

THE MARKOV CHAIN MARKET

BY

RAGNAR NORBERG

ABSTRACT

We consider a financial market driven by a continuous time homogeneous Markov chain. Conditions for absence of arbitrage and for completeness are spelled out, non-arbitrage pricing of derivatives is discussed, and details are worked out for some cases. Closed form expressions are obtained for interest rate derivatives. Computations typically amount to solving a set of first order partial differential equations. An excursion into risk minimization in the incomplete case illustrates the matrix techniques that are instrumental in the model.

KEYWORDS

Continuous time Markov chains, Martingale analysis, Arbitrage pricing theory, Risk minimization, Unit linked insurance.

INTRODUCTION

A. Prospectus

The theory of diffusion processes, with its wealth of powerful theorems and model variations, is an indispensable toolbox in modern financial mathematics. The seminal papers of Black and Scholes and Merton were crafted with Brownian motion, and so was the major part of the plethora of papers on arbitrage pricing theory and its ramifications that followed over the past good quarter of a century.

A main course of current research, initiated by the martingale approach to arbitrage pricing Harrison and Kreps (1979) and Harrison and Pliska (1981), aims at generalization and unification. Today the core of the matter is well understood in a general semimartingale setting, see e.g. Delbaen and Schachermayer (1994). Another course of research investigates special models, in particular Levy motion alternatives to the Brownian driving process, see e.g. Eberlein and Raible (1999). Pure jump processes have found their way into finance, ranging from plain Poisson processes introduced in Merton (1976) to fairly general marked point processes, see e.g. Björk et al. (1997). As a pedagogical exercise, the market driven by a binomial process has been intensively studied since it was proposed in Cox et al. (1979).

The present paper undertakes to study a financial market driven by a continuous time homogeneous Markov chain. The idea was launched in Norberg (1995) and reappeared in Elliott and Kopp (1998), the context being modeling of the spot rate of interest. These rudiments will here be developed into a model that delineates a financial market with a locally risk-free money account, risky assets, and all conceivable derivatives. The purpose of this exercise is two-fold: In the first place, there is an educative point in seeing how well established theory turns out in the framework of a general Markov chain market and, in particular, how and why it differs from the familiar Brownian motion driven market. In the second place, it is worthwhile investigating the potential of the model from a theoretical as well as from a practical point of view. Further motivation and discussion of the model is given in Section 5.

B. Contents of the paper

We hit the road in Section 2 by recapitulating basic definitions and results for the continuous time Markov chain. We proceed by presenting a market featuring this process as the driving mechanism and by spelling out conditions for absence of arbitrage and for completeness. In Section 3 we carry through the program for arbitrage pricing of derivatives in the Markov chain market and work out the details for some special cases. Special attention is paid to interest rate derivatives, for which closed form expressions are obtained. Section 4 addresses the Föllmer-Sondermann-Schweizer theory of risk minimization in the incomplete case. Its particulars for the Markov chain market are worked out in two examples, first for a unit linked life endowment, and second for hedging strategies involving a finite number of zero coupon bonds. The final Section 5 discusses the versatility and potential uses of the model. It also raises the somewhat intricate issue of existence and continuity of the derivatives involved in the differential equations for state prices, which finds its resolution in a forthcoming paper. Some useful pieces of matrix calculus are placed in the Appendix.

C. Notation

Vectors and matrices are denoted by boldface letters, lower and upper case, respectively. They may be equipped with top-scripts indicating dimensions, e.g. $\mathbf{A}^{n \times m}$ has n rows and m columns. We may write $\mathbf{A} = (a^{ef})_{\substack{e \in \mathcal{E} \\ f \in \mathcal{F}}}$ to emphasize the ranges of the row index e and the column index f . The transpose of \mathbf{A} is denoted by \mathbf{A}' . Vectors are taken to be of column type, hence row vectors appear as transposed (column) vectors. The identity matrix is denoted by \mathbf{I} , the vector with all entries equal to 1 is denoted by $\mathbf{1}$, and the vector with all entries equal to 0 is denoted by $\mathbf{0}$. By $\mathbf{D}_{e=1, \dots, n}(a^e)$, or just $\mathbf{D}(\mathbf{a})$, is meant the diagonal matrix with the entries of $\mathbf{a} = (a^1, \dots, a^n)'$ down the principal diagonal. The n -dimensional Euclidean space is denoted by \mathbb{R}^n , and the linear subspace spanned by the columns of $\mathbf{A}^{n \times m}$ is denoted by $\mathbb{R}(\mathbf{A})$.

The cardinality of a set \mathcal{Y} is denoted by $|\mathcal{Y}|$. For a finite set it is just its number of elements.

2. THE MARKOV CHAIN MARKET

A. The continuous time Markov chain

At the base of everything is some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $\{Y_t\}_{t \geq 0}$ be a continuous time Markov chain with finite state space $\mathcal{Y} = \{1, \dots, n\}$. We take the paths of Y to be right-continuous and Y_0 deterministic. Assume that Y is time homogeneous so that the transition probabilities

$$p_t^{ef} = \mathbb{P}[Y_{\tau+t} = f \mid Y_\tau = e]$$

depend only on the length of the time period. This implies that the transition intensities

$$\lambda^{ef} = \lim_{t \searrow 0} \frac{p_t^{ef}}{t}, \tag{2.1}$$

$e \neq f$, exist and are constant. To avoid repetitious reminders of the type “ $e, f \in \mathcal{Y}$ ”, we reserve the indices e and f for states in \mathcal{Y} throughout. We will frequently refer to

$$\mathcal{Y}^e = \{f; \lambda^{ef} > 0\},$$

the set of states that are directly accessible from state e , and denote the number of such states by

$$n^e = |\mathcal{Y}^e|.$$

Put

$$\lambda^{ee} = -\lambda^{e \cdot} = - \sum_{f; f \in \mathcal{Y}^e} \lambda^{ef}$$

(minus the total intensity of transition out of state e). We assume that all states intercommunicate so that $p_t^{ef} > 0$ for all e, f (and $t > 0$). This implies that $n^e > 0$ for all e (no absorbing states). The matrix of transition probabilities,

$$\mathbf{P}_t = (p_t^{ef}),$$

and the infinitesimal matrix,

$$\mathbf{\Lambda} = (\lambda^{ef}),$$

are related by (2.1), which in matrix form reads $\mathbf{\Lambda} = \lim_{t \searrow 0} \frac{1}{t} (\mathbf{P}_t - \mathbf{I})$, and by the forward and backward Kolmogorov differential equations,

$$\frac{d}{dt} \mathbf{P}_t = \mathbf{P}_t \mathbf{\Lambda} = \mathbf{\Lambda} \mathbf{P}_t. \tag{2.2}$$

Under the side condition $\mathbf{P}_0 = \mathbf{I}$, (2.2) integrates to

$$\mathbf{P}_t = \exp(\Lambda t). \tag{2.3}$$

The matrix exponential is defined in the Appendix, from where we also fetch the representation (A.3):

$$\mathbf{P}_t = \Phi \mathbf{D}_{e=1, \dots, n} \left(e^{\rho^e t} \right) \Phi^{-1} = \sum_{e=1}^n e^{\rho^e t} \boldsymbol{\varphi}^e \boldsymbol{\psi}^{e'}. \tag{2.4}$$

Here the first eigenvalue is $\rho^1 = 0$, and the corresponding eigenvectors are $\boldsymbol{\varphi}^1 = \mathbf{1}$ and $\boldsymbol{\psi}^{1'} = (p^1, \dots, p^n) = \lim_{t \rightarrow \infty} (p_t^{e1}, \dots, p_t^{en})$, the stationary distribution of Y . The remaining eigenvalues, ρ^2, \dots, ρ^n , have strictly negative real parts so that, by (2.4), the transition probabilities converge exponentially to the stationary distribution as t increases.

Introduce

$$I_t^e = 1[Y_t = e], \tag{2.5}$$

the indicator of the event that Y is in state at time t , and

$$N_t^{ef} = \left| \{ \tau; 0 < \tau \leq t, Y_{\tau-} = e, Y_{\tau} = f \} \right|, \tag{2.6}$$

the number of direct transitions of Y from state e to state $f \in \mathcal{Y}^e$ in the time interval $(0, t]$. For $f \notin \mathcal{Y}^e$ we define $N_t^{ef} \equiv 0$. The assumed right-continuity of Y is inherited by the indicator processes I^e and the counting processes N^{ef} . As is seen from (2.5), (2.6), and the obvious relationships

$$Y_t = \sum_e e I_t^e, \quad I_t^e = I_0^e + \sum_{f: f \neq e} (N_t^{fe} - N_t^{ef}),$$

the state process, the indicator processes, and the counting processes carry the same information, which at any time t is represented by the sigma-algebra $\mathcal{F}_t^Y = \sigma\{Y_\tau; 0 \leq \tau \leq t\}$. The corresponding filtration, denoted by $\mathbf{F}^Y = \{\mathcal{F}_t^Y\}_{t \geq 0}$, is taken to satisfy the usual conditions of right-continuity and completeness, and \mathcal{F}_0 is assumed to be trivial.

The compensated counting processes M^{ef} , $e \neq f$, defined by

$$dM_t^{ef} = dN_t^{ef} - I_t^e \lambda^{ef} dt \tag{2.7}$$

and $M_0^{ef} = 0$, are zero mean, square integrable, mutually orthogonal martingales with respect to $(\mathbf{F}^Y, \mathbb{P})$. We feel free to use standard definitions and results from counting process theory and refer to Andersen et al. (1993) for a background.

We now turn to the subject matter of our study and, referring to introductory texts like Björk (1998) and Pliska (1997), take basic concepts and results from arbitrage pricing theory as prerequisites.

B. The continuous time Markov chain market

We consider a financial market driven by the Markov chain described above. Thus, Y_t represents the state of the economy at time t , \mathcal{F}_t^Y represents the information available about the economic history by time t , and \mathbf{F}^Y represents the flow of such information over time.

In the market there are $m + 1$ basic assets, which can be traded freely and frictionlessly (short sales are allowed, and there are no transaction costs). A special role is played by asset No. 0, which is a “locally risk-free” *bank account* with state-dependent interest rate

$$r_t = r^{Y_t} = \sum_e I_t^e r^e,$$

where the state-wise interest rates r^e , $e = 1, \dots, n$, are constants. Thus, its price process is

$$S_t^0 = \exp\left(\int_0^t r_u du\right) = \exp\left(\sum_e r^e \int_0^t I_u^e du\right),$$

where $\int_0^t I_u^e du$ is the total time spent in economy state e during the period $[0, t]$. The dynamics of this price process is

$$dS_t^0 = S_t^0 r_t dt = S_t^0 \sum_e r^e I_t^e dt.$$

The remaining m assets, henceforth referred to as *stocks*, are risky, with price processes of the form

$$S_t^i = \exp\left(\sum_e \alpha^{ie} \int_0^t I_u^e du + \sum_e \sum_{f \in y^e} \beta^{ief} N_t^{ef}\right), \tag{2.8}$$

$i = 1, \dots, m$, where the α^{ie} and β^{ief} are constants and, for each i , at least one of the β^{ief} is non-null. Thus, in addition to yielding state-dependent returns of the same form as the bank account, stock No. i makes a price jump of relative size

$$\gamma^{ief} = \exp(\beta^{ief}) - 1$$

upon any transition of the economy from state e to state f . By the general Itô’s formula, its dynamics is given by

$$dS_t^i = S_{t-}^i \left(\sum_e \alpha^{ie} I_t^e dt + \sum_e \sum_{f \in y^e} \gamma^{ief} dN_t^{ef} \right). \tag{2.9}$$

(Setting $S_0^i = 1$ for all i is just a matter of convention; it is the relative price changes that matter.)

Taking the bank account as numeraire, we introduce the discounted asset prices $\tilde{S}_t^i = S_t^i/S_t^0$, $i = 0, \dots, m$. The discounted price of the bank account is $\tilde{S}_t^0 \equiv 1$, which is certainly a martingale under any measure. The discounted stock prices are

$$\tilde{S}_t^i = \exp\left(\sum_e (\alpha^{ie} - r^e) \int_0^t I_u^e du + \sum_e \sum_{f \in Y^e} \beta^{ief} N_t^{ef}\right), \quad (2.10)$$

with dynamics

$$d\tilde{S}_t^i = \tilde{S}_{t-}^i \left(\sum_e (\alpha^{ie} - r^e) I_t^e dt + \sum_e \sum_{f \in Y^e} \gamma^{ief} dN_t^{ef} \right), \quad (2.11)$$

$i = 1, \dots, m$.

C. Portfolios

A dynamic *portfolio* or *investment strategy* is an $m + 1$ -dimensional stochastic process

$$\theta_t = (\theta_t^0, \dots, \theta_t^m),$$

where θ_t^i represents the number of units of asset No i held at time t . The portfolio θ must be adapted to \mathbf{F}^Y and the shares of stocks, $(\theta_t^1, \dots, \theta_t^m)$, must also be \mathbf{F}^Y -predictable. For a sufficiently rigorous treatment of the concept of predictability, see Andersen et al. (1993). For our purposes it suffices to know that any left-continuous or deterministic process is predictable, the intuition being that the value of a predictable process at any time is determined by the strictly past history of the driving process Y . We will comment on these assumptions at a later point when the prerequisites are in place.

The *value* of the portfolio θ at time t is

$$V_t^\theta = \theta_t' \mathbf{S}_t = \sum_{i=0}^m \theta_t^i S_t^i.$$

Henceforth we will mainly work with discounted prices and values and, in accordance with (2.10), equip their symbols with a tilde. The discounted value of the portfolio at time t is

$$\tilde{V}_t^\theta = \theta_t' \tilde{\mathbf{S}}_t. \quad (2.12)$$

The strategy θ is *self-financing* (SF) if $dV_t^\theta = \theta_t' d\mathbf{S}_t$ or (recall $d\tilde{S}_t^0 = 0$)

$$d\tilde{V}_t^\theta = \theta_t' d\tilde{\mathbf{S}}_t = \sum_{i=1}^m \theta_t^i d\tilde{S}_t^i. \quad (2.13)$$

D. Absence of arbitrage

Let

$$\tilde{\Lambda} = (\tilde{\lambda}^{ef})$$

be an infinitesimal matrix that is equivalent to Λ in the sense that $\tilde{\lambda}^{ef} = 0$ if and only if $\lambda^{ef} = 0$. By Girsanov's theorem for counting processes (see e.g. Andersen et al. (1993)) there exists a measure $\tilde{\mathbb{P}}$, equivalent to \mathbb{P} , under which Y is a Markov chain with infinitesimal matrix $\tilde{\Lambda}$. Consequently, the processes M^{ef} , $e \in \mathcal{Y}$, $f \in \mathcal{Y}^e$, defined by

$$d\tilde{M}_t^{ef} = dN_t^{ef} - I_t^e \tilde{\lambda}^{ef} dt, \tag{2.14}$$

and $\tilde{M}_0^{ef} = 0$, are zero mean, mutually orthogonal martingales with respect to $(\mathbf{F}^Y, \tilde{\mathbb{P}})$. Rewrite (2.11) as

$$d\tilde{S}_t^i = \tilde{S}_{t-}^i \left[\sum_e \left(\alpha^{ie} - r^e + \sum_{f \in \mathcal{Y}^e} \gamma^{ief} \tilde{\lambda}^{ef} \right) I_t^e dt + \sum_e \sum_{f \in \mathcal{Y}^e} \gamma^{ief} d\tilde{M}_t^{ef} \right], \tag{2.15}$$

$i = 1, \dots, m$. The discounted stock prices are martingales with respect to $(\mathbf{F}^Y, \tilde{\mathbb{P}})$ if and only if the drift terms on the right vanish, that is,

$$\alpha^{ie} - r^e + \sum_{f \in \mathcal{Y}^e} \gamma^{ief} \tilde{\lambda}^{ef} = 0, \tag{2.16}$$

$e = 1, \dots, n$, $i = 1, \dots, m$. From general theory it is known that the existence of such an equivalent martingale measure $\tilde{\mathbb{P}}$ implies absence of arbitrage. The relation (2.16) can be cast in matrix form as

$$r^e \mathbf{1} - \alpha^e = \Gamma^e \tilde{\lambda}^e, \tag{2.17}$$

$e = 1, \dots, n$, where $\mathbf{1}$ is $m \times 1$ and

$$\alpha^e = (\alpha^{ie})_{i=1, \dots, m}, \quad \Gamma^e = (\gamma^{ief})_{i=1, \dots, m}^{f \in \mathcal{Y}^e}, \quad \tilde{\lambda}^e = (\tilde{\lambda}^{ef})_{f \in \mathcal{Y}^e}.$$

The existence of an equivalent martingale measure is equivalent to the existence of a solution $\tilde{\lambda}^e$ to (2.17) with all entries strictly positive. Thus, the market is arbitrage-free if (and we can show only if) for each e , $r^e \mathbf{1} - \alpha^e$ is in the interior of the convex cone of the columns of Γ^e .

Assume henceforth that the market is arbitrage-free so that (2.15) reduces to

$$d\tilde{S}_t^i = \tilde{S}_{t-}^i \sum_e \sum_{f \in \mathcal{Y}^e} \gamma^{ief} d\tilde{M}_t^{ef}. \tag{2.18}$$

Inserting (2.18) into (2.13), we find

$$d\tilde{V}_t^\theta = \sum_e \sum_{f \in \mathcal{Y}^e} \sum_{i=1}^m \theta_t^i \tilde{S}_{t-}^i \gamma^{ief} d\tilde{M}_t^{ef}, \quad (2.19)$$

which means that the value of an SF portfolio is a martingale with respect to $(\mathbf{F}^Y, \tilde{\mathbb{P}})$ and, in particular,

$$\tilde{V}_t^\theta = \tilde{\mathbb{E}}[\tilde{V}_T^\theta \mid \mathcal{F}_t] \quad (2.20)$$

for $0 \leq t \leq T$. Here $\tilde{\mathbb{E}}$ denotes expectation under $\tilde{\mathbb{P}}$. (The tilde, which in the first place was introduced to distinguish discounted values from the nominal ones, is also attached to the equivalent martingale measure because it arises from the discounted basic price processes.)

We remind of the standard proof of the result that the existence of an equivalent martingale measure implies absence of arbitrage: Under (2.20) one can not have $\tilde{V}_0^\theta = 0$ and at the same time have $\tilde{V}_T^\theta \geq 0$ almost surely and $\tilde{V}_T^\theta > 0$ with positive probability.

We can now explain the assumptions made about the components of the portfolio θ_t . The adaptedness requirement is commonplace and says just that the investment strategy must be based solely on the currently available information. Without this assumption it is easy to construct examples of arbitrages in the present and in any other model, and the theory would become void just as would practical finance if investors could look into the future. The requirement that $(\theta^1, \dots, \theta^m)$ be \mathbf{F}^Y -predictable means that investment in stocks must be based solely on information from the strict past. Also this assumption is omnipresent in arbitrage pricing theory, but its motivation is less obvious. For instance, in the Brownian world ‘predictable’ is the same as ‘adapted’ due to the (assumed) continuity of Brownian paths. In the present model the two concepts are perfectly distinct, and it is easy to explain why a trade in stocks cannot be based on news reported at the very instant where the trade is made. The intuition is that e.g. a crash in the stock market cannot be escaped by rushing money over from stocks to bonds. Sudden jumps in stock prices, which are allowed in the present model, must take the investor by surprise, else there would be arbitrage. This is seen from (2.19). If the θ_t^i , $i = 1, \dots, m$, could be any adapted processes, then we could choose them in such a manner that $d\tilde{V}_t^\theta \geq 0$ almost surely and strictly positive with positive probability. For instance, we could take them such that

$$\tilde{V}_t^\theta = \sum_e \sum_{f \in \mathcal{Y}^e} \int_0^t \tilde{M}_\tau^{ef} d\tilde{M}_\tau^{ef} = \frac{1}{2} \sum_e \sum_{f \in \mathcal{Y}^e} \left((\tilde{M}_t^{ef})^2 + \sum_{\tau \in (0, t]} (\Delta \tilde{M}_\tau^{ef})^2 \right).$$

Clearly, \tilde{V}_T^θ is non-negative and attains positive values with positive probability while $\tilde{V}_0^\theta = 0$, hence θ would be an arbitrage.

E. Attainability

A T -claim is a contractual payment due at time T . More precisely, it is an \mathcal{F}_T^Y -measurable random variable H with finite expected value. The claim is *attainable* if it can be perfectly duplicated by some SF portfolio θ , that is,

$$\tilde{V}_T^\theta = \tilde{H}. \tag{2.21}$$

If an attainable claim should be traded in the market, then its price must at any time be equal to the value of the duplicating portfolio in order to avoid arbitrage. Thus, denoting the price process by π_t and, recalling (2.20) and (2.21), we have

$$\tilde{\pi}_t = \tilde{V}_t^\theta = \tilde{\mathbb{E}}[\tilde{H} \mid \mathcal{F}_t], \tag{2.22}$$

or

$$\pi_t = \mathbb{E}\left[e^{-\int_t^T r} H \mid \mathcal{F}_t\right]. \tag{2.23}$$

(We use the short-hand $e^{-\int_t^T r}$ for $e^{-\int_t^T r_u du}$.)

By (2.22) and (2.19), the dynamics of the discounted price process of an attainable claim is

$$d\tilde{\pi}_t = \sum_e \sum_{f \in \mathcal{Y}^e} \sum_{i=1}^m \theta_t^i \tilde{S}_{t-}^i \gamma^{ief} d\tilde{M}_t^{ef}. \tag{2.24}$$

F. Completeness

Any T -claim H as defined above can be represented as

$$\tilde{H} = \tilde{\mathbb{E}}[\tilde{H}] + \int_0^T \sum_e \sum_{f \in \mathcal{Y}^e} \eta_t^{ef} d\tilde{M}_t^{ef}, \tag{2.25}$$

where the η_t^{ef} are \mathbf{F}^Y -predictable processes (see Andersen et al. (1993)). Conversely, any random variable of the form (2.25) is, of course, a T -claim. By virtue of (2.21) and (2.19), attainability of H means that

$$\begin{aligned} \tilde{H} &= \tilde{V}_0^\theta + \int_0^T d\tilde{V}_t^\theta \\ &= \tilde{V}_0^\theta + \int_0^T \sum_e \sum_{f \in \mathcal{Y}^e} \sum_i \theta_t^i \tilde{S}_{t-}^i \gamma^{ief} d\tilde{M}_t^{ef}. \end{aligned} \tag{2.26}$$

Comparing (2.25) and (2.26), we see that H is attainable iff there exist predictable processes $\theta_t^1, \dots, \theta_t^m$ such that

$$\sum_{i=1}^m \theta_t^i \tilde{S}_{t-}^i \gamma^{ief} = \eta_t^{ef},$$

for all e and $f \in \mathcal{Y}^e$. This means that the n^e -vector

$$\boldsymbol{\eta}_t^e = \left(\eta_t^{ef} \right)_{f \in \mathcal{Y}^e}$$

is in $\mathbb{R}(\boldsymbol{\Gamma}^{e'})$.

The market is *complete* if every T -claim is attainable, that is, if every n^e -vector is in $\mathbb{R}(\boldsymbol{\Gamma}^{e'})$. This is the case if and only if $\text{rank}(\boldsymbol{\Gamma}^e) = n^e$, which can be fulfilled for each e only if $m \geq \max_e n^e$, i.e. the number of risky assets is no less than the number of sources of randomness.

3. ARBITRAGE-PRICING OF DERIVATIVES IN A COMPLETE MARKET

A. Differential equations for the arbitrage-free price

Assume that the market is arbitrage-free and complete so that the price of any T -claim is uniquely given by (2.22) or (2.23).

Let us for the time being consider a T -claim that depends only on the state of the economy and the price of a given stock at time T . To simplify notation, we drop the top-script indicating this stock throughout and write just

$$S_t = \exp \left(\sum_e \alpha^e \int_0^t I_u^e du + \sum_e \sum_{f \in \mathcal{Y}^e} \beta^{ef} N_t^{ef} \right).$$

Thus, the claim is of the form

$$H = h^{Y_T}(S_T) = \sum_e I_T^e h^e(S_T). \quad (3.1)$$

Examples are a European call option defined by $H = (S_T - K)^+$, a caplet defined by $H = (r_T - g)^+ = (r^{Y_T} - g)^+$, and a zero coupon T -bond defined by $H = 1$.

For any claim of the form (3.1) the relevant state variables involved in the conditional expectation (2.23) are (S_t, t, Y_t) . This is due to the form of the stock price, by which

$$S_T = S_t \exp \left(\sum_e \alpha^e \int_t^T I_u^e du + \sum_e \sum_{f \in \mathcal{Y}^e} \beta^{ef} (N_T^{ef} - N_t^{ef}) \right), \quad (3.2)$$

and the Markov property, by which the past and the future are conditionally independent, given the present state Y_t . It follows that the price π_t is of the form

$$\pi_t = \sum_{e=1}^n I_t^e v^e(S_t, t), \quad (3.3)$$

where the functions

$$v^e(s, t) = \tilde{\mathbb{E}} \left[e^{-\int_t^T r} H \mid Y_t = e, S_t = s \right] \tag{3.4}$$

are the state-wise prices. Moreover, by (3.2) and the homogeneity of Y , we obtain the representation

$$v^e(s, t) = \mathbb{E} \left[h^{Y_{T-t}}(s, S_{T-t}) \mid Y_0 = e \right]. \tag{3.5}$$

The discounted price (2.22) is a martingale with respect to $(\mathbf{F}^Y, \tilde{\mathbb{P}})$. Assume that the functions $v^e(s, t)$ are continuously differentiable. Applying Itô to

$$\tilde{\pi}_t = e^{-\int_0^t r} \sum_{e=1}^n I_t^e v^e(S_t, t), \tag{3.6}$$

we find

$$\begin{aligned} d\tilde{\pi}_t &= e^{-\int_0^t r} \sum_e I_t^e \left(-r^e v^e(S_t, t) + \frac{\partial}{\partial t} v^e(S_t, t) + \frac{\partial}{\partial S} v^e(S_t, t) S_t \alpha^e \right) dt \\ &\quad + e^{-\int_0^t r} \sum_e \sum_{f \in \mathcal{Y}^e} \left(v^f(S_{t-}(1 + \gamma^{ef}), t) - v^e(S_{t-}, t) \right) dN_t^{ef} \\ &= e^{-\int_0^t r} \sum_e I_t^e \left(-r^e v^e(S_t, t) + \frac{\partial}{\partial t} v^e(S_t, t) + \frac{\partial}{\partial S} v^e(S_t, t) S_t \alpha^e \right) dt \\ &\quad + \sum_{f \in \mathcal{Y}^e} \left(v^f(S_{t-}(1 + \gamma^{ef}), t) - v^e(S_{t-}, t) \right) \tilde{\lambda}^{ef} dt \\ &\quad + e^{-\int_0^t r} \sum_e \sum_{f \in \mathcal{Y}^e} \left(v^f(S_{t-}(1 + \gamma^{ef}), t) - v^e(S_{t-}, t) \right) d\tilde{M}_t^{ef}. \end{aligned} \tag{3.7}$$

By the martingale property, the drift term must vanish, and we arrive at the non-stochastic partial differential equations

$$\begin{aligned} -r^e v^e(s, t) + \frac{\partial}{\partial t} v^e(s, t) + \frac{\partial}{\partial S} v^e(s, t) s \alpha^e \\ + \sum_{f \in \mathcal{Y}^e} \left(v^f(s(1 + \gamma^{ef}), t) - v^e(s, t) \right) \tilde{\lambda}^{ef} = 0 \end{aligned} \tag{3.8}$$

with side conditions

$$v^e(s, T) = h^e(s), \tag{3.9}$$

$e = 1, \dots, n$.

In matrix form, with

$$\mathbf{R} = \mathbf{D}_{e=1, \dots, n}(r^e), \quad \mathbf{A} = \mathbf{D}_{e=1, \dots, n}(\alpha^e)$$

and other symbols (hopefully) self-explaining, the differential equations and the side conditions are

$$-\mathbf{R}\mathbf{v}(s, t) + \frac{\partial}{\partial t} \mathbf{v}(s, t) + s\mathbf{A} \frac{\partial}{\partial s} \mathbf{v}(s, t) + \tilde{\mathbf{A}}\mathbf{v}(s(1 + \gamma), t) = \mathbf{0}, \quad (3.11)$$

$$\mathbf{v}(s, T) = \mathbf{h}(s).$$

There are other ways of obtaining the differential equations. One is to derive them from the integral equations obtained by conditioning on whether or not the process Y leaves its current state in the time interval $(t, T]$ and, in case it does, on the time and the direction of the transition. This approach is taken in Norberg (2002) and is a clue in the investigation of the assumed continuous differentiability of the functions v^e .

Before proceeding we render a comment on the fact that the price of a derivative depends on the drift parameters α^e of the stock prices as is seen from (3.8). This is all different from the Black-Scholes-Merton model in which a striking fact is that the drift parameter does not appear in the derivative prices. There is no contradiction here, however, as both facts reflect the paramount principle that the equivalent martingale measure arises from the path properties of the price processes of the basic assets and depends on the original measure only through its support. The drift term is a path property in the jump process world but not in the Brownian world. In the Markov chain market the pattern of direct transitions as given by the y^e is a path property, but apart from that the intensities \mathbf{F} do not affect the derivative prices.

B. Identifying the strategy

Once we have determined the solution $v^e(s, t)$, $e = 1, \dots, n$, the price process is known and given by (3.3).

The duplicating SF strategy can be obtained as follows. Setting the drift term to 0 in (3.7), we find the dynamics of the discounted price;

$$d\tilde{\pi}_t = e^{-\int_0^t r} \sum_e \sum_{f \in y^e} (v^f(S_{t-}(1 + \gamma^{ef}), t) - v^e(S_{t-}, t)) d\tilde{M}_t^{ef}. \quad (3.12)$$

Identifying the coefficients in (3.12) with those in (2.24), we obtain, for each time t and state e , the equations

$$\sum_{i=1}^m \theta_t^i S_{t-}^i \gamma^{ief} = v^f(S_{t-}(1 + \gamma^{ef}), t) - v^e(S_{t-}, t), \quad (3.13)$$

$f \in \mathcal{Y}^e$. The solution $(\theta_t^{i,e})_{i=1, \dots, m}$ certainly exists since $\text{rank}(\Gamma^e) \leq m$, and it is unique if $\text{rank}(\Gamma^e) = m$. It is a function of t and S_{t-} and is thus predictable.

Finally, θ^0 is determined upon combining (2.12), (2.22), and (3.6):

$$\theta_t^0 = e^{-\int_0^t r} \left(\sum_{e=1}^n I_t^e v^e(S_t, t) - \sum_{i=1}^m \theta_t^i S_t^i \right).$$

This function is not predictable.

C. The Asian option

As an example of a path-dependent claim, let us consider an Asian option, which is a T -claim of the form $H = \left(\frac{1}{T} \int_0^T S_\tau d\tau - K \right)^+$, where $K \geq 0$. The price process is

$$\begin{aligned} \pi_t &= \mathbb{E} \left[e^{-\int_t^T r} \left(\frac{1}{T} \int_0^T S_\tau d\tau - K \right)^+ \middle| \mathcal{F}_t^Y \right] \\ &= \sum_{e=1}^n I_t^e v^e \left(S_t, t, \int_0^t S_\tau d\tau \right), \end{aligned}$$

where

$$v^e(s, t, u) = \mathbb{E} \left[e^{-\int_t^T r} \left(\frac{1}{T} \int_t^T S_\tau d\tau + \frac{u}{T} - K \right)^+ \middle| Y_t = e, S_t = s \right].$$

The discounted price process is

$$\tilde{\pi}_t = e^{-\int_0^t r} \sum_{e=1}^n I_t^e v^e \left(t, S_t, \int_0^t S_\tau d\tau \right).$$

We are lead to partial differential equations in three variables.

D. Interest rate derivatives

A particularly simple, but important, class of claims are those of the form $H = h^{Y_T}$. Interest rate derivatives of the form $H = h(r_T)$ are included since $r_t = r^{Y_t}$. For such claims the only relevant state variables are t and Y_t , so that the function in (3.4) depends only on t and e . The differential equations (3.8) and the side condition (3.9) reduce to

$$\frac{d}{dt} v_t^e = r^e v_t^e - \sum_{f \in \mathcal{Y}^e} (v_t^f - v_t^e) \lambda^{ef}, \tag{3.14}$$

$$v_T^e = h^e. \tag{3.15}$$

In matrix form:

$$\begin{aligned}\frac{d}{dt} \mathbf{v}_t &= (\tilde{\mathbf{R}} - \tilde{\mathbf{\Lambda}}) \mathbf{v}_t, \\ \mathbf{v}_T &= \mathbf{h}.\end{aligned}$$

Similar to (2.3) we arrive at the explicit solution

$$\mathbf{v}_t = \exp\{(\tilde{\mathbf{\Lambda}} - \mathbf{R})(T-t)\} \mathbf{h}. \quad (3.16)$$

It depends on t and T only through $T-t$.

In particular, the zero coupon bond with maturity T corresponds to $\mathbf{h} = \mathbf{1}$. We will henceforth refer to it as the T -bond in short and denote its price process by $p(t, T)$ and its state-wise price functions by $\mathbf{p}(t, T) = (p^e(t, T))_{e=1, \dots, n}$;

$$\mathbf{p}(t, T) = \exp\{(\tilde{\mathbf{\Lambda}} - \mathbf{R})(T-t)\} \mathbf{1}. \quad (3.17)$$

For a call option on a U -bond, exercised at time $T (< U)$ with price K , \mathbf{h} has entries $h^e = (p^e(T, U) - K)^+$.

In (3.16)-(3.17) it may be useful to employ the representation (A.3),

$$\exp\{(\tilde{\mathbf{\Lambda}} - \mathbf{R})(T-t)\} = \tilde{\mathbf{\Phi}} \mathbf{D}_{e=1, \dots, n} \left(e^{\tilde{\rho}^e(T-t)} \right) \tilde{\mathbf{\Phi}}^{-1}. \quad (3.18)$$

4. RISK MINIMIZATION IN INCOMPLETE MARKETS

A. Incompleteness

The notion of incompleteness pertains to situations where there exist contingent claims that cannot be duplicated by an SF portfolio and, consequently, do not receive unique prices from the no arbitrage postulate alone. In Paragraph 2F we alluded to incompleteness arising from a scarcity of traded assets, that is, the discounted basic price processes are incapable of spanning the space of all martingales with respect to $(\mathbf{F}^Y, \tilde{\mathbb{P}})$ and, in particular, reproducing the value (2.25) of every financial derivative.

B. Risk minimization

Throughout this section we will mainly be working with discounted prices and values without any other mention than the tilde notation. The reason is that the theory of risk minimization rests on certain martingale representation results that apply to discounted prices under a martingale measure. We will be content to give just a sketchy review of some main concepts and results from the seminal paper of Föllmer and Sondermann (1986) on risk minimization.

Let \tilde{H} be a T -claim that is not attainable. This means that an *admissible* portfolio θ satisfying

$$\tilde{V}_T^\theta = \tilde{H}$$

cannot be SF. The *cost* by time t of an admissible portfolio θ is denoted by \tilde{C}_t^θ and is defined as that part of the portfolio value that has not been gained from trading:

$$\tilde{C}_t^\theta = \tilde{V}_t^\theta - \int_0^t \theta_\tau' d\tilde{S}_\tau.$$

The *risk* at time t is defined as the mean squared outstanding cost,

$$\tilde{R}_t^\theta = \mathbb{E} \left[(\tilde{C}_T^\theta - \tilde{C}_t^\theta)^2 \mid \mathcal{F}_t \right]. \tag{4.1}$$

By definition, the risk of an admissible portfolio θ is

$$\tilde{R}_t^\theta = \mathbb{E} \left[\left(\tilde{H} - \tilde{V}_t^\theta - \int_t^T \theta_\tau' d\tilde{S}_\tau \right)^2 \mid \mathcal{F}_t \right],$$

which is a measure of how well the current value of the portfolio plus future trading gains approximates the claim. The theory of risk minimization takes this entity as its objective function and proves the existence of an optimal admissible portfolio that minimizes the risk (4.1) for all $t \in [0, T]$.

The proof is constructive and provides a recipe for determining the optimal portfolio. One commences from the *intrinsic value* of \tilde{H} at time t defined as

$$\tilde{V}_t^H = \mathbb{E} [\tilde{H} \mid \mathcal{F}_t]. \tag{4.2}$$

This is the martingale that at any time gives the optimal forecast of the claim with respect to mean squared prediction error under the chosen martingale measure. By the Galchouk-Kunita-Watanabe representation, it decomposes uniquely as

$$\tilde{V}_t^H = \mathbb{E} [\tilde{H}] + \int_0^t \theta_t^{H'} d\tilde{S}_t + L_t^H, \tag{4.3}$$

where L^H is a martingale with respect to $(\mathbf{F}, \tilde{\mathbb{P}})$ which is orthogonal to the martingale \tilde{S} . The portfolio θ^H defined by this decomposition minimizes the risk process among all admissible strategies. The minimum risk is

$$\tilde{R}_t^H = \mathbb{E} \left[\int_t^T d \langle L^H \rangle_\tau \mid \mathcal{F}_t \right]. \tag{4.4}$$

C. Unit-linked insurance

As the name suggests, a life insurance product is said to be *unit-linked* if the benefit is a certain share of an asset (or portfolio of assets). If the contract stipulates a prefixed minimum value of the benefit, then one speaks of *unit-linked insurance with guarantee*. A risk minimization approach to pricing and hedging of unit-linked insurance claims was first taken by Møller (1998), who worked with the Black-Scholes-Merton financial market. We will here sketch how the analysis goes in our Markov chain market, which is a particularly suitable partner for the life history process since both are intensity-driven.

Let T_x be the remaining life time of an x years old who purchases an insurance at time 0, say. The conditional probability of survival to age $x + u$, given survival to age $x + t$ ($0 \leq t < u$), is

$$\mathbb{P}[T_x > u \mid T_x > t] = e^{-\int_t^u \mu_{x+s} ds}, \quad (4.5)$$

where μ_y is the mortality intensity at age y . Introduce the indicator of survival to age $x + t$, $I_t = 1[T_x > t]$, and the indicator of death before time t , $N_t = 1[T_x \leq t] = 1 - I_t$. The latter is a (very simple) counting process with intensity $I_t \mu_{x+t}$, and the associated (\mathbf{F}, \mathbb{P}) martingale M is given by

$$dM_t = dN_t - I_t \mu_{x+t} dt. \quad (4.6)$$

Assume that the life time T_x is independent of the economy Y . We will be working with the martingale measure \mathbb{P} obtained by replacing the intensity matrix Λ of Y with the martingalizing $\hat{\Lambda}$ and leaving the rest of the model unaltered.

Consider a unit-linked pure endowment benefit payable at a fixed time T , contingent on survival of the insured, with sum insured equal to the price S_T of the (generic) stock, but guaranteed no less than a fixed amount g . This benefit is a contingent T -claim,

$$H = (S_T \vee g) I_T.$$

The single premium payable as a lump sum at time 0 is to be determined. Let us assume that the financial market is complete so that every purely financial derivative has a unique price process. Then the intrinsic value of H at time t is

$$\tilde{V}_t^H = \tilde{\pi}_t I_t e^{-\int_t^T \mu},$$

where $\tilde{\pi}_t$ is the discounted price process of the derivative $S_T \vee g$, and we have used the somewhat sloppy abbreviation $\int_t^T \mu_{x+u} du = \int_t^T \mu$.

Using Itô together with (4.5) and (4.6) and the fact that M_t and $\tilde{\pi}_t$ almost surely have no common jumps, we find

$$\begin{aligned} d\tilde{V}_t^H &= d\tilde{\pi}_t I_{t-} e^{-\int_t^T \mu} + \tilde{\pi}_{t-} I_{t-} e^{-\int_t^T \mu} \mu_{x+t} dt + (0 - \tilde{\pi}_{t-} e^{-\int_t^T \mu}) dN_t \\ &= d\tilde{\pi}_t I_t e^{-\int_t^T \mu} - \tilde{\pi}_t e^{-\int_t^T \mu} dM_t. \end{aligned}$$

It is seen that the optimal trading strategy is that of the price process of the sum insured multiplied with the conditional probability that the sum will be paid out, and that

$$dL_t^H = -e^{-\int_t^T \mu} \tilde{\pi}_t dM_t.$$

Using $d\langle M \rangle_t = I_t \mu_{x+t} dt$ (see Andersen et al. (1993)), the minimum risk (4.4) now assumes the form

$$\tilde{R}_t^H = \mathbb{E} \left[\int_t^T e^{-2\int_t^\tau \mu} \tilde{\pi}_\tau^2 I_\tau \mu_{x+\tau} d\tau \mid \mathcal{F}_t \right] = I_t e^{-2\int_0^t r} \sum_e I_t^e R^e(S_t, t), \tag{4.7}$$

where

$$R^e(s, t) = \mathbb{E} \left[\int_t^T e^{-2\int_t^\tau \mu} e^{-2\int_t^\tau r} \tilde{\pi}_\tau^2 I_\tau \mu_{x+\tau} d\tau \mid S_t = s, Y_t = e, I_t = 1 \right].$$

Working along the lines of the proof of (3.8), this time starting from the martingale

$$\begin{aligned} M_t^R &= \mathbb{E} \left[\int_0^T e^{-2\int_t^\tau \mu} \tilde{\pi}_\tau^2 I_\tau \mu_{x+\tau} d\tau \mid \mathcal{F}_t \right] \\ &= \int_0^t e^{-2\int_t^\tau \mu} e^{-2\int_0^\tau r} \tilde{\pi}_\tau^2 I_\tau \mu_{x+\tau} d\tau + I_t e^{-2\int_0^t r} \sum_e I_t^e R^e(S_t, t), \end{aligned}$$

we obtain the differential equations

$$\begin{aligned} &(\pi_t^2 - R^e(s, t)) \mu_{x+t} - 2r^e R^e(s, t) + \frac{\partial}{\partial t} R^e(s, t) + \frac{\partial}{\partial s} R^e(s, t) s \alpha^e \\ &+ \sum_{f \in \mathcal{Y}^e} (R^f(s(1 + \gamma^{ef}), t) - R^e(s, t)) \tilde{\lambda}^{ef}. \end{aligned} \tag{4.8}$$

These are to be solved in parallel with the differential equations (3.8) and are subject to the conditions

$$R^e(s, t) = 0. \tag{4.9}$$

D. Trading with bonds: How much can be hedged?

It is well known that in a model with only one source of randomness, like the Black-Scholes-Merton model, the price process of one zero coupon bond will

determine the value process of any other zero coupon bond that matures at an earlier date. In the present model this is not the case, and the degree of incompleteness of a given bond market is therefore an issue.

Suppose an agent faces a contingent T -claim and is allowed to invest only in the bank account and a finite number m of zero coupon bonds with maturities $T_i, i = 1, \dots, m$, all post time T . The scenario could be that regulatory constraints are imposed on the investment strategy of an insurance company. The question is, to what extent can the claim be hedged by self-financed trading in these available assets?

An allowed SF portfolio θ has a discounted value process \tilde{V}_t^θ of the form

$$d\tilde{V}_t^\theta = \sum_{i=1}^m \theta_t^i \sum_e \sum_{f \in y^e} (\tilde{p}^f(t, T_i) - \tilde{p}^e(t, T_i)) d\tilde{M}_t^{ef} = \sum_e d\tilde{M}_t^{e'} \mathbf{Q}_t^e \theta_t,$$

where θ is predictable, $\tilde{M}_t^e = (\tilde{M}_t^{ef})_{f \in y^e}$ is the n^e -dimensional vector comprising the non-null entries in the e -th row of $\tilde{M}_t = (\tilde{M}_t^{ef})$, and

$$\mathbf{Q}_t^e = \mathbf{Y}^e \mathbf{Q}_t,$$

where

$$\mathbf{Q}_t = (\tilde{p}^e(t, T_i))_{e=1, \dots, n}^{i=1, \dots, m} = (\tilde{p}(t, T_i), \dots, \tilde{p}(t, T_m)), \tag{4.10}$$

and \mathbf{Y}^e is the $n^e \times n$ matrix which maps \mathbf{Q}_t to $(\tilde{p}^f(t, T_i) - \tilde{p}^e(t, T_i))_{f \in y^e}^{i=1, \dots, m}$. If e.g. $y^n = \{1, \dots, p\}$, then $\mathbf{Y}^n = (\mathbf{I}^{p \times p}, \mathbf{0}^{p \times (n-p-1)}, -\mathbf{1}^{p \times 1})$.

The sub-market consisting of the bank account and the m zero coupon bonds is complete in respect of T -claims iff the discounted bond prices span the space of all martingales with respect to $(\mathbf{F}^Y, \mathbb{P})$ over the time interval $[0, T]$. This is the case iff, for each e , $\text{rank}(\mathbf{Q}_t^e) = n^e$. Now, since \mathbf{Y}^e obviously has full rank n^e , the rank of \mathbf{Q}_t^e is determined by that of \mathbf{Q}_t in (4.10). We will argue that, typically, \mathbf{Q}_t has full rank. Thus, suppose $\mathbf{c} = (c_1, \dots, c_m)'$ is such that

$$\mathbf{Q}_t \mathbf{c} = \mathbf{0}^{n \times 1}.$$

Recalling (3.17), this is the same as

$$\sum_{i=1}^m c_i \exp\{(\tilde{\Lambda} - \mathbf{R})T_i\} \mathbf{1} = \mathbf{0},$$

or, by (3.18) and since $\tilde{\Phi}$ has full rank,

$$\mathbf{D}_{e=1, \dots, n} \left(\sum_{i=1}^m c_i e^{\tilde{p}^e T_i} \right) \tilde{\Phi}^{-1} \mathbf{1} = \mathbf{0}. \tag{4.11}$$

Since $\tilde{\Phi}^{-1}$ has full rank, the entries of the vector $\tilde{\Phi}^{-1}\mathbf{1}$ cannot be all null. Typically all entries are non-null, and we assume this is the case. Then (4.11) is equivalent to

$$\sum_{i=1}^m c_i e^{\tilde{\rho}^e T_i} = 0, \quad e = 1, \dots, n. \tag{4.12}$$

Using the fact that the generalized Vandermonde matrix has full rank (see Gantmacher (1959)), we know that (4.12) has a non-null solution \mathbf{c} if and only if the number of distinct eigenvalues $\tilde{\rho}^e$ is less than m . The role of the Vandermonde matrix in finance is the topic of a parallel paper by the author, Norberg (1999).

In the case where $\text{rank}(\mathbf{Q}_t^e) < n^e$ for some e we would like to determine the Galchouk-Kunita-Watanabe decomposition for a given \mathcal{F}_T^Y -claim. The intrinsic value process (4.2) has dynamics of the form

$$d\tilde{V}_t^H = \sum_e \sum_{f \in y^e} \eta_t^{ef} d\tilde{M}_t^{ef} = \sum_e d\tilde{\mathbf{M}}_t^{e'} \boldsymbol{\eta}_t^e, \tag{4.13}$$

where $\boldsymbol{\eta}_t^e = (\eta_t^{ef})_{f \in y^e}$ is predictable. We seek its appropriate decomposition (4.3) into

$$\begin{aligned} d\tilde{V}_t^H &= \sum_i \theta_t^i d\tilde{p}(t, T_i) + \sum_e \sum_{f \in y^e} \zeta_t^{ef} d\tilde{M}_t^{ef} \\ &= \sum_e \sum_{f \in y^e} \sum_i \theta_t^i (\tilde{p}^f(t, T_i) - \tilde{p}^e(t, T_i)) d\tilde{M}_t^{ef} + \sum_e \sum_{f \in y^e} \zeta_t^{ef} d\tilde{M}_t^{ef} \\ &= \sum_e d\tilde{\mathbf{M}}_t^{e'} \mathbf{Q}_t^e \boldsymbol{\theta}_t^e + \sum_e d\tilde{\mathbf{M}}_t^{e'} \boldsymbol{\zeta}_t^e, \end{aligned}$$

such that the two martingales on the right hand side are orthogonal, that is,

$$\sum_e I_t^e - \sum_{f \in y^e} (\mathbf{Q}_t^e \boldsymbol{\theta}_t^e)' \tilde{\Lambda}^e \boldsymbol{\zeta}_t^e = 0,$$

where $\tilde{\Lambda}^e = \mathbf{D}(\tilde{\lambda}^e)$. This means that, for each e , the vector $\boldsymbol{\eta}_t^e$ in (4.13) is to be decomposed into its $\langle \cdot, \cdot \rangle_{\tilde{\Lambda}^e}$ projections onto $\mathbb{R}(\mathbf{Q}_t^e)$ and its orthocomplement. From (A.4) and (A.5) we obtain

$$\mathbf{Q}_t^e \boldsymbol{\theta}_t^e = \mathbf{P}_t^e \boldsymbol{\eta}_t^e,$$

where

$$\mathbf{P}_t^e = \mathbf{Q}_t^e \left(\mathbf{Q}_t^{e'} \tilde{\Lambda}^e \mathbf{Q}_t^e \right)^{-1} \mathbf{Q}_t^{e'} \tilde{\Lambda}^e,$$

hence

$$\boldsymbol{\theta}_t^e = \left(\mathbf{Q}_t^{e'} \tilde{\Lambda}^e \mathbf{Q}_t^e \right)^{-1} \mathbf{Q}_t^{e'} \tilde{\Lambda}^e \boldsymbol{\eta}_t^e. \tag{4.14}$$

Furthermore,

$$\zeta_t^e = (\mathbf{I} - \mathbf{P}_t^e) \boldsymbol{\eta}_t^e, \quad (4.15)$$

and the minimum risk (4.4) is

$$\tilde{R}_t^H = \mathbb{E} \left[\int_t^T \sum_e \sum_{f \in \mathcal{Y}^e} I_\tau^e \lambda^{ef} (\zeta_\tau^{ef})^2 d\tau \middle| \mathcal{F}_t \right]. \quad (4.16)$$

The computation goes as follows: The coefficients η^{ef} involved in the intrinsic value process (4.13) and the state-wise prices $p^e(t, T_i)$ of the T_i -bonds are obtained by simultaneously solving (3.8) and (3.14), starting from (3.11) and (3.15), respectively, and at each step computing the optimal trading strategy $\boldsymbol{\theta}$ by (4.14) and the ζ from (4.15). The risk may be computed in parallel by solving differential equations for suitably defined state-wise risk functions. The relevant state variables depend on the nature of the T -claim as illustrated in the previous paragraph.

5. DISCUSSION OF THE MODEL

A. Versatility of the Markov chain

By suitable specification of \mathcal{Y} , Λ , and the asset parameters r^e , α^{ie} , and β^{ief} , we can make the Markov chain market reflect virtually any conceivable feature a real world market may have. We can construct a non-negative mean reverting interest rate. We can design stock markets with recessions and booms, bullish and bearish trends, and crashes and frenzies and other extreme events (not in the mathematical sense of the word, though, since the intensities and the jump sizes are deterministic). We can create forgetful markets and markets with long memory, markets with all sorts of dependencies between assets — hierarchical and others. In the huge class of Markov chains we can also find an approximation to virtually any other theoretical model since the Markov chain models are dense in the model space, roughly speaking. In particular, one can construct a sequence of Markov chain models such that the compensated multivariate counting process converges weakly to a given multivariate Brownian motion. An obvious route from Markov chains to Brownian motion goes via Poisson processes, which we will now elaborate a bit upon.

B. Poisson markets

A Poisson process is totally memoryless whereas a Markov chain recalls which state it is in at any time. Therefore, a Poisson process can be constructed by suitable specification of the Markov chain Y . There are many ways of doing it, but a minimalistic one is to let Y have two states $\mathcal{Y} = \{1, 2\}$ and intensities $\lambda^{12} = \lambda^{21} = \lambda$. Then the process N defined by $N_t = N_t^{12} + N_t^{21}$ (the total number

of transitions in $(0, t]$ is Poisson with intensity λ since the transitions counted by N occur with constant intensity λ .

Merton (1976) introduced a simple Poisson market with

$$S_t^0 = e^{rt},$$

$$S_t^1 = e^{\alpha t + \beta N_t},$$

where r , α , and β are constants, and N is a Poisson process with constant intensity λ . This model is accommodated in the Markov chain market by letting Y be a two-state Markov chain as prescribed above and taking $r^1 = r^2 = r$, $\alpha^1 = \alpha^2 = \alpha$, and $\beta^{12} = \beta^{21} = \beta$. The no arbitrage condition (2.17) reduces to $\tilde{\lambda} > 0$, where $\tilde{\lambda} = (r - \alpha)/\gamma$ and $\gamma = e^\beta - 1$. When this condition is fulfilled, $\tilde{\lambda}$ is the intensity of N under the equivalent martingale measure.

The price function (3.5) now reduces to an expected value in the Poisson distribution with parameter $\tilde{\lambda}(T - t)$:

$$v(s, t) = \mathbb{E} \left[e^{-r(T-t)} h \left(s e^{\alpha(T-t) + \beta N_{T-t}} \right) \right]$$

$$= e^{-(r + \tilde{\lambda})(T-t)} \sum_{n=0}^{\infty} \frac{(\tilde{\lambda}(T-t))^n}{n!} h \left(s e^{\alpha(T-t) + \beta n} \right). \tag{5.1}$$

A more general Poisson market would have stock prices of the form

$$S_t^i = \exp \left[\alpha^i t + \sum_{j=1}^n \beta^{ij} N_t^j \right],$$

$i = 1, \dots, m$, where the N^j are independent Poisson processes. The Poisson processes can be constructed by the recipe above from independent Markov chains $Y^j, j = 1, \dots, n$, which constitute a Markov chain, $Y = (Y^1, \dots, Y^n)$.

C. On differentiability and numerical methods

The assumption that the functions $v^e(s, t)$ are continuously differentiable is not an innocent one and, in fact, it typically does not hold true. An example is provided by the Poisson market in the previous paragraph. From the explicit formula (5.1) it is seen that the price function inherits the smoothness properties of the function h , which typically is not differentiable everywhere and may even have discontinuities. For instance, for $h(s) = (s - K)^+$ (European call) the function v is continuous in both arguments, but continuous differentiability fails to hold on the curves $\{(s, t); s e^{\alpha(T-t) + n\beta} = K\}, n = 0, 1, 2, \dots$. This warning prompts a careful exploration and mapping of the Markov chain terrain. That task is a rather formidable one and is not undertaken here. Referring to Norberg (2002), let it suffice to report the following: From a recursive system

of backward integral equations it is possible to locate the positions of all points (s, t) where the functions v^e are non-smooth. Equipped with this knowledge one can arrange a numerical procedure with controlled global error, which amounts to solving the differential equations where they are valid and gluing the piece-wise solutions together at the exceptional points where they are not. For interest rate derivatives, which involve only ordinary first order differential equations, these problems are less severe and standard methods for numerical computation will do.

ACKNOWLEDGMENTS

This work was partly supported by the Mathematical Finance Network under the Danish Social Science Research Council, Grant No. 9800335. A first draft, worked out during visits at the Centre for Actuarial Studies, University of Melbourne, and at the Department of Mathematics, Stanford University, was presented as invited lecture to the 1999 QMF Conference.

REFERENCES

- ANDERSEN, P.K., BORGAN, Ø., GILL, R.D. and KEIDING, N. (1993) *Statistical Models Based on Counting Processes*. Springer-Verlag.
- BJÖRK, T., KABANOV, Y. and RUNGALDIER, W. (1997) Bond market structures in the presence of marked point processes. *Mathematical Finance* **7**, 211-239.
- BJÖRK, T. (1998) *Arbitrage Theory in Continuous Time*, Oxford University Press.
- COX, J., ROSS, S. and RUBINSTEIN, M. (1979) Option pricing: A simplified approach. *J. of Financial Economics* **7**, 229-263.
- DELBAEN, F. and SCHACHERMAYER, W. (1994) A general version of the fundamental theorem on asset pricing. *Mathematische Annalen* **300**, 463-520.
- EBERLEIN, E. and RAIBLE, S. (1999) Term structure models driven by general Lévy processes. *Mathematical Finance* **9**, 31-53.
- ELLIOTT, R.J. and KOPP, P.E. (1998) *Mathematics of financial markets*, Springer-Verlag.
- FÖLLMER, H. and SONDERMANN, D. (1986) Hedging of non-redundant claims. In *Contributions to Mathematical Economics in Honor of Gerard Debreu*, 205-223, eds. Hildebrand, W., Mas-Colell, A., North-Holland.
- GANTMACHER, F.R. (1959) *Matrizenrechnung II*, VEB Deutscher Verlag der Wissenschaften, Berlin.
- HARRISON, J.M. and KREPS, D.M. (1979) Martingales and arbitrage in multi-period securities markets. *J. Economic Theory* **20**, 381-408.
- HARRISON, J.M. and PLISKA, S. (1981) Martingales and stochastic integrals in the theory of continuous trading. *J. Stoch. Proc. and Appl.* **11**, 215-260.
- KARLIN, S. and TAYLOR, H. (1975) *A first Course in Stochastic Processes*, 2nd. ed., Academic Press.
- MERTON, R.C. (1976) Option pricing when underlying stock returns are discontinuous. *J. Financial Economics* **3**, 125-144.
- MØLLER, T. (1998) Risk minimizing hedging strategies for unit-linked life insurance. *ASTIN Bulletin* **28**, 17-47.
- NORBERG, R. (1995) A time-continuous Markov chain interest model with applications to insurance. *J. Appl. Stoch. Models and Data Anal.*, 245-256.
- NORBERG, R. (1999) On the Vandermonde matrix and its role in mathematical finance. *Working paper No. 162*, Laboratory of Actuarial Math., Univ. Copenhagen.
- NORBERG, R. (2002) Anomalous PDEs in Markov chains: domains of validity and numerical solutions. Research Report, Department of Statistics, London School of Economics: <http://stats.lse.ac.uk/norberg>. (Submitted to *Finance and Stochastics*).
- PLISKA, S.R. (1997) *Introduction to Mathematical Finance*, Blackwell Publishers.

A. Appendix: Some useful matrix results

A diagonalizable square matrix $\mathbf{A}^{n \times n}$ can be represented as

$$\mathbf{A} = \mathbf{\Phi} \mathbf{D}_{e=1, \dots, n}(\rho^e) \mathbf{\Phi}^{-1} = \sum_{e=1}^n \rho^e \phi^e \psi^{e'}, \quad (\text{A.2})$$

where the ϕ^e are the columns of $\mathbf{\Phi}^{n \times n}$ and the $\psi^{e'}$ are the rows of $\mathbf{\Phi}^{-1}$. The ρ^e are the eigenvalues of \mathbf{A} , and ϕ^e and $\psi^{e'}$ are the corresponding eigenvectors, right and left, respectively. Eigenvectors (right or left) corresponding to eigenvalues that are distinguishable and non-null are mutually orthogonal. These results can be looked up in e.g. Karlin and Taylor (1975).

The exponential function of $\mathbf{A}^{n \times n}$ is the $n \times n$ matrix defined by

$$\exp(\mathbf{A}) = \sum_{p=0}^{\infty} \frac{1}{p!} \mathbf{A}^p = \mathbf{\Phi} \mathbf{D}_{e=1, \dots, n}(e^{\rho^e}) \mathbf{\Phi}^{-1} = \sum_{e=1}^n e^{\rho^e} \phi^e \psi^{e'}, \quad (\text{A.3})$$

where the last two expressions follow from (A.2). This matrix has full rank.

If $\mathbf{A}^{n \times n}$ is positive definite symmetric, then $\langle \eta_1, \eta_2 \rangle_{\Lambda} = \eta_1' \Lambda \eta_2$ defines an inner product on \mathbb{R}^n . The corresponding norm is given by $\|\eta\|_{\Lambda} = \langle \eta, \eta \rangle_{\Lambda}^{1/2}$. If $\mathbf{Q}^{n \times m}$ has full rank m ($\leq n$), then the $\langle \cdot, \cdot \rangle_{\Lambda}$ -projection of η onto $\mathbb{R}(\mathbf{Q})$ is

$$\eta_{\mathbf{Q}} = \mathbf{P}_{\mathbf{Q}} \eta, \quad (\text{A.4})$$

where the projection matrix (or projector) $\mathbf{P}_{\mathbf{Q}}$ is

$$\mathbf{P}_{\mathbf{Q}} = \mathbf{Q}(\mathbf{Q}' \Lambda \mathbf{Q})^{-1} \mathbf{Q}' \Lambda. \quad (\text{A.5})$$

The projection of η onto the orthogonal complement $\mathbb{R}(\mathbf{Q})^{\perp}$ is

$$\eta_{\mathbf{Q}^{\perp}} = \eta - \eta_{\mathbf{Q}} = (\mathbf{I} - \mathbf{P}_{\mathbf{Q}}) \eta.$$

RAGNAR NORBERG

London School of Economics and Political Science

Department of Statistics

Houghton Street, London WC2A 2AE

United Kingdom

E-mail: R.Norberg@lse.ac.uk