Applications of Markov Chains and Martingale Theory to Bond Pricing


1 Introduction

As evident by their frequent appearance in leading economics and finance journals, Markov chains and martingales are widely used within the field of finance and in particular, the sub-field of asset pricing. As an example, martingale theory is applied heavily in the pricing of stocks, bonds and derivatives, and stopping times underlie optimal trading strategies over these securities. In the paper to follow, we discuss the applications of Markov chains and martingale theory specific to the task of bond pricing, with a strong focus on the pricing of bonds where the risk of default (non-payment) is present. These securities are popularly known as ‘corporate debt’. We provide an overview of the seminal model of Jarrow and Turnbull [1995] (herein JT[1995],) to set the framework to discuss the extension proposed by Jarrow, Lando and Turnbull [1997] (herein JLT[1997]).

The paper is organized as follows: section 2 briefly outlines the concept of tree-structures and equivalent martingale measures to provide some background; section 3 describes the JT[1995] and JLT[1997] models, and discusses the results of JLT[1997]; section 4 provides a critique.

2 Economic Framework

2.1 Asset Pricing

The intuition behind pricing riskless securities (of which government bonds is a particular class,) is that if held to maturity, one would gain with certainty the sum of all cash flows stipulated in the contract of the security\(^1\). Thus, one would expect that its price

---

\(^1\) Throughout this paper, we will seem to use the terms bond, security and asset interchangeably. However, we note the following relationship: (1) bonds are a class of security, (2) securities are a class of asset.
should be some preference-adjusted mapping from its worth, (the sum of all future cash flows,) discounted by a measure of time preference. This thinking is inline with basic economic intuition, which dictates that a dollar today is worth more than a dollar tomorrow. We often (and exclusively for the purposes of this paper) choose the “risk-free” interest rate, \( r(t) \), as our measure of time preference, and thus always discount the value today of a dollar tomorrow by \( (1 + r(t))^{-1} \). A riskless bond such as a T-bill is priced via the equation:

\[
B(t, T) = \sum_{k=1}^{T} \frac{x_k}{(1 + r(k))^{k-t}}
\]

To price risky securities, we cannot forecast their future cash flows with certainty, as in different future states of the world the risky securities may earn very different cash flows. Thus, as an adjustment for this uncertainty, we take the expectation of the uncertain cash flows over all possible states of the world, and discount them by the risk-free rate. With this adjustment, the price of the asset, denoted \( v(t, T) \), is as given below:

\[
v(t, T) = \sum_{i=1}^{\theta} \sum_{k=1}^{T} p(i, k) \frac{x_{ki}}{(1 + r(k))^{k-t}}
\]

where \( p(i) \) denotes the probability of being in state \( i \) at time \( k \), and \( x_{ki} \) the cash flow at in state \( i \) at time \( k \). But, this price formula is not yet complete. If we price a risky security solely by the expectation of future cash flows, investors are not compensated for the variance associated with the future cash flows. To compensate for the risk taken by

\[\text{To see this, notice that we can always invest our dollar at time } t \text{ and earn a rate of return, } r, \text{ so that in the next period, } t+1, \text{ we would have } 1+r \text{ dollars.}\]

\[\text{We often define the risk-free interest rate as the interest rate offered by the government on it’s T-Bills.}\]
investment in these securities, we subtract a risk premium, $\mu_k$, from the price of the security.\(^4\) A risky bond would then have the value:

$$v(t,T) = \sum_{i=1}^{\theta} \sum_{k=1}^{T} p(i,k) \frac{x_{ki}}{(1 + r(k))^{k-t}} - \mu_k$$

From this, we can surmise that the price, $v(t,T)$, is a supermartingale. We deduce this through the notion of risk premium. Consider the value of a risky security at time $t$ to be $S_t$. As a risky security must earn more in one time period than a riskless security, the following should hold:

$$E[S_{t+1}] > (1 + r_t)S_t$$

To achieve equality, we must have that, for some positive $\mu$,

$$\frac{E[S_{t+1}]}{(1 + r_t)} = S_t(1 + \mu)$$

That is, the discounted expected future value of the security should be higher by some factor, which we call the risk premium. Thus, the value of the security is a submartingale. From a pricing standpoint, we could choose to, instead of adding a risk premium to the return, subtract a risk premium from the price of the asset (invariably doing the same thing), and thus the same logic would lead us to determine that the price of a risky asset is a supermartingale.

### 2.2 Risk-Neutral Pricing: Equivalent Martingale Probabilities

In asset pricing, the process is simplified greatly when we can convert the price of a risky asset to that of a martingale. To do so, there are two techniques often used. If the

---

\(^4\) Note here that the sum of the risk premia is smaller for securities closer to maturity.
drift is additive (in a sense), as we constructed above, we can use the technique known as “Doob-Meyer Decomposition”, which essentially removes the drift “by brute force”. The method is applicable when the returns (not the price) of the security satisfy the following theorem:

**Theorem:** If \( 0 \leq t \leq \infty \) is a right-continuous submartingale with respect to an information set, \( \{ I_t \} \), and if \( E[X_t] < \infty, \forall t \), then \( X_t \) admits the decomposition:

\[
X_t = M_t + A_t
\]

where \( M_t \) is a right-continuous martingale and \( A_t \) is a positive drift process measurable with respect to \( I_t \).\(^5\)

However, while this method may be applicable on numerous occasions,\(^6\) it may not be easily implemented. Often, it proves simpler to change the probability measure on the state space over which the asset price is defined. We define the state-probability space by the triple, \( (X, \Omega, Q) \). We now use the notion of an equivalent martingale measure to change \( Q \) such that under expectation, the price of a risky security is a martingale. An effective method for finding this equivalent martingale measure is known as “Girsanov's Theorem”.\(^7\) However, for the purposes of this paper, determining the equivalent martingale measure is not as important as the existence and uniqueness of one. To ensure this fact, JT[1995], and JLT[1997] invoke the following theorem without proof.\(^8\)

\(^6\) Many securities in the class known as ‘Exotic Options’ fail the right-continuous criterion.
\(^7\) See Neftci (2000) for a definition of Girsanov’s theorem and its application to asset pricing.
\(^8\) For a proof and an exposition of this assumption, see Harrison and Pliska, (1981).
Theorem: If the state space under the triple \((X, \Omega, Q)\) is finite, complete and arbitrage-free, then there exists a unique equivalent martingale measure, \(\tilde{Q}\) on \(X\).

By assuming that \(X\) is finite, complete and arbitrage-free, JT[1995] and JLT[1997] proceed with the arguments and results to follow. As a result of this theorem, the price of a risky bond becomes:

\[
v(t, T) = E^\tilde{Q}_t \left[ \sum_{k=0}^{T} \frac{x_k}{(1 + r(k))^{T-t}} \right] = E^Q_t \left[ \sum_{k=0}^{T} \frac{x_k}{(1 + r(k))^{T-t}} \right]
\]

(1.4)

with the negative drift term eliminated. As we can see, the price of the risky bond is tied to the expectation of the evolutionary process of the interest rate, which in many frameworks we assume is stochastic. In JT[1995] and JLT[1997] the authors describe the interest rate process as a simple binomial process.\(^9\)


3.1 Bond pricing under the existence of default

From (1.4) JT[1995] extends the model to account for the probability of bankruptcy by debt-issuing firms, and thus default on the outstanding debt security.\(^10\) JT[1995] define a stopping time, \(\tau^* = \min\{t > 0 | \text{bankruptcy occurs}\}\), and define the instance of default as a pair of indicator functions about the maturity time, \(T\).

---

\(^9\) The binomial process is described in the following way; at every future time-stage, the interest rate enters either an “up” state in which the interest rate rises, or a “down” state in which the interest rate declines. In these models, the up and down movements of the interest rate are not assumed to be constant for all time periods. We note here that, as is noted in both papers, that the interest rate processes can be chosen to follow any multinomial tree-structure.

\(^10\) As the extension to the continuous-time case in both JT[1995] and JLT[1997] are relatively straightforward applications of the Black-Scholes continuous-time asset pricing framework, we omit those extensions and consider only the discrete-time formulations of the models in both papers.
\[ v(t, T) = E^Q_t \left[ \sum_{k=t}^{T} \frac{x_k}{(1 + r(k))^{k-T}} \left( \delta 1_{[t^{*}(\tau^{*}) \leq T]} + 1_{[\tau^{*} > T]} \right) \right] \] (1.5)

Further, by assuming independence of the interest rate and default processes, the pricing formula simplifies to:

\[ v(t, T) = E^Q_t \left[ \sum_{k=t}^{T} \frac{x_k}{(1 + r(k))^{k-T}} \right] E^Q_t \left[ \delta 1_{[\tau^{*} > T]} + 1_{[\tau^{*} > T]} \right] \] (1.6)

As a result of this assumption, we see that we have a separation in pricing; that is, we can price a bond with default risk by the product of the default-free bond with the “price” of default. This makes our environment highly tractable, as pricing default separately from the security itself allows us to easily model the stochastic default process in a number of ways. There are three types of default processes that JLT[1997] propose, that is, an exogenous process where the default probability is constant, a default probability that evolves as a Markov process, and an endogenous default probability. The focus of JLT[1997] is strictly the Markov evolution of the default process. However, in discussing the Markov process one could easily deduce the constant special case. By taking the second expectation, (1.6) becomes,

\[ v(t, T) = E^Q_t \left[ \sum_{k=t}^{T} \frac{x_k}{(1 + r(k))^{k-T}} \right] \left( \delta + (1 - \delta) \tilde{Q}_t(\tau^{*} > T) \right) \] (1.7)

where \( \tilde{Q}_t(\tau^{*} > T) \) is the probability that the security is not in default at time \( t \).

### 3.2 Results

We can define the Markov process of default as a reducible 2-class Markov chain, (Fig. 1) where the state of solvency (non-default) is denoted by \( S \), and the state of default is denoted by \( \delta \).
JLT[1997] expand the transient class into various communicating states (the $S$ chain is thus irreducible within itself). They do so with the purpose that these states within $S$ can be likened to the various rating levels levied on corporate debt that are periodically produced by bond rating companies such as “Moody's” and “Standard and Poor's”.

Letting $\{1, 2, ..., K - 1\}$ be solvent states, and $K$ the absorbing state, they specify the following $K \times K$ transition matrix from time $t$ to $t+1$

$$
\tilde{Q}_{t,t+1} = \begin{pmatrix}
\tilde{q}_{11}(t,t+1) & \tilde{q}_{12}(t,t+1) & \cdots & \tilde{q}_{1K}(t,t+1) \\
\tilde{q}_{21}(t,t+1) & \tilde{q}_{22}(t,t+1) & \cdots & \tilde{q}_{2K}(t,t+1) \\
\vdots & \vdots & \ddots & \vdots \\
\tilde{q}_{K-1,1}(t,t+1) & \tilde{q}_{K-1,2}(t,t+1) & \cdots & \tilde{q}_{K-1,K}(t,t+1) \\
0 & 0 & \cdots & 1
\end{pmatrix}
$$

From this Markov matrix, JLT[1997] construct the probability of solvency. Given that the bond is in state $i$ at time $t$, its probability of not being in class $S$ is given by the following lemma;

**Lemma:** Let the firm be in state $i$ at time $t$, denoted by $\eta_t = i$, and given $\tau^*$ as previously defined, then the probability that default occurs after time $T$ is,

$$
\tilde{Q}^i(\tau^* > T) = \sum_{j \neq K} \tilde{q}_{ij}(t,T) = 1 - \tilde{q}_{iK}(t,T)
$$

where $K$ is the state of default, and $\tilde{q}_{ij}(t,T)$ represents the probability that the security goes from state $i$ to $j$ at time $t$. 
**Proof:** Since $K$ is an absorbing class, the event $(\tau^* > T)$ is equivalent to $\eta_i$ not being in state $K$ at time $T$, when starting in state $i$ at time $t$. Thus the probability of not being in default at time $T$ is equal to the sum of all states not $K$, or

$$1 - \left( \sum_{j \neq K} \tilde{q}_{ij}(t, T) \right)^C.$$ 

The use of this lemma allows us to formulate (1.7) as the price of the security with rating $i$ at time $t$ as,

$$v^i(t, T) = E^Q_t \left[ \sum_{k=1}^{T} \frac{x_k}{(1 + r(k))^k} \right] \left( \delta + (1 - \delta) \tilde{q}^i(\tau^* > T) \right)$$

(1.9)

This pricing equation is the main result of JLT[1997]. From this equation, JLT[1997] show how we can deduce formulae for forward rates and how the result can be used to hedge options written on risky bonds. Moreover, the authors show how one can estimate these Markov chain matrices by using historical volatility data and other methods such as Monte Carlo simulation.

### 4 A Critique

JT[1995] and JT[1997] provide a simplified and tractable framework for pricing corporate debt, which inherently carry with them a probability of default. However, some of the simplifying assumptions of the model and its extension prevent the applicability of the model to pricing certain types of debt. To induce a ‘separation in pricing’ between the security and the default process, the model assumes that the default and interest rate processes are independent. While independence of these processes may hold true for a number of cases, (or that the correlation effects are negligible,) the model may become
wildly inaccurate for corporate debt issued by firms whose revenue streams are tied to the level of the interest rate. One can find a wealth of example firms in the financial sector.

Further, the model assumes completeness of asset markets and the absence of arbitrage. When holding, the assumption provides us with the existence and uniqueness of the equivalent martingale measure that simplifies our pricing method to the derived formulae above. However, to hold, both completeness in markets and no-arbitrage must rely on the frequent updating of valuation information via market trading. That is, the securities in the market needed to be traded constantly to reveal current information about their prices. These properties hold reasonably well in the “investment grade”\textsuperscript{11} debt market, a class of securities which sees high daily trading volumes. Because of these high trading volumes, investment grade securities have up-to-date valuation information at every trading interval, and thus for every time step the market for investment grade securities can derive a “complete” set of prices. The assumption of no-arbitrage is satisfied in a similar way, where securities that are traded at high volumes have all opportunities for riskless profit ‘traded away’. The assumption fails, however, for the class of speculative securities. For this class, at every period, securities are generally thinly-traded, with some securities going untraded for multiple periods. Thus, for every security that is not traded in a certain time period, a price cannot be derived, and thus the market to which that security belongs is incomplete in prices, causing the assumption to

\textsuperscript{11} Investment grade debt is that which receives a grade of BBB or higher from credit rating agencies such as Standard and Poor’s or Moody’s. Any security rated below BBB is considered speculative grade.
fail. The failure of this assumption may lead the equivalent martingale measure to not be unique, or to fail to exist, resulting in a mispricing of the security.

Lastly, and perhaps the most pertinent issue to both the JLT[1997] use of a Markov chain and the current economic conditions, is the reliability of the bond ratings produced by firms. The problem is a true conflict-of-interest conundrum that calls into question the usefulness of bond ratings. While the firms may be independent of the bond-issuers themselves, it is often the bond-issuers who hire the bond-raters, and thus the bond-raters have an incentive to misrepresent the rating of the firm so that they may be hired again in the future. This is the major problem of “competitive bond rating”. A particular example pertains to the “Subprime Mortgage Crisis”, where various funds purchased mortgage debt that was rated to be very reliable investment grade, when in fact it was speculative grade. Because of the debt’s misrating, when the market for subprime securities crashed, it seemed as though securities for which \( \tilde{q}_{ik} = 0 \) were in fact reaching state \( K \) in large, unfathomable numbers.

In the final analysis, the model provides a fair and tractable approximation for pricing corporate debt, most realistically so in the case of investment grade debt. While much of the model remains to be empirically tested under its assumptions, the model provides a useful framework for thinking about the presence of default when pricing financial assets.

---

12 The economic intuition behind the claim that a security needs to be traded at a particular time-stage to be “price-able” at that time stage is that an asset is only worth what someone will pay for it, and thus a trade signals a willingness of an agent to pay for that asset at that time.
5. Bibliography


