint when parameters are constant because the threshold value for n can be determined once and for all. In the stochastic volatility model, a similar construction will cause the threshold value of n to be state contingent. In other words, for a given n-step lattice, some sub-lattices may not be consistent with the basic requirement for being a legitimate model for asset dynamics.

with other distributional features of the model as with complete markets.

Is stochastic volatility in complete markets the root cause for the breakdown of the standard risk-neutral valuation conclusion? The answer turns out to be a negative one. We now show why it is the case. If we replace equation (5) by

$$h^{(n)}_{\phi_{k,n}+1} = \beta_0 + \beta_1 h^{(n)}_{\phi_{k,n}} + \beta_2 \left[\ln\left(\frac{S^{(n)}_{\phi_{k,n}}}{S^{(n)}_{\phi_{k,n}-1}} \right) - \theta \sqrt{h^{(n)}_{\phi_{k,n}}} \right]^2, \tag{14}$$

we still have stochastic volatilities except that the conditional volatility is shocked by past returns rather than past unanticipated return innovations. This change still gives rise to a complete market characterized by a semi-recombined lattice. The resulting pricing system has a quite different theoretical feature, however. Specifically, $q^{(n)}_{\phi_{k,n}}$ is stochastic but is no longer a function of the risk premium $\left(2p^{(n)}_{\phi_{k,n}-1} - 1\right)\sqrt{nh^{(n)}_{\phi_{k,n}}}$, suggesting that the standard risk-neutral pricing conclusion prevails in this model. In summary, complete markets do not necessarily lead to the conclusion about the independence of risk premium (or expected return) of the underlying asset for pricing derivative contracts. Other features of the model are also important. If volatility is deterministic, the standard risk-neutral pricing conclusion is valid. If volatility is stochastic, in order to obtain the standard risk-neutral pricing conclusion, the volatility dynamic must not be shocked by the past unanticipated return innovations.

Denote the expectation and variance under the complete-market pricing system in equations (9)-(13) by $E^*(\cdot)$ and $Var^*(\cdot)$.

$$E^*\left[\ln\left(\frac{S^{(n)}_{\phi_{k,n}+1}}{S^{(n)}_{\phi_{k,n}}} \right) \middle| \mathcal{F}_{\phi_{k,n}} \right] = r - \frac{1}{2} h^{(n)}_{\phi_{k,n}+1} + O_1\left(\frac{1}{\sqrt{n}} \right) \text{ and} \tag{15}$$

$$Var^*\left[\ln\left(\frac{S^{(n)}_{\phi_{k,n}+1}}{S^{(n)}_{\phi_{k,n}}} \right) \middle| \mathcal{F}_{\phi_{k,n}} \right] = h^{(n)}_{\phi_{k,n}+1} + O_1\left(\frac{1}{n} \right). \tag{16}$$

Note that the derivations are similar to and, in fact, simpler than those for equations (7)-(8) because $0 < q^{(n)}_{\phi_{k,n}} < 1$ always. These results imply that the conditional mean under the complete-market pricing probability is changed to the risk-free rate but the conditional variance remains unchanged. The unanticipated shock (in terms of the complete-market pricing probability) becomes $\ln\left(S^{(n)}_{\phi_{k,n}+1}/S^{(n)}_{\phi_{k,n}}\right) - \left(r - \frac{1}{2}h^{(n)}_{\phi_{k,n}+1}\right)$. The asymmetry parameter governing the conditional volatility dynamic is thus changed from θ to $\theta + \sqrt{n}\left(2p^{(n)}_{\phi_{k,n}-1} - 1\right) + \frac{1}{2}\sqrt{h^{(n)}_{\phi_{k,n}}}$ (or loosely $\theta + \lambda$).

The complete-market pricing system and the system under the physical probability measure differ in their asymmetric response parameters. This difference has an important implication for the overall volatility level of the cumulative asset return. Let $\varphi = \beta_1 + \beta_2(1+\theta^2)$ and $\psi = \beta_1 + \beta_2 \left[1 + (\theta+\lambda)^2\right]$. Denote by $o(1)$ the term tending to 0 as n goes to ∞. We have the following results (see Appendix for details):

$$E\left\{\left[\ln\left(\frac{S_T^{(n)}}{S_0}\right) - \sum_{i=1}^{T} E\left[\ln\left(\frac{S_i^{(n)}}{S_{i-1}^{(n)}}\right)\middle| \mathcal{F}_{i-1}\right]\right]^2 \middle| \mathcal{F}_0\right\}$$
$$= \begin{cases} \frac{\beta_0 T}{1-\varphi} + \frac{1-\varphi^T}{1-\varphi}\left[h_1 - \frac{\beta_0}{1-\varphi}\right] + o(1) & \text{if } \varphi \neq 1 \\ \frac{\beta_0 T(T-1)}{2} + h_1 T + o(1) & \text{if } \varphi = 1 \end{cases} \tag{17}$$

and

$$E^*\left\{\left[\ln\left(\frac{S_T^{(n)}}{S_0}\right) - \sum_{i=1}^{T} E^*\left[\ln\left(\frac{S_i^{(n)}}{S_{i-1}^{(n)}}\right)\middle| \mathcal{F}_{i-1}\right]\right]^2 \middle| \mathcal{F}_0\right\}$$
$$= \begin{cases} \frac{\beta_0 T}{1-\psi} + \frac{1-\psi^T}{1-\psi}\left[h_1 - \frac{\beta_0}{1-\psi}\right] + o(1) & \text{if } \psi \neq 1 \\ \frac{\beta_0 T(T-1)}{2} + h_1 T + o(1) & \text{if } \psi = 1 \end{cases}. \tag{18}$$

Note that if θ and λ share the same sign, then $\psi \geq \varphi$. By ignoring the negligible terms, we have

$$E^*\left\{\left[\ln\left(\frac{S_T^{(n)}}{S_0}\right) - \sum_{i=1}^{T} E^*\left[\ln\left(\frac{S_i^{(n)}}{S_{i-1}^{(n)}}\right)\middle| \mathcal{F}_{i-1}\right]\right]^2 \middle| \mathcal{F}_0\right\}$$
$$\geq E\left\{\left[\ln\left(\frac{S_T^{(n)}}{S_0}\right) - \sum_{i=1}^{T} E\left[\ln\left(\frac{S_i^{(n)}}{S_{i-1}^{(n)}}\right)\middle| \mathcal{F}_{i-1}\right]\right]^2 \middle| \mathcal{F}_0\right\}, \tag{19}$$

if $sign(\theta) = sign(\lambda)$. Since equity returns are known to exhibit a negative asymmetric volatility response to return shocks, we expect $\theta > 0$. This in conjunction with a positive risk premium, i.e., $\lambda > 0$, suggests that the complete-market pricing system should use a cumulative volatility that is higher than the cumulative volatility implied by the physical system. A similar conclusion was first established by Duan (1995) in the GARCH framework, although his model constitutes an incomplete market. The difference in the cumulative volatility levels between the complete-market pricing system and the physical system is of course due to our modeling of conditional volatility as a function of the unanticipated component of the price change instead of the total price change. This modeling approach is, in fact, the essence of the GARCH model, and has a great deal to do with its empirical success. Intuitively, it is nonsensical to let the conditional volatility respond to the anticipated component of the price change because the anticipated component can hardly be viewed as a shock (or surprise).

A logical implication of (19) is that the prices of derivative contracts are direct functions of the risk premium for the underlying asset. This complete-market pricing theory further suggests that the historical and/or realized volatility of the underlying asset should be lower than the Black-Scholes implied volatilities for the options traded on this asset. The higher is the risk premium, the larger is the discrepancy. Interestingly, the empirical evidence shows that the historical volatility is indeed substantially lower than the Black-Scholes implied volatilities.

4 Limiting models

In this section, we study the limit of the semi-recombined binomial lattice model and that of its companion pricing system. We first define a family of continuous-time, right-continuous stochastic processes over $[0, T]$, denoted by $\left\{S^{(n)}; n = 1, 2, \cdots\right\}$. They are constructed from the semi-recombined lattice by setting $S_t^{(n)} = S_{ks}^{(n)}$ for $ks \le t < (k+1)s$. Clearly, the sample paths of $S^{(n)}$ belong to $D[0, T]$, the space of functions on $[0, T]$ that are right-continuous and have left-hand limits. Let $\mathcal{D}$ denote the σ-field of Borel sets in $D[0, T]$ for the Skorohod topology. The distribution on $(D[0, T], \mathcal{D})$ induced by $S^{(n)}$ in accordance with equations (1)-(5) is denoted by P_n and the one induced by $S^{(n)}$ in accordance with equations (9)-(13) is denoted by Q_n. Denote weak convergence in measure by $\Rightarrow$.

The following result implies that the limiting model of the semi-recombined binomial lattice is the continuous-time version of the NGARCH(1,1) model, a slightly more general than the continuous-time version first used in Kallsen and Taqqu (1998).

Theorem 1 *For $S^{(n)}$ in equations (1)-(5), $S^{(n)} \xrightarrow{D} S$ (i.e., $P_n \Rightarrow P$) as $n \to \infty$, where P is the distribution on $(D[0, T], \mathcal{D})$ induced by S, a stochastic process over $[0, T]$ evolving as*

$$d \ln S_t = \left(r + \lambda\sigma_{t_-} - \frac{1}{2}\sigma_{t_-}^2\right) dt + \sigma_{t\_} dW_t \tag{20}$$

$$\sigma_t^2 = h_i \quad for\ i - 1 \le t < i \tag{21}$$

$$h_i = \beta_0 + \beta_1 h_{i-1} + \beta_2 h_{i-1}\left[(W_{i-1} - W_{i-2}) - \theta\right]^2 \tag{22}$$

with the known initial value (S_0, h_1) and W_t is a Wiener process.

Proof: see Appendix.

Note that the limiting model, as a result of introducing parameter θ, allows for the asymmetric volatility responses to positive and negative return innovations. The limiting model is essentially the NGARCH(1,1) model of Engle and Ng (1993) connected over discrete time points piecewise by geometric Brownian motions. Under the limiting model, the asset price process has continuous sample paths but the volatility takes jumps at those pre-specified discrete time points. Because of this feature, the limiting model constitutes a continuous-time complete market and the standard martingale pricing conclusion of Harrison and Pliska (1981) should therefore directly apply. Indeed, that is the conclusion of Kallsen and Taqqu (1998).

The relationship between the semi-recombined binomial lattice and the continuous-time GARCH model can be likened to the one between the standard binomial lattice and the geometric Brownian motion. They are complete market models of discrete-time and continuous-time counterparts. It is therefore natural to expect that the pricing conclusion of the discrete-time model will carry over to its continuous-time equivalent model. Indeed, there is a weak convergence result for the pricing system.

Theorem 2 *For $S^{(n)}$ in equations (9)-(13), $S^{(n)} \xrightarrow{D} S$ (i.e., $Q_n \Rightarrow Q$) as $n \to \infty$, where Q is the distribution on $(D[0,T], \mathcal{D})$ induced by S, a stochastic process over $[0,T]$ evolving as*

$$d \ln S_t = \left(r - \frac{1}{2}\sigma_{t-}^2\right) dt + \sigma_{t-} dW_t^* \tag{23}$$

$$\sigma_t^2 = h_i \quad \text{for } i-1 \leq t < i \tag{24}$$

$$h_i = \beta_0 + \beta_1 h_{i-1} + \beta_2 h_{i-1} \left[(W_{i-1}^* - W_{i-2}^*) - \theta - \lambda\right]^2 \tag{25}$$

with the known initial value (S_0, h_1) and W_t^* is a Wiener process.

Proof: similar to Theorem 1.

As expected, our limiting result agrees with the pricing result of Kallsen and Taqqu (1998), which was directly derived using the continuous-time version of the GARCH model. If we restrict the attention to the discrete time points $\{0, 1, 2, \cdots, T\}$, the pricing conclusion is then the same as that of Duan (1995), where the theoretical importance of the risk premium for the underlying asset, i.e., λ, was first established. The results in this paper and Kallsen and Taqqu (1998) simply show that the conclusion of Duan (1995) is not simply due to incompleteness of the discrete-time GARCH model. As we have discussed in the preceding section, complete markets *per se* cannot render the risk premium of the underlying asset irrelevant for the purpose of derivatives pricing. In fact, the irrelevancy typically viewed in the options literature as an inherent feature of complete markets has much to do with the specific distributional feature of the model rather than with complete markets.

5 Appendix

5.1 $p^{*(n)}_{\phi_{k,n}}$ converges to 1/2 in probability

First, we note that by the Mean Value Theorem,

$$\begin{aligned}
\exp\left(\frac{r+\lambda\sqrt{h^{(n)}_{\phi_{k,n}+1}}}{n}\right) &= 1+\frac{r+\lambda\sqrt{h^{(n)}_{\phi_{k,n}+1}}}{n}+\frac{1}{2}e^{x}\frac{\left(r+\lambda\sqrt{h^{(n)}_{\phi_{k,n}+1}}\right)^2}{n^2} \\
\text{where } x &\in \left[0,\frac{r+\lambda\sqrt{h^{(n)}_{\phi_{k,n}+1}}}{n}\right] \text{ if } \frac{r+\lambda\sqrt{h^{(n)}_{\phi_{k,n}+1}}}{n}\geq 0 \text{ or} \\
x &\in \left[\frac{r+\lambda\sqrt{h^{(n)}_{\phi_{k,n}+1}}}{n},0\right] \text{ if } \frac{r+\lambda\sqrt{h^{(n)}_{\phi_{k,n}+1}}}{n}< 0; \\
\exp\left(\frac{r}{n}+\sqrt{\frac{h^{(n)}_{\phi_{k,n}+1}}{n}}\right) &= 1+\frac{r}{n}+\sqrt{\frac{h^{(n)}_{\phi_{k,n}+1}}{n}}+\frac{1}{2}\left(\frac{r}{n}+\sqrt{\frac{h^{(n)}_{\phi_{k,n}+1}}{n}}\right)^2+\frac{1}{6}e^{y}\left(\frac{r}{n}+\sqrt{\frac{h^{(n)}_{\phi_{k,n}+1}}{n}}\right)^3 \\
\text{where } y &\in \left[0,\frac{r}{n}+\sqrt{\frac{h^{(n)}_{\phi_{k,n}+1}}{n}}\right]; \\
\exp\left(\frac{r}{n}-\sqrt{\frac{h^{(n)}_{\phi_{k,n}+1}}{n}}\right) &= 1+\frac{r}{n}-\sqrt{\frac{h^{(n)}_{\phi_{k,n}+1}}{n}}+\frac{1}{2}\left(\frac{r}{n}-\sqrt{\frac{h^{(n)}_{\phi_{k,n}+1}}{n}}\right)^2+\frac{1}{6}e^{z}\left(\frac{r}{n}-\sqrt{\frac{h^{(n)}_{\phi_{k,n}+1}}{n}}\right)^3 \\
\text{where } z &\in \left[\frac{r}{n}-\sqrt{\frac{h^{(n)}_{\phi_{k,n}+1}}{n}},0\right] \text{ if } \frac{r}{n}<\sqrt{\frac{h^{(n)}_{\phi_{k,n}+1}}{n}} \text{ or} \\
z &\in \left[0,\frac{r}{n}-\sqrt{\frac{h^{(n)}_{\phi_{k,n}+1}}{n}}\right] \text{ if } \frac{r}{n}\geq\sqrt{\frac{h^{(n)}_{\phi_{k,n}+1}}{n}}.
\end{aligned}$$

Thus,

$$p^{*(n)}_{\phi_{k,n}} = \frac{\exp\left(\frac{r+\lambda\sqrt{h^{(n)}_{\phi_{k,n}+1}}}{n}\right)-d^{(n)}_{\phi_{k,n}}}{u^{(n)}_{\phi_{k,n}}-d^{(n)}_{\phi_{k,n}}}$$

$$= \frac{\frac{1}{2} + \frac{1}{2\sqrt{n}}\left(\lambda - \frac{1}{2}\sqrt{h^{(n)}_{\phi_{k,n}+1}}\right) + A_n}{1 + B_n}.$$

where

$$\begin{aligned}
A_n &= \frac{r}{2n} - \frac{r^2}{4n^{3/2}\sqrt{h^{(n)}_{\phi_{k,n}+1}}} + \frac{1}{4}e^x \frac{\left(r + \lambda\sqrt{h^{(n)}_{\phi_{k,n}+1}}\right)^2}{n^{3/2}\sqrt{h^{(n)}_{\phi_{k,n}+1}}} - \frac{1}{12n\sqrt{h^{(n)}_{\phi_{k,n}+1}}}e^z \left(\frac{r}{\sqrt{n}} - \sqrt{h^{(n)}_{\phi_{k,n}+1}}\right)^3 \\
B_n &= \frac{1}{4\sqrt{nh^{(n)}_{\phi_{k,n}+1}}}\left[\left(\frac{r}{\sqrt{n}} + \sqrt{h^{(n)}_{\phi_{k,n}+1}}\right)^2 - \left(\frac{r}{\sqrt{n}} - \sqrt{h^{(n)}_{\phi_{k,n}+1}}\right)^2\right] \\
&\quad + \frac{1}{12n\sqrt{h^{(n)}_{\phi_{k,n}+1}}}e^y \left(\frac{r}{\sqrt{n}} + \sqrt{h^{(n)}_{\phi_{k,n}+1}}\right)^3 - \frac{1}{12n\sqrt{h^{(n)}_{\phi_{k,n}+1}}}e^z \left(\frac{r}{\sqrt{n}} - \sqrt{h^{(n)}_{\phi_{k,n}+1}}\right)^3 \\
&= \frac{r}{n} + \frac{1}{12n\sqrt{h^{(n)}_{\phi_{k,n}+1}}}e^y \left(\frac{r}{\sqrt{n}} + \sqrt{h^{(n)}_{\phi_{k,n}+1}}\right)^3 - \frac{1}{12n\sqrt{h^{(n)}_{\phi_{k,n}+1}}}e^z \left(\frac{r}{\sqrt{n}} - \sqrt{h^{(n)}_{\phi_{k,n}+1}}\right)^3.
\end{aligned}$$

We now compute

$$\begin{aligned}
&\left\| p^{*(n)}_{\phi_{k,n}} - \frac{1}{2} - \frac{1}{2\sqrt{n}}\left(\lambda - \frac{1}{2}\sqrt{h^{(n)}_{\phi_{k,n}+1}}\right)\right\|_{L^2} \\
\leq\ &\left\| \left[\frac{1}{1+B_n} - 1\right]\left[\frac{1}{2} + \frac{1}{2\sqrt{n}}\left(\lambda - \frac{1}{2}\sqrt{h^{(n)}_{\phi_{k,n}+1}}\right)\right]\right\|_{L^2} + \left\|\frac{A_n}{1+B_n}\right\|_{L^2}. \qquad (26)
\end{aligned}$$

By Holder's inequality, we can bound the first part of the right-hand side of the above inequality as follows:

$$\begin{aligned}
&\left\| \left[\frac{1}{1+B_n} - 1\right]\left[\frac{1}{2} + \frac{1}{2\sqrt{n}}\left(\lambda - \frac{1}{2}\sqrt{h^{(n)}_{\phi_{k,n}+1}}\right)\right]\right\|_{L^2} \\
\leq\ &\left\|\frac{1}{1+B_n} - 1\right\|_{L^4} \left\|\frac{1}{2} + \frac{1}{2\sqrt{n}}\left(\lambda - \frac{1}{2}\sqrt{h^{(n)}_{\phi_{k,n}+1}}\right)\right\|_{L^4}.
\end{aligned}$$

Moreover, for $\epsilon > 0$,

$$\begin{aligned}
&n^{1-\epsilon}\left\|\frac{1}{1+B_n} - 1\right\|_{L^4} \\
=\ &n^{1-\epsilon}\left\|\frac{B_n}{1+B_n}\right\|_{L^4} \to 0 \text{ as } n \to \infty.
\end{aligned}$$

This is true due to two facts. First, B_n is in the order of $\frac{1}{n}$. Second $h^{(n)}_{\phi_{k,n}+1}$ is bounded below by $\beta_0 > 0$ and is a quadratic function of $\ln\left(\frac{S^{(n)}_{\phi_{k,n}}}{S^{(n)}_{\phi_{k,n}-1}}\right)$, which has values in the form of $\left(r + i\sqrt{\frac{h^{(n)}_{\phi_{k,n}}}{n}} - (n-i)\sqrt{\frac{h^{(n)}_{\phi_{k,n}}}{n}}\right)$ and a corresponding probability of $\left(\begin{array}{c} n \\ i \end{array}\right)\left(p^{(n)}_{\phi_{k,n}-1}\right)^i\left(1-p^{(n)}_{\phi_{k,n}-1}\right)^{n-i}$ for $i = 0, 1, 2, \cdots, n$. Therefore all terms in B_n involving $h^{(n)}_{\phi_{k,n}+1}$ have finite fourth moments. Similarly, $\left\|\frac{1}{2} + \frac{1}{2\sqrt{n}}\left(\lambda - \frac{1}{2}\sqrt{h^{(n)}_{\phi_{k,n}+1}}\right)\right\|_{L^4} < \infty$. The second part of the right-hand side of (26) is in order of $\frac{1}{n}$ because $n^{1-\epsilon}\left\|\frac{A_n}{1+B_n}\right\|_{L^2} \to 0$ as $n \to \infty$ for the same reason as before. Putting them together, we have for any $\epsilon > 0$,

$$n^{1-\epsilon}\left\|p^{*(n)}_{\phi_{k,n}} - \frac{1}{2} - \frac{1}{2\sqrt{n}}\left(\lambda - \frac{1}{2}\sqrt{h^{(n)}_{\phi_{k,n}+1}}\right)\right\|_{L^2} \to 0 \quad \text{as } n \to \infty.$$

In other words,

$$p^{*(n)}_{\phi_{k,n}} = \frac{1}{2} + \frac{1}{2\sqrt{n}}\left(\lambda - \frac{1}{2}\sqrt{h^{(n)}_{\phi_{k,n}+1}}\right) + O_2\left(\frac{1}{n}\right). \tag{27}$$

Finally, we compute

$$\begin{aligned}
\left\|p^{*(n)}_{\phi_{k,n}} - \frac{1}{2}\right\|_{L^2} &= \left\|\frac{1}{2\sqrt{n}}\left(\lambda - \frac{1}{2}\sqrt{h^{(n)}_{\phi_{k,n}+1}}\right) + O_2\left(\frac{1}{n}\right)\right\|_{L^2} \\
&\leq \frac{1}{2\sqrt{n}}\left\|\left(\lambda - \frac{1}{2}\sqrt{h^{(n)}_{\phi_{k,n}+1}}\right)\right\|_{L^2} + O\left(\frac{1}{n}\right) \\
&\to 0 \quad \text{as } n \to \infty \quad (\text{because} \left\|\left(\lambda - \frac{1}{2}\sqrt{h^{(n)}_{\phi_{k,n}+1}}\right)\right\|_{L^2} < \infty).
\end{aligned}$$

Thus, $p^{*(n)}_{\phi_{k,n}}$ converges to $\frac{1}{2}$ in L^2-norm and consequently converges in probability.

5.2 Derivations for the conditional mean and variance of the semi-recombined binomial lattice

The conditional mean can be computed as follows:

$$\begin{aligned}
& E\left[\ln\left(\frac{S^{(n)}_{\phi_{k,n}+1}}{S^{(n)}_{\phi_{k,n}}}\right)\middle| \mathcal{F}_{\phi_{k,n}}\right] \\
= \; & E\left\{\sum_{i=n\phi_{k,n}+1}^{n\phi_{k,n}+n} \ln\left[u^{(n)}_{\phi_{k,n}} 1_{\{\epsilon_i \geq c(\phi_{k,n})\}} + d^{(n)}_{\phi_{k,n}} 1_{\{\epsilon_i < c(\phi_{k,n})\}}\right]\middle| \mathcal{F}_{\phi_{k,n}}\right\}
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=0}^{n}\left(r + i\sqrt{\frac{h^{(n)}_{\phi_{k,n}+1}}{n}} - (n-i)\sqrt{\frac{h^{(n)}_{\phi_{k,n}+1}}{n}}\right)\binom{n}{i}\left(p^{(n)}_{\phi_{k,n}}\right)^{i}\left(1-p^{(n)}_{\phi_{k,n}}\right)^{n-i} \\
&= r - \sqrt{nh^{(n)}_{\phi_{k,n}+1}} + 2\sqrt{\frac{h^{(n)}_{\phi_{k,n}+1}}{n}}\sum_{i=0}^{n} i\binom{n}{i}\left(p^{(n)}_{\phi_{k,n}}\right)^{i}\left(1-p^{(n)}_{\phi_{k,n}}\right)^{n-i} \\
&= r + \left(2p^{(n)}_{\phi_{k,n}} - 1\right)\sqrt{nh^{(n)}_{\phi_{k,n}+1}} \qquad (28) \\
&= \begin{cases} r + \lambda\sqrt{h^{(n)}_{\phi_{k,n}+1}} - \frac{1}{2}h^{(n)}_{\phi_{k,n}+1} + O_1\left(\frac{1}{\sqrt{n}}\right) & \text{if } v < p^{(n)}_{\phi_{k,n}} < 1-v \\ r + (1-2v)\sqrt{nh^{(n)}_{\phi_{k,n}+1}} & \text{if } p^{(n)}_{\phi_{k,n}} = 1-v \\ r - (1-2v)\sqrt{nh^{(n)}_{\phi_{k,n}+1}} & \text{if } p^{(n)}_{\phi_{k,n}} = v \end{cases} . \text{ (by equation (27))}
\end{aligned}
$$

The conditional variance can be derived as

$$
\begin{aligned}
& E\left\{\left[\ln\left(\frac{S^{(n)}_{\phi_{k,n}+1}}{S^{(n)}_{\phi_{k,n}}}\right) - r - \left(2p^{(n)}_{\phi_{k,n}} - 1\right)\sqrt{nh^{(n)}_{\phi_{k,n}+1}}\right]^2 \middle| \mathcal{F}_{\phi_{k,n}}\right\} \text{ (by equation (28))} \\
&= E\left\{\left[\sum_{i=n\phi_{k,n}+1}^{n\phi_{k,n}+n}\ln\left[u^{(n)}_{\phi_{k,n}}1_{\{\epsilon_i \geq c(\phi_{k,n})\}} + d^{(n)}_{\phi_{k,n}}1_{\{\epsilon_i < c(\phi_{k,n})\}}\right] - r - \left(2p^{(n)}_{\phi_{k,n}} - 1\right)\sqrt{nh^{(n)}_{\phi_{k,n}+1}}\right]^2 \middle| \mathcal{F}_{\phi_{k,n}}\right\} \\
&= \sum_{i=0}^{n}\left(i\sqrt{\frac{h^{(n)}_{\phi_{k,n}+1}}{n}} - (n-i)\sqrt{\frac{h^{(n)}_{\phi_{k,n}+1}}{n}} - \left(2p^{(n)}_{\phi_{k,n}} - 1\right)\sqrt{nh^{(n)}_{\phi_{k,n}+1}}\right)^2\binom{n}{i}\left(p^{(n)}_{\phi_{k,n}}\right)^{i}\left(1-p^{(n)}_{\phi_{k,n}}\right)^{n-i} \\
&= \frac{4h^{(n)}_{\phi_{k,n}+1}}{n}\sum_{i=0}^{n}\left(i - np^{(n)}_{\phi_{k,n}}\right)^2\binom{n}{i}\left(p^{(n)}_{\phi_{k,n}}\right)^{i}\left(1-p^{(n)}_{\phi_{k,n}}\right)^{n-i} \\
&= 4p^{(n)}_{\phi_{k,n}}\left(1-p^{(n)}_{\phi_{k,n}}\right)h^{(n)}_{\phi_{k,n}+1} \\
&= \begin{cases} h^{(n)}_{\phi_{k,n}+1} + O_1\left(\frac{1}{n}\right) & \text{if } v < p^{(n)}_{\phi_{k,n}} < 1-v \\ 4v(1-v)h^{(n)}_{\phi_{k,n}+1} & \text{if } p^{(n)}_{\phi_{k,n}} = v \text{ or } 1-v \end{cases} . \text{ (by equation (27))}
\end{aligned}
$$

5.3 Derivations for two volatilities of the cumulative return

First, we consider

$$
\begin{aligned}
& \left\| 1_{\{v<p^{(n)}_{i-1}<1-v\}}h^{(n)}_i + 4v(1-v)1_{\{p^{(n)}_{i-1}=v \text{ or } 1-v\}}h^{(n)}_i - h^{(n)}_i \right\|_{L^1} \\
&= \left\| \left(1_{\{v<p^{(n)}_{i-1}<1-v\}} + 4v(1-v)1_{\{p^{(n)}_{i-1}=v \text{ or } 1-v\}} - 1\right)h^{(n)}_i \right\|_{L^1}
\end{aligned}
$$

$$
\begin{aligned}
&\le \left\| 1_{\{v<p_{i-1}^{(n)}<1-v\}} + 4v(1-v)1_{\{p_{i-1}^{(n)}=v \text{ or } 1-v\}} - 1 \right\|_{L^2} \left\| h_i^{(n)} \right\|_{L^2} \text{ (by Holder's inequality)} \\
&= |4v(1-v)-1| \left\| h_i^{(n)} \right\|_{L^2} \sqrt{\Pr\{p_{i-1}^{(n)} = v \text{ or } 1-v\}} \\
&\to 0 \text{ as } n \to \infty.
\end{aligned}
$$

Thus,

$$
1_{\{v<p_{i-1}^{(n)}<1-v\}} h_i^{(n)} + 4v(1-v)1_{\{p_{i-1}^{(n)}=v \text{ or } 1-v\}} h_i^{(n)} = h_i^{(n)} + o_1(1). \tag{29}
$$

Then,

$$
\begin{aligned}
& E\left\{ \left[\ln\left(\frac{S_T^{(n)}}{S_0}\right) - \sum_{i=1}^{T} E\left[\ln\left(\frac{S_i^{(n)}}{S_{i-1}^{(n)}}\right) \middle| \mathcal{F}_{i-1} \right] \right]^2 \middle| \mathcal{F}_0 \right\} \\
= \; & E\left\{ \left[\sum_{i=1}^{T} \left[\ln\left(\frac{S_i^{(n)}}{S_{i-1}^{(n)}}\right) - E\left[\ln\left(\frac{S_i^{(n)}}{S_{i-1}^{(n)}}\right) \middle| \mathcal{F}_{i-1} \right] \right] \right]^2 \middle| \mathcal{F}_0 \right\} \\
= \; & E\left\{ \begin{array}{c} \sum_{i=1}^{T} \sum_{j=1}^{T} \left[\ln\left(\frac{S_i^{(n)}}{S_{i-1}^{(n)}}\right) - E\left[\ln\left(\frac{S_i^{(n)}}{S_{i-1}^{(n)}}\right) \middle| \mathcal{F}_{i-1} \right] \right] \\ \times \left[\ln\left(\frac{S_j^{(n)}}{S_{j-1}^{(n)}}\right) - E\left[\ln\left(\frac{S_j^{(n)}}{S_{ij1}^{(n)}}\right) \middle| \mathcal{F}_{j-1} \right] \right] \end{array} \middle| \mathcal{F}_0 \right\} \\
= \; & \sum_{i=1}^{T} \sum_{j=1}^{T} E\left\{ E\left\{ \begin{array}{c} \left[\ln\left(\frac{S_i^{(n)}}{S_{i-1}^{(n)}}\right) - E\left[\ln\left(\frac{S_i^{(n)}}{S_{i-1}^{(n)}}\right) \middle| \mathcal{F}_{i-1} \right] \right] \\ \times \left[\ln\left(\frac{S_j^{(n)}}{S_{j-1}^{(n)}}\right) - E\left[\ln\left(\frac{S_j^{(n)}}{S_{ij1}^{(n)}}\right) \middle| \mathcal{F}_{j-1} \right] \right] \end{array} \middle| \mathcal{F}_{\max(i,j)-1} \right\} \middle| \mathcal{F}_0 \right\} \\
= \; & \sum_{i=1}^{T} E\left\{ E\left\{ \left[\ln\left(\frac{S_i^{(n)}}{S_{i-1}^{(n)}}\right) - E\left[\ln\left(\frac{S_i^{(n)}}{S_{i-1}^{(n)}}\right) \middle| \mathcal{F}_{i-1} \right] \right]^2 \middle| \mathcal{F}_{i-1} \right\} \middle| \mathcal{F}_0 \right\} \\
= \; & \sum_{i=1}^{T} E\left\{ \left(1_{\{v<p_{i-1}^{(n)}<1-v\}} h_i^{(n)} + 4v(1-v)1_{\{p_{i-1}^{(n)}=v \text{ or } 1-v\}} h_i^{(n)} \right) \middle| \mathcal{F}_0 \right\} + O\left(\frac{1}{n}\right) \\
& \text{(by equation (8))} \\
= \; & \sum_{i=1}^{T} E\left\{ h_i^{(n)} \middle| \mathcal{F}_0 \right\} + o(1). \text{ (by equation (29))}
\end{aligned}
$$

By equations (5), (7) and (8), we have

$$
\begin{aligned}
& E\left\{ h_i^{(n)} | \mathcal{F}_{i-2} \right\} \\
= \; & \beta_0 + \beta_1 h_{i-1}^{(n)} + \beta_2 \left[\left(1_{\{v<p_{i-2}^{(n)}<1-v\}} h_{i-1}^{(n)} + 4v(1-v)1_{\{p_{i-2}^{(n)}=v \text{ or } 1-v\}} h_{i-1}^{(n)} \right) + \theta^2 h_{i-1}^{(n)} + O_1\left(\frac{1}{n}\right) \right] \\
= \; & \beta_0 + \beta_1 h_{i-1}^{(n)} + \beta_2 \left(1+\theta^2 \right) h_{i-1}^{(n)} + o_1(1) \quad \text{(by equation (29))} \\
= \; & \beta_0 + \varphi h_{i-1}^{(n)} + o_1(1).
\end{aligned}
$$

Recall that $\varphi = \beta_1 + \beta_2(1+\theta^2)$. This in turn allows us to obtain

$$\begin{aligned} E\left\{h_i^{(n)}|\mathcal{F}_0\right\} &= \beta_0 + \varphi E\left\{h_{i-1}^{(n)}|\mathcal{F}_0\right\} + o(1) \\ &= \beta_0 + \beta_0\varphi + \cdots + \beta_0\varphi^{i-2} + \varphi^{i-1}h_1 + o(1) \\ &= \begin{cases} \frac{\beta_0\varphi^{i-1}}{1-\varphi} + \varphi^{i-1}h_1 + o(1) & \text{if } \varphi \neq 1 \\ (i-1)\beta_0 + h_1 + o(1) & \text{if } \varphi = 1 \end{cases}. \end{aligned}$$

We thus have

$$\begin{aligned} & E\left\{\left[\ln\left(\frac{S_T^{(n)}}{S_0}\right) - \sum_{i=1}^{T} E\left[\ln\left(\frac{S_i^{(n)}}{S_{i-1}^{(n)}}\right)\middle|\mathcal{F}_{i-1}\right]\right]^2\middle|\mathcal{F}_0\right\} \\ = & \begin{cases} \sum_{i=1}^{T}\left(\frac{\beta_0\varphi^{i-1}}{1-\varphi} + \varphi^{i-1}h_1\right) + o(1) & \text{if } \varphi \neq 1 \\ \sum_{i=1}^{T}\{(i-1)\beta_0 + h_1\} + o(1) & \text{if } \varphi = 1 \end{cases} \\ = & \begin{cases} \frac{\beta_0 T}{1-\varphi} + \frac{1-\varphi^T}{1-\varphi}\left[h_1 - \frac{\beta_0}{1-\varphi}\right] + o(1) & \text{if } \varphi \neq 1 \\ \frac{\beta_0 T(T-1)}{2} + h_1 T + o(1) & \text{if } \varphi = 1 \end{cases}. \end{aligned}$$

We have therefore established equation (17).

By equations (13), (15) and (16), we similarly have (except that $0 < q_i^{(n)} < 1$ always and yields a simpler derivation)

$$\begin{aligned} E^*\left\{h_i^{(n)}|\mathcal{F}_{i-1}\right\} &= \beta_0 + \beta_1 h_{i-1}^{(n)} + \beta_2\left[\left(1+(\theta+\lambda)^2\right)h_{i-1}^{(n)} + o_1(1)\right] \\ &= \beta_0 + \psi h_{i-1}^{(n)} + o_1(1). \end{aligned}$$

Recall that $\psi = \beta_1 + \beta_2\left[1+(\theta+\lambda)^2\right]$. By an analogous argument, we obtain

$$\begin{aligned} & E^*\left\{\left[\ln\left(\frac{S_T^{(n)}}{S_0}\right) - \sum_{i=1}^{T} E^*\left[\ln\left(\frac{S_i^{(n)}}{S_{i-1}^{(n)}}\right)\middle|\mathcal{F}_{i-1}\right]\right]^2\middle|\mathcal{F}_0\right\} \\ = & \begin{cases} \frac{\beta_0 T}{1-\psi} + \frac{1-\psi^T}{1-\psi}\left[h_1 - \frac{\beta_0}{1-\psi}\right] + o(1) & \text{if } \psi \neq 1 \\ \frac{\beta_0 T(T-1)}{2} + h_1 T + o(1) & \text{if } \psi = 1 \end{cases}. \end{aligned}$$

Equation (18) is thus established.

5.4 Proof of Theorem 1

It suffices to show two things to complete the weak convergence proof for the processes over $[0,T]$. First, if $(\ln S_i^{(n)}, h_{i+1}^{(n)}) \xrightarrow{D} (\ln S_i, h_{i+1})$, then any finite-dimensional distribution of

$\left\{\ln S_t^{(n)}, i \le t \le i+1\right\}$ converges to that of $\{\ln S_t; i \le t \le i+1\}$. This result can then be extended to the entire interval $[0,T]$. Second, we need to show tightness over $[0,T]$.

Let c_ϵ be some positive value such that $\Pr\{h_{i+1} > c_\epsilon\} \le \epsilon$. Denote the k-dimensional distribution of $\left\{\left(\ln\left(S_{t_1}^{(n)}/S_i^{(n)}\right), \cdots, \ln\left(S_{t_k}^{(n)}/S_{t_{k-1}}^{(n)}\right)\right); t_j \in [i, i+1]\right\}$ conditional on $(\ln S_i, h_{i+1})$ by $F^{(n)}(\mathbf{X}|\ln S_i, h_{i+1})$. The marginal distribution of $(\ln S_i^{(n)}, h_{i+1}^{(n)})$ is denoted by $G^{(n)}(\cdot,\cdot)$. Their corresponding distributions under the target process are denoted by $F(\mathbf{X}|\ln S_i, h_{i+1})$ and $G(\cdot,\cdot)$. We examine the distance between the two joint distribution functions evaluated at any $(k+1)$-dimensional point $\mathbf{Y} = (y_0, y_1, \cdots, y_k)$ as follows:

$$\begin{aligned}
& \left|\int_{z\le y_0} F^{(n)}\left((y_1,\cdots,y_k)\,|z,w\right) dG^{(n)}(z,w) - \int_{z\le y_0} F\left((y_1,\cdots,y_k)\,|z,w\right) dG(z,w)\right| \\
\le\ & \int_{z\le y_0} \left|F^{(n)}\left((y_1,\cdots,y_k)\,|z,w\right) - F\left((y_1,\cdots,y_k)\,|z,w\right)\right| dG^{(n)}(z,w) \\
& + \left|\int_{z\le y_0} F\left((y_1,\cdots,y_k)\,|z,w\right) dG^{(n)}(z,w) - \int_{z\le y_0} F\left((y_1,\cdots,y_k)\,|z,w\right) dG(z,w)\right|.
\end{aligned}$$

Since $F\left((y_1,\cdots,y_k)\,|z,w\right)$ is a bounded function,

$$\left|\int_{z\le y_0} F\left((y_1,\cdots,y_k)\,|z,w\right) dG^{(n)}(z,w) - \int_{z\le y_0} F\left((y_1,\cdots,y_k)\,|z,w\right) dG(z,w)\right| \to 0 \text{ as } n\to\infty$$

if $(\ln S_i^{(n)}, h_{i+1}^{(n)}) \xrightarrow{D} (\ln S_i, h_{i+1})$. It remains to consider

$$\begin{aligned}
& \int_{z\le y_0} \left|F^{(n)}\left((y_1,\cdots,y_k)\,|z,w\right) - F\left((y_1,\cdots,y_k)\,|z,w\right)\right| dG^{(n)}(z,w) \\
=\ & \int_{z\le y_0} 1_{\{w\le c_\epsilon\}} \left|F^{(n)}\left((y_1,\cdots,y_k)\,|z,w\right) - F\left((y_1,\cdots,y_k)\,|z,w\right)\right| dG^{(n)}(z,w) \\
& + \int_{z\le y_0} 1_{\{w> c_\epsilon\}} \left|F^{(n)}\left((y_1,\cdots,y_k)\,|z,w\right) - F\left((y_1,\cdots,y_k)\,|z,w\right)\right| dG^{(n)}(z,w).
\end{aligned}$$

Note that h_{i+1} is bounded below by $\beta_0 > 0$. By constraining $h_{i+1} \le c_\epsilon$, we restrict to a compact set of h_{i+1}. Moreover, both $F^{(n)}(\mathbf{X}|\ln S_i, h_{i+1})$ and $F(\mathbf{X}|\ln S_i, h_{i+1})$ do not depend on $\ln S_i$ once h_{i+1} is known. In other words, there exists n_1 such that for $n \ge n_1$,

$$1_{\{w\le c_\epsilon\}} \left|F^{(n)}\left((y_1,\cdots,y_k)\,|z,w\right) - F\left((y_1,\cdots,y_k)\,|z,w\right)\right| \le \epsilon,$$

because the standard binomial lattice weakly converges to the target process for every time segment $[i, i+1]$ if $(\ln S_i^{(n)}, h_{i+1}^{(n)}) = (\ln S_i, h_{i+1})$. It follows that

$$\int_{z\le y_0} 1_{\{w\le c_\epsilon\}} \left|F^{(n)}\left((y_1,\cdots,y_k)\,|z,w\right) - F\left((y_1,\cdots,y_k)\,|z,w\right)\right| dG^{(n)}(z,w) \le \epsilon.$$

Note that

$$\left|F^{(n)}\left((y_1,\cdots,y_k)\,|z,w\right) - F\left((y_1,\cdots,y_k)\,|z,w\right)\right| \le 1$$

because they are distribution functions. This then implies

$$\begin{aligned} & \int_{z\leq y_0} 1_{\{w>c_\epsilon\}} \left| F^{(n)}\left((y_1,\cdots,y_k)\,|z,w\right) - F\left((y_1,\cdots,y_k)\,|z,w\right)\right| dG^{(n)}(z,w) \\ \leq & \int_{z\leq y_0} 1_{\{w>c_\epsilon\}} dG^{(n)}(z,w). \\ \leq & \Pr\left\{h_{i+1}^{(n)} > c_\epsilon\right\}. \end{aligned}$$

If $(\ln S_i^{(n)}, h_{i+1}^{(n)}) \xrightarrow{D} (\ln S_i, h_{i+1})$, there exists n_2 such that for $n \geq n_2$,

$$\left|\Pr\left\{h_{i+1}^{(n)} > c_\epsilon\right\} - \Pr\left\{h_{i+1} > c_\epsilon\right\}\right| \leq \epsilon.$$

Therefore, for $n \geq \max(n_1, n_2)$, we have

$$\int_{z\leq y_0} \left| F^{(n)}\left((y_1,\cdots,y_k)\,|z,w\right) - F\left((y_1,\cdots,y_k)\,|z,w\right)\right| dG^{(n)}(z,w) \leq 3\epsilon$$

and then

$$\left|\int_{z\leq y_0} F^{(n)}\left((y_1,\cdots,y_k)\,|z,w\right) dG^{(n)}(z,w) - \int_{z\leq y_0} F\left((y_1,\cdots,y_k)\,|z,w\right) dG(z,w)\right| \to 0$$

as $n\to\infty$ if $(\ln S_i^{(n)}, h_{i+1}^{(n)}) \xrightarrow{D} (\ln S_i, h_{i+1})$.

Since the finite-dimensional distribution of $\left(\ln S_i^{(n)}, \ln S_{t_1}^{(n)}, \cdots, \ln S_{t_k}^{(n)}\right)$ is a continuous function of $\left(\ln S_i^{(n)}, \ln\left(S_{t_1}^{(n)}/S_i^{(n)}\right), \cdots, \ln\left(S_{t_k}^{(n)}/S_{t_{k-1}}^{(n)}\right)\right)$, we have

$$\left(\ln S_i^{(n)}, \ln S_{t_1}^{(n)}, \cdots, \ln S_{t_k}^{(n)}\right) \xrightarrow{D} \left(\ln S_i, \ln S_{t_1}, \cdots, \ln S_{t_k}\right)$$

if $(\ln S_i^{(n)}, h_{i+1}^{(n)}) \xrightarrow{D} (\ln S_i, h_{i+1})$.

We now turn to tightness. By Theorem 15.5 of Billingsley (1968), two conditions need to be checked. First, for each positive η, there exist an a such that $\Pr\left\{\ln S_0^{(n)} > a\right\} \leq \eta$ for $n \geq 1$. Since $\ln S_0^{(n)} = \ln S_0$, a fixed value, this condition is trivially met. Second, for each positive ϵ and η, there exists $0 < \delta < T$ and an integer n^* such that $\Pr\left\{\sup_{|\tau-t|<\delta}\left|\ln S_\tau^{(n)} - \ln S_t^{(n)}\right| \geq \epsilon\right\} \leq \eta$ for $n \geq n^*$. It is clearly possible to find some $c_\eta > 0$ and a positive integer n_1 such that for $n \geq n_1$, $\Pr\left\{\max_{i=1,2,\cdots,T} h_i^{(n)} > c_\eta\right\} \leq \eta/2$, The reason is similar to the first part of the proof. By our construction of the semi-recombined binomial lattice, the step size is bounded by $\max\left(\left|\frac{r}{n} + \sqrt{\frac{h_i^{(n)}}{n}}\right|, \left|\frac{r}{n} - \sqrt{\frac{h_i^{(n)}}{n}}\right|\right)$ over $[i-1, i]$. It is clear that under $\max_{i=1,2,\cdots,T} h_i^{(n)} \leq c_\eta$, the step size will be bounded by $\max\left(\left|\frac{r}{n}+\sqrt{\frac{c_\eta}{n}}\right|, \left|\frac{r}{n}-\sqrt{\frac{c_\eta}{n}}\right|\right)$ over $[0,T]$ with the binomial probability approaching 1/2. It is therefore always possible to find δ and n_2 such that for $n \geq n_2$,

$$\Pr\left\{\left(\sup_{|\tau-t|<\delta}\left|\ln S_\tau^{(n)} - \ln S_t^{(n)}\right| \geq \epsilon\right) \cap \left(\max_{i=1,2,\cdots,T} h_i^{(n)} \leq c_\eta\right)\right\} \leq \frac{\eta}{2}.$$

Finally, for $n \geq n^* = \max(n_1, n_2)$,

$$\begin{aligned}
& \Pr\left\{\sup_{|\tau-t|<\delta}\left|\ln S_\tau^{(n)} - \ln S_t^{(n)}\right| \geq \epsilon\right\} \\
= & \Pr\left\{\left(\sup_{|\tau-t|<\delta}\left|\ln S_\tau^{(n)} - \ln S_t^{(n)}\right| \geq \epsilon\right) \cap \left(\max_{i=1,2,\cdots,T} h_i^{(n)} \leq c_\eta\right)\right\} \\
& + \Pr\left\{\left(\sup_{|\tau-t|<\delta}\left|\ln S_\tau^{(n)} - \ln S_t^{(n)}\right| \geq \epsilon\right) \cap \left(\max_{i=1,2,\cdots,T} h_i^{(n)} > c_\eta\right)\right\} \\
\leq & \eta.
\end{aligned}$$

Tightness is thus established.

References

[1] Billingsley, P., 1968, Convergence of Probability Measures, John Wiley & Sons.

[2] Black, F. and M. Scholes, 1973, "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy* 81, 637-659.

[3] Bollerslev, T., 1986, "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics* 31, 307-327.

[4] Brennan, M., 1979, "The Pricing of Contingent Claims in Discrete Time Models," *Journal of Finance* 34, 53-68.

[5] Cox, J. and S. Ross, 1976, "The Valuation of Options for Alternative Stochastic Processes," *Journal of Financial Economics* 3, 145-166.

[6] Cox, J., S. Ross and M. Rubinstein, 1979, "Option Pricing: A Simplified Approach," *Journal of Financial Economics* 7, 229-263.

[7] Duan, J.-C., 1995, "The GARCH Option Pricing Model," *Mathematical Finance* 5, 13-32.

[8] Engle, R. and V. Ng, 1993, "Measuring and Testing the Impact of News on Volatility," *Journal of Finance* 48, 1749-1778.

[9] Harrison, M. and D. Kreps, 1978, "Martingales and Arbitrage in Multiperiod Securities Markets," *Journal of Economic Theory* 20, 381-408.

[10] Harrison, M. and S. Pliska, 1981, "Martingales and Stochastic Integrals in the Theory of Continuous Trading," *Stochastic Processes and Applications* 11, 215-260.

[11] Kallsen, J. and M. Taqqu, 1998, "Option pricing in ARCH-type models," *Mathematical Finance* 8, 13-26.

[12] Luenberger, D., 1998, Investment Science, Oxford University Press.

[13] Merton, R., 1973, "The Theory of Rational Option Pricing," *Bell J. Econ. Management Sci.* 4, 141-183.

[14] Rubinstein, M., 1976, "The Valuation of Uncertain Income Streams and the Pricing of Options," *Bell J. Econ. Management Sci.* 7, 407-425.

[15] Taylor, S., 1986, Modelling Financial Time Series, John Wiley & Sons.

[16] Wilmott, P., 1998, Derivatives - The Theory and Practice of Financial Engineering, John Wiley & Sons.

Figure 1. A diagram of the relationship among various models.

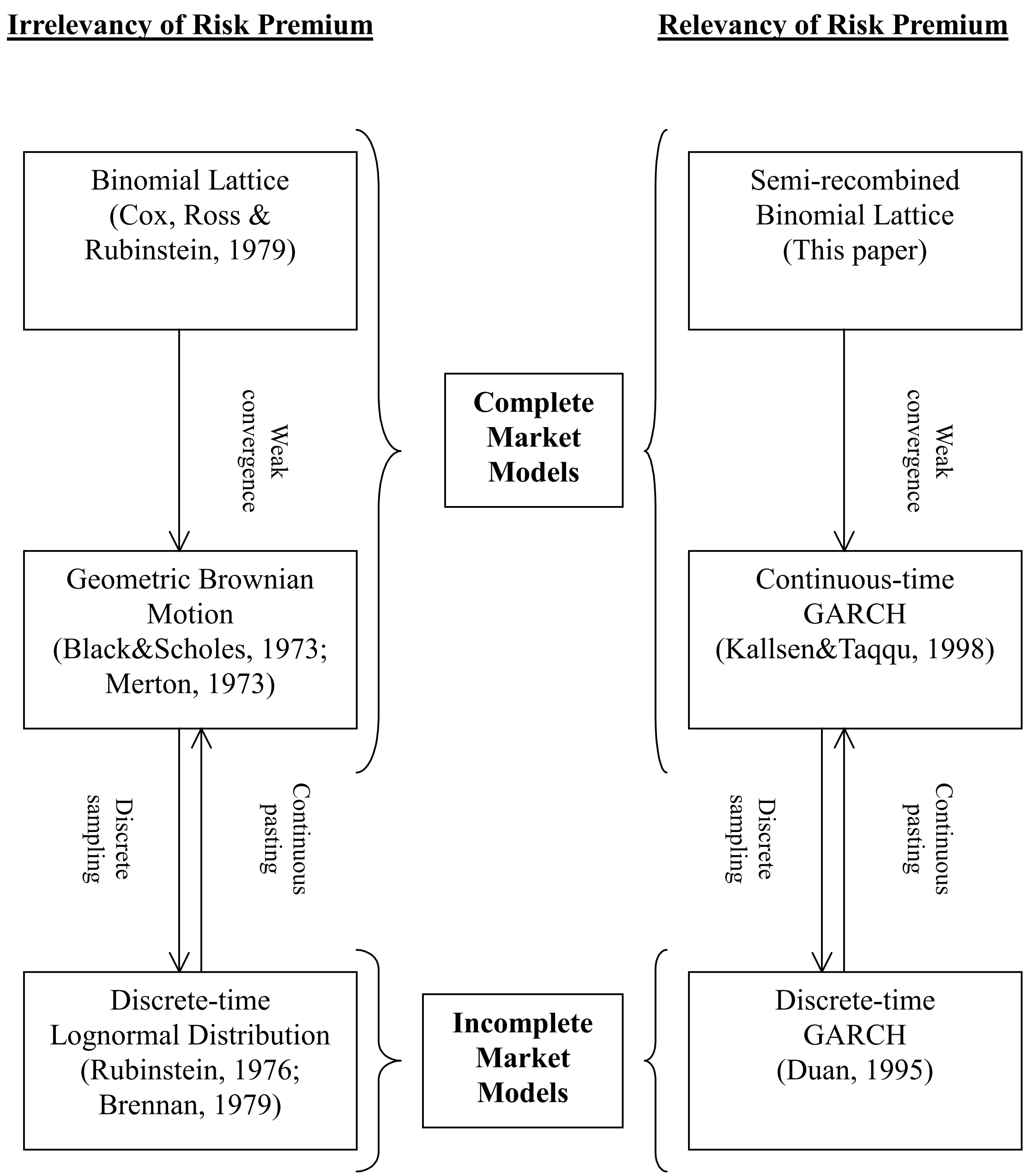
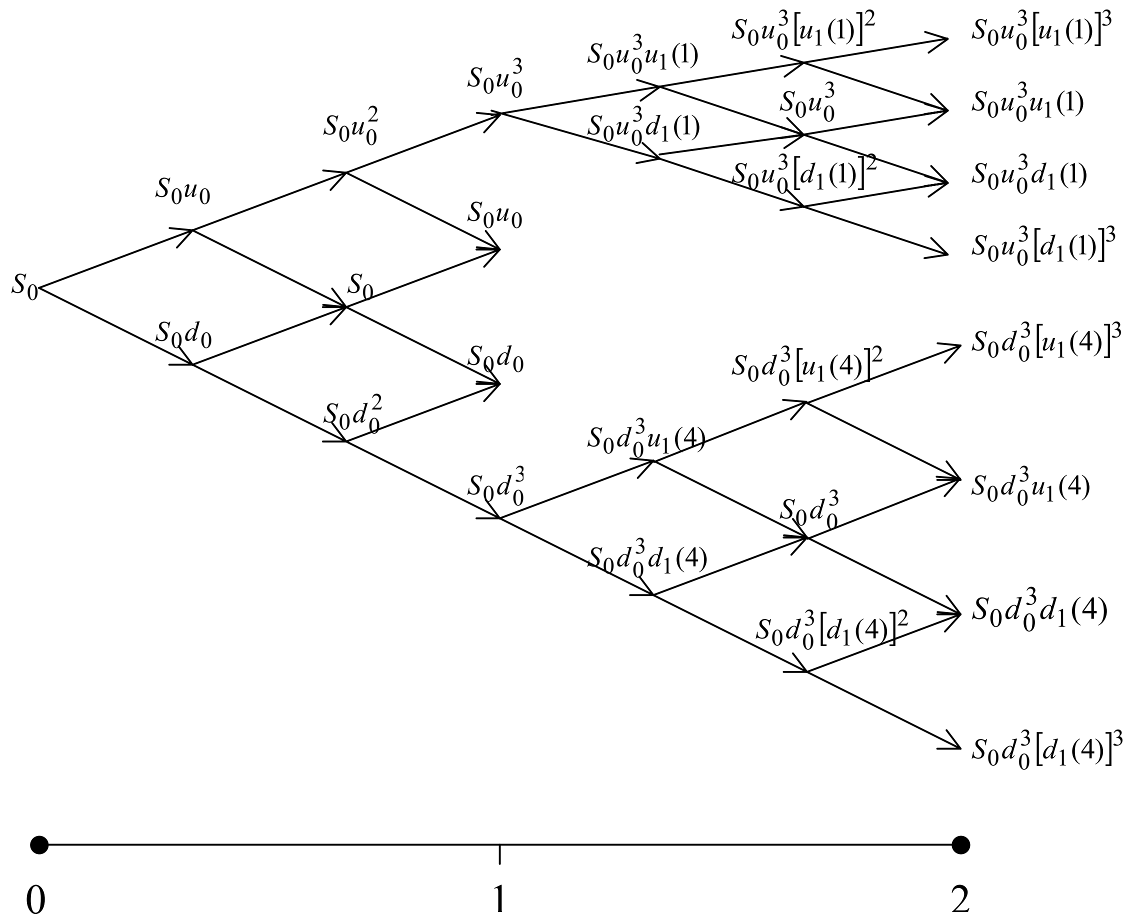


Figure 2. A semi-recombined binomial lattice (*n* = 3, i.e., 6 steps over [0,2]; *r* = 0) with the two middle sub-branches omitted.



Note: $u_0 \neq u_1(1) \neq u_1(4)$