Pricing Defaultable Bonds

A Study on the Performance of a Hazard Rate Model on US Bond Indices

The Aarhus School of Business
September 2002
# TABLE OF CONTENTS

## INTRODUCTION AND CONCLUSION

PART 1  EVIDENCE OF CREDIT RISK

1.1 The Market for Corporate Bonds

1.2 Historical Evidence of Credit Risk

1.3 Explaining the Credit-spread Dynamics

   1.3.1 Dimensions Underlying Credit-Spread Changes

   1.3.2 US Corporate Bond Indices - A Preliminary Study

      1.3.2.1 Data Description

      1.3.2.2 The Short-Run Dynamics

      1.3.2.3 The Long-Run Dynamics

1.4 Credit Risk Modeling

   1.4.1 The Theoretical Approaches to the Pricing Problem

   1.4.2 Intensity-Based Models

PART 2  THE PRICING OF DEFAULTABLE BONDS

2.1 The Fundamental Framework and Assumptions

   2.1.1 Preliminaries

   2.1.2 Markovian Settings and the General Pricing Model

   2.1.3 The Equivalent Martingale Measure

   2.1.4 Analytical Solutions in Affine Settings

   2.1.5 The Multifactor CIR Model

      2.1.5.1 Deriving the ODE’s and Analytical Solutions in the CIR Setup

2.2 Deriving the Generic Hazard Rate Model

   2.2.1 The Survival and Hazard Function

   2.2.2 Constructing Default Times

   2.2.3 The PDE Representation

   2.2.4 Different Recovery Schemes

      2.2.4.1 Recovery at Treasure (RT)

      2.2.4.2 Recovery of Face Value (RFV)

      2.2.4.3 Recovery at Market Value (RMV)
2.2.4.4 Implementing Recovery Considerations .......................................................... 57
2.2.5 Accounting for Coupon Payments ...................................................................... 59
2.2.6 Area of Application and Previous Empirical Studies ......................................... 61
2.3 The Duffee (1999) Model ................................................................................... 63
2.4 Interpreting the Properties of the Model .............................................................. 67
  2.4.1 Fundamental Properties .................................................................................... 67
    2.4.1.1 Cross-section Properties ........................................................................ 68
    2.4.1.2 Fitting the Observed Term Structures and Excess Yields ....................... 72
  2.4.2 Specification of the SDE in the CIR setup ....................................................... 75
    2.4.2.1 Investigating the Drift Function .............................................................. 75
    2.4.2.2 Investigating the Diffusion Function ...................................................... 76

PART 3 EMPIRICAL INVESTIGATION ............................................................................. 79
3.1 Estimation Methodology ..................................................................................... 79
  3.1.1 The Modified Chen & Scott Maximum Likelihood Estimator ......................... 82
  3.1.2 The Default-free Process ............................................................................... 83
  3.1.3 The Default-risk Process ............................................................................... 85
  3.1.4 Implementation in GAUSS ............................................................................ 87
  3.1.5 Evaluation of the Models .............................................................................. 89
3.2 Estimation Results ............................................................................................... 91
  3.2.1 The two-factor Treasury Model .................................................................... 91
  3.2.2 The one-factor Credit-Spread Model ............................................................ 97
3.3 Robustness and Specification Tests ..................................................................... 104
  3.3.1 Alternative Error Structures ....................................................................... 104
  3.3.2 Finite-sample Consistency .......................................................................... 108
    3.3.2.1 The CMLBoot in GAUSS .................................................................... 110
    3.3.2.2 Revision of the critical values .............................................................. 110
  3.3.3 Specification Tests ......................................................................................... 112
3.4 Pricing Accuracy and Model Evaluation .............................................................. 115

References
Appendix
Introduction and Conclusion

Credit risk has long been recognized as an important aspect of fixed income securities, but it is not until recent years the quantitative modeling of credit risk has received attention. With the increased popularity of corporate issued debt, the emergence of various over-the-counter derivative contracts, the attention to new contracts based explicitly on credit risk, and the very important new BIS (Bank of International Settlement) regulations, the interest in credit risk and credit risk modeling has increased considerably and will presumably gain further attention in the future. To address the appearance of credit risk, practitioners have adopted a variety of models of more or less sophistication, all with the essential objective of measuring credit exposure such that decision-making, pricing, and hedging is made on more confident grounds.

In addition, considerable new knowledge about the mechanisms that lead to default and what happens in the event of default has enhanced the pricing of bonds subject to credit risk. However, major problems still exist for the adequate modeling of credit risk as well as the estimation of the parameters in the models. So, even today the vast majority of market participants use pricing models that do not implement credit risk, or at least only account for the possibility of default in more pragmatic ways.

In this thesis we examine the no-arbitrage model presented by Duffee (1999). In particular the purpose is to evaluate the model’s ability to capture the dynamic behavior of US corporate bond indices by analyzing the model’s theoretical properties, and by testing the performance of the model against market prices. Bearing in mind that no tractable model can fully capture the complexities of the dynamics in the market, the design is naturally a compromise between economical realism, computational tractability, and flexibility. The Duffee (1999) model makes reasonable compromises among these counteractive objectives. Indeed, the model treats the default event as a one-jump process governed by exogenous specified variables. Hence, the time of default occurs at a surprise stopping time, or stated otherwise, the model utilizes the intensity-based framework whereat the technical difficulties are reduces to those faced when modeling default-free bonds. However, recognizing that the risk profile of corporate and Treasury bonds are very different, it might seem strange that ordinary term-structure models can resemble the dimensions underlying corporate bonds. In order to clarify this aspect and ultimately possible misspecifications we therefore address the issue of disclosing the elements of credit risk using the information contained in the term structure of credit spreads. This enables us to explain the results on more confident
grounds, and accordingly which features of the data that are well-described in an affine framework and which are not. The discussion is general and therefore applies independently of the specified model. Furthermore, we investigate the consistency of the results. This involves both the specific econometric procedure as well as the significance of the estimates in our finite sample.

Applying intensity-based models for pricing defaultable bonds provide a very flexible framework and allow extensions in several respects. Basically, the underlying philosophy of the models is that default is presumed characterized by a simple binomial process, default or not default, which occurs completely unexpected. The trick, however, is that instead of addressing the discontinuous two-stage default process we let the process be represented by a probability (intensity or instantaneous default probability). The likelihood of default is then governed entirely by the specific probability process, which entails the desirable property that defaultable securities can be priced using an adjusted discount rate. In the Cox-methodology developed by Lando (1998) this provides tractable mathematics and enables the intensity and possible recovery to be stochastic, fluctuating with the financial health of the firm. Assuming no recovery in the event of default, the only thing that distinguishes defaultable and default-free bonds is exclusively the probability of default. However, in the event of default bondholders usually receive some fraction of the value of their original claims. So to motivate a fully operational specification of the bond price zero recovery is not tenable. We therefore address three generic recovery schemes: the recovery at Treasury, recovery at face, and recovery at market value, where the latter entails that the discount rate consolidates time value, default probability, and recovery considerations. As the emphasis of this thesis is investment grade bonds recovery considerations are of second order. Moreover, as we show, the schemes are approximately identical for reasonable parameters.

Given the framework, the success of the intensity-based models entirely depends on the parameterization of the intensity and recovery process. In the Duffee (1999) model the dynamics of the defaultable bond prices are separated into dynamics resulting from market and credit risk, respectively. This is done by letting a translated two-factor CIR model characterize the default-free term structure and the default intensity process by a translated one-factor CIR model. Employing a constant recovery at market value scheme, and two linear terms to account for the empirically verified inverse relation between default risk and interests rates, we estimate the parameters in the model using the maximum likelihood inversion method. The method involves a panel data approach whereat cross-sectional and time-series properties are exploited simultaneously. As a joint estimation of the model is computationally infeasible we employ a two-step procedure. In the first
step the parameters of the Treasury term structure are estimated. Keeping these constant, we then estimate the parameters of the corporate bond dynamics. Because the model is estimated on bond indices we expect all non-systematic risks to be diversified, which justifies that we only incorporate systematic risk factors in the intensity formulation. Using Monte Carlo simulation to generate pseudo data based on our models we find that the MLE programs in general provide non-biased estimators. However, the speed of adjustment coefficients exhibit small biases. The most likely explanation is the ad hoc specifications of the true yields and measurement errors in the likelihood functions. As a result we investigate two alternative model designs and find that the derived ML-models are sensitive to the applied error structures. This has the effect that the alternatives not only provide different mean log-likelihood values but seemingly different parameter estimates. This reveals an obstacle with the consistency of the applied ML-models, and ultimately inadequate estimation methodology, implying that the framework does not manage to incorporate all the information in the yields. Duffee (1996) encounter similar problems with the MLE-inversion method and employ a Kalman Filter instead (Duffee, 1999), which alleviates the problem.

On average we find that the model fits the yields of corporate bond indices fairly for maturities less than 7 years. The average root mean squared error is generally less than 20 basis points but the fit deteriorates considerably for longer maturities. Not surprisingly we find that the inherent two-factor Treasury model is the most accurate, with root mean squared errors less than 10 basis points. In accordance with Chen and Scott (1993) we find that the first factor exhibits strong mean reversion and that the second is close to non-stationary under the pricing measure. This result entails that the model is capable of capturing the dynamics of long maturity Treasury yields. Economic interpretations of the state variables reveal that the first factor is strongly related to the slope of the term structure whereas the second factor is strongly related to the level.

The single-factor formulation of the instantaneous default probability is in theory capable of replicating the observed term structures of credit spreads. However, this is at the cost of negative mean reversion under the pricing measure. This has several important implications. First of all, since the intensity process implies an explosive drift structure the instantaneous default probability will be positively related to its level. Fitting the cross-sectional properties will therefore be at the cost of not capturing the excess yields as they become explosive. These aspects are both economically and intuitively implausible, and substantiated by the fact that we typically find that the derived default probabilities are stationary. In practice this has the discouraging consequence that we cannot use the model for pricing contingent claims when it is necessary to simulate the payout distri-
bution of the claim at maturity. Secondly, negative mean reversion makes the model very sensitive to changes in the intensity and induces fairly steep credit-spread curves even for low levels of the intensity. Thus, given a set of parameter values, the implied term structures will quickly become inconsistent with empirical findings once the business climate changes. Furthermore, negative mean reversion restricts the range in which the intensity can vary without implying downward-sloped credit curves at long maturities. This is in contrast to the common findings of investment grade bonds, and the model is therefore a priori limited to bonds of short and intermediate maturities.

Testing the one-factor model’s ability to describe credit risk we find that the model on average produces credit spreads consistent with empirical observations i.e. when the level rises, so do the slope, and credit curves become steeper for lower credit quality. Moreover, we find that the minimum spread increases for lower credit quality, although it was not possible to confirm statistically. The negative relation to the level of the Treasury term structure is well supported, or stated in an economical sense, investors require a lower credit adjustment when interest rates are high as this indicate improvements in the business climate. Unfortunately, we are not able to verify the empirically observed stronger correlation for lower ratings. The correlation with the slope of the Treasury term structure is also questionable although a preliminary regression analysis revealed significant negative relationship. Compared to the Treasury term-structure model the fit is not surprisingly less accurate and only acceptable for maturities less than 7 years where the root mean squared errors generally are below 20 basis points.

To avoid the likely finite-sample bias we test the significance of the parameters using their empirical distributions. In fact we find noticeable differences between the empirical and asymptotic distribution functions, and consequently many of the parameters become insignificant. Whether this indicates misspecification is revealed by two separate investigations. First of all, some of the limitations owe to persistent volatility clustering and non-linearities, which we observe in the data. As the CIR model is not adequately designed to capture these aspects we get biased and eventually inefficient parameter estimates. Secondly, the model does not imply that the time series of normalized innovations of the instantaneous default probability is white noise. However, recognizing that most models encounter such problems, we find that the most basic failure is the model’s inability to simultaneously fit both the level and the slope of the credit curves. This indicates a weakness with the fundament of our model. To explain the steepness of the curves a relatively large mean aversion is necessary, which is inconsistent with the flat term structures that are
observed when the instantaneous default probability is low. Thus, it seems crucial to incorporate a stochastic long-term level like in the double decay model. We believe, when a stochastic level is incorporated into the drift function, then revising the volatility process is of second order although it seems plausible that higher volatility elasticity should be integrated. It should of cause be noted that implementing such a model does not come without sacrifices. Especially, it will be at the cost of the inherent mathematical tractability, and to some extend the economical interpretation of the results.

Although the model is conceptually elegant, we emphasize its limitations in practical applications. Instead we believe it is very useful in giving theoretical insights in the underlying dimensions of corporate bond dynamics. In particular it describes the fundamental relationships in a simple and intuitive manner, and reveals, which features that are well-supported in an affine framework.

The thesis is organized into three main parts. In the first part we discuss the elements of credit risk and present the data on a preliminary basis. In the second part we introduce the state-space framework and the basic underlying assumptions wherein we shall work. Subsequently, we present and derive the generic model for pricing bonds subject to credit risk. Finally we test the model’s ability to price defaultable bond indices in the third part.

To clarify some formulation we use the term hazard rate, intensity and instantaneous default rate or probability synonymously.
Part I: Evidence of Credit Risk

In order to formulate usable models for defaultable securities, it is essential to recognize the stochastic nature of credit spreads and to understand the components affecting it. Thus, to motivate the functional form of the pricing model derived later and to explain possible misspecifications, we need to know how default occurs and what happens in event of default. Therefore, in this part we evaluate the historical evidence of credit risk in the American market for corporate bonds, and shed light on the nature of default behavior, covering both the likelihood and severity of default. We focus on discussing the different aspects of credit risk, with the aim of addressing the question of how to integrate these in pricing of corporate bonds. Eventually, we review the different theoretical attempts to model credit risk, and present the main refinements within the intensity-based approach.

1.1 The Market for Corporate Bonds

The American market for corporate bonds is by far the largest in the world, with a total market value of outstanding bonds of approximately $3.68 ($2.35) trillions in 2001 (1995), and new issues of approximately $0.67 ($0.33) trillions. In contrast, the European market for corporate issued debt is fairly undeveloped in the amount outstanding (approximately 1/5 of the American market), rating and liquidity. The reasons for the differences between the two markets are undoubtedly comprehensive, but to a large extend a consequence of tradition in the credit culture. Whereas European corporations tend to finance their activities through traditional bank loans and equity issues, US corporations have by tradition issued bonds to finance their activities. Apparently, as it appears form Figure 1-1, the market has grown steadily since the 80’es, and from the mid 90’es the market experienced very significant growth.
Part I: Evidence of Credit Risk

Figure 1-1: Evolution of the US Corporate Bond Market ($ Billion), 1986 – 2001

The figure depicts the evolution in the corporate and Treasury bond market, 1985-2001, with segmented information about the value of new issues and the total debt outstanding.


The growth was primarily driven by the improved business climate, where corporate bonds were used to provide working capital for growing companies. Essentially the market has increased from a ratio of 0.71 to 1.08 of the size of the US Treasury public outstanding debt from 1995 to 2001.¹ Moreover the corporate market has typically been dominated by long-term debt, but in recent years the short-term issues have increased considerably. As a result the short-term market is now larger than the corresponding Treasury market, while the corporate bond market for long-term obligations has been the largest since the early 90’es.²

1.2 Historical Evidence of Credit Risk

The basic paradigm confronting bonds associated with credit risk is that the yield should exceed the yield on an otherwise identical default-free bond, because risk-averse investors require compensation for the inherent risk of default. Stated differently, the credit spread (the difference between the corporate bond yield and the corresponding Treasury yield) contains a risk premium to reimburse bondholders for the possibility that the issuer fails to fulfill its obligations, in which case bondholders only receive partial payment.

Using Moody’s credit database as proxy for the overall credit quality of US outstanding corporate debt, the quality has continually improved during the beginning and mid 90’es. From Figure 1-2 it

¹ Public Treasury debt includes all non-intergovernmental debt.
² The information is obtained from Moody’s Investors Service, Special Comments 2002 and the homepage of The Bond Market Association under Research (statistics).
Part I: Evidence of Credit Risk

appears that average default rates have almost relentlessly decreased from the peak in June 1991 of approximately 5.3% (13.0% Spec. grade) to approximately 0.5% (1.4% Spec. grade) in April 1995. The trend is an obvious reflection of the robust economic growth and continually worldwide activity dominating this period.

**Figure 1-2: Historical Default Rates of Corporate Bonds, 1991-2001**

The figure illustrates the 12-month moving average of historical default rates of all corporate bonds together with the average rate of speculative grade bonds, 1991-2001. Note that default rates are calculated with the issuers as the unit of study. The numerator represents the particular number of issuers that defaulted within a 12-month period, and the denominator represents the number of issuers that potential could have defaulted. Hence, withdrawn issuers are subtracted from the denominator.


Considering the late 90’es the trend has, however, revolved. Although worldwide economic activity continued its steady pace of expansion, 1999 showed a sharp increase in the default activity, and as it appears from the figure the credit quality has continued to deteriorate in 2000 and 2001, revealing the change in business climate and general market skepticism.³

Since the spread serves as a compensator for the inherent credit risk it is obvious that credit spreads must be highly correlated with default rates. Figure 1-3 depicts the aggregate yield spread for corporate bonds of speculative grade versus the average default rate for these companies.

The average spread in the period is approximately 410 bp, and it appears that high correlation is predominantly characterizing the period. The positive spread is analogous to saying that corporate

³ We emphasize that some of the upward tendency reported in the late 90’es growth is caused by slowdown in the growth of the issuer pool, which is directly reflected in the default-rates denominator. Hypothetically, assume the numbers of defaults are constant through time, and that the growth in the issuer pool is zero – i.e. the denominator is only reduced by the number of defaults – the default ratio will then increase exponential towards one. See Moody’s Investor Service – Special Comment, January 2000 for further elaboration on historical default rates of corporate bonds.
Part I: Evidence of Credit Risk

bonds trade at a discount compared to treasuries, which is essential and necessary for precluding arbitrage. The inherent credit risk is therefore clearly reflected in the yields, however, given the default rates, adverse movements in the credit spread do occur. As seen in the figure, the correlation deteriorated in the fall of 1998 as low liquidity and shifting investor preferences shocked the speculative market and increased the spread significantly.4

Figure 1-3: Speculative-Grade 12-Month Moving Average Default Rate vs. Average one-year Speculative-Grade Spread over Treasuries, 1991-1999

The figure shows the 12-month moving average of historical default rates of speculative grade bonds together with the average credit spread measured as the number of basis points over Treasuries.

Although not completely obvious from the figure it also seems that the correlation decreases when spreads widen, indicating that in times of high uncertainty in the credit market the credit-spread dynamics become more sensitive to other factors than those describing default risk.

Not disclosed in the figure is how the spread evolves across maturity. Like the Treasury term structure shows how yields of zero-coupon bonds vary with maturities, the term structure of credit spreads depicts the relationship between the credit spreads and their respective maturities. The spreads are most likely not constant across maturity, and would certainly reveal very distinct cross-sectional properties compared to the Treasuries. In normal situations the term structure of credit spreads is expected to increase, which intuitively seems logical as one would expect the default probability of an e.g. highly rated firm to be low in the near future, but to increase in the distant, as the effect of possible worsening of the firm’s competitive state comes into perspective. From an economical point of view, upward sloping credit curves can therefore be seen as a consequence of an increased probability of competitors entering the market, shifting consumer prefer-

4 For further details the reader is referred to Moody’s Investors Service, Special Comments (2002).
Part I: Evidence of Credit Risk

ences etc., as we look further into the future. Empirical studies verify these reflections, although, at the same time disclose that credit curves appear in different shapes. They are not only upward sloped but downward as well as humped. These observations can be seen in the credit spread curves of US investment grade bonds in the period 1991-2002 provided in Appendix A (Treasury term structures are also depicted in the appendix). As illustrated the credit curves are typical upward sloped, though in some years displays humped tendencies.

Researchers have verified these findings, emphasizing that credit risk and time to maturity is not independent, and that successful model specification must be capable of replicating a verity of term-structure shapes. However, in contrast to upward sloped credit curves, downward or humped credit curves can be difficult to give an economical interpretation, and for that reason they are often argued to be a consequence of some sort of bias in the data. Fons (1994) notes that credit spreads increase with maturity, and finds evidence which support that spreads of lower rated bonds is slightly curved or downward-sloped. He explains these findings by the fact that IPO issues, which typically involve high uncertainty in the introduction stage, dominate the speculative grade market. If the firms survive the first stage their credit risk is expected to decrease. Helwege and Turner (1999) argue that the downward sloping credit curves derived from indices might be a consequence of using average credit spreads of bonds within the same rating class. Safer firms tend to issue longer matured debt, causing the average spread to decline with maturity even though for a single company the spread increases.

Although their arguments seem plausible and undoubtedly will explain some of the downward tendency, we also find it sensible to expect that downward-sloped credit curves in fact are a consequence of lower expected losses. Imagine for example, that the markets currently are in recession, i.e. default probabilities are high, but leading economical indicators reveal expected future optimism. Consequently we would expect the mean losses to decrease.

1.3 Explaining the Credit-spread Dynamics

Throughout history a large body of econometric literature covers the explanation of credit-spread changes. Even so the actual stochastic behavior of credit spreads is still not well-understood. Unlike market risk, credit risk is very difficult to measure correctly as it is not just a simple manifestation of one single source or driver of the risky event. In order to capture every aspects of credit risk, it seems necessary to evaluate all factors that may influence the firm’s future earnings

Part I: Evidence of Credit Risk

potential. This is obviously a complicated task, involving examination of a range of factors covering general economic conditions, industrial trends and specific issuer factors like the issuer’s financial well-being, degree of leverage, market value, capital structure, and less tangible factors, such as reputation and management skills etc. The risk profile of corporate bonds therefore differs significantly compared to the risk profile of other financial debt obligations such as government bonds. In the absence of an adequate framework for modeling credit risk, we have to decide which elements of credit risk that in general are the most important i.e. giving us the most information. The choice is indeed non-trivial, and is directly related to the specific area of application. On the other hand, we might instead attend to the available information provided by external rating agencies, which are specialized in exploiting and evaluating the credit risk components. Using their contribution will to some extend provide the necessary information, and, accordingly, from a pricing perspective, this has been the motivation for several theoretical researchers. Nevertheless, this is not without problems, which we elaborate later in this chapter.

In our specific application it is worthwhile to recognize the knowledge gained from econometric analyses of credit-spread changes. In such, it has been verified that a large part of the variance is adequately captured by a limited number of systematic risk factors so, in practice, the dimensionality of the analyses can generally be reduced considerably (at least on an aggregate level). Throughout this chapter we therefore discuss the elements characterizing credit risk, and how these elements translate into credit-spread changes, in order to i) build a fundament for examining whether these properties are consistent with the factor model framework presented later, and ii) to reveal the problems encountered when investigation corporate yield data.

1.3.1 Dimensions Underlying Credit-Spread Changes

The elements determining the credit premia contained in the spreads are from a theoretical perspective usually decomposed into three key variables - the default probabilities as mentioned above, the severity of loss in the event of default, and the exposure level at default. The severity of loss in the event of default (commonly denoted loss given default, LGD) is specified by the recovery function of the issuers, i.e. the fraction of the promised payment, which the defaulting entities are able to pay. Credit spreads should then, at least theoretically, be driven by the mean loss rate, \( \mathbb{E}_i \{ L \} \), defined as

\[
\mathbb{E}_i \{ L \} = \mathbb{P}_i \times (1 - \phi_i) \times \xi_i
\]

where \( \mathbb{P}_i \) is the default probability, \( \phi_i \) is the recovery function, and \( \xi_i \) is the exposure at default. Equation (1-1) reveals why defaultable bonds, despite the strong analogy, differ significantly from
Part I: Evidence of Credit Risk

default-free bonds when it comes to expected returns. To compute the expected return on defaultable bonds one has to know the mean loss rate, i.e. the product between the likelihood of default and how much of the value is expected lost in default, \( (1 - \phi_i) \xi \). However, due to the non-stochastic credit exposure of bonds the complexity is somewhat lower than e.g. derivative contracts. In the case of corporate bonds the exposure is known to the size of the principal and for fixed-rate bonds the coupon payments, while the exposure of a derivative contract is not known in advance. As a consequence we can disregard the last term - the exposure at default - in equation (1-1).

The likelihood of default, measured by the default probability process, is then the key variable in measuring credit risk. Albeit its usually small size, the impact of default influences the return on an investment considerably. Consequently, if \( \mathbb{P}_i > 0 \) the distribution of possible returns becomes skewed i.e. the profit potential is limited compared to the possible loss. It is of course very difficult to predict that a given firm defaults within e.g. the next year with a given probability. Even so, the likelihood of default is usually not entirely unexpected, since default is often initiated by a deterioration of the firm’s financial position, and as illustrated in Figure 1-3 default rates fluctuate systematically with changes in business climate and shifting macroeconomic trends. Hence, we can define \( \mathbb{P}_i \) as

\[
\mathbb{P}_i = f \left( \text{firm specific variables, macroeconomic variables} \right)
\]

where firm specific variables cover key drivers such as asset value and asset volatility, leverage, solvency etc. and the non-diversifiable macroeconomic variables cover GDP and other leading indicators. In principle an adequate modeling of default probabilities should implement these variables, which presumably are distinct across industry segments and ratings.

In the particular field of default dynamics corporations like KMV, Moody’s and Standard and Poor’s have achieved progressive insight, and expanded the understanding of the causal mechanisms leading to default. Nevertheless, defaults are still characterized by noteworthy randomness. Take for example the resent defaults by Arthur Andersen, Enron, and WorldCom where concepts such as event- and business risk consequently become synonymous with credit risk.

Turning the attention to the recovery function and what happens in the event of default the knowledge is somewhat limited compared to the insight in default risk. In fact, much of the quantitative analyses of credit risk have been devoted to modeling \( \mathbb{P}_i \), whereas severity considerations have
Part I: Evidence of Credit Risk

largely been neglected. The reason is that the factors affecting the recovery function are rather complex and information about these factors is not easily obtained. Most important to consider when analyzing the recovery function is the seniority and the quality of collateral (if any). Both variables have positive impacts on the bond price as they increase the expected recovery in case of default. Moreover, Altman, Resti and Sironi (2001) and Fridson (2000) suggest that the recovery rate is a contemporaneously implied function of default, and suggest significant variations in recovery though time, and between different segments. Their findings reveal that default probability and recovery is inversely related, hence recovery payouts tend to be low when defaults rise and vice versa. This negative phenomenon may, however, just be derived from the two variables mutual dependence on the business climate. The intuition behind this observation is that if a bond defaults the value depends on the value of the collateral, which again depends on macroeconomic conditions. If the economy experiences a recession, recovery rates may decrease just as default rates tend to increase, which give rise to the negative correlation. Nevertheless, the recovery can be expressed as the function

\[ \phi_t = f(\text{seniority, collateral, macroeconomic variables}) \]

although we emphasize that the literature does not entirely agree on these considerations. Carey and Gordy (2001) for example find that the correlation is close to zero. Still they recognize the systematic relationship between economic performance measures and expected default rates but find that macroeconomic forces and their changes are less important in explaining recovery rates.

Combining these results reveal why the loss process (1-1) is difficult to decompose and rather complicated to clarify in detail. Nevertheless, if we consider the variance of the spread, it is obvious that it must be dictated by the markets perception of the issuer’s general creditworthiness. Using the logic of equation (1-1), we are capable of establishing the link between the theoretical loss rate and the actual observable spreads. In order to disclose the components in equation (1-1), credit spreads are often investigated through the relation

\[ \Delta CS_t = \alpha + \sum_{j=1}^{n} \beta_j \Delta f_{ij} + \varepsilon_t \]  

(1-2)

to reveal the variables causing the yield spread to increase. CS is the credit spread, and \( f_i \) are different candidates for explaining the spread.

However, to focus solely on the spread could lead to erroneous conclusions about the issuer’s credit quality since the spread has embedded not only credit information but also components
non-related to default risk. Stated differently, premiums due to expected default losses are insufficient in explaining the corporate bond spread. Most important to consider are components such as tax- and coupon-induced biases, call-ability, and liquidity-effects. Furthermore, several researchers suggest that the corporate bond market is considerably segmented driven by local supply/demand shocks (see Duffee (1996), Elton and Gruber (1997), and Collin-Dufresne, Goldstein, and Martin (2000)). In effect they report that although considering a large number of variables that should measure changes in default probabilities or recovery rates, it is generally not possible to explain more than 25 percent of the observed credit-spread changes on individual bonds. So, while credit quality is undoubtedly important, and has a straightforward impact on spreads, it is definitely difficult to measure with the necessary precision.

For our present objective, the aspects outlined above may have the effect, even if the pricing model is well-specified, that the parameters become insignificant or spurious in the empirical estimation. A simple way to circumvent some of these problems is by using AAA-rated bonds as proxy for default-free bonds. However, as we wish to estimate the factor model to observed yields, which include components non-related to default risk, we do not pursue this here. Even though the actual decomposition of the spread is not specified in the bond pricing model we derive later, it is fundamental to understand the different components and the characteristics of spread changes when conducting the empirical study of the model. Otherwise it will be difficult to explain potential mispricing patterns.

1.3.2 US Corporate Bond Indices - A Preliminary Study

In the subsequent study, we exclusively focus on systematic non-diversifiable credit risk factors using risk-free interest rate variables as proxies. This approach may seem harsh but nevertheless, Bakshi, Madan, and Zhang (2001) show that firm specific distress factors are not redundant but only of second-order importance. Once systematic distress factors, measured by interest rate changes, are taken into account, the pricing improvements of using firm specific factors are only marginally. Furthermore, as we investigate corporate bond indices, diversifiable risk are expected to be eliminated.

We investigate the short-run dynamics of the credit spreads through the regression

\[ \Delta CS_t = \alpha + \Delta Level ST_t + \Delta Level LT_t + \Delta Slope_t + Non LE_t \]  \hspace{1cm} (1-3)

where the credit spread is expressed as a function of the short and long term level of the Treasury term structure (Level ST and Level LT, respectively), the slope, and finally any non-linear rela-
Part I: Evidence of Credit Risk

tionship between the Treasury term structure and credit spreads is modeled by the square of the first difference in the long-term interest rate (NonLE). The reason why we include both the short and long term level as explanatory variables is the discrepancy between the empirical studies covering this area, where Duffee (1998) uses a 3-months Treasury rate whereas Collin-Dufresne et al. (2000) uses a 10-year Treasury rate to describe the level of the Treasury term structure.  

We proceed as Collin-Dufresne et al. (2000) among others and analyze the relationship through a regression model in first differences indicating non-stationary variables. Since a non-stationary process implies an explosive volatility structure over time, it seems implausible that the level of Treasury yields, the slope of the Treasury term structure, or credit spreads are non-stationary over longer periods of time, because they all somehow are bounded by regulation. However, since most pricing models have a relatively short time horizon, it seems plausible that these yields could be non-stationary, which is labeled I(1).

Table 1-1: Test of Unit-Roots, 01/1991-01/2002

<table>
<thead>
<tr>
<th>ADF-test</th>
<th>CS</th>
<th>Level ST</th>
<th>Level LT</th>
<th>Slope</th>
<th>NonLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>t-stat</td>
<td>-2.78</td>
<td>-1.76</td>
<td>-3.41</td>
<td>-0.01</td>
<td>-22.4</td>
</tr>
<tr>
<td>Critical value</td>
<td>-2.87</td>
<td>-2.87</td>
<td>-3.42</td>
<td>-1.94</td>
<td>-3.42</td>
</tr>
</tbody>
</table>

This table presents the ADF-Unit-Root test based on weekly observations in the period January 1991-2002. The critical values are obtained at the 5% significance level. The credit spread is defined as the difference between the corporate bond yields and the corresponding zero-coupon yields with equal duration. The shown spread in this table is based on 2-years duration of AA-rated bonds. Furthermore the level variables are measured as the 1-year and 10-year zero-coupon yields, respectively. The slope is defined as the difference between the 20-year and 1-year yield and finally the square of the first difference of the 10-year yield is used as a proxy of non-linearity effects.

As seen from Table 1-1 we were only able to reject the hypothesis of a unit-root in NonLE indicating a stationary variable. This conclusion applies for the credit spread independent of the chosen maturity (short, intermediate or long) and rating, which is seen from Appendix B. Turning more explicitly to the Treasury variables we observe a t-statistic of Level LT close to the critical value at the 5% significance level. Therefore we applied the alternative Phillips-Perron Unit-Root test with the same result supporting that the variable is I(1). Although it is commonly known that there exist several pitfalls in such unit-root analyses we proceed with the estimation of equation (1-3).  

---

6 We are fully aware of this may induce multicollinearity in our model, which we inspect among other possible problems of model misspecifications.
7 These results are presented in appendix B.
8 The authors are aware of the pitfalls in using unit-root tests in connection with interest rate series as these series involve time varying volatility as indicated by Chan, Karolyi, Longstaff and Sanders (1992).
1.3.2.1 Data Description

The corporate yield data is compiled from Merrill Lynch Fixed Income Database, and contains observations in the period January 1991 to January 2002. The database provides daily information on all investment grade ratings of US corporate bond indices, categorized into four industrial groupings: Industrials, Utilities, Financial, and Yankees with maturities covering a span of two to 15+ years divided into six subgroups. However, we do not consider the analysis of industrial subgroups, and instead aggregate yields into the respective rating groups. Furthermore, the average duration for the indices vary somewhat across time, so, in order to keep the cross-sectional points fixed we generated new series of 2, 4, 6, 8, and 10 years duration by linear interpolation. It is then straightforward to compute the credit spread by matching the equivalent Treasury STRIPS. As an alternative to fixing the cross-sectional points we could have matched the duration of the corporate bonds with the Treasury yields of same duration. The advantage of this alternative is that the larger cross-section of Treasury bonds raises the hope of a more accurate interpolation, and secondly it gives an extra maturity on the corporate cross-section. The drawback is the method requires information of the maturities of the specific bonds, which is not easily incorporated in the preliminary regression study. We therefore use the former method in the following, and the latter method when estimating the parameters of the arbitrage model.

Before turning to the modeling of the spread dynamics we present some statistics of the series. Table 1-2 shows descriptives for the period January 1991 to January 2002, and two subperiods. Comparing the term structure of credit spreads and Treasury term structure, there are, not strangely, several notable differences. Combined with the plots in Appendix A, the table reveals an important difference; the Treasury term structure is steep when the short rate is low, and the volatility is larger for the shorter maturity yields. In contrast credit spreads typically exhibit steep tendencies when the short spread is high, and the volatility increases with maturity. The conclusions are generally consistent across all rating classes except for the indices of BBB-rated bonds where the standard deviation decreases with maturity as depicted in Appendix A.
Part I: Evidence of Credit Risk

Table 1-2: Summary Statistics for the US Treasury Yields, 1991 - 2002

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>STD</td>
<td>Mean</td>
</tr>
<tr>
<td>1-year</td>
<td>5.05</td>
<td>1.09</td>
<td>4.90</td>
</tr>
<tr>
<td>3-year</td>
<td>5.68</td>
<td>0.96</td>
<td>5.90</td>
</tr>
<tr>
<td>5-year</td>
<td>6.01</td>
<td>0.91</td>
<td>6.45</td>
</tr>
<tr>
<td>7-year</td>
<td>6.27</td>
<td>0.91</td>
<td>6.82</td>
</tr>
<tr>
<td>10-year</td>
<td>6.54</td>
<td>0.91</td>
<td>7.20</td>
</tr>
<tr>
<td>15-year</td>
<td>6.85</td>
<td>0.89</td>
<td>7.58</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>AA Credit Spreads, yield to maturity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-year</td>
</tr>
<tr>
<td>4-year</td>
</tr>
<tr>
<td>6-year</td>
</tr>
<tr>
<td>8-year</td>
</tr>
<tr>
<td>10-year</td>
</tr>
</tbody>
</table>

This table presents the sample means and standard deviations (STD) of the zero-coupon yields and credit spreads of AA-rated corporate bond indices. The sample contains 546 weekly observations in the period between January 1991 and January 2002. The two subperiods are January 1991 to December 1995 and January 1996 to January 2002, with 236 and 310 observations, respectively.

Moreover the Treasury yields are more volatile in the beginning of the 90’es, which is in contrast to the credit spreads. But more importantly, there seems to be a regime shift, which is indicated in the table by the differences in statistics of the two subperiods. This shift is in line with the evolution of the defaults rates around 1994, cf. Figure 1-2, where the prior period was characterized by high and volatile default rates. As a consequence the credit spreads in Appendix A exhibit awkward looking term structures from 91-93, but from 1994 the term structures are typically upward sloped although they in some years display humped tendencies. This observation is formalized in high Chow-test statistics indicating significant differences in the estimated equations in the two subperiods.\(^9\) As a result of the structural change the estimates become inefficient, and in addition to the factor model presented later this cannot capture such shifts and therefore suffers the same effect. To alleviate this problem we conduct our analysis on the period January 1994 to January 2002.

\(^9\) We divided the data into two subsamples – before and after January 1994 – then conducted the test for structural changes using Chow’s test. Generally we observed \(F\)-statistics above 10 (\(F\) critical is approximately 3) for medium and longer maturities, but for shorter maturities it was not always possible to reject the null-hypothesis of no structural change.
Part I: Evidence of Credit Risk

1.3.2.2 The Short-Run Dynamics

Table 1-3 presents the results of the regression for intermediate maturity (6-years duration) AA-rated bonds. The analysis is furthermore conducted on short (2-years duration) and long maturities (10-years duration) of corporate bonds in risk classes AAA to BBB (results not shown here are reported in Appendix B).

Table 1-3: Regression in first differences – AA rated intermediate maturity, 1994-2002

<table>
<thead>
<tr>
<th>Variables</th>
<th>Coefficient</th>
<th>STD</th>
<th>t-stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.000</td>
<td>0.004</td>
<td>0.25</td>
</tr>
<tr>
<td>ΔLevel ST</td>
<td>-0.104</td>
<td>0.019</td>
<td>-5.20</td>
</tr>
<tr>
<td>ΔLevel LT</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ΔSlope</td>
<td>-0.068</td>
<td>0.023</td>
<td>-2.85</td>
</tr>
<tr>
<td>Non-LE</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

The table presents the results of the estimation of (1-3) on AA-rated intermediate maturity bonds, 1994-2002. We only report the significant parameters, and therefore the cell is empty if the variable is not significant at a 5% level. The standard errors reported in the table are adjusted for heteroscedasticity and autocorrelation using Newey-West ($R^2 = 0.15$).

We generally find that credit spreads move inverse with the level and slope of the Treasury term structure, but we find relatively low values of $R^2$. The success of prediction generally increases with maturity and there is furthermore a tendency that the relationship is more prevalent for lower rated bonds. The poor performance of the model is depressing, but in line with Longstaff and Schwartz (1995), Duffee (1996), and Collin-Dufresne et al. (2000), and reveals some basic stylish relationships.

The negative relation to the level of Treasuries in general holds for every combination of rating and maturity, and as Appendix B shows the correlation increases slightly as the credit quality falls. From an economical point of view the sign of the correlation is not surprising as it is interpreted in line of the business cycle, i.e. when interest rates rise, the economy is expanding, hereby improving the economical environment for the firms, hence, we would expect the default probabilities to fall inducing lower credit spreads. What is more doubtful is the magnitude of the correlation, and indirectly the sensitivity of the implied default rates to changes in business climate. Longstaff and Schwartz (1995) find that credit spreads are strongly negative correlated to the level of the Treasury term structure. However, since corporate bonds often are callable, Duffee (1996, 1999) argues that the strong negative relation owes to variations in the embedded option components (call-options). This argument is in line with findings by Iwanowski and Chandre (1995), who avoided the call-option bias, and found only minor negative relations. The data used for our investigation contains both callable as well as non-callable bonds so, the results are a priori expected to be upward biased.
Part I: Evidence of Credit Risk

As mentioned we found a negative relation to the slope of the Treasury term structure albeit this relationship is only significant at the intermediate maturities. This result is supported by the findings of Alessandrini (1999) and is interpreted as future economic growth thereby narrowing credit spreads.

1.3.2.3 The Long-Run Dynamics

While focusing on credit spreads in first difference eliminates the problem of spurious regression, it also results in loss of information on the long-run interaction of variables, because this sort of econometric procedure has the potential disadvantage of only capturing the short-run dynamics. To overcome this problem, and entail a long-run relation in our model, we test for cointegration between the above-mentioned variables. The implication of a cointegrating vector is, that while the variables are individually I(1), a linear combination of the variables is stationary, resulting in a stationary error structure.

As Appendix B shows, we find a cointegrating relation in the majority of our series. The tendency is very clear as the null hypothesis of no cointegration is more strongly rejected at lower ratings. Furthermore we generally find it more difficult to reject the null as the maturity increases except for the series of AAA-rated bonds. We interpret the cointegrating relation as an equilibrium expression for the determinants of credit spreads and for AA-rated bonds we find the following cointegrating equations

\[
CS_{AA_{ST}} = -0.117\text{Level ST} - 0.145\text{Slope}
\]
\[
CS_{AA_{INT}} = -0.129\text{Level ST} - 0.115\text{Slope}
\]
\[
CS_{AA_{LT}} = -0.179\text{Level LT} - 0.062\text{Slope}
\]

The short-term AA credit-spread relation suggests that a 100 bp change in the level of Treasuries leads to a 11.7 bp decrease in the spread. This tendency is intensified, as the credit spread is more sensitive to changes in the level of Treasuries as the maturity increases and this applies independently of the chosen rating class. Moreover the relationship is contemporaneously magnified for bonds of lower credit quality. If we focus on the Treasury slope we see that the long-run equilibrium of the credit spreads decreases when the slope of the Treasury term structure increases. In contrast to the level variable this relationship evolves in the opposite direction as the influence from the slope becomes almost negligible at the long maturity. This is a very interesting observation as it indicates that the credit spreads on bonds of long maturities, in equilibrium, are unaffected when the Treasury term structure becomes steeper. We interpret this as a signal of, that the markets assess high uncertainty with the long-run expectations of the future short interest rate, as
the expected improvement in business climate is not reflected in the credit spreads. Once again this result is independent of ratings and in accordance with the findings of Duffee (1996). The lesson from this analysis is the necessity of incorporating the relationship between the Treasury variables and the credit spreads in the factor model derived later.

1.4 Credit Risk Modeling

Building credit risk models as the basis for evaluating default exposure is, as mentioned previously, of fundamental importance to financial economists. The risk profile of corporate bonds is as mentioned very different from the risk profile of ordinary bonds, and investment in these markets therefore requires pricing tools capable of incorporating these aspects. Consistent with this objective, various theoretical researchers continue to shed light on the nature of credit sensitive securities.

As illustrated in the previous section the determination of defaults rates and recoveries is a complex matter. No tractable model can therefore capture all of the underlying aspects so, in practice, all models involve tradeoffs regarding how various aspects of credit risk are captured.

1.4.1 The Theoretical Approaches to the Pricing Problem

The research, in finding tractable analytical ways of modeling the term structure of credit sensitive securities, has in general evolved along two main lines

\[ i) \textit{structural models, and} \]
\[ ii) \textit{intensity-based models.} \]

The structural model approach was developed in the early path-breaking article by Merton (1974) as an application of the contingent claims analysis pioneered by Black and Scholes (1973). Merton’s work initiated a series of theoretical studies all relying on the philosophy that the likelihood of default can be derived endogenously from the variability of the firm’s assets relative to its liabilities.\(^{10}\) The ambition is to model a given process for the evolution of the firm’s assets, and led the default be triggered if the process reaches a given lower boundary. The intuition behind these models is undoubtedly appealing form a theoretical perspective, and provides a rich parameteriza-

\(^{10}\) Hence, the name \textit{structural models}, as they model the evolution of the firm’s capital structure. Occasionally, they are also referred to as \textit{firm-value based models} or \textit{predictable stopping time models}. In Merton’s model default can only occur at maturity of the debt in contrary to the approaches in the barrier option framework. See Black and Cox (1976), Brennan and Schwartz (1980), and Longstaff and Schwartz (1995) among others. Sometimes the latter type is referred to as first passage time models.
Part I: Evidence of Credit Risk

...tion of the price of credit sensitive securities. However, these models involve certain characteristics, which, form a practical point of view, makes them limited for quantitative purpose.

Generally, the lack of success arises from two critical issues. First, the parameters in the model are very difficult to estimate. The difficulties are primarily related to i) modeling the evolution of the value of the firm, ii) the modeling of realistic boundary conditions which triggers the default, and, iii) the capital structure of a firm is generally far too complex to specify the time of default and recovery in the event of default. Estimation therefore requires detailed information about every issuer, but the accessibility of such information is not easily obtained since the majority of the involved quantities are neither tradable nor observable. Secondly, the mathematics underlying these models can be very complex, and the numerical procedures are generally not very tractable.\(^\text{11}\)

The more recently introduced intensity-based models are less complete from a theoretical perspective but more manageable in practice. The approach is entirely different from that of the structural models, as an exogenous process specifies the time of default. The process is usually defined as a one-jump process, which entails the possibility of jumping from no-default to default. The probability of a jump in a given time interval cannot be foretold on the basis of information available today, and is governed by the default intensity based on various state variables.\(^\text{12}\) These variables can in principle be chosen arbitrary and without any functional restrictions, though, in order to derive closed-form solutions various conditions have to be obeyed, and, in an economical perspective the variables must give meaning.

From a theoretical perspective the intensity-based models are less appealing, as they are further from reality than structural models. However, despite the less preferable setup, the approach is indeed suitable for quantitative use, and provides a large degree of freedom for model formulation. Intensity-based models focus on the instantaneous likelihood of default, without conditioning default on the value of the firm. As a result, the approach is simpler to use and possess a crucial empirical advantage by avoiding the parameterization of this value. Furthermore, the approach imply the important feature (probably the most important), that under the equivalent Martingale measure the price of the defaultable claim may be derived by replacing the instantaneous default-free short-rate with a default-adjusted process. As a result, standard term-structure models for default-free debt are directly applicable.

\(^{11}\) Recent articles, which address these problems, are Longstaff and Schwartz (1995) and Zhou (1997). However from an empirical perspective the models perform poorly. The credit spreads implied by the models are consistently lower than the observed market spreads. See Bakshi, Madan and Zhang (2001) and Duffee (1999).
Part I: Evidence of Credit Risk

Despite the attractiveness of the intensity-based models, we emphasize the importance of understanding the underlying philosophy of the structural models as they provide rich information about the concept of credit risk. Thus, it should be noted, that the distinction between the two approaches is not clear-cut, as different authors have proposed various forms of hybrid models mixing the two concepts. Structural models can be extended to cover aspects of the intensity-based approach, by describing the value of the firm as a Jump-diffusion process (see Zhou 1997). Similar the intensity-based models can be extended to cover aspects of the structural approach, by involving the value of the firm as a state variable governing the default intensity (see e.g. Bakshi, Madan and Zhang, 2001).

1.4.2 Intensity-Based Models

Numerous authors have in the past decade applied the intensity-based approach. The critical modeling problem in intensity-based models is how to describe the dynamics of the intensity. To enclose the brief introduction to credit risk modeling, we therefore end this chapter with a quick review on the treatment of the intensity and briefly discuss its applications.

Basically, the models rely on the same fundamental construction that the time of default is modeled exogenously as the time of the first jump in a Poisson process. We recognize, from the properties of the Poisson process, that the probability of no jump in a given time interval \([0;T]\) is

\[
P(N_\tau = 0) = \exp\left(-\int_0^T l(s) ds\right)
\]

assuming that \(N_0\) is zero. \(P(\cdot)\) is then the probability that default will not occur over the entire life of the bond, and can be interpreted as the survival probability from time zero to \(T\). Dropping the technical details for a moment, merely focusing on the underlying intuition, the ambition is to specify the intensity \(l(t)\) in accordance with the observed properties of credit spreads. Apart, from some technical requirements, the setup allows for rich dynamics in the intensity process, specifically, it enables the modeling of stochastic credit spreads.\(^{14}\)

---

\(^{12}\) The default intensity is in the litterateur also referred to as hazard rate or the instantaneous rate of default.

\(^{13}\) Zhou (1997) proposed a framework where the value process of the firm is assumed to follow a jump diffusion of the form: 

\[
dv = (\mu - \lambda v) dt + \sigma dZ + (\Pi - 1) dY .
\]

\(^{14}\) Formally there are three ways of measuring the intensity: (1) in context of an homogeneous Poisson process, i.e. assuming constant intensity, (2) using an in-homogeneous Poisson process where \(l(t)\) is a deterministic function of time, or (3) one can choose to describe the time of default as the first jump of a Poisson process with random intensity, i.e. a Cox process.
Jarrow and Turnbull (1995) can be seen as the pioneers in this particular field.\textsuperscript{15} In their continuous time approach they assume that the time of default $\tau$ is exponentially distributed with parameter $\lambda$, modeled by a homogenous Poisson process. Essentially, this setup is very simple and rather unrealistic for practical use. The main criticism is that the default intensity is independent of any state variable, and that default is always equally probable during the entire lifetime of the bond.

To overcome some of the shortcomings of the Jarrow and Turnbull (1995) model, Jarrow, Lando and Turnbull (1997) proposed a slightly different model, which relates default probabilities to credit ratings. The ratings were used as state variables for the default intensity, and an inhomogeneous continuous time Markov chain, specified in terms of a generator matrix, governed the rating transition through time. This type of model has the advantage of relying on readily observable data, and in addition to the Jarrow and Turnbull model, it eliminates the assumption that default intensities are constant through time. On the other hand the model is limited because it does not allow for stochastic spread dynamics within the rating classes, as it implies that credit spreads can only change whenever a rating transition occurs. This may not reflect reality, as ratings have shown to be quite unreactive to new information.\textsuperscript{16} In addition to this, Lando (1998) proposed a generalization of the model of Jarrow et al. (1997) that allowed for credit-spread changes without simultaneous changes in credit ratings. Lando (1998) also relaxed the independence assumption and extended Jarrow et al.’s model so that default probabilities could depend on the level of interest rates.

Lando (1998) further developed the general Cox-process methodology with Markov state space settings in line with Duffie (1998), Duffie and Singleton (1997, 1999) among others. Where Duffie and Singleton (1997, 1999) assumed recovery proportional to the pre-defaultable market value of the debt, Lando used the fractional recovery of Treasury formulation. To induce further realism into the modeling technique Madan and Unal (1998) proposed an intensity model with stochastic hazard and recovery rate.

\textsuperscript{15}This class of models includes Duffie, Schroder, and Skidas (1996), Jarrow, Lando, and Turnbull (1997), Lando (1997, 1998), Madan and Unal (1994) among others.
\textsuperscript{16}See Lando (2000).
Part II: The Pricing of Defaultable Bonds

In this part we examine the factor model framework for pricing defaultable bonds, which we evaluate empirically in the following part of this thesis. In chapter 2.1 and 2.2 we present and derive the theoretical framework for the pricing problem. More specifically, since the pricing is both economical and conceptually related to the principles and dynamics of the default-free term structure, the theoretical framework for pricing default-free bonds is evaluated in chapter 2.1. This chapter serves both the presentation of theoretical techniques, and mathematical results, which become very useful in the later work. In chapter 2.2 we derive, with direct application of the results made in 2.1, the hazard rate framework for pricing defaultable bonds, and some useful extensions. In continuation chapter 2.3 and 2.4 implement the framework, specify, and evaluate the model for our present application.

2.1 The Fundamental Framework and Assumptions

This chapter elaborates on various aspects of the term structure theory within the Markovian settings, and presents a consistent continuous time model for the stochastic evolution of bond prices. The examination covers the derivation of the general bond pricing equation, some aspects of the risk-neutral pricing, and the derivation of tractable solutions within the exponential affine setup. These elements serve as the conceptual framework for the later work, and are derived independent of the term-structure model used, and therefore apply in a more general framework than the scope of this thesis. At this point we exclude the possibility of default and merely concentrate on the behavior of the default-free term structure. Later we extend the model to cover the existence of default risk, and utilize the results derived here.

2.1.1 Preliminaries

The standard assumption when modeling theoretical prices is that the efficient market hypothesis (EMH) is upheld. When a capital market is referred to as being efficient, it normally means that the asset prices and returns are determined in equilibrium between supply and demand in a competitive market, peopled by rational risk adverse traders with homogeneous expectations.\(^{17}\) These rational traders rapidly assimilate any information, which in this thesis is described by the instantaneous value of a specific number of state variables, and adjust prices accordingly. Moreover it is assumed that the market trade without any imperfections such as taxes, transaction costs etc. It is,

---

\(^{17}\)This means that all traders are assumed to know the true probability distribution of the stochastic returns at all points in time. In accordance, this entails all traders in the market to unanimously agree that the state variables are characterized by given stochastic processes.
of course, not tenable to believe that all of these assumptions are strictly upheld but it is not within
the scope of this thesis to discuss the validity of these assumptions.

In the context of term-structure modeling the essential building blocks are the default-free, pure
discount bonds. We refer to this bond as a financial claim, which, with certainty pays one unit at
maturity. The price of the pure default-free discount bond is in the following denoted as \( P(t,T) \).
This definition leads to the fundamental relationship between the bond price and the per annum
continuous yield to maturity as

\[
Y(\tau) = \frac{-\log P(t,T)}{\tau}, \quad \text{where } \tau = T - t \quad 0 \leq t < T
\]

and the spot interest rate, \( r(t) \), is defined as the limit of (2-1) as \( \tau \to 0 \).

In addition, the existence of the limit of (2-1) as \( \tau \to 0 \) enables an investment, at any time \( t \),
of one unit of account in default-free deposits and “rolling over” the proceeds until a later time \( s \),
gives a deterministic yield. The value, measured in terms of the yield of a roll-over strategy in the risk-
free money-market / bank account, evolves according to the following ordinary differential equation

\[
dB(t) = r(t)B(t)dt, \quad \text{where } B(0) = 1.
\]

Equation (2-2) is easily solved as\(^\text{18}\)

\[
B(t) = \exp\left(\int_0^t r(s)ds\right), \quad t \in [0,T].
\]

In contrast to other stochastic yields in the bond market, we refer to \( r(t) \) in (2-2) as the instantaneous
risk-free interest rate, and to \( B(t) \) as the market value of the bank account. Moreover, \( D(t,T) \)
denotes the discount factor in the time interval \([t;T]\), and is given as

\[
D(t,T) = \frac{B(t)}{B(T)} = \exp\left(-\int_t^T r(s)ds\right).
\]

which reflects the amount at time \( t \) that is equivalent to one unit payable at time \( T \).

\[^{18}\text{ } dB(t) \frac{d}{dt} B(t) = r(t)dt \Rightarrow \log B(t) = \int r(t)dt \Leftrightarrow B(t) = \exp\left(\int_0^t r(s)ds\right)\]
2.1.2 Markovian Settings and the General Pricing Model

Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a complete probability space, where \(\Omega\) denotes the sample space, the \(\sigma\)-field \(\mathcal{F}\) represents the set of events in which we work, and \(\mathbb{P}\) is the objective probability measure on the sample space \(\Omega\). Moreover \(\mathcal{F}_t = \{ \cdot : 0 \leq t \leq T_{\text{max}} \}\) is the natural filtration, representing the arrival of information through time, such that \(\mathcal{F}_t \subseteq \mathcal{F}_u \subseteq \mathcal{F}_{T_{\text{max}}} \equiv \mathcal{F}\) for all \(t < u\), meaning that the information increases with time and never exceeds the whole set of events. In addition, we use the short hand notation \(\mathbb{E}_t[\cdot]\) for the \(\mathcal{F}_t\)-conditional expectation operator \(\mathbb{E}[\cdot | \mathcal{F}_t]\), except when conditioning on other sub-filtrations, and we explicitly indicate when the expectations are taken under other measures than \(\mathbb{P}\).

Based on the above probability space we assume that the dynamics of the \(n\)-dimensional vector of state variables \(x_t' = [x(t)_1, x(t)_2, \ldots, x(t)_n]\) are governed according to the traditional time-invariant stochastic differential equation (SDE), of the form\(^\text{19}\)

\[
\begin{align*}
    dx_t &= u(x_t) dt + \Sigma(x_t) dw_t \\
    &\equiv u dt + \Sigma dw_t.
\end{align*}
\]

(2-5)

The infinitesimal increment \(dx_t\) on the left hand side consists of a deterministic drift term \(udt\) and a stochastic noise term \(\Sigma dw_t\), where \(u\) denotes the \(n\)-dimensional instantaneous drift vector, \(\Sigma\) is the \(n \times n\) diagonal diffusion matrix, and \(w_t\) denotes the \(n\)-dimensional vector of correlated Wiener processes (i.e. \(\text{cov}(w_t) = \Omega_t\)).\(^\text{20}\)

The design of (2-5) is general in the sense, that there are no restrictions on the specification of the functional forms of \(u\) and \(\Sigma\). It allows bond prices to depend on an arbitrary number of state variables following arbitrary specified Ito processes. Of course, to develop an explicit bond pricing model, the functional form of the parameters involved must be theoretically or empirically specified (more on this later). In essence this setup implies, that the dynamics of the yield curve are

---

\(^{19}\)This SDE is referred to as an \(n\)-dimensional Ito process. Formally, the differential equation in (2-5) has no independent meaning. It is only a shorthand version of the stochastic integral

\[
X(t) = X(s) + \int_s^t u(X(u), u) du + \int_s^t \sigma(X(u), u) dw(u).
\]

\(^{20}\) We use the assumption of correlated Wiener processes for ease of exposition, while it is general enough for our present needs. \(dw\) is a random number drawn from a Normal distribution with mean zero and standard derivation \(dt^{1/2}\) so that \(\text{E}[dw] = 0\) and \(\text{E}[dw^2] = dt\), and \(\text{E}[dw_t dw_u] = \rho dt\). It is, though, sometimes convenient to apply models based upon standard independent Wiener processes. The choice is just a question of which term includes the information, but in fact the system represented by standard independent Wiener processes is slightly more general as it can be defined with non-quadratic dimensions, \(m > n\), which enables the opportunity of having more independent Wiener processes than state variables.
Part II: The Pricing of Defaultable Bonds

explained by a finite set of exogenous state variables, which describe all the uncertainty and evolution of all information in the economy. Stated differently, the prices are endogenously determined using the exogenous state variables.

Consequently, the price of any default-free zero-coupon bond $P(t, T)$ can be described as a function of the $n$ state variables, the current time $t$, and maturity $T$. Applying Ito’s Formula, showing that $P(t, T)$ is sufficiently differentiable and using stochastic variability and covariability rules, it follows that the process for the bond price is likewise an Ito process, given by

$$dP_t = P(t, T)u_p(t, T)dt + P(t, T)\sigma_p(t, T)dw_t,$$  \hspace{1cm} (2-6)

where the drift and diffusion coefficients $u_p(t, T)$ and $\sigma_p(t, T)$ are defined as

$$u_p(t, T) P(t, T) = \frac{\partial P}{\partial t} + \frac{\partial P'}{\partial x} u + \frac{1}{2} tr \Sigma \Sigma' \frac{\partial^2 P}{\partial x \partial x'}$$ \hspace{1cm} (2-7a)

$$\sigma_p(t, T) P(t, T) = \frac{\partial P'}{\partial x} \Sigma$$ \hspace{1cm} (2-7b)

for all maturity dates.

To derive the general finit-dimensional parabolic PDE for bonds and other interest rate contingent claims, we go along the classical indirect / absence of arbitrage approach initially presented by Vasicek (1977).\(^{22}\) The underlying assumption of this approach is that, at any time, the values of the set of exogenous specified state variables fully characterize the state of nature at that particular point in time. In contrast to the complete market model derived by Black and Scholes (1973) for option contracts, the Markov model derived here is not complete in the sense that it, for every state variable, induces time and state-dependent risk premiums unspecified by the no-arbitrage argument.\(^{23}\)

\(^{21}\)If A and B are symmetric square $m \times m$ matrices, trace (tr) are a convenient way to write the sum of the products of all elements in A and B:

$$tr(AB) = \sum_{i=1}^{m} \sum_{j=1}^{m} A_{ij}B_{ij}.$$

\(^{22}\)In addition to the indirect approach, the direct approach models the prices of the individual bonds directly as state variables. The advantage of this approach, as noted by Brennan and Schwartz (1979), is that the risk premiums are determined endogenously in the model, as the state variables are traded assets themselves.

\(^{23}\)In order to determine these premiums, we have to incorporate more structure into the model. For example, as Cox, Ingersoll and Ross (1985), to embed the model in a general equilibrium setup hereby enclosing the correspondence to the utility function of the representative investor. Cox, Ingersoll and Ross (1985) derived the PDE in a general equilibrium model by solving a representative investor’s intertemporal consumption and investment decision problem. It is then possible to specify the risk premiums, but at the cost of assumptions concerning the investor’s utility function – see Miltersen (1993), pp. 14-15.
Part II: The Pricing of Defaultable Bonds

As our independent state vector \( x_t \) consist of non-traded assets, we have to construct a portfolio of \( n + 1 \) bonds of different maturities in order to hedge the underlying uncertainty. Recall that in the Black-Scholes equation we perfectly hedge the randomness by constructing a portfolio of a long position in the option and a short position in the underlying. The reason why the portfolio is risk-free and risk premiums are excluded is that both the option and the underlying asset are affected by the same source of uncertainty. However, the observation made when dealing with bonds is, that the underlying randomness, i.e. the yield \( r \), is not a traded asset. Essentially, this is why we hedge the randomness using other bonds of different maturities.

The infinitesimal yield on the portfolio is then given as

\[
d\Pi = \omega' dP/P = \omega' u^k_r dt + \omega' \sigma^k_r dw
\]

where \( \omega \) is a \( k \)-dimensional vector containing the weights, such that \( \omega' = 1 \) and \( k = 1, 2, \ldots, n + 1 \) bonds. \( u^k_r \) is the \( k \) dimensional drift vector and \( \sigma^k_r \) is the \( k \times n \) dimensional matrix of diffusions for the \( k \) bonds. Furthermore \( dP/P \) defines the vector with elements \( \{dP_i/P_i\} \).

By constructing the portfolio \( \Pi \) of \( k \) bonds, it is possible to apply a strategy (assuming unrestricted short sales) such that the yield on the portfolio is unaffected by the \( n \) stochastic elements – i.e. the second term of (2-8) equals zero \((\omega' \sigma^k_r dw = 0)\). All sources of uncertainty are hereby eliminated from the portfolio value, making the infinitesimal yield deterministic (this is possible due to the fact that there are only \( n \) sources of risk but \( k \) bonds in the portfolio). Absence of arbitrage then implies that the portfolio must earn the risk-free rate. Defining \( r_i = r; i \) we get

\[
\omega' u^k_r dt = \omega' r dt \iff \omega'(u^k_r - r) = 0 \iff \omega' q = 0
\]

Combining these results we have the following no-arbitrage conditions
\[ \Lambda \omega = 0, \quad \text{where } \Lambda = \begin{bmatrix} q' \\ \sigma_p' \end{bmatrix} \]  

(2-10)

which is equivalent to

\[
\begin{bmatrix}
\mu_{p,1}(T_i) - r & \cdots & \mu_{p,k}(T_i) - r \\
\sigma_{p,1}(T_i) & \cdots & \sigma_{p,k}(T_i) \\
\vdots & \ddots & \vdots \\
\sigma_{p,n}(T_i) & \cdots & \sigma_{p,n}(T_k)
\end{bmatrix}
\begin{bmatrix}
\omega_1 \\
\omega_2 \\
\vdots \\
\omega_k
\end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}
\]

Provided that the matrix on the left hand side (\( \Lambda \)) is singular, this homogeneous system of equations provides non-trivial solutions. Conversely, if the matrix is non-singular – i.e. the inverse relationship exists – the unique solution is the null-vector – implying that, in order to make the portfolio risk-free we should buy neither of the bonds. Since at least one element of the vector \( \omega \) is non-zero, the rows of \( \Lambda \) must be linear dependent.

If we then write \( q \) as a linear combination of \( \sigma_{p}^{k'} \), we get the following relationship for the \( k' \)th bond

\[ u_p - r = \sigma_{p}^{k'} \lambda \]  

(2-11)

where \( \lambda \) is an \( n \)-dimensional column vector containing the information of the risk induced by the \( n \) state variables. Since the portfolio of bonds is chosen arbitrary, this relation should hold for any bond, hence, the vector could be chosen independent of the specific choice of maturity. In (2-11) \( \lambda_0 \) may be interpreted as the amount of excess return the investors get in equilibrium per unit of risk induced by the \( i \)th state variable, hence, it is called the market price of risk. Recall, that the infinitesimal yield on the money-market account is deterministic, in contrast to an investment in a zero-coupon bond. Consequently an investor, in return for taking on the extra risk, requires a risk premium.\(^{24}\)

Substituting the expressions for \( \mu_p \) and \( \sigma_p' \) from equation (2-7) into equation (2-11) and rearranging, we obtain the fundamental bond pricing equation\(^{25}\)

---

\(^{24}\)Note, because of the presence of risk premiums the equilibrium prices depend on investor’s preferences. In contrast the Black-Scholes equation is derived independent of these preferences.

\(^{25}\)In the equation \( \text{tr} \) (trace) is taken over the entire matrix product \( \Sigma \Omega \Sigma H \), where \( H \) is the Hessian matrix.
This equation must hold for any asset which value depends solely on the state variable vector \( x(t) \) and the time to maturity \( \tau \), so in order to derive equilibrium prices, \( P(t,T) \), must satisfy the PDE above. Notice, that there are at no time in the derivation of (2-12) used contract specific conditions. This means that the PDE should hold for any solely interest rate dependent assets, with an appropriate boundary condition. In addition to our specific objective, the PDE must be solved subject to \( P(T,T)=1 \), which is the boundary condition for zero-coupon bonds.

Through a simple rearranging of (2-12) we are able to imply a more intuitive and economical interpretation of the fundamental price equation

\[
\frac{\partial P}{\partial t} + \frac{1}{2} tr \Sigma \Sigma \frac{\partial^2 P}{\partial x \partial x'} + \frac{\partial P'}{\partial x} \left[ u - \Sigma \lambda \right] - rP = 0 \tag{2-13}
\]

where the left hand side of (2-13) is, from Ito’s Formula, the expected price change for the bond. Hence, this relation discloses the expected rate of return on the bond as

\[
r + \frac{1}{P} \frac{\partial P'}{\partial x} \Sigma \lambda \tag{2-14}
\]

and so the return consists of the risk-free rate and a risk premium or excess return. We see that the excess return is proportional to the variance of the bond, where the proportionality factor \( \lambda \) is constant and independent of the bond. Employing an economical intuition the equation reveals the obvious relation to CAPM. Furthermore the equation provides important information in relation to the adequate term-structure model, which is depicted in the necessity of a valid and satisfactory specification of the market price of risk function.

An important difference, compared to the Black-Scholes equation, is as mentioned the presence of risk preferences, as the market price of risk is incorporated into the equation. The prices in equilibrium are therefore not unambiguously determined, but reflect the agents’ attitudes towards risk. It may seem surprising, that the risk premiums do not have any functional relationship with the initial term structure on interest rates. Thus the risk premiums are uniquely related to the uncertainty generated by the stochastic state variables.
Part II: The Pricing of Defaultable Bonds

Turning to the more practical side, there are some drawbacks with applying the factor models described above. For instance to determine the market price of risk an \( n \)-factor model requires \( n \) a priori given bond prices. Not until the market prices of risk are determined is it possible to compute the theoretical prices for the remaining market. In practice one will therefore experience that the theoretical prices differ from the observed prices. In application, this implies that another choice of a priori given bonds leads to different estimates of the market prices of risk and consequently the equilibrium bond prices are therefore only relative prices.

2.1.3 The Equivalent Martingale Measure

From (2-5) we see that in order to compute the price of the bond in this \( n \)-dimensional space, we have to solve the PDE given in (2-12) with \( n \) state variables and for a general case this has to be done by numerical methods. The corresponding probability theory enables another, and sometimes simpler, way of solving the fundamental PDE. The method exploits the theory of risk-neutral pricing in which prices of contingent claims can be solved by discounting the expected payoff at time \( T \) at the risk-free rate.

In the following we present the basic techniques underlying the equivalent Martingale measure and the applicability in relation to the pricing of non-defaultable zero-coupon bonds. Later we utilize the measure to account for the pricing of defaultable bonds.\(^{26}\)

Basically, the application of risk-neutral pricing can be traced back to Black and Scholes' work in option pricing. Applying delta hedging and no arbitrage arguments, the fundamental hedge portfolio would generate the same return no matter what state of the world, and which, in the absence of arbitrage, is the risk-free rate. Despite Black and Scholes early work, the term risk-neutral pricing is usually associated with Cox and Ross (1976) and the later formalization and extension by Harrison and Kreps (1979) and Harrison and Pliska (1981, 1983) through the so-called Martingale measure.

As stressed by Lund (1993) and others, a key result in Harrison and Kreps (1979) is the established connection between the economic concept of absence of arbitrage, and the mathematical property of the existence of an equivalent Martingale measure. They proved that, the existence of an equivalent Martingale measure \( Q \), under technical conditions, is equivalent to the absence of arbitrage, and the uniqueness of the measure is determined by its ability to replicate. Hence, mul-

\(^{26}\)It is beyond the scope of this thesis to justify the technicalities behind this measure. The two key references that rigorously relate derivative prices and results from probability theory are Harrison and Kreps (1979) and Harrison and Pliska (1981).
The results are in general based on the concept of self-financed strategies. A trading strategy, \{\theta_t\}, is said to be self-financing if its value changes only due to changes in the asset prices. Stated differently, except for the final withdraw, no additional cash inflows or outflows occur after the initial time \( t \). This can be expressed as

\[ \theta_t P_s = \theta_0 P_t + \int_0^t \theta_u dP(u), \quad t < s \leq T. \]  

(2-15)

where \( P_t \) is the price process governed by an Ito process and \( \theta_0 P_t \) is the initial value of the portfolio. The trading strategy is said to be an arbitrage if the value at time \( t \) is non-positive while the closing value at time \( T \) is non-negative. In fact, this entails that the cash inflows from the strategy are non-negative in all states of nature i.e., a free lunch. As shown by Duffie (1996a, chapter 6), if the price process is a martingale under the measure \( \mathbb{Q} \), arbitrage is precluded. The stochastic integral on the right hand side of equation (2-15), usually denoted as the gain process \( G(t,s) \), will then equals zero. The self-financing condition therefore implies that

\[ \theta_t P_T = \mathbb{E}_t^\mathbb{Q} \left[ \theta_0 P_T - \int_0^T \theta_u dP(u) \right] = \mathbb{E}_t^\mathbb{Q} \left[ \theta_0 P_T \right] \]  

(2-16)

By this, it appears that a non-negative closing portfolio value can only occur if, and only if, the initial value is non-negative, which provides that arbitrage opportunities cannot exist.

Whether or not arbitrage opportunities exist depend on the stochastic process governing \( P_t \). But, instead of focusing on the price in the context of an Ito process, we consider the deflated price process. To be more specific, we are interested in the probability measure under which the deflated price processes are martingales. This is accomplished by introducing the normalized/deflated price process \( Z_t \) defined as

\[ Z_t = \frac{P_t}{B_t}. \]  

(2-17)

---

27Roughly speaking, absence of arbitrage is equivalent to the impossibility of investing zero today and receiving a non-negative amount with positive probability tomorrow.

28\( \theta_t \) and \( P_T \) are stochastic variables at time \( t \).
where $B_t$ is the non-decreasing deflator defined by the money-market account. From Duffie (1996a) the numeraire inversion theorem implies that the no-arbitrage property of $P_t$ will only hold if it holds for $Z_t$.

Using Ito’s Formula on (2-17) together with (2-2) and (2-6) (the definitions of the process for the money-market account and the bond price, one-dimensional) we get

$$
\frac{dZ_t}{Z_t} = \frac{1}{P_t} dP_t - \frac{1}{B_t} dB_t
= (\mu - r_t) dt + \sigma dW_t.
$$

(2-18)

This enables us to write the diffusion coefficient as a linear combination of the drift coefficients by introducing $\eta_t$

$$
\sigma \eta_t = \mu - r_t, \quad t \in [0, T].
$$

(2-19)

In the $n$-dimensional setting this amounts to solving a set of linear equations, and hence we assume that $dZ_t/Z_t$ is reducible, that is $\sigma^{-1}$ exists. Rearranging in terms of $\eta$ we see, that the solution is identical to the definition of the market price of risk in section 2.1.2. We therefore denote the solution by $\lambda$ instead of $\eta$. We then apply Girsanov’s Theorem, which states that a change in the Brownian motion process (which changes the future probability distribution for the asset) allows us to adjust the drift of the diffusion process to almost any desired level. Girsanov’s Theorem hereby changes the mean of the SDE by changing the probability measure but leaves the volatility structure intact. We do this by using the standard Brownian motion $W^Q$, which is a martingale under $Q$ and defined as $W^Q_t = W_t + \int_0^t \lambda(s) ds$, such that

$$
dW^Q_t = \lambda_t dt + dW_t.
$$

(2-20)

We hereby see that the process for $dZ/Z$, under the equivalent Martingale measure, is a martingale defined by

---

29Given the processes for the state variables under the objective probability measure, Girsanov’s Theorem defines the processes of the variables under the equivalent Martingale measure. Under technical assumptions (Novikov condition is satisfied, and the Radon-Nikodym derivative has finite variance), which are beyond the scope of this thesis, the equivalent Martingale measure $Q$ is then defined by the Radon-Nikodym derivative of $Q$ with respect to $P$, $dQ/dP = \xi(P)$. 

33
Part II: The Pricing of Defaultable Bonds

\[
\frac{dZ_t}{Z_t} = \sigma dW^Q_t \tag{2-21}
\]

and by application of Itô’s Formula on \( P_t (P_t = Z_t B_t) \), just substitute (2-20) into (2-6), we see that under \( \mathbb{Q} \) the SDE for \( \frac{dP_t}{P_t} \) is

\[
\frac{dP_t}{P_t} = \left( \mu - \lambda, \sigma \right) dt + \sigma dW^Q_t \\
= rdt + \sigma dW^Q_t \tag{2-22}
\]

To investigate whether this is equivalent to the results found in the previously section, substitute the explicit formulae for \( \mu_p \) and \( \sigma_p \) from equation (2-7) into the first line of (2-22). Taking expectations, knowing it must earn the risk-free rate of return, gives the PDE in (2-12).

We conclude this section by considering the application of the risk-neutral valuation property in addition to zero-coupon bonds. We saw that the absence of arbitrage implies the existence of an equivalent Martingale measure, under which all normalized price processes are martingales. This is equivalent to that the process has the risk-free rate as its rate of return, which in addition means, that the process has the risk-neutral valuation property. To realize this, recall that

\[
P(t,T) = B_t \mathbb{E}^\mathbb{Q}_t \left[ \frac{P_T}{B_T} \right] = \mathbb{E}^\mathbb{Q}_t \left[ D(t,T) P(T,T) \right]
\]

so

\[
P(t,T) = \mathbb{E}^\mathbb{Q}_t \left[ \exp \left( -\int_t^T r(s) ds \right) P(T,T) \right]. \tag{2-23}
\]

And because the risk-free rate is independent of the final payoff \( (P(T,T) = 1) \) we have

\[
P(t,T) = \mathbb{E}^\mathbb{Q}_t \left[ \exp \left( -\int_t^T r(s) ds \right) \right] = \mathbb{E}^\mathbb{Q}_t \left[ D(t,T) \right] \tag{2-24}
\]

which is also known as the probabilistic solution to the fundamental PDE. From a computational point of view, equation (2-23) reveals, that we can price a bond by taking the expected value of the discounted payoff, where the expectations are taken under the measure \( \mathbb{Q} \), and the discounting is made at the risk-free rate. It is important to notice, that the risk-neutral expectation operator is only employed for valuation purpose, i.e. the defined appearance of the state variables under \( \mathbb{Q} \) is
Part II: The Pricing of Defaultable Bonds

a result of arbitrage regularity conditions. The process for the state variables is still given by definition (2-5) under the objective probability measure.

2.1.4 Analytical Solutions in Affine Settings

As the intension of this thesis is to conduct an empirical survey, it is convenient if we obtain a simple closed-form expression for the bond price. Most multifactor models with known analytical solutions belong to the exponential-affine (linear) class, as the PDE simplifies to \( n + 1 \) ordinary differential equations (ODE). In particular, for exponential-affine models, we guess the solution to the PDE to have the following tractable form

\[
P(t,T) = \exp\left( A(\tau) + B(\tau)'x_\tau \right), \quad \tau = T - t. \tag{2-25}
\]

where \( A(\tau) \) is a scalar function and \( B(\tau) \) is a \( n \times 1 \) column vector.

The functions \( A(\tau) \) and \( B(\tau) \) can either be found analytically or numerically by solving a series of ordinary differential equations. Whether analytical closed-form solutions exist or not depend on the stochastic process governing the state variables. Duffie and Kan (1996) proposed a general class of term-structure models that includes the Gaussian and the Cox, Ingersoll and Ross model. We refer to Duffie and Kan (1996) for a thorough discussion of the necessary conditions for (2-25) to be a solution to the PDE, which in summary can be expressed as finite differential and affine regularity conditions. Hereby Duffie and Kan (1996) state, that if the general SDE in (2-5) is a compatible term-structure factor model where the drift and variance functions, besides the interest rates, are affine in the state variables and there is a finite solution to the necessary ordinary differential equations, then (2-25) is an affine solution to the PDE.

Bearing the partial derivatives of (2-25) in mind

\[
\frac{\partial P}{\partial \tau} = -\frac{\partial P}{\partial t} = -\left( -\frac{\partial A(\tau)}{\partial \tau} - \frac{\partial B(\tau)'x_\tau}{\partial \tau} \right) P(t,T) \tag{2-26a}
\]

\[
\frac{\partial P}{\partial x_\tau} = B(\tau)P(t,T) \tag{2-26b}
\]

\[
\frac{\partial^2 P}{\partial x_\tau'\partial x_\tau} = B(\tau)'B(\tau)'P(t,T) \tag{2-26c}
\]

and substituting these expressions into the general PDE in equation (2-12), dividing by \( P(t,T) \) and rearranging in terms of constants and \( x_i \) gives us the general equations to solve. If we exploit the multifactor affine-model setup and apply the terminology in Duffie & Kan (1996) the \( n+1 \) ordi-
nary differential equations, known as the Ricatti equations, are (for intermediate steps see Appendix J)

\[
\frac{\partial A(\tau)}{\partial \tau} = \theta'K'B(\tau) - \sum_{i=1}^{n} \chi_i \alpha_i B(\tau) + \frac{1}{2} \sum_{i=1}^{n} \left[ \Sigma'B(\tau) \right] \chi_i \delta_i \\
\frac{\partial B(\tau)}{\partial \tau} = -K'B(\tau) - \sum_{i=1}^{n} \chi_i \beta_i B(\tau) + \frac{1}{2} \sum_{i=1}^{n} \left[ \Sigma'B(\tau) \right] \beta_i \delta_i .
\]

(2-27a)  
(2-27b)

The reason for imposing an affine structure now becomes clear. PDE’s are difficult to solve, even numerically. But the structure of (2-27) gives the exponential-affine models mathematical tractability.

2.1.5 The Multifactor CIR Model

Given the affine regularity conditions, we now exploit the affine framework and build the general structure of our model for pricing defaultable bonds. We derive the analytical solutions to the multifactor CIR model, in which we assume that the instantaneous rate is expressed as a sum of \( n \) independent square root processes. We see that this particular setup is very attractive in the context of empirical studies, as it provides reasonable tradeoffs between computational tractability and economical realism.

The convenience of the assumption of independence is easily seen if we consider the expression of the bond price

\[
P(t, T) = \mathbb{E}^Q_t \left[ \exp \left\{ -\int_t^T \left( \sum_{i=1}^{n} \chi_i(s) \right) ds \right\} \right].
\]

(2-28)

As shown by Duffie (1996b), if the state variables are \( \mathbb{Q} \)-independent, we may change the order of integration and summation. Equation (2-28) can therefore be rewritten as

\[
P(t, T) = \mathbb{E}^Q_t \left[ \exp \left\{ -\sum_{i=1}^{n} \left( \int_t^T \chi_i(s) ds \right) \right\} \right] \\
= \mathbb{E}^Q_t \left[ \prod_{i=1}^{n} \exp \left\{ -\int_t^T \chi_i(s) ds \right\} \right] \\
= \prod_{i=1}^{n} \mathbb{E}^Q_t \left[ \exp \left\{ -\int_t^T \chi_i(s) ds \right\} \right]
\]

(2-29)

where we in the last line once again exploit the assumption of independence, and move the expectation operator inside the product operator. This fact is very convenient and shows that the bond
price equation decomposes into the product of $n$ independent “one-factor” models, resulting in writing (2-25) as

$$P(t,T) = \exp \left( \sum_{i=1}^{n} A_i(\tau) + \sum_{i=1}^{n} B_i(\tau)\gamma_i \right)$$

$$= \prod_{i=1}^{n} \exp \left( A_i(\tau) + B_i(\tau)\gamma_i \right)$$

(2-30)

where every $A_i$ and $B_i$ are derived the exact same way. This is a very tractable result indeed. Now we only have to focus on deriving $n$ exactly alike solutions to the ODE’s, which decomposes into deriving the ODE’s for only one state variable. This is what Cox, Ingersoll & Ross did in 1985, and this setting provides the attractive feature of allowing the possibility of giving the factors an economical interpretation, which are of particular interest in our present area of application.

### 2.1.5.1 Deriving the ODE’s and Analytical Solutions in the CIR Setup

Analogous with the arguments stated in section 2.1.2 we can derive the PDE for interest sensitive securities, when the state variables are assumed governed by the CIR model

$$dx_t = \kappa (\theta - x_t) dt + \sigma \sqrt{x_t} dw_t$$

(2-31)

where (2-31) illustrates the dynamics of the state variables under the physical probability measure and $\kappa$, $\theta$, and $\sigma$ are scalars and time-invariant. For $\kappa, \theta > 0$, this corresponds to a continuous time first-order autoregressive process where the randomly moving state variable is elastically pulled toward the unconditional mean, $\theta$. The parameter $\kappa$ determines the speed of adjustment also known as the mean-reversion parameter.

As mentioned above we assume that the state variables are independent under the equivalent Martingale measure, $\mathbb{E}^Q(dw_t dw_{\tau}) = 0$, since it implies closed-form solutions for zero-coupon bond prices. The dynamics under the $\mathbb{Q}$-measure are given according to Girsanov’s theorem (cf. section 2.1.3), and entail that the generic process for the state variables is rewritten as

$$dx_t = \left[ \kappa (\phi - x_t) - \sigma \sqrt{x_t} \lambda(x) \right] dt + \sigma \sqrt{x_t} dw^0_t$$

$$= \left[ \kappa \theta - \left( \kappa + \lambda(x)\sigma / \sqrt{x_t} \right)x_t \right] dt + \sigma \sqrt{x_t} dw^0_t$$

This is an important result noted in Chen & Scott (1993) and Duffie & Kan (1996).
such that, if the risk premium factor, \( \lambda(x) \), is defined as \( \lambda x^{\alpha}/\sigma \) we can write the state variable dynamics as\(^{31}\)

\[
dx_t = (\kappa \theta - (\kappa + \lambda) x_t) dt + \sigma \sqrt{x_t} dw^Q_t.
\]

Now the PDE in the single-factor CIR framework amounts to

\[
\frac{\partial p}{\partial t} + \frac{1}{2} \sigma^2 x \frac{\partial^2 p}{\partial x^2} + \left[ \kappa \theta - (\kappa + \lambda) x \right] \frac{\partial p}{\partial x} - xp = 0
\]

(2-33)

with the boundary conditions \( P(T,T) = 1 \). Suggesting that the solution to (2-33) is of the general affine form, we can obtain the Ricatti equations. Inserting the partial derivatives in (2-33) and dividing by \( P \) yields

\[
\frac{1}{2} \sigma^2 x B(\tau)^2 + \left[ \kappa \theta - (\kappa + \lambda) x \right] B(\tau) + \left( - \frac{\partial A(\tau)}{\partial \tau} - \frac{\partial B(\tau)}{\partial \tau} x \right) - x = 0
\]

(2-34)

Rearranging in terms of \( x \), and noting that for the solution to be valid for every \( x \), following two ordinary differential equations arise

\[
\frac{\partial A(\tau)}{\partial \tau} = \kappa \theta B(\tau)
\]

(2-35a)

\[
\frac{\partial B(\tau)}{\partial \tau} = -B(\tau)(\kappa + \lambda) + \frac{1}{2} \sigma^2 B(\tau)^2 - 1
\]

(2-35b)

These equations could also be derived directly from (2-27) bearing the restrictions of the one-factor CIR process in mind (\( \alpha = \delta_0 = 0 \), \( \beta = \delta = 1 \), \( \Sigma = \sigma \) - see Appendix J).

Now our initial interest is to find the solution to (2-35b). Inserting this into (2-35a) ends the derivation of \( A(\tau) \) and \( B(\tau) \) (for intermediate steps of the derivation the reader is referred to Appendix K).

Recognizing that (2-35b) is a separable differential equation, manipulating, bearing the boundary condition \( B(0) = 0 \) in mind, renders the solution

\[^{31}\text{The choice of lambda ensures that the drift function under Q is linear in the state variable as under P.}\]
Part II: The Pricing of Defaultable Bonds

\[
B(\tau) = \frac{-2(1 - e^{-\gamma \tau})}{2\gamma + (\kappa + \lambda - \gamma)(1 - e^{-\gamma \tau})} = \frac{-2(e^{\tau \gamma} - 1)}{2\gamma + (\kappa + \lambda + \gamma)(e^{\tau \gamma} - 1)} \tag{2-36}
\]

This is the first part of equation 23 in Cox, Ingersoll & Ross (1985). There is an obvious difference between (2-36) and equation 23 in CIR, which can be traced back to the affine form of the bond price. In this thesis we follow (2-25) where CIR in contrast apply a minus in front of \( B \) in the bond price equation.

In the derivation of \( A(\tau) \) it is convenient to define the denominator of (2-36) as \( h(\tau) \) and rewriting the equation. After some manipulation and the utilization of the boundary condition \( A(0) = 0 \) the expression for \( A(\tau) \) is

\[
A(\tau) = \frac{2\kappa \theta}{\sigma^2} \log \left[ \frac{2\gamma e^{(\kappa + \lambda - \gamma)\tau}}{2\gamma + (\kappa + \lambda - \gamma)(1 - e^{-\gamma \tau})} \right]
= \frac{2\kappa \theta}{\sigma^2} \log \left[ \frac{2\gamma e^{(\kappa + \lambda + \gamma)\tau}}{2\gamma + (\kappa + \lambda + \gamma)(e^{\tau \gamma} - 1)} \right]. \tag{2-37}
\]

The price of the discount bond in the multifactor CIR model is then obtained by substituting the expressions of \( A(\tau) \) and \( B(\tau) \) into (2-30). The full expression for the price is then given by

\[
P(t, T) = \prod_{i=1}^{n} \left[ \frac{2\gamma e^{(\kappa_i + \lambda_i + \gamma_i)T}}{2\gamma + (\kappa_i + \lambda_i + \gamma_i)(e^{\tau \gamma} - 1)} \right]^{2\kappa_i \theta \sigma_i^2 \tau} \exp \left( \frac{-2(e^{\gamma \tau} - 1)}{2\gamma + (\kappa_i + \lambda_i + \gamma_i)(e^{\tau \gamma} - 1)} \right) \tag{2-38}
\]

or equivalently expressed in terms of yields

\[
Y(\tau) = \frac{-\log P(t, T)}{\tau} = -\frac{\sum_{i=1}^{n} A_i(\tau) + \sum_{i=1}^{n} B_i(\tau) \lambda_i}{\tau}, \quad \tau = T - t. \tag{2-39}
\]

Later we evaluate the yield curves produced in this setting, and address the question of why/how it can be used to describe both interest rate as well as credit-spread dynamics. In order to avoid repetition, we do not consider these aspects before Section 2.4.
2.2 Deriving the Generic Hazard Rate Model

When faced with the probability that the issuing party may fail to meet its obligations on the promised payments, one has to account for the uncertainty surrounding the future cash flows. In this chapter we review the fundamental methodologies in the area of intensity-based modeling, and derive the model for pricing defaultable bonds. Special attention is devoted to the technical issues of the role of conditioning information in computations involving random times. The unexpected jump representing the time of default is assumed triggered by the first jump in a Poisson process with stochastic intensity (Cox-process), by making the intensity a general function of the state variables describing the economy. Therefore our approach in modeling survival is similar to that of Duffie and Singleton (1999), and Lando (1998).

Before getting into the details about the derivation we recall, that a bond is fundamentally a stream of cash flows consisting typically of a promised face value, $H$, paid at maturity and a stream of coupon payments paid in the interim. The difference between the default-free discount bond and the defaultable bond lie in the existence of a time where default may occur, i.e. the sample space, $\Omega$, now contains states representing defaults. Although it is a matter of judgment when a bond has a significant exposure to default, and that it can be difficult to isolate this effect from other effects, there is a clear difference in the probabilistic definition. Given the probability of full payment is below one, the bond is exposed to credit risk. Essentially, this probability should at any time reflect the issuer’s general capability of fulfilling its obligations to the bondholders. In our context the default may be represented by a unit step function $x(t) = 1_{\{Z \leq t\}}$, normally referred to as an indicator function, with the random default time $Z$ such that

$$x(t) = \begin{cases} 1 & \text{If } Z \leq t \\ 0 & \text{Otherwise} \end{cases}$$

The function $x(t)$ represent the status of the bond, and takes the value 0 if no default and 1 if defaulted. In the case of default the remaining cash flows from the defaultable bond amounts to the single recovery, $\psi(Z)$, generally far below the remaining promised payments. The price of the defaultable bond $\tilde{P}(t,T)$ can now be defined by these entities.

If we assume the market trade without any imperfections (except credit risk) and absence of arbitrage is ensured by the existence of an equivalent Martingale measure $Q$, then the price of the defaultable zero-coupon bond is given as
Part II: The Pricing of Defaultable Bonds

\[
\tilde{P}(t,T) = \mathbb{E}_T^\mathbb{Q} \left[ \frac{B(t)}{B(T)} (1 - x(T)) H + \int_t^T \frac{B(t)}{B(u)} (1 - x(u)) \psi(u) dx(u) \right]
\]  

(2-41)

where \(x\) and \(\psi\) are the default and recovery processes respectively.

The first term of (2-41) accounts for the promised face value payment given no default, and the second term accounts for the single recovery at the default time, where \(dx(u)\) is unity for at most one time \(u\), so there is at most one default payout.

The difficulties in evaluating (2-41) in contrast to ordinary default-free bonds is how to address the specification of the discontinuous random process \(x(t)\) and the recovery function \(\psi(t)\). The expression of the price of the defaultable bond is very difficult to compute, and will even for very simple specifications of the recovery require numerical solutions. In the following we i) show how to compute default times i.e. address the specification of the unit step function \(x(t)\). Basically we are interested in construction a tractable expression for the conditional probability of no default prior to maturity, \(\mathbb{Q}(Z > T | Z > t)\), and the conditional probability of default before maturity of the bond, \(\mathbb{Q}(Z \leq T | Z > t)\), conditional on the bond has not defaulted up to time \(t\), ii) elaborate on the functional form of the recovery function and how to address this specific issue, and iii) eventually extend the model to account for coupon payments.

But before getting into technical details and arbitrage requirements consider first the intuition of the survival and hazard function under the physical probability measure, \(\mathbb{P}\).

2.2.1 The Survival and Hazard Function

As mentioned, we are employing a stopping time to characterize default time. More precisely we are working with an intensity-based model which entails modeling the default time as a surprise stopping time. This triggers the question about how to model the likelihood of a default within a short time interval, given that the bond has not defaulted yet. In order to address this question, consider the random variable \(Z\) defining the default time. Suppose it has a continuous density function \(f(t)\), the cumulative probability is then

\[
F(t) = \int_0^t f(s)ds = \mathbb{P}(Z \leq t)
\]

such that the probability of no event occurring within the time length \(t\), is given by the survival function.
Part II: The Pricing of Defaultable Bonds

\[ S(t) = 1 - F(t) = \mathbb{P}(Z > t). \]

We can now formally define the probability of default within the next short time interval \( dt \), conditional on survival up to time \( t \)

\[ \mathbb{P}(t \leq Z \leq t + dt \mid Z > t). \]

A useful way of characterizing this aspect is through the hazard function (or intensity function). In order to establish the general relationship between the probability of no default, default in a given time interval and the hazard rate process we derive the survival function in our modeling framework. Following Greene (2000) the hazard function is formally defined as

\[
h(t) = \lim_{dt \to 0} \frac{\mathbb{P}(t \leq Z \leq t + dt \mid Z \geq t)}{dt} \quad (2-42)
\]

which is a strictly positive conditional process, interpreted as the instantaneous likelihood of default. Intuitively, given the firm has survived up to time \( t \), the probability of a default occurring in the next short time interval \([t; t+dt]\) is given as \( h(t)dt + o(dt) \). By the law of conditional probability, (2-42) can be written as

\[
h(t) = \lim_{dt \to 0} \frac{\mathbb{P}(t \leq Z \leq t + dt \cap Z \geq t)}{dt} / \mathbb{P}(Z \geq t)
\]

\[
= \lim_{dt \to 0} \frac{\mathbb{P}(t \leq Z \leq t + dt \cap Z \geq t)}{dt} \frac{1}{\mathbb{P}(Z \geq t)}
\]

Because the common values of \( t \leq Z \leq t + dt \) and \( Z \geq t \) is just \( t \leq Z \leq t + dt \) we can write the above as

\[
h(t) = \lim_{dt \to 0} \frac{\mathbb{P}(t \leq Z \leq t + dt)}{dt} \frac{1}{\mathbb{P}(Z \geq t)} \quad (2-43)
\]

This expression can be further manipulated if we recognize that the random default time \( Z \) is distributed with the cumulative distribution function \( F(t) \) and probability density function \( f(t) \)

\[ \text{It is commonly referred to as the forward default probability see Litterman and Ibel (1991).} \]
Part II: The Pricing of Defaultable Bonds

\[ h(t) = \lim_{dt \to 0} \frac{F(t + dt) - F(t)}{dt} \frac{1}{1 - F(t)} \]

\[ = \frac{\partial F(t)}{\partial t} \frac{1}{1 - F(t)} = \frac{f(t)}{1 - F(t)} \quad \text{or} \quad \frac{f(t)}{S(t)} \]  

(2-44)

where \( S(t) \) is the survival function as defined above. It now appears that the hazard is a function of time, the density \( f(t) \), and the survival function \( S(t) \). Furthermore the last fraction of (2-44) reveals an intuitive relation between the hazard and survival function because an increase in the cumulative probability of survival \( S(t) \) is exactly offset by a decrease in the cumulative probability of default \( F(t) \). This indicates how to derive an expression of \( S(t) \) as a function of the hazard rate. Utilizing a property of the logarithm, \( \ln(x)' = x'/x \), we differentiate \( \ln S(t) \) with respect to \( t \), which gives the useful result

\[
\frac{d \ln S(t)}{dt} = \frac{S'(t)}{S(t)} = -\frac{f(t)}{S(t)} = -h(t)
\]

(2-45)

As stressed by Greene (2000), this equation explains the relationship between the survival function and the hazard rate. Integrating both sides of (2-45) and taking exponentials yields

\[
S(t) = \exp \int_0^t \frac{d \ln S(s)}{ds} ds = \exp \left( -\int_0^t h(s) ds \right),
\]

(2-46)

where the term in brackets is referred to as the integrated hazard function. Equation (2-46) reveals a very tractable result, as we are now able to express the survival function as a discount rate implied by the hazard function. And, intuitively it states that in order to survive from time \( 0 \) to time \( t \) the bond has to survive in each little time interval \( s+ds \) for all \( s \in [0;t] \), conditional on its survival until \( s \). In addition, the cumulative distribution function of default is obviously also an implied function of the hazard and given as

\[
F(t) = 1 - S(t) = 1 - \exp \left( -\int_0^t h(s) ds \right)
\]

\[
= \int_0^t h(s) \exp \left( -\int_0^s h(s) ds \right) ds
\]

(2-47)
where we in the last expression specify the cumulative distribution function as the integrated probability density function.33

Equation (2-46) and (2-47) are the key relations in understanding the principles of the hazard rate methodology. In fact various forms of hazard rate formulations, analyses of survival times or duration data exploit these two relations in a wide scope of applications. The hazard rate is often more interesting than the survival rate or the cumulative distribution function. Rather than specifying the survival function, it is desirable to specify the hazard rate and then integrate “backward” to obtain the cumulative distribution function. Note that being a probability the hazard function cannot take on negative values. Therefore the function governing the hazard rate typically has to contain non-negative characteristics.

The convenience of the formulation is moreover obvious, as it enables us to describe the evolution of the probability of an event occurring without defining the exact event. For the present purpose, we wish to model the probability of default of the bond across maturity. Considering (2-46) it should then be noted that if $h_t$ is assumed constant, $z$ (the arrival of default) is simply the date of the first jump of a standard Poisson process. Pricing defaultable bonds in this setting has the intuitive construction that in a given time interval, (2-46) and (2-47), is the probability of no arrival and exactly one arrival, respectively. However, as discussed earlier, any parameterization of the hazard function must, of course, take into account the fact that default is not equally probable during the lifetime of the bond. Although a constant hazard rate makes the calculations very simple it is of poor resemblance with reality, and the implementation of a more general model comes with little extra complexity. We therefore choose a hazard model in the Cox-process framework, as proposed by Lando (1998), where the hazard/intensity is allowed to be an arbitrary function of the state variables $h_t = h(X_t)$.34

---

33 The equivalence between the first and second line in (2-47) is obtained if we recall the fact that $f(t) = \frac{\partial F(t)}{\partial t} = -\frac{\partial S(t)}{\partial t} = h(s)\exp\left(-\int_0^t h(s)\,ds\right)$. Calculating the indefinite integral of $f$ and employing the appropriate boundary condition that $F(0) = 0$, i.e. that at time zero, no default has occurred, $S(0) = 1$, we find that $F(t) = -S(t) + c = -\exp\left(-\int_0^t h(s)\,ds\right) + c = 1 - S(t)$.

34 A process $N$ with intensity allowed being random defined as $h_t = h(X_t)_{t \geq 0}$ is called a Cox-process, and is therefore a generalization of the usual Poisson process, with the restriction that conditional on the realization of $h$, $N$ is an inhomogeneous Poisson process. See Lando (1998) for thoroughly discussion.
2.2.2 Constructing Default Times

Utilizing the framework above we are now capable of addressing the question of how to model the properties of the unit step function under the equivalent Martingale measure, as this enables us to write the price equation in a more tractable way. Again our objective is to construct the link between the unit step function and the hazard rate where the hazard rate is adapted to the subfiltration of continuous evolving information \( \mathcal{G}_t \). The derivation are however somewhat more technical under the pricing measure, as it must be done in a way that satisfies the necessary arbitrage conditions. Stated otherwise, our objective is to provide the martingale characterization of \( h \).

Lando (1998) defines the information set we work in as \( \mathcal{F}_t \) considering two subfiltrations \( \mathcal{F}_t = \mathcal{G}_t \cup \mathcal{H}_t \), where \( \mathcal{G}_t = \sigma \{ X_t; t \in [0;T]\} \) is the filtration generated by the Markov diffusion processes describing the instantaneous interest rate and the hazard rate of default, and \( \mathcal{H}_t = \sigma \{ 1_{\{ z \leq T \} }; z \in [0;T] \} \) holds the information of whether a default has occurred at time \( t \). This definition entails that the information set of the state variables is independent of the information set of the unit step function characterizing default. The intuition behind this division of the information set is that we merely want to condition on \( \mathcal{G}_t \), i.e. the usual state variables, where the state variable governing the hazard rate holds the link to the default process. In order to construct the link, we need two important tools: i) the usual Ito Formula, and ii) the semi-martingale representation. Taking these as given, we take the mathematics from here and show how to get the components we need for the pricing of defaultable bonds (the technical details are beyond the scope of this thesis).

Conferring section 2.1.3 we know that the expectation has to be taken under the equivalent Martingale measure \( \mathbb{Q} \), where all discounted price processes as well as the numeraire must be martingales. From the definition of the default process \( x(t) \) we saw that it essentially takes values in the set \( \{0,1\} \) that starts at zero and jumps at most once to the value one, at some random time. Clearly the default process is an increasing function, and is therefore not a martingale under the measure \( \mathbb{Q} \). By the semi-martingale representation, the key is therefore the existence of a positive hazard rate process \( h(t) \), with the property under the equivalent Martingale measure, that the one-jump process \( x(t) = 1_{\{ z \leq T \}} \) has an absolute continuous and predictable increasing compensator process \( A(t) \) with \( A(0) = 0 \), such that
Part II: The Pricing of Defaultable Bonds

\[ m(t) = x(t) - A(t) \]
\[ = x(t) - \int_0^t (1 - x(u)) h(u) du \]  \hfill (2-48)

is a martingale difference. Hereby is meant that \( \mathbb{E} [m(t)] \) equals zero, which can be derived through the equivalence of showing that the expectations of the Doléans-Dade exponential of (2-48) equals one. Madan (2000) explicitly write this exponential as

\[ M(t) = \exp[-m(t)] = \exp\left(\int_0^t (1 - x(u)) h(u) du\right)(1 - x(t)) \]

which is an \( \mathcal{F}_t \) adapted martingale. If we observe that we may delete the jump process from the exponential as it is unity while \( x(u) = 0 \) and once \( x(t) \) is one then \( 1 - x(t) \) is zero and therefore the exponential is no longer relevant. Taking expectations result in

\[ \mathbb{E}^Q \left[ M(t) \mid \mathcal{G}_t \right] = \exp\left(\int_0^t h(u) du\right) \mathbb{E}^Q \left[ (1 - x(t)) \mid \mathcal{G}_t \right] = 1 \]

which Madan (2000) argues is an \( \mathcal{G}_t \) adapted martingale, and is only possible if\(^{35}\)

\[ \mathbb{E}^Q \left[ (1 - x(t)) \mid \mathcal{G}_t \right] = \exp\left(-\int_0^t h(u) du\right). \] \hfill (2-49)

Equation (2-49) gives us the necessary fundamental in implementing the survival function under \( \mathbb{Q} \). To obtain this expression for the probability of no default prior to the maturity date \( T \), we note that the conditional probability of no default in the interval \( [t; T] \) given no default by time \( t \) i.e. \( \mathbb{Q}(z > T \mid Z > t) \) is

\[ \mathbb{E}^Q \left[ (1 - x(T)) \mid \mathcal{F}_t \right] = \mathbb{E}^Q \left[ (1 - x(T)) \mid \mathcal{G}_t \land \mathcal{H}_t \right] \]
\[ = \mathbb{E}^Q \left[ Z > T \mid \mathcal{G}_t, Z > t \right]. \] \hfill (2-50)

As the intention is exclusively to condition on \( \mathcal{G}_t \) we utilize Madan’s results and manipulate (2-50) using the laws of conditional and iterated expectations such that

\(^{35}\)See Madan (2000) for the technicalities. He states that it is also required that \( M(t) \) is a predictable and bounded process of integrable variation.
where the common values of $Z > T$ and $Z > t$ is just $Z > T$. The above result draws crucially on
the existence of the defined semi-martingale relationship between $(1 - x(t))$ and $h(t)$ from equation
(2-48) as it is utilized in the last step of (2-51). The existence of this martingale is explicitly justified in Madan and Unal (1994) and Madan (2000). It should be emphasized that under the equivalent Martingale measure, $h(t)$ is referred to as the risk-neutral hazard rate process and will not equal the objective instantaneous probability of default as long as the market price of risk associated with the jump process is non-zero. Henceforth, $\mathbb{E}_t[ \cdot ]$ refers to the $\mathcal{G}_t$-conditional expectation operator.

The above equation allows the rewriting of the first part of (2-41) and in order to address the last part of the price equation we need to compute the density function for $Z$. Stated differently, we need to compute the probability density distribution of default within the interval $[t; T]$ given no default prior to time $t$. The simplest way of doing so is by using the relation from (2-44) where

$$f(t) = S(t)h(t)$$

overall with the statement for the survival function derived in (2-51)

$$f(t) = \frac{\partial}{\partial s} \mathbb{Q}(Z \leq s \mid Z > t, \mathcal{G}_t) = \exp \left(- \int_t^s h(u)du \right) h(s). \quad (2-52)$$

Conditioning on information available at time $t$ and integrating over the entire maturity of the bond, we get the expression we need for the price equation.\(^{36}\)

Using the results in (2-51) and (2-52), we are now able to establish a joint expression for the discounting term in the price equation, which consolidates time value and the default rate. Shown merely for the no default term, (2-53) states this adjustment

\(^{36}\) Now the definition of the compensator process in equation (2-48) becomes clear as

$$E^Q \left( m(t) \mid \mathcal{G}_t \right) = E^Q \left( x(t) \mid \mathcal{G}_t \right) - E^Q \left( \int_0^t (1-x(s))h(s)ds \mid \mathcal{G}_t \right)$$

$$= E^Q \left( x(t) \right) - E^Q \left( \int_0^t h(s)ds \mid \mathcal{G}_t \right) = E^Q \left( x(t) \right) - E^Q \left( 1 - \exp \left( \int_0^t h(s)ds \right) \right) = 0.$$  

This relationship proves that under the equivalent Martingale measure the expectations of the unit step function equals the cumulative distribution function of the random default time.
Having specified the default-adjusted discount rate, it immediately follows that the expression for the defaultable zero-coupon bond can be rewritten as

\[
P(t, T) = \mathbb{E}_t^Q \left[ \exp \left( -\int_t^T R(s) ds \right) H + \int_t^T \exp \left( -\int_u^T R(s) ds \right) h(u) \psi(u) du \right].
\]  

Unlike (2-41) the above equation eliminates all reference to the discontinuous process \( x(t) \), and therefore provides a very tractable result since it reduces the problem of pricing defaultable bonds to that of pricing non-defaultable bonds. The first term in (2-54) is the value of a claim that pays \( H \) on survival at maturity. The expression inside the outer integral of the second term is the price density of a claim that pays \( \psi \) if default occurs in an incremental interval, thus the whole term is the price of the proceeds given default before maturity.

At this point we have not made any assumptions about the recovery function. If we for simplicity assume zero recovery in the case of default, the expression of the zero-coupon defaultable bond reduces to

\[
\tilde{P}(t, T, 0) = \mathbb{E}_t^Q \exp \left( -\int_t^T R(s) ds \right)
\]

where

\[ R(t) = r(t) + h(t). \]

The implications of the hazard process \( h(t) \) are very straightforward. With increasing probability of default the level of \( h(t) \) rise, hereby increasing the discounting and lowering the bond price, and changing expectations to the likelihood of default of the bond are captured by the stochastic properties of the hazard rate. The variables describing these properties are usually applied in the Markov diffusion framework since it provides solutions entirely in line with standard term-structure models for default-free debt. Closed-form solutions are attainable if the conditions stated by Duffie and Kan (1996), are obeyed for the hazard rate as well.
As the only thing that distinguishes the defaultable and default-free bonds in the above Cox-framework is the probability of default measured by \( h(t) \), the credit spread provided by the model exactly measures this probability. To verify this, note that the continuously compounded credit spread is defined as the yield difference between the defaultable zero-coupon bond and an equivalent default-free bond

\[
s(t, T) = -\frac{1}{T-t} \ln \frac{\hat{P}(t, T)}{P(t, T)}
\]  

for a bond at time \( t \) with maturity \( T-t \). It is therefore seen that the implied spread is

\[
s(t, T) = -\frac{1}{T-t} \ln \frac{\mathbb{E}_t^Q \exp \left( -\int_t^T r(s) + h(s) ds \right)}{\mathbb{E}_t^Q \exp \left( -\int_t^T r(s) ds \right)}
\]

\[
= -\frac{1}{T-t} \ln \mathbb{E}_t^Q \exp \left( -\int_t^T h(s) ds \right)
\]

\[
= -\frac{1}{T-t} \ln \mathbb{Q}(Z > T)
\]

\[
= -\frac{1}{T-t} \ln (1 - \mathbb{Q}(Z \leq T))
\]  

assuming no default prior to \( t \). For small default probabilities the above expression can reasonably be approximated by \( \mathbb{Q}(Z \leq T) / \tau \), i.e. the spread is approximately the average default probability of the remaining maturity of the bond.

Indeed the valuation model is very flexible concerning the formulation of credit spreads, as it allows a variety of state dependent processes to describe the hazard rate. Yet, applications of the valuation model still implicitly rely on several very important assumptions. For a given credit spread process and default-free term-structure formulation, the crucial assumption underlying the valuation equation, is that the processes involved are exogenously specified, i.e. \( r(t) \) and \( h(t) \) does not depend on the value of the defaultable claim itself. It is of course not sensible to think of \( r(t) \) as being determined endogenously, but it is plausible to imagine situations where the probability of default of the issuing firm depends on the price of the claim. In the case where the payments on the bonds outstanding constitute a major part of a particular firm’s operation, it is generally invalid to assume exogenous specified hazard rates. In this situation the debt payments will presumably influence the firms probability of default, why we have to account for this in the price equation by allowing the risk-neutral default probability, \( h(t, x, \hat{P}) \), to depend endogenously on the
Part II: The Pricing of Defaultable Bonds

price of the claim itself. As a result, the linear valuation model becomes non-linear in the promised payoff.\footnote{Whether the default rate would become dependent of the market value of the bond, such that we have to include non-linear effects is generally not clear-cut. For further details see Duffie and Singleton (1999) and Lando (1998).}

As stressed by Duffie and Singleton (1999), another essential issue to consider is the parameterization of the default-adjusted short rate. Depending on the purpose of the analysis, one can either parameterize $R$ directly, or parameterize the components $r$ and $h$ separately. For a given defaultable bond market, parameterizing $R$ directly, as opposed to separate parameterization, provides less information concerning the default rate. On the other hand, this formulation permits empirical characterizations of the default-adjusted short-rate without the need to commit to a formulation of the credit spread. Indeed this strategy may be attractive if the main part of the credit-spread dynamics owes to non-default factors, as it is less exposed to misspecification errors. Apparently, the benefit of this econometric formulation is only attractive if one is merely concerned with characterizing the distribution of $R$, say for the purpose of computing the zero-coupon yield curve implied by the defaultable bonds. Conversely, pursuing this approach, nothing can be learned directly about the default and recovery processes, so, if the intension is to study these two components individually the parameterization must be separated.

2.2.3 The PDE Representation

Although the above model is based on simple desirable assumptions the derivation is somewhat technical. Furthermore the framework will, except under the affine settings, not permit closed-form solutions. Application of the Cox-framework therefore requires numerical methods in order to solve the price expression, and hence, the expectation expression in (2-55) is often rewritten into its equivalent partial differential equation. Having specified the PDE, several finite-difference schemes are then straightforward to apply.

From the Feynman-Kac representation it is intuitively seen that with little modification of the “discounting” term (2-55) is equivalent to the usual parabolic partial differential equation. However, deriving the PDE has some appealing intuitive interpretations, and will shed light on the pricing concepts in a more intuitive way. For simplicity we only present the one-dimensional case and adopt a slightly different way of deriving the PDE compared to the method used in chapter 2.1. Otherwise matrix notation will quickly blur the appealing characteristics, and the idea of the derivation will be minimal.
Part II: The Pricing of Defaultable Bonds

As Wilmott (1998), consider a portfolio of a long position in a risky bond \( \tilde{P}(t, T) \) and \( \Delta \) short in a risk-free bond \( P(t, T) \)

\[
\Pi = \tilde{P} - \Delta P
\]  

(2-58)

Employing Ito’s Formula, and as previously assume time invariant parameters, the choice of \( \Delta \) to hedge the risky \( dr \) term entails that we can write the change in portfolio value as

\[
d\Pi = \left[ \frac{\partial \tilde{P}}{\partial t} + \frac{1}{2} \sigma^2 \frac{\partial^2 \tilde{P}}{\partial r^2} - \Delta \left( \frac{\partial P}{\partial t} + \frac{1}{2} \sigma^2 \frac{\partial^2 P}{\partial r^2} \right) \right] dt
\]

\[\equiv (d\tilde{P} - \Delta dP)dt\]  

(2-59)

where \( \Delta = \frac{\partial \tilde{P}}{\partial r} \cdot \frac{\partial P}{\partial r} \).

Now recall that there exist two different states indicating if the risky bond has defaulted or not, and in contrast to the fundamental hedge portfolio for default-free bonds, (2-59) will therefore not be deterministic in the next incremental time interval, and will with the probability of \( hdt \) be in the state of default and with the probability \( 1 - hdt \) be in the state of no default. We therefore have to take expectations with regards to the jump, giving us the result

\[
\mathbb{E}^Q (d\Pi) = (1 - hdt)(d\tilde{P} - \Delta dP)dt - h\tilde{P}dt + O(dt^2)
\]  

(2-60)

The last part of the equation tells us that if default has occurred, then the change in the value of the portfolio is dominated by the loss of the risky bond. Recall that \( \mathbb{E}(dt|dt) = 0 \), as the error of \( O(dt^2) \) is negligible, we manipulate (2-60) to give

\[
\mathbb{E}^Q (d\Pi) = (d\tilde{P} - \Delta dP - h\tilde{P})dt.
\]  

(2-61)

As the portfolio is locally non-random under the pricing measure \( Q \), absence of arbitrage is precluded if the portfolio yields the risk-free spot rate, i.e. in the next time interval \( dt \) the expected value is \( \mathbb{E}^Q (d\Pi) = r\Pi dt \) such that

\[
(d\tilde{P} - \Delta dP - h\tilde{P})dt = r(\tilde{P} - \Delta P)dt \quad \Leftrightarrow \quad (d\tilde{P} - (r + \eta)\tilde{P})dt = (\Delta dP - r\Delta P)dt.
\]

Dropping \( dt \) and rearranging in terms of \( \tilde{P} \) and \( P \) yields
This is the very important difference compared to the normal differential equation, and the part
where the default adjustment in the discounting term becomes clear. Equation (2-62) is easily
simplified because we know the default-free bond, $P$, satisfies the basic bond price equation. The
right-hand side therefore equals $\sigma \lambda$ giving us the essential result

$$\frac{d\hat{P} - (r + h) \hat{P}}{\partial \hat{P} / \partial t} = \sigma \lambda - u.$$  

The pricing equation for the risky bond is therefore

$$d\hat{P} + (u - \lambda \sigma) \frac{\partial \hat{P}}{\partial r} - (r + h) \hat{P} = 0$$

and subsequently inserting the expression for $d\hat{P}$ gives the final result

$$\frac{\partial \hat{P}}{\partial t} + \frac{1}{2} \sigma^2 \frac{\partial^2 \hat{P}}{\partial r^2} + (u - \lambda \sigma) \frac{\partial \hat{P}}{\partial r} - (r + h) \hat{P} = 0.$$  

We see that the only difference compared to the usual PDE for default-free bonds is the presence of $h$ in the discounting term.

### 2.2.4 Different Recovery Schemes

The last variable to address in the pricing equation is the recovery payout upon default. In light of
the empirical evidence discussed earlier, different aspects emerge when inferring recovery rate
issues, and for the same reason a range of proposals with different applications have been pre-
sented. Besides the fact that companies have very distinct priority structures, the determination of
recoveries to creditors during bankruptcy proceedings is a complex process that typically involves
substantial negotiation and litigation. No tractable economical model therefore captures all as-
pects of this process so, in practice, all models involve tradeoffs regarding how various aspects of
default are captured.

Our motivation for addressing recovery issues are of second-order compared to addressing aspects
regarding the default probabilities. The main argument underlying this concern is that our empiri-
Part II: The Pricing of Defaultable Bonds

cal study is solely based on investment grade bonds where default probabilities are expected to be low, entailing that bond prices are less sensitive to recovery issues.

In the previous sections, we saw that default is implemented in the pricing equation by adjusting the discounting term. When considering the payout structure upon default, three assumptions are typically made about the recovery:

i. Recovery at Treasury (RT):

\[ \psi(u) = \phi(u) P(u, T) \]

where the creditor receives an exogenous fraction, of an otherwise equivalent default-free bond \( P(u, T) \), specified by the recovery function \( \phi(u) \).

ii. Recovery of Face Value (RFV):

\[ \psi(u) = \phi(u) H \]

where the creditor receives an exogenous fraction of face value immediately upon default (no recovery of coupons) specified by the recovery process \( \phi(u) \).

iii. Recovery at Market Value (RMV):

\[ \psi(u) = \phi(u) \tilde{P}(u, T) \]

where the creditor receives an exogenous fraction \( \phi(u) \), of the immediate pre-defaultable market value of the bond (which involves recovery of coupons). In contrast to the above two schemes this formulation describes an endogenously defined recovery function.

At first the three schemes seems very alike, but when it comes to computation the techniques required are very different. In the following we briefly comment on the different schemes, outline their basic properties, and verify our motivation for selecting a particular recovery function. To understand the underlying concept of recovery considerations we refer to the earlier discussion in section 1.3.1.

2.2.4.1 Recovery at Treasury (RT)

One very important aspect in all of the above approaches is that \( \phi \) can either be deterministic or stochastic. However, under the RT formulation the computational burden of evaluating the price of the defaultable bond, in the case of a stochastic fractional recovery process, is unfortunately substantial cf. Duffie & Singleton (1999). Furthermore, even when the recovery is deterministic
problems arise. To understand this aspect notice that under the RT formulation the price equation for a unit face defaultable zero-coupon bond is written as

\[
\tilde{P}(t,T,\phi) = E^Q_t\left[ \exp\left( -\int_t^T R(s)ds \right) H + \int_t^T \exp\left( -\int_u^t R(s)ds \right) h(u)\phi(u)P(u,T)du \right].
\]

When the recovery process is deterministic, but time varying, the payoff at maturity will vary and then, not only is the time of default relevant, but an estimation of the joint distributions of \( \phi(t) \), \( h(t) \) and \( r(t) \) is necessary in the estimation of the price of the defaultable bond. Indeed this plays a very computational challenge, and largely for this reason, the recovery process is usually assumed constant when using the RT formulation.

With \( \phi(t) = \phi \), the payoff at maturity in the event of default is constant regardless of when default occurred. Assuming constant recovery the zero-coupon bond may conveniently be written as

\[
\tilde{P}(t,T,\phi) = E^Q_t\left[ S, P(t,T,0)H + (1 - S)\psi(t) \right]
\]

where \( S \) is the conditional probability of survival under the \( Q \)-measure (the survival function). As shown by Lando (2000), employing the RT formulation and assuming that the default-free interest and hazard rates are \( Q \)-independent, we can rearrange the equation using the results from (2-51), (2-52) and (2-55)

\[
\tilde{P}(t,T,\phi) = E^Q_t\left[ D(t,T)S, H + \int_t^T D(t,u)\exp\left( -\int_u^t h(s)ds \right) h(s)\phi D(u,T)Hdu \right]
\]

\[
= E^Q_t\left[ S, P(t,T) + \int_t^T \exp\left( -\int_u^t h(s)ds \right) h(s)\phi D(t,u)P(u,T)du \right]
\]

\[
= E^Q_t\left[ S, P(t,T) + \int_t^T \exp\left( -\int_u^t h(s)ds \right) h(s)\phi P(t,T)du \right]
\]

\[
= E^Q_t\left[ P(t,T)S, + P(t,T)(1 - S)\phi \right]
\]

\[
= \phi P(t,T) + (1 - \phi)\tilde{P}(t,T,0). \tag{2-64}
\]

It appears that the absence of arbitrage implies that the price of the unit face defaultable zero-coupon bond with recovery of \( \phi \), of an otherwise equivalent default-free bond \( P(t,T) \), equals a weighed sum of two distinct payments. Namely it equals the sum of receiving \( \phi \) of a default-free zero-coupon bond and \( 1-\phi \) of a default-risky no recovery zero-coupon bond. The assumption of constant recovery is of course not very tenable or realistic. But as Duffee (1996) explicitly states, it should be thought of as an approximation to a model in which there is very little information revealed about the recovery value until the issuer is close to default.
2.2.4.2 Recovery of Face Value (RFV)

As under the RT assumption implementing stochastic recovery and hazard simultaneously require numerical solutions. But even if recovery is assumed constant another drawback arises in the RFV, as the formulation does not allow for analytical solutions without the use of very complex tools. Our interest in the RFV formulation is therefore limited. To understand this hypothesis substitute the RFV formulation into equation (2-54), which entails

\[
\hat{P}(t,T,\phi) = E^Q_t\left[ \exp\left(-\int_t^T R(s)ds\right)H + \int_t^T \exp\left(-\int_u^T R(s)ds\right)h(u)\phi(u)Hdu \right].
\]

(2-65)

The problem with this recovery scheme, from a computational perspective, is that it cannot be rearranged to avoid numerical computation of the second term. In contrast to the RT formulation the recovery payment, \(\phi H\), is not discounted from expiry to the time of default as indicated by \(D(u,T)\) in (2-64). Therefore the only possibility is to manipulate the expectation operator (still, we assume that the variables are independent under the pricing measure), such that

\[
\hat{P}(t,T,\phi) = E^Q_t\exp\left(-\int_t^T R(s)ds\right)H + E^Q_t\int_t^T \exp\left(-\int_u^T R(s)ds\right)h(u)\phi(u)Hdu.
\]

Given the affine regularity conditions are satisfied the two expectations can be solved explicitly, so the computation of \(\hat{P}\) calls for one numerical integration of the second term. Bakshi et al. (2002) and Dai & Singleton (2002) propose other methods, which give analytical solutions. Both paths are, however, difficult and cumbersome and beyond the scope of this thesis.

2.2.4.3 Recovery at Market Value (RMV)

In contrast to the two other formulations the RMV allows not only the intensity but also the recovery rate to enter the default adjusted spot rate process (Lando uses the notation; thinning the intensity). The tractability and convenience of this formulation arise because it enables a simple reduction of the discount rate corresponding to the proportion \(\phi(t)\) of the hazard. To see this, recall the price under the RMV

\[
\hat{P}(t,T,\phi) = E^Q_t\left[ \exp\left(-\int_t^T R(s)ds\right)H + \int_t^T \exp\left(-\int_u^T R(s)ds\right)h(u)\phi(u)\hat{P}(u, T)du \right].
\]

(2-66)

From Madan (2000) and section 2.1.3, we know that the associated discounted gains process, which consists of the price and the recovery
Part II: The Pricing of Defaultable Bonds

\[ G(t) = \int_0^t \tilde{D}(u) h(u) \phi(u) \tilde{P}(u, T) du + \tilde{D}(t) \tilde{P}(t, T) \]  

(2-67)

where \( \tilde{D}(t) \) is the adjusted discount function

\[ \tilde{D}(t) = \exp \left( -\int_0^t \left( r(s) + h(s) \right) ds \right) \]

is a martingale under \( \mathbb{Q} \) and therefore \( \mathbb{E}^\mathbb{Q} \left[ \Delta G(t) \right] \) equals zero. Employing Ito’s Formula and manipulating gives us the following relationship

\[ dG(t) = \tilde{D}(t) \left[ h(t) \phi(t) \tilde{P}(t, T) + \frac{\partial \tilde{P}}{\partial t} - (r(t) + h(t)) \tilde{P}(t, T) \right] dt \]

\[ = \tilde{D}(t) \left[ h(t) \phi(t) \tilde{P}(t, T) - (r(t) + h(t)) \tilde{P}(t, T) \right] dt + \tilde{D}(t) d\tilde{P} \]

(2-68)

As the expectations of equation (2-68) under \( \mathbb{Q} \) equals zero we obtain a very tractable result\(^{38}\)

\[ \mathbb{E}^\mathbb{Q} \left[ \frac{d\tilde{P}}{dt} \right] = (r(t) + (1 - \phi(t))h(t)) \tilde{P}(t, T) \]

(2-69)

By utilizing Ito’s Formula once again we obtain a new expression for the discounted gains martingale in (2-67)\(^{39}\)

\[ L(t) = \tilde{D}(t, \phi) \tilde{P}(t, T) \]

(2-70)

where

\[ \tilde{D}(t, \phi) = \mathbb{E}^\mathbb{Q}_t \exp \left( -\int_0^t \left( r(s) + h(s) \left[ 1 - \phi(s) \right] \right) ds \right) \]

is the price of an unit face defaultable bond. Using the boundary condition \( \tilde{P}(T, T) = H \) we obtain the desired result, which is a price equation for the zero-coupon defaultable bond where the recovery is incorporated in the discount rate

\[ \text{\footnotesize \cite{38} When deriving this result it is crucial to remember which part of equation (2-68) is deterministic as } E(xdt) = xdt. \]

\[ \text{\footnotesize \cite{39} The easiest way to verify the derivation is by differentiating (2-70), and noting that } \mathbb{E}^\mathbb{Q}[\Delta L(t)] = 0, \text{ whereby the result in (2-69) is obtained.} \]
Part II: The Pricing of Defaultable Bonds

\[ \hat{P}(t, T, \phi) = \mathbb{E}_t^Q \left[ \exp \left( - \int_t^T \left( r(s) + h(s)[1 - \phi(s)] \right) ds \right) H \right]. \]  \hspace{1cm} (2-71)

Now the discount rate implements time value, hazard rate and recovery consideration.

2.2.4.4 Implementing Recovery Considerations

The choice of recovery scheme will always be pragmatic but should of course be based on the ability to replicate the true density function. Besides discriminating between the schemes using mathematical tractability and model flexibility as criterions, one should also address the more qualitative issues as the legal structure of the claim to be priced.

Regarding mathematical tractability it is already mentioned that the RFV formulation is less advantageous compared to the RT and RMV. Choosing between the two latter assumptions we take a closer look at the implied credit spreads.

Utilizing the RT formulation and the definition of the credit spread in (2-56) it is easily seen that the spread is computed as

\[ \hat{P}(t, T, \phi) = P(t, T) \left( \phi + (1 - \phi) \mathbb{E}_t^Q \exp \left( - \int_t^T h(s) ds \right) \right) \Rightarrow \]

\[ s(t, T) = - \frac{1}{T - t} \ln \left( \phi + (1 - \phi) \mathbb{E}_t^Q \exp \left( - \int_t^T h(s) ds \right) \right) \]  \hspace{1cm} (2-72)

and for the RMV formulation the credit spread is

\[ s(t, T) = - \frac{1}{T - t} \ln \frac{\mathbb{E}_t^Q \exp \left( - \int_t^T r(s) + (1 - \phi) h(s) ds \right)}{\mathbb{E}_t^Q \exp \left( - \int_t^T r(s) ds \right)} \]

\[ = - \frac{1}{T - t} \ln \mathbb{E}_t^Q \exp \left( - \int_t^T (1 - \phi) h(s) ds \right). \]  \hspace{1cm} (2-73)

In (2-72) the effect of the recovery and the hazard rate is separated in contrast to the implied spread of the RMV formulation in (2-73). This is an obvious weakness of the RMV formulation, as it is virtually impossible to separate the effect of recovery and hazard rates empirically, whereas the RT leads to identifiable impacts on credit spreads. Furthermore, when the hazard and interest rates are assumed independent, the market expectations of default probability and recovery can be inverted from the term structures, which is easily seen by rearranging (2-72) as
Part II: The Pricing of Defaultable Bonds

\[ \mathbb{Q}(Z > T) = \mathbb{E}_t^Q \exp \left( - \int_t^T h(s) ds \right) = \frac{\bar{P}(t,T,\phi) - \phi P(t,T)}{(1 - \phi) P(t,T)}. \]

The price of the otherwise equivalent defaultable and Treasury bonds on the right hand side are known, and the implied default probability and recovery is therefore easily identified solving a system of equations. In the case of the RMV formulation the above derivation is not possible even when interest and hazard rates are assumed independent. We only know that

\[ \mathbb{Q}(Z > T)^{(1-\phi)} = \mathbb{E}_t^Q \exp \left( - \int_t^T h(s) ds \right)^{(1-\phi)} \]

\[ = \mathbb{E}_t^Q \exp \left( - \int_t^T h(s)(1-\phi) ds \right) \]

when the hazard rate is constant (or deterministic) and otherwise not. In the general case where analytical solutions for the hazard process are unattainable it will not be possible to compute the implied default probabilities. This is clearly not attractive and an important limitation of the RMV assumption, as it eliminates all interesting aspects of stochastic hazard rates. In the special case though, it will be possible to compute the implied default rates. As stressed by Lando (1997) one can do so, when an analytical expression for the right hand side is specified. It will then be possible to separate the individual effects of recovery and hazard. The problem is, as always, to specify an expression for the hazard that permits analytical solutions and still reflects some sort of reality. Therefore it is difficult to discriminate between the RT and RMV on mathematical grounds although the RT seems slightly more pragmatic.

On the other hand, if we use the legal structure as criterion for choosing between the recovery schemes, Duffie and Singleton (1999) note that if liquidation occurs at default and absolute priority applies, the RFV is the most realistic, since it implies equal recovery for bonds of equal seniority. However Duffie and Singleton at the same time found that the prices under the RMV and RFV formulations were rather similar in empirical testing. Relying on these findings, we additionally need to ask what the consequence is of applying the RMV compared to the RT formulation, and in the end is there a significant difference between the formulations at all?

To address this issue assume that the hazard and recovery are exogenously specified functions, where the hazard rate follows a one-factor CIR model, and the recovery is assumed constant at 0.5. Furthermore as we consider investment grade bonds where default probabilities are expected to be low we assume that the credit spread essentially is zero for \( \tau \to 0 \), and that the Treasury term structure is flat. The results are shown in Figure 2-1.
Part II: The Pricing of Defaultable Bonds

Figure 2-1: The Influence of the $RT$ and $RMV$ formulation on Credit Spreads

The figure illustrates the implied credit spreads of the $RMV$ and $RT$ formulation together with the credit spread is no recovery (i.e. the pure hazard rate process). The calculations are made for reasonable parameters of the hazard rate process ($\kappa = 0.3$, $\theta = 0.03$, $\lambda = -0.1$ and $\sigma = 0.2$).

Notice that the implied credit spreads are remarkable identical for shorter maturities, and the difference is only 13 bp at 10 years to maturity. This conclusion is verified by Duffie and Singleton (1999) for a variety of reasonable parameters, and not surprisingly the difference between the implied spreads is enlarged with increasing recovery. In addition the recovery formulations imply different impacts of the hazard rate as well. Duffie and Singleton (1999) find that the price of the defaultable bond is higher under $RFV$ as the last part of (2-64) increases the price more than the reduced discounting under the $RMV$ although the difference is negligible. A possible misspecification of our model is therefore ultimately not that sensitive to the chosen scheme.

2.2.5 Accounting for Coupon Payments

As stressed by Duffee (1996) much empirical work involves modeling coupon bonds instead of zeros. Therefore the interest of this section involves extending the knowledge gained in the sections 2.2.1 - 2.2.4. Using the same notation for the defaultable bond price, we accommodate both discrete and continuous coupon payment by $c(t)$, which is a non-decreasing function of cumulated coupon payments until time $t$. Following Bakshi et al. (2001) the bond price equation (2-41) can then be rewritten as
Part II: The Pricing of Defaultable Bonds

\[ \tilde{P}(t, T, C, \phi) = \mathbb{E}_t^Q \left[ \int_t^T \frac{B(t)}{B(u)} (1-x(u)) c(u)du + \frac{B(t)}{B(T)} (1-x(T)) H \right. \\
\left. + \int_t^T \frac{B(t)}{B(u)} (1-x(u)) \phi(u) dx(u) \right] F_t \]  

(2-74)

The first two terms in (2-74), accounts for the full payment if default does not occur. More specifically the first term discounts the stream of coupon payments received as long as there is no default and stops at the default time. The second term accounts for the payment of promised face value given no default. Finally, the third term accounts for the single recovery at the default time. Furthermore \( dx(u) \) is unity for at most one time \( u \) so there is at most one default payout.

Because the above equation is non-operational, as the unit step function is in the expression, we employ the relationship between \( 1-x(t) \) and the hazard rate as given in (2-51) and remember the functionality of the probability distribution function of default time (2-52), whereby the bond price equation yields

\[ (t, T, C, \phi) = \mathbb{E}_t^Q \left[ \int_t^T \exp \left( -\int_t^R R(s)ds \right) c(u)du + \exp \left( -\int_t^R R(s)ds \right) H \right. \\
\left. + \int_t^T \exp \left( -\int_t^u R(s)ds \right) h(u)\phi(u)du \right] \]  

(2-75)

where

\( R(t) = r(t) + h(t). \)

As we concluded above the RFV formulation did not apply analytical solutions and is therefore disregarded here. Likewise the RT did not provide analytical solutions under stochastic recovery, but possible when imposing constant recovery. As emphasize by Duffee (1996) this is easily extended to coupon bonds, since the price is simply the sum of the individual cash flows. Turning to the RMV we know from Madan (2000) and Bakshi et al. (2001) that this formulation enables us to expand the gains process from (2-67) to accommodate coupon payments as

\[ G(t) = \int_0^T \tilde{D}(u) C(u)du + \int_0^T \tilde{D}(u) h(u)\phi(u)\tilde{P}(u,T)du + \tilde{D}(u)\tilde{P}(t,T) \]  

(2-76)

Using Ito’s Formula, remembering that the discounted gains process in (2-76) is a martingale under \( Q \), we isolate the first derivative of the defaultable bond price with regards to the time
Part II: The Pricing of Defaultable Bonds

\[ \mathbb{E}^{Q}\left[ \frac{d\tilde{P}}{dt} \right] = -C(t) + \left( r(t) + (1 - \phi(t))h(t) \right) \tilde{P}(t,T). \]  

(2-77)

By utilizing Ito’s Formula once again and employing the boundary condition \( \tilde{P}(T,T) = H \) we obtain the price equation for the defaultable coupon bond

\[ \tilde{P}(t,T,C,\phi) = \mathbb{E}^{Q}\left[ \int_t^T \tilde{D}(t,u)C(u)du + \tilde{D}(t,T)H \right] \]  

(2-78)

where

\[ \tilde{D}(t,T,\phi) = \exp\left( -\int_t^T \left( r(s) + h(s)[1 - \phi(s)] \right) ds \right) \]

Equation (2-78) reveals that once the discount rate has been adjusted to reflect the mean loss rate one may essentially treat defaultable bonds as if they were risk-free and in particular value coupons as portfolios of zeros.

2.2.6 Area of Application and Previous Empirical Studies

We conclude this chapter with a brief discussion of the area of application of the models, and the knowledge gained from previous empirical studies. The recent years overwhelming interest in the quantitative modeling of credit risk must be seen in light of the growing market of credit derivatives. Virtually all credit derivatives require a model for the default/mean loss process of the underlying assets. Historically, such models have not been available, and, in addition, the hazard rate framework outlined above is clearly a reflection of this.

We saw that the hazard rate framework simplified the default process in two-states, default or not default, which occur completely unexpected. In reality, the default process is not a simple binomial process. There is usually a deterioration of credit quality up to the default event. Hence, this framework clearly lacks economical interpretation and offers no guidance on the presence of a structural change in firm specific variables. The framework is therefore far from ideal, but in recognition of the difficulties of capturing every aspect of the mechanisms leading a firm in distress, this is perhaps the closest we get. Moreover, other qualities justify the use of the hazard rate models. If we believe we can measure credit risk through the mean loss rate given in equation (1-1), ordinary term-structure theory provides a rich array of models to cope with the dynamics of the implied elements. And finally, the framework involves a key advantage compared to usual Merton model as they generate realistic short maturity credit spreads.
Part II: The Pricing of Defaultable Bonds

The problem is of course that we do not know the likelihood of default until it is too late, so in reality the parameters in the models are impossible to measure. To circumvent this researchers usually assume that the market knows the parameters, and then fit or calibrate the models such that the theoretical prices match the market prices. In essence, the real question to be asked is therefore whether the market is efficiently pricing credit risk. Unfortunately empirical studies of the performance of hazard rate models are rather limited.

Duffee (1996, 1999) examined US corporate bond prices based on the Duffie and Singleton (1995 and 1997) framework. Using the MLE inversion method proposed by Chen and Scott (1993), he parameterized a model based on Markovian settings, assuming independent state variables and constant recovery. On average, his model fits corporate bond prices reasonably well, but the model clearly lacks on its cross-sectional properties, as it cannot simultaneously capture low, flat term structures and high, steeper term structures. Consistent with the evidence that changes in credit spreads move inverse with changes in the default-free term structure Duffee (1999) extended his earlier work to cover this particular aspect by implementing negative correlated increments still keeping the state variables independent. Duffee (1999) verifies several key features of the term structure of credit spreads (discussed in section 1.3) both in sign and magnitude. However, not surprisingly the model experiences difficulties matching certain features of the data. Most importantly, Duffee argues that there appear to be persistent fluctuations in the volatilities of the yields (GARCH-like effects) that are not captured by the model, and parameter estimations are instable as the firm’s credit quality changes.

Bakshi et al. (2000) proposed a model consistent with the theoretical justification of the structural models, as they involve firm specific distress factors to describe the hazard function. Measured by both out of sample and hedging errors, they find that their proposed credit risk model performs relatively better for high-grade bonds and for bonds with longer maturities. In accordance with prior research their model reveals that firm specific factors such as leverage and book-to-market ratios account for some of the cross-sectional variations in corporate bonds and that these variables are positively correlated with yields. However their main conclusion is that the use of firm specific distress factors provides only marginal improvements. Additionally, the credit-spread analysis made by Bakshi et al. (2000) examines the determinants of credit-spread changes on an individual bond-level. Their contribution is appealing, although concluding that variables, which from a theoretical perspective ought to determine credit spreads, lack considerably explanatory power in empirical testing.
2.3 The Duffee (1999) Model

Following the empirical observations of credit spreads made in Part I we now shift our attention to the specification of the term-structure model describing the corporate yield dynamics. Applying the framework outlined in the previous chapter we specify the price of the defaultable coupon bonds under the recovery at market value formulation, such that the default and severity adjusted discount rate is expressed as $R_t = r_t + h_t(1-\phi)$. The recovery $\phi$ is assumed constant, but various across ratings (AAA, AA, A and BBB). This structure is with little modification in line with Duffee (1999). In the following we explain and derive the model.

It is well-known that in order to satisfactory capture the dynamics of the term structure, Treasury or corporate, we generally need more than one-factor. Moreover, we need a functional specification of the state variables such that the model adequately captures key features observed in the market prices. In this perspective we present a three-factor model for the prices of defaultable bonds, where we separate the dynamics of the defaultable bond prices into dynamics resulting from market risk and credit risk, respectively.

We assume that the instantaneous default-free interest rate $r_t$ equals the sum of a constant and two state variables

$$r_t = \alpha + x_{1,t} + x_{2,t}$$

(2-79)

where the underlying unobservable stochastic process of each state variable $x_{i,t}$ is assumed to follow a square-root diffusion CIR model. The price of the default-free zero-coupon bond, and the equivalent yield to maturity, is

$$P(t,T) = \prod_{i=1}^{2} \exp \left( A_i(\tau) + B_i(\tau)x_i - \alpha_i\tau \right) \Rightarrow$$

$$Y(t,T) = -\frac{\ln P(t,T)}{T-t} = -\sum_{i=1}^{2} \frac{A_i(\tau) + B_i(\tau)x_i - \alpha_i\tau}{\tau}, \quad \tau = T-t$$

(2-80)

We emphasize that the ability to fit the observed yield curves of course increase with the number of imbedded state variables (at least in in-sample analyses). However this will be at the cost of extra parameters to estimate. The parameterization of the default-free term structure, using two factors, is in line with the evidence presented by Chen and Scott (1993) and the findings of Pearson and Sun (1994). They conclude that one-factor models do not adequately characterize the actual yield curves. Using only one-factor to describe the term-structure dynamics implies that changes in the yield curve at different maturities are perfectly correlated, which obviously is a
Part II: The Pricing of Defaultable Bonds

highly restrictive approximation. The assumption implies that these models cannot capture the
dynamics of both the curvature as well as the slope. In addition, two- and three-factor models both
perform reasonable well, although the fit using three-factor models are marginally better.

Including two additional terms, $\beta_1$ and $\beta_2$, to account for the possible relation to the Treasury term
structure, the process for the credit spread $h_t$ is assumed to follow a translated single-factor
square-root process given by

$$h_t = \alpha_h + \tilde{h}_t + \beta_1 x_{1,t} + \beta_2 x_{2,t}$$  \hspace{1cm} (2-81)

Employing the RMV recovery rule (Duffee uses the RT rule), the price of corporate zero-coupon
bonds are then expressed as

$$\tilde{P}(t, T, 0, \phi) = \mathbb{E}_t^Q \left[ \exp \left( -\int_t^T (r(s) + h(s)(1-\phi))ds \right) \right]$$

$$= \mathbb{E}_t^Q \left[ \exp \left( -\int_t^T \left( (\alpha_r + x_r(s)) + x_2(s) + (\alpha_h + \tilde{h}(s) + \beta_1 x_{1,t} + \beta_2 x_{2,t})(1-\phi) \right) ds \right] .$$  \hspace{1cm} (2-82)

The third and fourth argument refers to coupons and recovery, respectively. One remark is in or-
der. The dependence, measured in terms of the two beta coefficients, between the hazard and the
interest rates is restricted such that it is only possible for the interest rates to influence the hazard
and not the opposite, through a two step estimation procedure. To avoid repetition we do not sub-
stantiate this aspect until later.

If we rearrange in terms of constants, hazard and default-free state variables, equation (2-82) can
be rewritten as

$$\tilde{P}(t, T, 0, \phi) = \mathbb{E}_t^Q \left[ \exp \left( -\int_t^T (\alpha_r + \alpha_h(1-\phi))ds \right) \right] \prod_{i=1}^n \mathbb{E}_t^Q \left[ \exp \left( -\int_t^T (1 + \beta_i(1-\phi))x_i(s)ds \right) \right]$$

$$\mathbb{E}_t^Q \left[ \exp \left( -\int_t^T \tilde{h}(s)(1-\phi)ds \right) \right].$$  \hspace{1cm} (2-83)

Now define
Part II: The Pricing of Defaultable Bonds

\[ g_t = (1 - \phi) \tilde{h}_t \]  
\[ z_{i,t} = (1 + \beta_i (1 - \phi)) x_{i,t}, \quad i \in 1, 2. \]  

(2-84a)  
(2-84b)

Utilizing Ito’s Formula, \( g_t \) and \( z_t \) may be transformed into

\[
dg_t = d(1 - \phi) \tilde{h}_t \\
= (1 - \phi) d\tilde{h}_t \\
= (k\tilde{\theta}^z - (k + \lambda) g_t) dt + \sigma^z \sqrt{g_t} d\tilde{w}^0_t
\]

\[
dz_{i,t} = d(1 + \beta_i (1 - \phi)) x_{i,t} \\
= (1 + \beta_i (1 - \phi)) dx_{i,t} \\
= (1 + \beta_i (1 - \phi)) \left[ (k_i \tilde{\theta}_i - (k_i + \lambda_i) x_{i,t}) dt + \sigma_i \sqrt{x_{i,t}} d\tilde{w}^0_{i,t} \right] \\
= (k_i \tilde{\theta}_i - (k_i + \lambda_i) z_{i,t}) dt + \sigma_i \sqrt{z_{i,t}} d\tilde{w}^0_{i,t}, \quad i \in 1, 2.
\]

where the adjusted \( \theta \) and \( \sigma \) are defined as

\[
\tilde{\theta}_i^z = \theta_i (1 + \beta_i (1 - \phi)), \quad \sigma_i^z = \sigma_i \sqrt{(1 + \beta_i (1 - \phi))}, \quad i \in 1, 2
\]

\[
\tilde{\theta}^z = \theta (1 - \phi), \quad \sigma^z = \sigma \sqrt{(1 - \phi)}.
\]

The expression of the bond price is then written accordingly to these new processes

\[
P(t, T, \phi) = \delta \prod_{i=1}^{2} E^Q_i \left[ \exp \left( -\int_t^T z_i(s) ds \right) \right] E^Q_i \left[ \exp \left( -\int_t^T g(s) ds \right) \right]
\]

(2-85)

where

\[
\delta = \exp \left[ -\tau (\alpha_\tau + \alpha_\delta (1 - \phi)) \right].
\]

For tractability we do not pursue the investigation of the double integral in (2-78). Instead we approximate the price, assuming that the coupons fall discretely, and that coupon payments are paid semi-annually, amortized after a serial loan principle. Using the same logic as when evaluating default-free bonds, arbitrage arguments reveal that the price of a coupon bond is the sum of the prices of the individual coupon payments and the principal. This is a very important assumption that exploits the condition of no default prior to coupon payments (this is fulfilled in the current framework).
The price of the coupon paying defaultable bond is expressed as

\[ \tilde{P}(t,T,C,\phi) = \mathbb{E}_t^{\phi} \left[ \int_t^T \tilde{P}(t,u,0,\phi)c(u)du + \tilde{P}(t,T,0,\phi)H \right] \]

\[ \approx \mathbb{E}_t^{\phi} \left[ \sum_{i=1}^{m} \tilde{P}(t,t_i,0,\phi)C(t_i) + \tilde{P}(t,T,0,\phi)H \right] \]

where

\[ \tilde{P}(t,t_i,0,\phi) = \delta \exp \left( -\int_{t_i}^{t_i+\Delta} (g(s) + z_i(s) + z_1(s))ds \right), \Delta = t_i - t_{i-1} \]

and \( m \) is the number of coupon payments. As all variables are affine in the state variables the price of the defaultable coupon bond is now computed by inserting the expressions for \( A(\tau) \) and \( B(\tau) \) into (2-86). Thus, the full expression of the bond price is written as

\[ \tilde{P}(t,T,C,\phi) = \sum_{i=1}^{k} \tilde{P}(t,T_i,0,\phi)C_i + \tilde{P}(t,T,0,\phi) \]

(2-87)

where

\[ \tilde{P}(t,T_i,0,\phi) = \delta \prod_{i=1}^{k} \exp \left( A_i^c(\tau_i) + B_i^c(\tau_i)z_{i_j} \right) \exp \left( A_i^g(\tau_i) + B_i^g(\tau_i)g_{i_j} \right) \]

where the superscript \( z \) and \( g \) refers to the respective processes.

In addition to recovery considerations we saw from chapter 1.3 that proper management of recovery requires as a minimum a stochastic recovery function in the bond price model. Because of the problems emerging when assuming stochastic recovery (discussed in section 2.2.4.4) we are drawn to the computational appealing nature of a static recovery restriction. However, when employing this restriction it certainly only applies within a certain rating class. Hereby we hypothesize that the optimal setting would be to state that the constant recovery at least varies between ratings. As an indication hereof Moody’s Investors Service (2001) illustrate that the mean recoveries for AAA-BBB rated corporate bonds from 1981 to 2000 were between 30% and 60% of the US$ par value.\(^{40}\)

To the authors knowledge there is no public available literature on recovery at market values, so the recoveries for the model will to a wide extent be guessing. Nevertheless, we expect the differences in RF estimates across ratings to be present in the RMV as well, and we furthermore believe that the market value of bonds just prior to default on average are lower that par value. So, we pragmatic assume that RMV amounts to AAA (70%), AA (65%), A (60%) and BBB (55%).

\(^{40}\) See Elton, Gruber, Agrawal and Mann (2001).
Part II: The Pricing of Defaultable Bonds

2.4 Interpreting the Properties of the Model

With small modifications the state-space design outlined has become very popular in the econometric literature - see e.g. Duffee (1996, 1999), Duffie and Singleton (1999) and Bakshi et al. (2000, 2001). By considering the model’s time-series and cross-section properties separately, the intension of this chapter is essentially to give a deeper understanding of the model. We explain why such term-structure models are attractive in describing the dynamics of credit spreads discussed in Part I. Thus, we primarily focus on the intensity formulation, and discuss how the data used for our analyses support this specification. Besides discussing the fundamental properties and design of the model, we explain the importance of the constant in both the interest and intensity formulation, and why it is expected to be negative in the former and positive in the latter. Furthermore, we explain the rationality of using two state variables in replicating the interest rate dynamics and only one for the intensity.

2.4.1 Fundamental Properties

The specification of $h_t$ is designed to capture various important features of the credit spreads. First of all the hazard rate is adopted in a Cox-process context. The obtained yield spreads will therefore be stochastic fluctuating with the financial health of the firm (hopefully), which, as argued elsewhere is very important for realistic default models. Secondly, the process for the hazard rate allows for a minimum spread given by the constant term $\alpha$. This implies that, even at the short end of the yield curve, and regardless of the financial health of the firm, there is some level below which credit spreads cannot fall. Thirdly, the model allows the default process to incorporate the important correlation with the innovations in the default-free term structure, through the coefficients $\beta_1$ and $\beta_2$, and still provides closed-form solutions. The negative values of $\beta_1$ and $\beta_2$ imply that investor requires a lower credit adjustment when higher levels of the Treasury term structure are observed as it indicates improvements in business climate.

The dynamics of the state variables under the physical and equivalent Martingale measure are as mentioned earlier specified as

\[
\mathbb{P}: \quad dh_t = \kappa\left(\theta - \bar{h}\right)dt + \sigma \sqrt{h_t}dw_t, \\
\mathbb{Q}: \quad \tilde{d}h_t = \kappa^*\left(\theta^* - \tilde{h}\right)dt + \sigma \sqrt{h_t}dw^*_t
\]

where all parameters are time invariant scalars, and $\kappa^* = \kappa + \lambda$ and $\theta^* = \kappa \theta / (\kappa + \lambda)$ i.e. both processes are mean reverting, and entail the attractiveness of precluding negative state variables. Stated differently, with the square root, whenever $\bar{h}$ gets near zero, the volatility term vanishes,
transferring the whole power to the mean-reversion behavior of the drift term. Thus, the behavior of the state variable implied by this structure has the following empirically comfortable properties:

- Negative values of the state variable, and thereby the interest rates and credit spreads, are precluded.
- If the state variable reaches zero, it can subsequently become positive.
- The absolute variance of the state variable increases when the state variable itself increases.

These features are very important arguments for using the CIR model instead of e.g. the standard Gaussian model. Negative hazard rates imply negative credit spreads, indicating that defaultable bonds are priced above Treasuries. This is of course inconsistent with the no-arbitrage condition. However, in the present specification of the hazard rate negative values are not precluded, because $\alpha_\kappa$ and $\beta_i$ are not bounded by zero. Hence the attractive feature of the CIR model vanishes. It may therefore be seen as a poor model specification, as well as, the choice of two factors for describing the Treasury dynamics and only one factor for the credit spread also appears pragmatic. Intuitively, it is therefore tempting to exclude the constant and apply two factors for the credit spread as well. Yet, the use of two factors for the Treasury dynamics has to be evaluated in relation to the empirically observed properties. Even though it might seem worthwhile to apply two factors for describing the credit spreads, several well-founded arguments support the current use of only one factor.

2.4.1.1 Cross-section Properties

The state-variable process obviously determines the curves, and formally we can divide the shapes of the yield curves accordingly to the sign of $\kappa + \lambda$. To address the properties of the CIR model consider the yield curves generated in situations of positive $\kappa + \lambda$ and situations of negative $\kappa + \lambda$. Figure 2-2 depicts the continuously compounded yield produced by the one-factor CIR model in the normal situation - that is $\kappa + \lambda > 0$.

---

41 See Cox, Ingersoll & Ross (1985) for the details.
Figure 2-2: Yield Curves Generated in a One-Factor Square Root Model, \(\kappa + \lambda > 0\)

Starting in the front of the figure, it appears that the yield curves are monotonically increasing across maturity. In contrast the yield curves in the back of the figure are monotonically decreasing, and humped for specific values in between. Furthermore the figure reveals that the term structures converge to a long term limit independent of the initial value of the state variable.

In particular it can be shown, that there exists an asymptotic limit of \(Y(\tau)\) as \(t \to \infty\), specified as

\[
Y(\infty) \equiv \lim_{t \to \infty} Y(x, t) = \frac{2\kappa \theta}{\kappa + \lambda + \gamma} \quad \text{or} \quad \frac{2\kappa \theta}{\kappa + \gamma} \tag{2-88}
\]

Combined with an upper limit defined as \(\kappa \theta / (\kappa + \lambda)\), the asymptotic limit in relation to the level of the state variable defines the yield curves in Figure 2-2. Specifically, the term structure will be monotonically increasing if the state variable is below \(Y(\infty)\), while it is monotonically decreasing if it is above \(\kappa \theta / (\kappa + \lambda)\), and humped for values in between.\(^\text{43}\) Equation (2-88) further reveals why the long-term yield is independent of the state variable, as it is determined by the parameters of the state variable, and not the current value itself. As a result a change in the state variable is therefore more pronounced in the short end and will gradually diminish with maturity. In contrast, the effect of changes in \(\theta\) will have greater impact on the longer maturities. E.g. using the same parameters as above, a 50 bp increase in the current rate will cause the yield curve to increase

---

\(^\text{42}\)Here we focus on the original CIR model and not the translated version of Duffie i.e. this relationship it stated without the influence of the constant.

\(^\text{43}\)See Cox, Ingersoll & Ross (1985) for further details.
Part II: The Pricing of Defaultable Bonds

approximately 45 bp in the short end (0.5 years) while only 7 bp in the long end (15 years). In contrast, changes in the unconditional mean have approximately the opposite effect.

These features of the CIR model have long been popular in describing interest rate dynamics, and combined with additional factors it will be possible to produce a wider scope of shapes including double humped term structures. It is not the objective here to examine the joint properties of the multifactor models – see Chen and Scott (1993) for a thorough discussion. However it is important to notice that in contrast to Treasury term structures, credit spreads for investment grade bonds are normally expected to increase with maturity and very infrequently expected to invert. In relation to this thesis one could therefore argue that the essential objective of the factor describing the credit spread is to capture the inherent term premia. The factors describing the Treasury dynamics must additionally be capable of producing inverting term structures. In this context Duffie and Singleton (1997) argue that in the affine setup the two-factor model cannot simultaneously capture inverting Treasury term structures and the level of zero-coupon yields without the existence of a negative constant. To clarify this aspect, remember the yields generated by the affine model cf. (2-39). Even when the state variables are zero the positive \( A_i(\tau) \) constitute to an upward sloping term structure. Matching inverting term structures therefore requires a large value of \( B_i(\tau)x_i \) to circumvent the inherent upward slope. Duffie and Singleton (1997) state that this subsequently large value tends to overstate the level of zero-coupon yields therefore a negative constant is in order.

In the same sense, a negative constant is not required when focusing on the term structure of credit spreads of investment grade bonds as they typically are non-inverting. However, it is important to emphasize that credit spreads in the above settings are poorly described. On the other hand if we consider the situation where \( \kappa + \lambda \leq 0 \) then \( \kappa \theta / (\kappa + \lambda) \) becomes negative, and the state variable will for reasonable parameters (\( \kappa, \theta > 0 \)) always be above its unconditional mean \( \theta^* = \kappa \theta / (\kappa + \lambda) \). The resulting term structures will then either be increasing or humped, and the slope of the yield curve will in contrast to the situation where \( \kappa + \lambda > 0 \) increase with the state variable (at least in the beginning) - see Figure 2-3.
Part II: The Pricing of Defaultable Bonds

Once again, the figure depicts low values of the state variable in the front and high in the back. Translated into an economical understanding, the front reflects times of low credit risk and the back reflects time of recession.

It appears that the model generates essentially flat yield curves across maturities when default rates are low, and as we move towards times of high credit risk, the slope of the yield curves gradually increases.

Figure 2-3: Yield Curves Generated in a One-Factor Square Root Model, $κ + λ < 0$

![Yield Curves Figure]

Note: $κ = 0.1$, $θ = 0.001$, $σ = 0.1$, $λ = -0.2$

If we recall the empirical properties of credit spreads, these features are very important, hereby favoring the use of the CIR model as the process that governs the credit spreads. But, even though $κ + λ < 0$ is important in capturing the variety of credit-spread slopes there are two aspects to be aware of. Firstly, the negativity of $κ + λ$ induces fairly steep credit curves even for low levels of the state variable. Although this is not a problem in times of optimistic business climates this aspect verifies the role of the constant term $αₗ$. As credit curves do typically not diminish completely as time to maturity goes to zero, neglecting the constant term implies that $\hat{h}_j$ and $θ$ would have to absorb the size of the minimum spread in order to keep $h_n$ and expectations of future values of $h$, fixed. Even small increases in either $\hat{h}_j$ or $θ$ do, however, affect the yield slope significantly and the yield curves will quickly become inconsistent with empirical findings. The role of a positive constant, $αₗ$, is therefore essential, as it dampens the overall steepness of the credit curve. Combined with $κ + λ < 0$ it enables the model to simultaneously generate low, nearly flat term structures and steeper credit curves when default rates are high.

Secondly, empirical studies have shown that using CIR models in describing the credit-spread dynamics typically involves $κ$ parameters close to zero. From (2-88) we see that the long-term rate goes to zero as $κ$ goes to zero, which causes a potential problem in both the cross-section and time-series properties of the model. In worst case the model will become inconsistent with the
observed credit curves at the long end as they become downward sloped in this setting. So, we do not find comfort in the model when explaining the dynamics of long maturity bonds. Therefore, we preliminary expect the model to be limited to investment grade bonds of shorter maturities. The problem concerning the time-series properties it that the instantaneous default rate will not only be non-stationary but explosive if \( \kappa + \lambda < 0 \). This is essentially problematic in relation to the derived inference and also a potential problem when simulating bond prices. We return to these aspects when conducting the empirical investigation in Part III.

2.4.1.2 Fitting the Observed Term Structures and Excess Yields

As noted by Duffee (2000) a successful term-structure model should not only be consistent with the shapes observed in the data, but also with the empirically observed time-series patterns in expected returns to bonds (excess yields). In theory we know that the cross-sectional and time-series behavior of the term structure must be linked in an internally consistent way to avoid arbitrage opportunities. Unfortunately this restriction entails that fitting the observed term structures will to a wide extend be at the cost of not capturing the empirical behavior of the excess yields. In the following we address this trade-off. In summery this concerns the aspect of precluding negative state variables and the model’s ability to capture the volatility in both the short and long end (fitting the term structure) and producing the appropriate excess yields.

The key restriction in the CIR model is that investors’ compensation for risk is inherently related to the underlying uncertainty i.e. it is proportional to the variance of the state variable. This structure ensures exclusion of arbitrage as the risk compensation goes to zero if the risk goes to zero. But this also implies that the excess yield cannot switch sigh over time. In order to uphold the empirically observed fraction of mean excess yield to its standard deviation the one-factor model therefore places restrictions on the admissible values of the state variable. To see this clearly remember the definitions of the short rate, \( r = \alpha + x \), and the market price of risk, \( \lambda(x) = \lambda \sqrt{x/\sigma} \), in the one-factor CIR model.

From (2-14) we know that the expected instantaneous return under the pricing measure is specified as

\[
E^Q \left[ \frac{dP}{P} \right] = \left[ r + \alpha(x) \right] dt
\]

\[\text{---}\]

\[44\text{Despite the dispersion in the empirical findings, especially when default rates are high, humped yield-spread curves are only reported in ratings lower than BBB.}\]
Part II: The Pricing of Defaultable Bonds

where \( \alpha_t(\tau) \) denotes the instantaneous excess yield, \( B(\tau) \sqrt{\lambda(x)} \). Using the definition of the market price of risk, we can rewrite the excess yield as

\[
\alpha_t(\tau) = B(\tau) \lambda x
\]

(2-89)

which is an affine function of \( x \). If we as Duffee (2002) derive the mean over standard deviation of the excess yield, the restrictions become clear. Manipulate the fraction as follows

\[
\frac{E[\alpha_t(\tau)]}{\sqrt{Var[\alpha_t(\tau)]}} = \frac{B(\tau) \lambda E(x_t)}{\sqrt{\left[E(B(\tau) \lambda x_t - B(\tau) \lambda E(x_t))^2\right]}} = \frac{B(\tau) \lambda E(x_t)}{\sqrt{\left[E(x_t - E(x_t))^2\right]}}
\]

(2-90)

In the same manner we derive the mean over standard deviation for the excess credit spread, which is analogous to (2-90), and Table 2-1 summarizes the descriptive statistics for AA-rated corporate bonds. As expected, the mean excess credit spread increases with maturity, which is interpreted as increasing term premia. More interesting is the relatively steady fraction of mean over standard deviation across maturity (approximately 1.37), which reveals important information about the one-factor model.

Table 2-1: Summary Statistics for Excess Credit Spread, AA-rated bonds, 1994 – 2002

<table>
<thead>
<tr>
<th>Excess Yield</th>
<th>Mean</th>
<th>STD</th>
<th>Mean/STD</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha(2) )</td>
<td>0.113</td>
<td>0.090</td>
<td>1.26</td>
</tr>
<tr>
<td>( \alpha(4) )</td>
<td>0.117</td>
<td>0.096</td>
<td>1.22</td>
</tr>
<tr>
<td>( \alpha(6) )</td>
<td>0.188</td>
<td>0.131</td>
<td>1.43</td>
</tr>
<tr>
<td>( \alpha(8) )</td>
<td>0.246</td>
<td>0.173</td>
<td>1.42</td>
</tr>
<tr>
<td>( \alpha(10) )</td>
<td>0.347</td>
<td>0.231</td>
<td>1.50</td>
</tr>
</tbody>
</table>

This table presents the sample means and standard deviations (STD) of the excess credit spread of AA-rated corporate bonds defined as the excess spread of 2, 4, 6, 8 and 10-years credit spread over the credit spread of a one-year bond. The sample contains 412 weekly observations in the period between January 1994 and January 2002.

Recall that the variance of the instantaneous credit spread and the state variable are equal in the one-factor model. Using the unconditional mean (0.619%) and standard deviation (0.212%) of the instantaneous credit spread of AA-rated bonds, which we approximate by the one-year credit spread, reveal information about the state variable. Inserting the standard deviation into (2-90) gives an estimate of \( \theta \) of 0.290%, which corresponds to a positive minimum instantaneous credit spread of \( \alpha_h = 0.329\% \). As our sample of the instantaneous credit spread varied from a low of
Part II: The Pricing of Defaultable Bonds

0.282% to a high of 1.293% the implied range of \( \tilde{h} \) is -0.05 to 0.96. This is very positive as it almost obeys the non-negative restrictions in the CIR model. We interpret this as an indication of that the one-factor CIR model is capable of capturing the observed term structures and the excess credit-spread dynamics while including a positive constant, \( \alpha \). Similar analyses are conducted for the credit spreads of bonds of AAA-BBB ratings and the results appear in Appendix C. The conclusions are unanimous except for the AAA-rated bonds but still these results are encouraging for the further investigation.

Considering the analysis of excess Treasury yield dynamics Table 2-2 summarizes the descriptive statistics. Again we find evidence of increasing term premia and steady fractions of mean over standard deviation across maturity (about 1.34).

**Table 2-2: Summary Statistics for Selected Excess Treasury Yields, 1994 – 2002**

<table>
<thead>
<tr>
<th>Excess Yield</th>
<th>Mean</th>
<th>STD</th>
<th>Mean/STD</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha(3) )</td>
<td>0.46</td>
<td>0.34</td>
<td>1.35</td>
</tr>
<tr>
<td>( \alpha(5) )</td>
<td>0.64</td>
<td>0.48</td>
<td>1.33</td>
</tr>
<tr>
<td>( \alpha(7) )</td>
<td>0.79</td>
<td>0.59</td>
<td>1.34</td>
</tr>
<tr>
<td>( \alpha(10) )</td>
<td>1.01</td>
<td>0.77</td>
<td>1.32</td>
</tr>
</tbody>
</table>

This table presents the sample means and standard deviations (STD) of the excess zero-coupon yields defined as the excess yields of 3, 5, 7 and 10-year bonds over the one-year bond i.e. \( \alpha(10) \) is calculated as the difference between the yield of the ten-year zero-coupon bond and the one-year zero. The sample contains 412 weekly observations in the period between January 1994 and January 2002.

By using the unconditional mean (5.06%) and standard deviation (1.04%) of the instantaneous interest rate, which we approximate by the one-year zero-coupon yield, we disclose following information. Inserting the standard deviation into (2-90) gives an estimate of \( \theta \) of 1.39%, which corresponds to a minimum instantaneous interest rate of \( \alpha \_r = 3.67\% \). But in our sample the short interest rate varied from a low of 1.80% to a high of 7.18%, which implies a range of \( x \) from -1.87 to 3.51. If we focus on capturing the observed term structures, and apply a negative \( \alpha \_r \), then the state variable will always be positive but this is at the cost of ignoring the excess yield dynamics. So, as a result the one-factor Treasury model cannot simultaneously capture the properties of the short rate volatility, the excess yields and the observed term-structure shapes. Although this problem carries over to multifactor models, Duffee (2002) argues that a two-factor model is better at fitting the behavior of expected excess bond yields. Anyhow, the two-factor model does not to a satisfactory extend capture this behavior so we are faced with a choice. The answer lies in the inherent properties of the excess yield formulation cf. (2-89). When \( \kappa + \lambda < 0 \) then the \( B(\tau) \) quickly becomes very negative, and the excess yields subsequently become unrealistic high. We
saw that under the pricing measure an explosive structure of the hazard rate was necessary in capturing the cross-section properties of the credit-spread curves. Therefore we a priori know that we cannot capture the excess yields of corporate bonds. As a consequence we focus on capturing the term structure in both the Treasury and hazard rate formulation.

2.4.2 Specification of the SDE in the CIR setup

We saw earlier that using the CIR model as the process that governs the state variables places tight restriction on their possible forms, as the CIR model implies a linearly drift with stochastic shocks driven by a square root diffusion. Furthermore we saw that these restrictions in theory renders mathematical tractability but this is useless if the underlying data does not support this structure. As a consequence this section investigates whether the conditions for applying the CIR model is reasonably obtained in our data. This is of course the main essence of Part III. However, the forthcoming analyses reveal important information, which serves as inputs when interpreting the results in Part III.

2.4.2.1 Investigating the Drift Function

To address the question we proceed as Chapman and Pearson (1999) and employ a weighted least squares procedure\textsuperscript{45} to estimate the drift function. To conduct an isolated analysis we restrict the diffusion to that of the CIR model and estimate the simple discretization

\[
dr_{t+1} = (\alpha_0 + \alpha_1 r_t + \alpha_2 r_t^2 + \alpha_3 / r_t) dt + \beta_0 r_t^\beta dw.
\]

(2-91)

Recognizing the possible problem of multicollinearity the regressors in (2-91) are transformed so that they are the incremental effect of adding the non-linear terms.\textsuperscript{46}

\textsuperscript{45} Conduct a first-stage ordinary least squares to obtain the residuals from the drift function and use non-linear least squares to estimate the diffusion function. Then, given this fitted function, re-estimate the drift parameters using weighted least squares.

\textsuperscript{46} E.g. if we want to obtain the incremental effect of adding another variable we have to exclude its correlation with the other variables i.e. we regress the new variable on the present regressors and obtain the residuals. These residuals constitute the transformed regressor, which is independent.
Part II: The Pricing of Defaultable Bonds

Table 2-3: WLS Estimation of the Drift Function, Treasury and Credit Spreads, 1994-2002

<table>
<thead>
<tr>
<th>Parameters</th>
<th>α₀</th>
<th>α₁</th>
<th>α₂</th>
<th>α₃</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treasury</td>
<td>0.0167</td>
<td>-0.0050</td>
<td>-0.0059</td>
<td>0.0161</td>
</tr>
<tr>
<td></td>
<td>(0.61)</td>
<td>(-0.93)</td>
<td>(-1.09)</td>
<td>(3.12)</td>
</tr>
<tr>
<td>Credit Spread</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AAA</td>
<td>0.064</td>
<td>-0.123</td>
<td>0.004</td>
<td>0.049</td>
</tr>
<tr>
<td></td>
<td>(4.98)</td>
<td>(-4.97)</td>
<td>(0.15)</td>
<td>(2.00)</td>
</tr>
<tr>
<td>AA</td>
<td>0.037</td>
<td>-0.057</td>
<td>-0.009</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>(3.18)</td>
<td>(-3.11)</td>
<td>(-0.50)</td>
<td>(2.27)</td>
</tr>
<tr>
<td>A</td>
<td>0.020</td>
<td>-0.024</td>
<td>-0.014</td>
<td>0.042</td>
</tr>
<tr>
<td></td>
<td>(1.77)</td>
<td>(-1.46)</td>
<td>(-0.89)</td>
<td>(1.38)</td>
</tr>
<tr>
<td>BBB</td>
<td>0.019</td>
<td>-0.012</td>
<td>-0.023</td>
<td>0.015</td>
</tr>
<tr>
<td></td>
<td>(1.11)</td>
<td>(-0.74)</td>
<td>(-1.52)</td>
<td>(0.56)</td>
</tr>
</tbody>
</table>

The table illustrates the WLS-regression of the interest rate and credit spread processes estimated on the sample from January 1994 to 2002, where the instantaneous interest rate is approximated by the yield on the one-year zero-coupon Treasury and the instantaneous probability of default by the one-year credit spread of the respective ratings. Newey-West standard errors are used in calculating the reported t-statistics in parentheses.

As Table 2-3 shows the Treasury STRIP data indicates possible non-linearity in the short yield. The results are in agreement with Chapman and Pearson and there is some, albeit little evidence against a nonzero drift i.e. α₀ through α₂ are not significantly different from zero. In contrast the table reveals more desirable results when examining the credit spreads of A and BBB-rated bonds as this analysis indicates that the underlying credit spread data might support a linear drift structure.

2.4.2.2 Investigating the Diffusion Function

Turning to the question if the data supports the form of time varying volatility implied by the CIR model we exploit the discrete transformation in Chen, Karolyi, Longstaff and Sanders (1992).[^1]

\[
\Delta r_{i+1} = \alpha_0 + \alpha_1 r_i + \varepsilon_{i+1},
\]

\[
\mathbb{E}_i (\Delta r_{i+1}) = \alpha_0 + \alpha_1 r_i,
\]

\[
\text{Var}_i (\Delta r_{i+1}) = \mathbb{E}_i (\varepsilon_{i+1})^2 = \beta r_i.
\]

If the CIR model is rich enough to capture the dynamics found in the data we expect[^2]

[^1]: This enables us to consistently replicate the moments of the general SDE at our advantage. The α₀ and α₁ are not equivalent with θ and κ respectively because of our discrete observations. Recall that \( E_i [r_{i+\Delta} - r_i] = \theta \left( 1 - e^{-\kappa \Delta} \right) + \left( e^{-\kappa \Delta} - 1 \right) r_i \). Therefore \( \alpha_0 = \theta \left( 1 - e^{-\kappa \Delta} \right) \) and \( \alpha_1 = \left( e^{-\kappa \Delta} - 1 \right) \).

[^2]: This is argued in Aït-Sahalia (1996). To obtain the standardized residuals we conduct a feasible generalized least squares where the residuals are computed from a first-stage ordinary least squares of the discrete-time transformation. The squared residuals from the first-stage OLS are then regressed by non-linear least squares, with the discretization of the variance function, which gives the fitted diffusion function.
Part II: The Pricing of Defaultable Bonds

\[ \mathbb{E}_t \left( \frac{e_{t+1}}{\sigma(\Delta t_{t+1})} \right)^2 \]

to be homoscedastic or, as Chen and Scott (1993) hypothesize, that the sample variance of the standardized residuals should be close to one if the diffusion function is well specified. Figure 2-4 illustrates the time series of the standardized residuals for the Treasury and AA credit spread (results not shown are depicted in Appendix D). Obviously the series exhibit volatility clustering, and are therefore heteroscedastic.

Figure 2-4: Time Series of Scaled Residuals from the CIR Model

A) The CIR Diffusion (Treasury)  
B) The CIR Diffusion (credit spread)

The figure shows the time series for the scaled residuals of the one-year zero-coupon yield and the one-year AA credit spread. The estimations are based on weekly observations from January 1994-2002.

Still a more thorough analysis of the series in Table 2-4 unveils significant GARCH-effects in both series indicating a possible misspecification.

Table 2-4: GARCH-analysis of the scaled residuals, January 1994-2002.

<table>
<thead>
<tr>
<th>Volatility model</th>
<th>Treasury</th>
<th>Credit Spread</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARCH (1)</td>
<td>0.105</td>
<td>0.157</td>
</tr>
<tr>
<td></td>
<td>(4.50)</td>
<td>(2.91)</td>
</tr>
<tr>
<td>GARCH (1)</td>
<td>0.877</td>
<td>-0.007</td>
</tr>
<tr>
<td></td>
<td>(32.65)</td>
<td>(-0.06)</td>
</tr>
</tbody>
</table>

The table shows the volatility analysis (GARCH analysis) of the time series for the scaled residuals of the one-year zero-coupon yield and the one-year AA credit spread. The numbers in parentheses are the z-statistics.

Comparing the results across rating Appendix D discloses the general tendency of more significant volatility clustering as the credit quality decrease. This tendency is not only obtained from the GARCH-analyses as the different plots visualize this quite clearly. This analysis indicates that the time varying volatility found in the Treasury STRIP and the credit spread data is not adequately described by the square-root diffusion in the CIR model.
Part II: The Pricing of Defaultable Bonds

The outcomes of the analyses conducted in this section have to be considered when interpreting the results of the empirical investigation in Part III. Especially, this is important when considering the adequacy of the model through its accuracy performance and the inherent misspecifications analyzed in section 3.3.3.
Part III: Empirical Investigation

In this part we estimate the parameters of the model described in chapter 2.3. Although we are working in a theoretical pseudo economy constructed in a no-arbitrage framework it is crucial that this economy resembles the actual. If not, any sophisticated model based on this theoretical framework is bound to fail in replicating the true observed behavior of the term structures. In order to replicate the observed yields to the highest extend it is necessary to implement the estimation procedure on panel data, which is a combination of time series and cross sectional data. This choice is a reflection of i) using pure time series (e.g. CKLS (1992)) Ball and Torous (1996) shows that the volatility parameter is estimated accurately but the speed of adjustment coefficient is estimated with significant upward bias implying that the estimated yield curves converges far too quickly to the corresponding long-term rates. ii) Using pure cross-section Brown and Dybvig (1986) shows that κθ is estimated precisely but it is impossible to separately identify the involved parameters. Furthermore neither of the approaches are capable of estimating the market price of risk parameter, λ, and as a result the identified models cannot be used in pricing problems. If the two approaches are used in combination then it is possible to efficiently identify all the parameters, and therefore we estimate the model using panel data.

3.1 Estimation Methodology

In the previous part, we uncovered the link between the state variables and zero-coupon bond prices at a single point in time. In this section, we use the dynamics of the state variables to establish the link between yield curves through time. Recall the relationship between zero-coupon yields and the n state variables as

\[
Y(t,T) = \frac{-\ln P(t,T)}{\tau} = -\sum_{i=1}^{2} A_i(\tau) + B_i(\tau) x_i - \alpha \tau, \quad \tau = T - t
\]  

(3-1)

Then, if we select a set of k distinct maturities (\(\tau_1, \ldots, \tau_k\)) where \(k = n\), with the corresponding vector of yields at time t

\[
Y(t, T) = \begin{pmatrix}
Y(t, \tau_1 + \tau_1) \\
\vdots \\
Y(t, \tau_1 + \tau_k)
\end{pmatrix}
\]  

(3-2)

the relationship between the vector of yields and the state variables is given by
Part III: Empirical Investigation

\[ Y_k(t) = \begin{pmatrix} A \end{pmatrix} + \begin{pmatrix} B \end{pmatrix} X(t) \]  

(3-3)

where \( \begin{pmatrix} A \end{pmatrix} \) is a \( k \)-dimensional column vector and \( \begin{pmatrix} B \end{pmatrix} \) is a square \( k \times n \) matrix defined as

\[
\begin{align*}
\begin{pmatrix} A \end{pmatrix} &= (\begin{pmatrix} A - \alpha, I \end{pmatrix}) \begin{pmatrix} i \end{pmatrix}, \\
\begin{pmatrix} A \end{pmatrix} &= \begin{pmatrix}
-A_1(\tau_1)/\tau_1 & \cdots & -A_n(\tau_1)/\tau_1 \\
\vdots & \ddots & \vdots \\
-A_1(\tau_k)/\tau_k & \cdots & -A_n(\tau_k)/\tau_k
\end{pmatrix}
\end{align*}
\]

and

\[
\begin{align*}
\begin{pmatrix} B \end{pmatrix} &= \begin{pmatrix}
-B_1(\tau_1)/\tau_1 & \cdots & -B_n(\tau_1)/\tau_1 \\
\vdots & \ddots & \vdots \\
-B_1(\tau_k)/\tau_k & \cdots & -B_n(\tau_k)/\tau_k
\end{pmatrix}
\end{align*}
\]

where \( \begin{pmatrix} i \end{pmatrix} \) is the \( n \)-dimensional column vector of ones.

Given the observed yields, \( Y_k(t) \), and that \( \begin{pmatrix} B \end{pmatrix} \) is non-singular, equation (3-3) can be solved for the state variables

\[ X(t) = \begin{pmatrix} B \end{pmatrix}^{-1} \left[ Y_k(t) - \begin{pmatrix} A \end{pmatrix} \right]. \]  

(3-5)

This transformation is very convenient because the non-observable state variables are now described by the observable yields.

More specific, if the conditional distribution of the state variables is known and given by

\[ f_s \left( X_s \big| X_t \right) \quad \forall \ 0 \leq t \leq s, \]  

(3-6)

then the conditional distribution of the yields follows directly from a transformation of \( f_s \) (Greene 2000, pp. 76-77)

\[ f_s \left( Y_s \big| Y_t \right) = \mathcal{J} \ f_s \left( X(Y_s) \big| X(Y_t) \right) \quad \forall \ 0 \leq t < s \]  

(3-7)

where \( X(Y_s) \) describes the transformation from \( Y(t) \) to \( X(t) \) and \( \mathcal{J} \) is referred to as the Jacobian of transformation i.e.
Part III: Empirical Investigation

\[ J = \left| \frac{\partial X(Y_t)}{\partial Y_t} \right| \]

which is simply \( |B^{-1}| \) in the case of zero-coupon bonds. Given (3-7) and a set of observations at times \((t, ..., T)\) the log-likelihood function follows from the Markov property of the state variables

\[
\log L(Y(t) \mid \psi) = \sum_{i=1}^{T} \log f_{Y_i} (Y_i \mid Y_{i-1})
\]

(3-8)

where the vector of parameter solutions, \( \psi \), is obtained by maximizing this equation.

If this log-likelihood model is well specified and the yields are observed without error, the choice of maturities would not matter because any choice would imply the same time series for the state variables cf. Fisher & Gilles (1996). But it is not strictly tenable to assume that the yields are observed without errors and the model is not necessarily well specified. This could lead to different results when different maturities are used. Since there is information contained in each of the yields, efficiency is upheld by using more yields than necessary to infer the state variables. Chen & Scott (1993) implement this by assuming that additional yields are measured with error, which provides a convenient framework.

Suppose there are \( m \) additional yields that are measured with error, besides the \( k \) yields there are observed without error, then, if the conditional distribution of measurement errors is given by

\[
f_u(U_s \mid U_t) \quad \forall \ t \leq s
\]

(3-9)

where \( U_t \) is the vector of measurement errors, the new log-likelihood function is

\[
\log L(Y(t) \mid \Psi) = \sum_{i=1}^{T} \log f_{Y_i} (Y_i \mid Y_{i-1}) + \sum_{i=1}^{T} \log f_{U_i} (U_i \mid U_{i-1})
\]

(3-10)

where the first density function depends on the specific affine model and the second density function depends on the assumptions imposed on the measurement errors. When including measurement errors it is important to acknowledge that if the sum of state variables and errors exceeds the
Part III: Empirical Investigation

number of observed bond yields, then the resulting joint density function does not entail a convenient closed-form solution i.e. we implicitly employ \( m + n = k \) in the following derivation.\(^{49}\)

### 3.1.1 The Modified Chen & Scott Maximum Likelihood Estimator

To explicitly develop the ML estimator we have to employ the density function of the state variables and the measurement errors. From the assumption of independence between the involved state variables the conditional density distribution for the yields in (3-7) can be rewritten as

\[
f_{y}(Y_{s} | Y_{t}) = \mathbb{E}^{-1} \prod_{i=1}^{n} f_{x_{i,s}}(x_{i,s} | x_{i,t}) \quad \forall \ t \leq s
\]

where the conditional distribution of each state variable is the non-central \( \chi^2 \) -distribution

\[
f_{x_{i,s}}(x_{i,s} | x_{i,t}) = c_{i} \exp\left(-a_{i} - b_{i}\right) \left(\frac{a_{i}}{b_{i}}\right)^{q_{i} \theta_{i}} I_{q_{i}}\left(2\sqrt{a_{i} b_{i}}\right), \quad \text{for } i = 1, 2, \ldots, n \ \forall \ t \leq s
\]

where

\[
a_{i} = c_{i} x_{i,s},
\]

\[
b_{i} = c_{i} x_{i,t} \exp\left(-\kappa_{i} (s-t)\right),
\]

\[
c_{i} = \frac{2\kappa_{i}}{\sigma_{i}^{2} \left(1 - \exp\left(-\kappa_{i} (s-t)\right)\right)} \quad \text{and} \quad q_{i} = \frac{2\kappa_{i} \theta_{i}}{\sigma_{i}^{2}} - 1
\]

and \( I_{q_{i}}(\cdot) \) is the modified Bessel function of the first kind of order \( q_{i} \).

In deriving the estimator Chen & Scott (1993) allowed for serial correlation in, and contemporaneous correlation between the measurement errors. The serial correlation was modeled as the first-order autoregressive process

\[
u_{i,s} = \rho_{i} u_{i,t} + e_{i,s}, \quad e_{s} \sim N(0, \Omega) \quad \text{for } i = 1, \ldots, m \quad \forall \ t \leq s
\]

where \( m \) is the number of measurement errors, and \( \Omega \) is the covariance matrix for the joint normal distributed \( e_{i} \). Assuming correlation in and between the measurement errors is indeed tenable if our objective is statistical approval, since it is not plausible that these errors are independent in any model. However, implementing a correlated error structure does not come without a certain disadvantage. Remember, justifying the use of \( m \) errors require an estimate of the \((m \times m)\) covariance matrix. This entails the problem of estimating about twice as many parameters compared to

\(^{49}\) See Chen & Scott (1993).
Part III: Empirical Investigation

an assumption of independent errors. Therefore we base our estimation methodology with the
strict assumption of an independent and normally distributed error structure where

\[
f(u_{ij}) = \frac{1}{(2\pi\hat{\sigma}_{e,i,j}^2)^{1/2}} \exp\left(\frac{-1}{2\hat{\sigma}_{e,i,j}^2} e_{i,j}^2\right), \quad \text{for } i = 1, \ldots, m
\]  

(3-13)

and therefore the log-likelihood function can be written as

\[
\log L(Y(t) | \Psi) = \sum_{i=2}^{T} \sum_{j=i}^{m} \log f_i(x_{i,j} | x_{i-1,j}) - (T-1) \log |\mathbf{B}|
\]

\[
-\frac{(T-1)}{2} \left( m \log(2\pi) + \sum_{i=1}^{m} \log \hat{\sigma}_{e,i}^2 \right) - \frac{1}{2} \sum_{i=2}^{T} \sum_{j=i}^{m} \frac{e_{i,j}^2}{\hat{\sigma}_{e,i,j}^2}
\]  

(3-14)

Although it may not seem so obvious, this assumption eases our estimation considerably. The
choice between the two suggestions is a pragmatic reflection, and is therefore a fine balance be-
tween tractability, both mathematical and in the sense of larger degrees of freedom in the model,
and statistical approval. A final remark on the log-likelihood function in (3-14) compared to Chen
& Scott’s, is that we use the observations from \( t = 2 \) to \( T \) thereby excluding the first observation,
and the evaluation of the unconditional density distribution. As Duffee (1996) states, this involves
the property of not assuming stationary state variables i.e. we allow the \( \kappa + \lambda \) to become non-positive.

3.1.2 The Default-free Process

To keep the estimation relatively simple, we have chosen six different yields. The choice reflects
considerations about employing a reduced number of yields and still involving as much informa-
tion as possible while keeping a wide focus on the whole term structure. The final structure of the
time to maturity of the chosen yields was one, two, four, six, ten and fifteen years. This selection
gave manifestation to the following measurement error design
Part III: Empirical Investigation

\[
\log P(t, T_i) = -Y(t, T_i) \pi_k + \varepsilon_{1j},
\]

\[
\log P(t, T_2) = -Y(t, T_2) \pi_2 - \varepsilon_{1j},
\]

\[
\log P(t, T_4) = -Y(t, T_4) \pi_4 + \varepsilon_{2j},
\]

\[
\log P(t, T_6) = -Y(t, T_6) \pi_6 - (\varepsilon_{2j} + \varepsilon_{3j}),
\]

\[
\log P(t, T_{10}) = -Y(t, T_{10}) \pi_{10} + \varepsilon_{3j},
\]

\[
\log P(t, T_{15}) = -Y(t, T_{15}) \pi_{15} + \varepsilon_{4j}
\]

(3-15)

where

\[
Y(t, T_k) \pi_k = - \left( \sum_{i=1}^{k} A_i(t, T_k) + B_i(t, T_k) x_{1j} - \alpha \tau_k \right), \quad \tau_k = T_k - t
\]

is analogous to (3-1), and \( k \) is the chosen yield i.e. \( T_i \) refers to the one-year yield, etc.

In the above system of equations we assume that all yields are measured with error. This is in contrast to Chen and Scott (1993) who fit some points of the cross section precisely, i.e. choosing yields without error. This choice is a pragmatic weighing of either fitting one or two points on the Treasury curve precisely or fitting more points to the highest extent.

Although this description of measurement errors is ad hoc it is extremely convenient because it enables us to define two new yield-measures without error

\[
\begin{bmatrix}
Y_{1j} \\
Y_{2j}
\end{bmatrix} = - \left[ \frac{\sum_{k=[1,3]} \log P(t, T_k) \pi_{1j}}{\sum_{k=[4,6,10]} \log P(t, T_k)} \right] \left[ \frac{\sum_{k=[1,3]} Y(t, T_k) \pi_{1j}}{\sum_{k=[4,6,10]} Y(t, T_k) \pi_{1j}} \right]
\]

\[
= \left[ \frac{\sum_{k=[1,3]} A_k(t, T_k) + A_k(t, T_k) - \alpha \tau_k - t}{\sum_{k=[4,6,10]} A_k(t, T_k) + A_k(t, T_k) - \alpha \tau_k - t} \right] - \left[ \frac{\sum_{k=[1,3]} B_k(t, T_k) \sum_{k=[4,6,10]} B_k(t, T_k)}{\sum_{k=[1,3]} B_k(t, T_k) \sum_{k=[4,6,10]} B_k(t, T_k)} \right] \begin{bmatrix}
x_{1j} \\
x_{2j}
\end{bmatrix}
\]

(3-16)

As seen from (3-15) and (3-16) the fifteen-year yield does not enter any of the new yield-measures and as a consequence the yield only influences the parameter estimation through the contribution from the measurement errors. The reason why this yield is excluded is that the maturities of the corporate bonds, examined later in this thesis, are concentrated in the one to ten year ranges, hence to accurately extract default probabilities, it is more important to accurately fit the default-free term structure in this range than in the long end. But the long yield still holds valuable
information of the future term structure and, hence, cannot totally be neglected. With this relation between the yields and the state variables in hand it is straightforward to follow the path defined in the previous section to derive the log-likelihood function for the default-free parameter estimation.

3.1.3 The Default-risk Process
With little modification, the estimation of the process governing credit spreads is based on the methodology described in the previous section. The estimation procedure is carried out on the assumption that the parameters describing the default-free process are given, and therefore constant in the estimation of the credit-spread parameters.\(^{50}\)

From (2-57) we know that the credit spread is computed as

\[
s(t, T) = -\frac{1}{T-t} \ln \mathbb{E}^Q \left[ \exp \left( -\int_t^T h(s) ds \right) \right] = -\frac{1}{T-t} \ln (1 - Q(Z \leq T))
\]

However, when conducting the estimation there is a very important difference compared to the estimation of the Treasury term structure. The data used for estimating the default-free parameters exploit a panel of Treasury STRIPS, for which reason the bond price equation can be inverted to produce a closed-form solution for the state variables. This is not possible when estimating the parameters of the defaultable term structure, as the log of the bond price no longer is a linear function of the state variables. In particular, the yield to maturity on the defaultable coupon bond, \(\tilde{Y}(T, t)\), must be derived by solving the non-linear equation

\[
P_{\text{obs}}(t, T, C, \psi) - P_{\text{calc}}(t, T, C, \psi) = 0 \quad \Leftrightarrow (3-17)
\]

\[
P_{\text{obs}}(t, T, C, \psi) - \sum_{i=1}^k \exp \left( -\tilde{Y}(t, T) \tau_i \right) CF_i = 0
\]

\[
P_{\text{obs}}(t, T, C, \psi) - \sum_{i=1}^k \exp \left( -\left[ s(t, T) + Y(t, T) \right] \tau_i \psi \right) CF_i = 0
\]

given the coupon bond price and the promised cash flow stream \((CF_i)\) conditional on the default-free parameter vector \(\psi = \{\alpha, \kappa_1, \phi_1, \lambda_1, \sigma_1, \kappa_2, \phi_2, \lambda_2, \sigma_2\}\).

---

\(^{50}\)The estimates are rating specific, i.e. they are not restricted to be constant across the evaluated rating classes \((AAA - BBB)\).
Part III: Empirical Investigation

The available data on defaultable bonds are more limited, so we are somewhat restricted in defining an error structure. The available cross sectional elements consist of only six maturity segments c.f. section 1.3.2.1. However, as we only apply one-factor for the estimation it is of course not necessary to have as many cross-section observations in comparison to the estimation of the Treasury parameters. Nevertheless, it is problematic that we do not have any information of the short (< 1 year) end as well as the long end (> 15 years). The shortest maturities are very important in capturing the innovations of the instantaneous default probabilities, and long-term bond prices are important because they reflect expectations about the hazard rates into the distant future. Bearing these obstacles in mind the chosen system of equations for the credit spread is

\[
\log \tilde{P}_{\text{obs}}(t, T_1, C, \phi) = \log \tilde{P}_{\text{calc}}(t, T_1, C, \phi) \\
\log \tilde{P}_{\text{obs}}(t, T_3, C, \phi) = \log \tilde{P}_{\text{calc}}(t, T_3, C, \phi) + \epsilon_{1t} \\
\log \tilde{P}_{\text{obs}}(t, T_5, C, \phi) = \log \tilde{P}_{\text{calc}}(t, T_5, C, \phi) + \epsilon_{2t}
\]  

(3-18)

where \(k\) is the chosen yield i.e. \(T_1, T_3\) and \(T_5\) refers to the one-to-three, five-to-seven, and ten-to-fifteen years maturity segments. This approach in deriving the parameters of the implied hazard rate is in line with Duffee (1996). As it appears, we model the shortest maturity segment as a true yield, i.e. without measurement error, and the other log-bond prices are observed with \(N_{i\text{iid}}\) measurement errors. To examine the robustness of this assumption other error structures have been applied and will be commentated later.

As explained above, the price expression is non-linear, hence the default intensity, \(h_t\), cannot be found analytically. The transition density for the default intensity is therefore adjusted to cope with the presence of coupons, and is written as

\[
f_t(s | s_t) = J f_h(h_t | h_s) \quad \forall \ t \leq s
\]  

(3-19)

where \(f_h(h_t | h_s)\) is the appropriate analogue of (3-12), and the Jacobian of transformation \(J\) is computed as\(^{51}\)

---

\(^{51}\)Normally the Jacobian of transformation is the determinant of the matrix of first derivatives but as we only use one state variable in deriving the true yield \(\bar{J}\) is the absolute value of \(\partial h / \partial \log \tilde{P}\). As we follow Duffee \(\bar{J}\) is absolute value of \(\partial h / \partial \log \tilde{P}\).
The log-likelihood function is again partitioned into two separate parts expressed by the sum of the transition density of the hazard and the measurement errors

\[
\log L(\phi) = \sum_{t=2}^{T} \log f_{g_t|g_{t-1}} + (T-1) \log J
\]

\[
= -\frac{(T-1)}{2} \left( m \log(2\pi) + \sum_{j=1}^{m} \log \sigma_{\varepsilon_{ij}}^2 \right) - \frac{1}{2} \sum_{j=2}^{T} \sum_{i=1}^{m} \frac{\varepsilon_{ij}^2}{\sigma_{\varepsilon_{ij}}^2}
\]

Analogous to the estimated Treasury term structure, we do not preclude non-stationarity of \( g \), thus we exclude the first observation. The Maximum Likelihood estimates can then be found as the parameters \((\alpha, \kappa, \theta_1, \lambda_2, \sigma_\alpha, \beta_1, \beta_2, \sigma_{\varepsilon})\) that maximizes (3-21).

3.1.4 Implementation in GAUSS

The estimation procedures are implemented in GAUSS (Appendix H and Appendix I) using the module for constrained maximum likelihood (CML). Although it is not of primary interest to impose restrictions on the parameters, the CML module provides several options very convenient for the present needs.

Generally we encountered a convergence problem, which was mainly caused by difficulties in computing the modified Bessel function and the covariance matrix of the parameters i.e. the Hessian. To alleviate these problems we employed several techniques and numerical approximations to ensure that the Gradients did not diverge and to optimize iteration time. These technical applications are briefly commented below.

As mentioned above the estimation procedures are based on the true conditional densities of the state variables, hereby exploiting the full information of the probability density of the yields. However, the method proved very difficult to apply primarily due to the complexity in calculating the modified Bessel function. We therefore applied a Quasi Maximum Likelihood method (QML) as an approximation to the exact solution when the inherent conditions for evaluating the Bessel
function were not obtained. This was foremost the case when the parameter solutions were far from optimum, which in summary entailed that the QML-procedure was utilized in the beginning of the optimization when very little was known about the true estimates. As the solutions approached optimum the conditions for evaluating the modified Bessel function were obeyed and the program switched to the ML-procedure.

Our definition of the QML is particularly easy to implement as it approximates the true densities of the yields, by exploiting the information in the first and second conditional moments of the state variables and employing the normal distribution as the true density approximation. The moments are as given in Cox, Ingersoll and Ross (1985), which discloses following approximation

\[ f(x_i \mid x_{i-1}) = \frac{1}{(2\pi v_i)^{1/2}} \exp\left(-\frac{1}{2v_i} (x_i - \mu_i)^2\right), \quad \text{for } i = 1, 2 \]  
\[ \mu_i = E_{t-1}(x_i) = x_{i-1} \exp(-\kappa \Delta t) + \theta_i (1 - \exp(-\kappa \Delta t)) \]  
\[ v_i = Var_{t-1}(x_i) = x_{i-1} \left(\frac{\sigma_i^2}{\kappa_i}\right) \left(\exp(-\kappa \Delta t) - \exp(-2\kappa \Delta t)\right) + \theta_i \left(\frac{\sigma_i^2}{2\kappa_i}\right) (1 - \exp(-\kappa \Delta t))^2 \]  

This method is less exact but still a useful alternative when the conditions for evaluating the pure ML-method are violated.

Another aspect directly associated with the convergence of the derived log-likelihood expressions in (3-14) and (3-21) is the behavior of the Jacobian. If the MLE-inversion method does not render a full rank matrix then the inverse does not exist and the program terminates. To implement a simple solution we bounded the determinant of the inverted Jacobian to a small number \(10^{-3}\) if the inverse did not exist. Furthermore, if for some reason, the Jacobian provided negative state variables we imposed the value 0.001 as negative values conflict with the modified Bessel function in GAUSS.

Besides these specific technicalities a more general problem emerged as the likelihood surface revealed several local maxima whereat the iterative methods not necessarily converged to the true MLE. We therefore used several different choices of initial starting points, but unfortunately the programs proved very sensitive with regards to the choice of initial parameter values. In addition, we found the usual Newton-Rapson method inefficient; depending on the initial values the method either showed disproportional large fluctuations in the parameter estimates and gradients,
Part III: Empirical Investigation

or \(| \hat{\psi}^{(k+1)} - \hat{\psi}^{(k)} |\) never became sufficiently small to declare convergence – even for fairly high gradient tolerance \((10^{-3})\). A reasonable explanation of the failure of the Newton-Rapson method is that the criterion function, which is to be maximized, is not approximately quadratic or globally concave.\(^{53}\) As an alternative we employed the BFGS algorithm as gradient method. The BFGS algorithm eliminates the direct computation of the Hessian at each iteration, and even though this method implements more iterations than the Newton method we found it to converge in less overall time. The underlying methodology of the method is that it approximates the Hessian using an update algorithm hereby reducing the computational requirements considerably.\(^{54}\) However, two problems remain. As it accumulates an estimate of the Hessian, a small error has the possible effect of amplifying through the iterations, and although the method is very efficient, it therefore sometimes converges to a less precise estimate of the covariance matrix.\(^{55}\)

Eventually, we found it appropriate to employ a switching algorithm between the DFP and BFGH convergence algorithms whenever the change in the criterion function, or the line step length, was less than 0.0001. In retrospective the choice of switching convergence algorithms proved to be of less importance, when the initial starting values was closer to the true MLE. Finally, to stabilize the program we furthermore implemented a grid search method, which main objective was for GAUSS to find a new direction from which to continue the optimization when the line search methods failed.

3.1.5 Evaluation of the Models

Before turning to the results of the estimations we evaluate the convergence of the models. The intention is to reveal whether the developed maximum likelihood models produce biased parameter estimates. To do so, two Monte Carlo studies are performed, based on 1,000 data set simulations containing 500 observations each. To take the model sensitivity towards the initial parameter values into account, the estimation procedures were divided into batches of 20 with different initial parameter values. We hereby isolate the investigation of the ML-model’s bias, without noise from the model sensitivity towards the choice of the initial values.

In Table 3-1 the true (defined) and the mean estimated parameters are displayed for the translated two-factor model.

\(^{52}\)We could have chosen the value 1, which renders neutrality when calculating the natural logarithm, but then the program is not punished when deriving unusable Jacobians.

\(^{53}\)A very simple quadratic function is \(F(\theta) = a + b'\theta - 1/2\theta'C\theta\), where \(C\) is a positive definite matrix. Bearing this in mind it is probably not very tenable to expect that the investigated function is strictly quadratic.

\(^{54}\)For more details on the methods see Greene (2000), or the CML manual for GAUSS.

Table 3-1: Monte Carlo Estimates of the True Translated Two-Factor CIR-process

<table>
<thead>
<tr>
<th>Parameter</th>
<th>True</th>
<th>Mean</th>
<th>Std. Error</th>
<th>t test</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\kappa_1$</td>
<td>0.6000</td>
<td>0.5909</td>
<td>0.0322</td>
<td>-8.92</td>
</tr>
<tr>
<td>$\kappa_2$</td>
<td>0.1500</td>
<td>0.1525</td>
<td>0.0170</td>
<td>4.70</td>
</tr>
<tr>
<td>$\theta_1$</td>
<td>0.0400</td>
<td>0.0401</td>
<td>0.0052</td>
<td>0.31</td>
</tr>
<tr>
<td>$\theta_2$</td>
<td>0.0100</td>
<td>0.0106</td>
<td>0.0122</td>
<td>1.68</td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>-0.2500</td>
<td>-0.2489</td>
<td>0.0218</td>
<td>1.25</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>-0.1000</td>
<td>-0.0960</td>
<td>0.0494</td>
<td>2.57</td>
</tr>
<tr>
<td>$\sigma_1^2$</td>
<td>0.0200</td>
<td>0.0193</td>
<td>0.0096</td>
<td>-1.78</td>
</tr>
<tr>
<td>$\sigma_2^2$</td>
<td>0.0100</td>
<td>0.0095</td>
<td>0.0051</td>
<td>-2.79</td>
</tr>
<tr>
<td>Constant</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>-</td>
</tr>
</tbody>
</table>

The table displays the mean estimates from the Monte Carlo analysis of the translated two-factor model together with the true values. To evaluate the accuracy we report the standard errors and subsequently the t-statistics. To calculate the t-statistics remember to divide the standard errors by the square root of the number of data set simulations i.e. $\sqrt{1000}$ to get the standard error of the mean estimate.

As seen from the table the true and the mean estimates are not that far apart in absolute values although the difference is significant for half the parameters. The most interesting tendencies are revealed when focusing on the mean-reversion and volatility parameters as the other parameters are within a fair range of the true values. When looking at the speed of mean reversion the model significantly underestimate the coefficient for the first state variable and significantly overestimate the coefficient for the second state variable. In the same sense both volatility coefficients are underestimated. These problems arise because of the ad hoc specification of the true yields and measurement errors. If we instead defined the one-year yield as the first true yield in the model, we would expect a higher $\kappa_1$ and $\sigma_1^2$ as the shorter yields typically exhibit stronger mean reversion and higher volatility. In the same sense if we incorporated the fifteen-year yield in the second true yield, we would a priori expect a lower $\kappa_2$ but unfortunately also a lower $\sigma_2^2$.

In Table 3-2 an analogous investigation is conducted on the translated one-factor model, on simulated credit spreads. As illustrated the model gives a significantly downward biased estimate of the mean-reversion parameter. This fact is possibly due to the a priori chosen low value of $\kappa$, which inflicts the model with some difficulty, because the series is close to being non-stationary under the $\mathcal{Q}$-measure as $\kappa + \lambda$ is close to zero. On the other hand this bias could also originate from the implemented measurement errors where utilizing the longer matured yields as calibration points could inflict the mentioned downward bias. Furthermore we once again show that the ML-model experiences some struggles in replicating the inherent volatility.
Part III: Empirical Investigation

Table 3-2: Monte Carlo Estimates of the True Translated One-Factor CIR-process

<table>
<thead>
<tr>
<th>Parameter</th>
<th>True</th>
<th>Mean</th>
<th>Std. Error</th>
<th>t test</th>
</tr>
</thead>
<tbody>
<tr>
<td>κ</td>
<td>0.1500</td>
<td>0.1412</td>
<td>0.0412</td>
<td>-6.75</td>
</tr>
<tr>
<td>θ</td>
<td>0.0100</td>
<td>0.0110</td>
<td>0.0079</td>
<td>3.82</td>
</tr>
<tr>
<td>λ</td>
<td>-0.1000</td>
<td>-0.1019</td>
<td>0.0299</td>
<td>-1.95</td>
</tr>
<tr>
<td>σ²</td>
<td>0.0400</td>
<td>0.0382</td>
<td>0.0217</td>
<td>-2.54</td>
</tr>
<tr>
<td>β₁</td>
<td>0.1000</td>
<td>0.0899</td>
<td>0.0127</td>
<td>-1.49</td>
</tr>
<tr>
<td>β₂</td>
<td>-0.1000</td>
<td>-0.1095</td>
<td>0.0438</td>
<td>-1.79</td>
</tr>
<tr>
<td>Constant</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>-</td>
</tr>
</tbody>
</table>

The table displays the mean estimates from the Monte Carlo analysis of the translated one-factor model together with the true values. To evaluate the accuracy we report the standard errors and subsequently the t-statistics. To calculate the t-statistics remember to divide the standard errors by the square root of the number of data set simulations i.e. \( \sqrt{1000} \) to get the standard error of the mean estimate.

As a consequence of the chosen methodology the biases have to be taken into account, as they are important inputs and necessary when interpreting the forthcoming estimations.

3.2 Estimation Results

In this chapter we present the estimated parameters in the three-factor model, comment on, and evaluate the performance by use of standard statistics such as correlation, maximum and mean prediction error, as well as the more specific accuracy measures root mean squared error (RMSE) and Theil U. The latter is reported as the RMSE measure suffers from an obvious scaling problem, whereas the U statistic does not. Furthermore, to account for possible spurious correlations, both the actual and the correlation in first difference are displayed.

3.2.1 The two-factor Treasury Model

The results from the maximum likelihood optimization are illustrated in Table 3-3. As expected the parameter estimates of κ and θ in the drift function for \( x₁ \) are larger than for the second state variable. The estimates of κ imply a mean half-life of the state variables of 1.14 and 3.09 years respectively. These figures give an indication of the duration for the processes to return halfway to their long run average, θ, and as noted in Chen and Scott (1993) the second state variable is constrained to exhibit slow mean reversion if the model should be capable of capturing the variation in the long maturity yields. This aspect is considered obeyed in our model and the result of

---

56 Theil U is defined as (Greene 2000, p. 310): \( U = \text{RMSE} / \sqrt{E\{\sum y_i^2\}} \).

57 The mean half-life is the expected time for the process to return halfway to its long-run average θ. The mean half-life is calculated from the theoretical autocorrelation function as \( \exp(-\kappa t) = 0.5 \Rightarrow t = \ln(2)/\kappa \), where \( t \) is the mean half-life.
the Monte Carlo study in the previous section does not conflict with this conclusion as it only entails a shorter mean-life for the first state variable and a larger mean-life for the second state variable.

Table 3-3: Maximum Likelihood Estimates of the Two-Factor Model, Weekly Data, 1994 to 2002

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Panel A: Parameters of Diffusion x₁</th>
<th>Panel B: Parameters of Diffusion x₂</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>κ₁</td>
<td>θ₁</td>
</tr>
<tr>
<td>Estimate</td>
<td>0.60550</td>
<td>0.0337</td>
</tr>
<tr>
<td>Std. error</td>
<td>0.0248</td>
<td>0.00129</td>
</tr>
<tr>
<td></td>
<td>(0.0593)</td>
<td>(0.0080)</td>
</tr>
</tbody>
</table>

The table displays the ML-estimates of the translated two-factor Treasury model, January 1994-2002. The standard errors are based on the Hessian of the likelihood function and the values in parentheses are errors based on the cross product of the Jacobian.

The fairly low value of κ₂ makes the second state variable a martingale under the Q-measure as the estimate of κ₊λ is approximately zero, whereas the first state variable exhibits a positive value. Stated differently, we therefore expect the first state variable to be stationary and the second state variable to be non-stationary. However both the Augmented Dickey-Fuller test and the Phillips-Perron test indicate that x₁ and x₂ is I(1). Our findings are similar to Chen and Scott (1993) and we interpret these test results as a cause of the low power of these test statistics and are therefore confident in our ML-estimates of the mean-reversion parameters. However, most important is that both risk premium coefficients λ₁ and λ₂ are significantly negative, which implies that the term premiums are, on average, positive as maturity increases.

58The test results were ADF (PP): −1.0 (−1.1) and −2.0 (−1.9) for the two state variables respectively. The critical value at 5% is −2.87.
Part III: Empirical Investigation

Turning to the volatility function we surprisingly find the estimate of $\sigma^2$ to be larger for the second state variable. This is in discrepancy with the common findings and also the observed term structures. Our results indicate that the volatility is higher on long maturity bonds than short maturity bonds. We find this result unfortunate, but certainly a consequence of the specified error structure reported in section 3.1.2. Furthermore the significances of the parameters are also doubtful. As indicated by the values in parentheses the standard errors are very sensitive to the choice of estimation procedure as a calculation based on the cross product of the Jacobian is likely to alter the conclusions. Yet, Duffie and Singleton (1997) state, that this observation is common in finite samples, and in our case it entails that the parameters in the second state variable are insignificant, and the insignificant volatility parameter in $x_1$ provides a piecewise linear function as the best fit for the state variable. As this is highly unlikely we investigate the estimated parameters in details in the subsequent chapter, but for now we continue the use of the current standard errors based on the Hessian.

The estimates of $\theta$ are both significant and together with the estimate of the scale factor ($\alpha_r$) the long-run average of the state variables give an indication of the long-run mean of the instantaneous interest rate. The estimate of the constant, $\alpha_r$, at the final iteration was – 0.001 but the data was unable to pin down the exact value with any reliability. In light of the arguments in section 2.4.1 this result was to some extend expected as the arguments of Duffie and Singleton (1997), about the necessity of a negative constant, conflict with the empirical behavior of excess yields. Anyway this is a common phenomenon as Pearson and Sun (1994) and Duffee (1996) among others encounter the same problem. Bearing this in mind we estimate the long-run mean to 3.84%, which is probable in light of the data description in section 1.3.2.1.

To give the state variables an economical interpretation, it is informative to examine their relation to the Treasury term structure. The sample correlation of $\Delta x_1$ with changes in the slope of the Treasury term structure (15 year yield – 1 year yield) is approximately –0.92, whereas the correlation of $\Delta x_2$ with changes in the one-, five-, ten- and fifteen-year yields are 0.62, 0.87, 0.94 and 0.95 respectively. Thus the first factor behaves like the negative of the slope of the Treasury term structure, and the second factor behaves like the level of the fifteen-year yield. This is illustrated in Figure 3-1 where the state variables are plotted with their relevant counterparts.
Part III: Empirical Investigation

Figure 3-1: The Estimated State Variables and the Slope and Level of the Treasury Term Structure, 1994 – 2002

Panel A: $x_1$ and the negative slope

Panel B: $x_2$ and the level

The figure illustrates the time series of the estimated state variables from the two-factor Treasury model, 1994-2002. Panel A discloses the first state variable and the negative slope of the Treasury term structure (1 year zero-coupon yield – 15 year zero-coupon yield). Panel B discloses the level of the Treasury term structure (15 year zero-coupon yield) and the second state variable.

These conclusions are a consequence of the properties of the affine model. This is revealed when examining $A_i$ and $B_i$ from the zero-coupon bond yields. Across maturity $B_2(\tau)/\tau$ is hovering just above unity and $A_2(\tau)/\tau$ is essentially zero. It follows that the second factor represents a parallel shift in the entire zero-coupon yield curve induced by the change in $x_2$. In light of the correlations with the Treasury yields, this parallel-shift factor is well approximated empirically by the fifteen-year zero-coupon yield. Furthermore $B_1(\tau)/\tau$ decline with maturity such that positive shifts in $x_1$ induce flattening in the slope of the Treasury term structure. The finding that $x_1$ and the slope of the term structure are highly negatively correlated and that $B_1(\tau)/\tau$ is large for small values of $\tau$, suggest that slope changes are associated with greater variation in short-maturity yields during the sample period. The rapid decline in $B_1(\tau)/\tau$ as the maturity increases is induced in the CIR model largely by a fast rate of mean reversion of the first factor.

The above interpretation of the parameters makes no sense if the model lacks explanatory power, or stated otherwise, if the model’s ability to fit the observed yields is limited any economical interpretation of the parameters would be erroneous. We therefore conclude this section with an evaluation of the model’s pricing accuracy. To set the performance into perspective we compare the results with that of a naive martingale model. A priori the CIR model is expected to outperform the martingale model, and opposite findings are therefore interpreted as evidence against the usefulness of the model.

---

59 This martingale model assumes that the expected future yields one-week ahead are the current yields.
Part III: Empirical Investigation

Before turning to the results, consider the four measurement errors. Table 3-4 generally illustrates small errors even though they increase with maturity. This indicates precise predictions although the model experiences difficulties in fitting the prices of longer matured bonds. This conclusion seems to consistently reflect the entire sample, since the standard errors are very small.

Table 3-4: Estimates of the Measurement Error Standard Deviation, Weekly Data, 1994 to 2002

<table>
<thead>
<tr>
<th>Measurement Errors</th>
<th>1-year</th>
<th>4-year</th>
<th>10-year</th>
<th>15-year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ω</td>
<td>0.0015</td>
<td>0.0050</td>
<td>0.0064</td>
<td>0.0214</td>
</tr>
<tr>
<td>Std. Error</td>
<td>0.0001</td>
<td>0.0002</td>
<td>0.0002</td>
<td>0.0008</td>
</tr>
</tbody>
</table>

The table illustrates the fitted measurement errors implied by the two-factor Treasury Model and their respective standard errors, 1994-2002.

Supporting these observations Table 3-5 reports high correlations between the observed and predicted yields and together with the low RMSE and $U$ statistics, this indicates rather precise model performance. Not surprisingly the best forecasts are on the intermediate yields. The reason is found in the higher volatility in the shorter yields and the partly exclusion of the long yields in the derivation of the state variables. The 15-year yield is only used as a calibration point, by which is meant, that it is only utilized through the measurement errors and as a consequence the model will perform less well in predicting long-term yields.

Table 3-5: Pricing Accuracy of the Two-Factor Model, 1994 to 2002

<table>
<thead>
<tr>
<th>Maturity</th>
<th>Correlation</th>
<th>Correlation (Δ)</th>
<th>Max. Deviation</th>
<th>Mean Dev.</th>
<th>RMSE</th>
<th>Theil U</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-year</td>
<td>0.995</td>
<td>0.977</td>
<td>59.6 bp</td>
<td>10.5 bp</td>
<td>13.1 bp</td>
<td>2.44%</td>
</tr>
<tr>
<td>2-year</td>
<td>0.999</td>
<td>0.991</td>
<td>47.6 bp</td>
<td>5.7 bp</td>
<td>7.4 bp</td>
<td>1.32%</td>
</tr>
<tr>
<td>5-year</td>
<td>0.998</td>
<td>0.995</td>
<td>21.6 bp</td>
<td>4.7 bp</td>
<td>6.2 bp</td>
<td>1.05%</td>
</tr>
<tr>
<td>6-year</td>
<td>0.999</td>
<td>0.998</td>
<td>21.7 bp</td>
<td>2.9 bp</td>
<td>4.0 bp</td>
<td>0.66%</td>
</tr>
<tr>
<td>7-year</td>
<td>0.999</td>
<td>0.998</td>
<td>20.0 bp</td>
<td>3.2 bp</td>
<td>4.2 bp</td>
<td>0.68%</td>
</tr>
<tr>
<td>10-year</td>
<td>0.999</td>
<td>0.999</td>
<td>25.0 bp</td>
<td>7.2 bp</td>
<td>8.7 bp</td>
<td>1.20%</td>
</tr>
<tr>
<td>15-year</td>
<td>0.984</td>
<td>0.994</td>
<td>48.4 bp</td>
<td>10.0 bp</td>
<td>13.8 bp</td>
<td>2.10%</td>
</tr>
</tbody>
</table>

The table illustrates the pricing accuracy of the Two-factor model, 1994-2002. In relation to the utilized treasury yields, the table presents a general view of the model’s performance through the correlation, maximum and mean deviation and the more specific measures RMSE (root mean squared error) and Theil $U$. Yields not used in the estimation are depicted in Appendix F together with the performance of the naive martingale model.
The average RMSE ($U$ statistic) in the short (1-4Y), intermediate (5-9Y) and long maturity segments (10-15Y) are about 10.7 bp (1.9%), 4.8 bp (0.85%) and 11.2 bp (2.0%) respectively. The corresponding measures from the naive martingale model are 15.1 bp (2.7%), 14.7 bp (1.7%) and 13.4 bp (2.1%). So, in comparison, the two-factor model out-performs the martingale model at short and intermediate maturities, and perform relatively better at long maturities. However, considering maturities beyond 15 years the errors of the two-factor model increase whereas the martingale model improves (correlation increases monotonically from 28% for the 1-year yield to 84% for the 20-year yield and the RMSE ($U$ statistic) decrease with maturity from 14.2 bp (2.65%) to 12.1 bp (1.81%) – see Appendix F). Still we do not interpret this as evidence against the CIR model but rather the ad hoc specification of the true yields and measurement errors. The translated two-factor CIR model therefore appears a fairly useful tool when predicting the empirically observed short- and intermediate Treasury yields, although not perfect.

In review, it can be difficult to fully understand the performance of the model. To enlighten when the model performs successfully and when it encounters problems, it is very informing to consider the observed and predicted cross sections. Figure 3-2 depicts two different situations observed in the data – Panel A shows a normal situation, and Panel B a worst case scenario.

**Figure 3-2: The Observed and Predicted Treasury Term Structure**

*Panel A: Monotonically upward sloped*  
*Panel B: Inverted*

The figure shows two observed and predicted term structures in the two-factor Treasury model. Panel A discloses the normal situation with an upward slope and panel B discloses a worst-case scenario where the term structure is relatively flat while inverting.

Not surprisingly, the figure reveals that the model performs accurately when the term structures behave nicely, i.e. monotonically upward sloped. In contrast, the model performs less accurate when the term structure inverts although the RMSE ($U$ statistic) of the cross sections are in the range 5-10 bp (1-2%).
3.2.2 The one-factor Credit-Spread Model

Table 3-6 summarizes the parameter estimates of the default intensity processes. The mean log-likelihood values of the four estimations are in the range 9.5 to 14 where AAA accounts for the lowest and all other values are above 11. We do not report t-statistics on all the estimates, as some of the computed standard errors were unrealistic high.

Table 3-6: Maximum Likelihood Estimates of the Hazard Rate Process, Weekly Frequency, 1994 - 2001

<table>
<thead>
<tr>
<th>Rating</th>
<th>Estimate</th>
<th>κ</th>
<th>θ</th>
<th>λ</th>
<th>σ</th>
<th>α₉</th>
<th>β₁</th>
<th>β₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>Estimate</td>
<td>0.0595</td>
<td>0.0035</td>
<td>-0.1504</td>
<td>0.0765</td>
<td>0.0034</td>
<td>0.0032</td>
<td>-0.094</td>
</tr>
<tr>
<td></td>
<td>t-stat</td>
<td>(2.91)</td>
<td>(3.10)</td>
<td>(-5.37)</td>
<td>(3.56)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>AA</td>
<td>Estimate</td>
<td>0.1452</td>
<td>0.0055</td>
<td>-0.2003</td>
<td>0.0853</td>
<td>0.0031</td>
<td>0.0026</td>
<td>-0.0812</td>
</tr>
<tr>
<td></td>
<td>t-stat</td>
<td>(5.73)</td>
<td>(5.09)</td>
<td>(-6.37)</td>
<td>(7.01)</td>
<td>-</td>
<td>(1.46)</td>
<td>(-2.40)</td>
</tr>
<tr>
<td>A</td>
<td>Estimate</td>
<td>0.1731</td>
<td>0.00861</td>
<td>-0.2401</td>
<td>0.1113</td>
<td>0.0055</td>
<td>-0.0094</td>
<td>-0.0676</td>
</tr>
<tr>
<td></td>
<td>t-stat</td>
<td>(7.39)</td>
<td>(5.01)</td>
<td>(-7.67)</td>
<td>(6.91)</td>
<td>-</td>
<td>(2.86)</td>
<td>(-2.24)</td>
</tr>
<tr>
<td>BBB</td>
<td>Estimate</td>
<td>0.1873</td>
<td>0.00807</td>
<td>-0.1912</td>
<td>0.1015</td>
<td>0.0079</td>
<td>0.0037</td>
<td>-0.1412</td>
</tr>
<tr>
<td></td>
<td>t-stat</td>
<td>(6.39)</td>
<td>(4.43)</td>
<td>(-6.10)</td>
<td>(5.76)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

The table displays the ML-estimates of the translated one-factor hazard rate models, January 1994-2002. The standard errors are based on the Hessian of the likelihood function, and the recovery rates used are 70% (AAA), 65% (AA), 60% (A), and 55% (BBB).

In Table 3-6 it is seen that the negative correlation to the level of the Treasury term structure is well supported by the β₁ estimates, although we are not able to verify the general empirical pattern of greater sensitivity for lower-rated firms. Considering the negative slope of the Treasury term structure, the results were somewhat ambiguous. As the intensities are expected to mean revert about their long rate, given the positive estimates of kappa, we expect that an increase in the slope of the Treasury curve would narrow the credit spreads, as it indicates that the short-term yield is expected to increase in the future, which entails improvement in business climate. Nevertheless, the estimates of β₁ were both positive and negative.

Furthermore both the minimum default rate, measured by α₉, and the unconditional long-term mean increases with lowering credit quality. This is intuitively logical and empirically a very convenient result as it reconciles prior empirical research and theoretical considerations. The fitted

---

60The slope as a predictor of future changes in the short interest rate is empirically supported by the findings of Campbell and Schiller (1991).
Part III: Empirical Investigation

$\alpha_+$ for single-A firms is about 20 bp above AAA firms and $\theta$ is about 60 bp above. Moreover, the estimated $\alpha_+$ is in contrast to the Treasury estimate positive across all ratings, indicating that the credit spread does not converge to zero as time to maturity goes to zero, or equivalently, the minimum credit spread will even for the healthiest firm always be positive. The same tendency is revealed when focusing on the volatility parameter as it increases with lower rating. This indicates that the higher uncertainty associated with bonds of lower ratings make investors require a higher risk premium or equivalently a lower price.

The estimate of $\lambda$ is negative for all ratings, and analogous to the interest rate model this negative sign implies that investors require compensation for the variation in default risk. More intuitively, a bond that pays off in times of high uncertainty, with greater certainty, is valued higher. Combined with this logic the size of the estimates seems very plausible, since $\lambda$ becomes more negative with lowering credit quality, indicating that investors require additional compensation for investing in lower rated bonds. The tendency is however not unambiguous since the BBB-estimate is less negative than the estimates of A and AA firms. Combined with the size of $\kappa$ we notice, as Duffee (1996, 1999), the very important feature discussed in section 2.4. In contrast to the earlier Treasury results, $\lambda$ is larger than the size of $\kappa$. As a consequence the mean-reversion parameter $\kappa + \lambda$ under the pricing measure $\mathbb{Q}^+$ is negative, hereby implying that the instantaneous default risk has an explosive drift structure. As a result, the drift of $h$ will independently of its level always be positive, and a higher level imposes a further increase in $h$. However, ADF-tests do not verify this aspect as the time series of instantaneous default probabilities are, except for the BBB-rating, stationary.\footnote{All test statistics were close to the non-rejection area, which imply that the derived conclusions are sensitive towards the significance level.} The negative sign of $\kappa + \lambda$ has several very important implications, but before getting into details about explaining these aspects recall for now that a key component in determining the term premia is the $\lambda$ parameter. Interpreted alone $\lambda$ therefore tells us that investors price bonds based on expectations that default probabilities are both higher than the observed default probabilities and expected to increase with time.

Combining the results the four models generally produce credit spreads consistent with several empirical observations, but also reveal properties inconsistent with both theoretical and empirical considerations. Figure 3-3 depicts three charts. The top chart shows the yield spread generated by the model in a normal situation. The middle and lower chart shows yield spreads of the lower and upper 2.5 percentile where default rates are low and high, respectively.
Part III: Empirical Investigation

Figure 3-3: Estimated Term Structure of Credit Spreads

Note: (a) the upper figure displays the median default intensity, (b) the middle displays the lower 2.5 percentile, and (c) the lower figure displays the 97.5 percentile produced by the translated one-factor hazard rate models.

If we focus on the normal situation (the top chart) it is seen that the spread curves are consistent with the stylized facts of increasing yield spread with maturity, and except for the BBB-rating, that the slope increases with lowering credit quality. Moreover AAA and AA firms disclose about the same estimated $\alpha_h$ and mean $h$, indicating that investors, over short horizons, view default risk of the two rating categories about equal. Over longer periods of time, though, the spreads increase to 110 and 140 bp or equivalently approximately 3 and 5 bp per additional year to maturity. The A
and BBB curves increases about 6-8 bp per year, which together with the intersection at the 14 years maturity seems implausible.

Compared to the finding of Fons (1994) our results are significantly higher. He finds that AAA curves are essentially flat across maturities, while yield spreads for BBB-rated firms increase about 2 bp per each additional year to maturity. The difference is not surprisingly a reflection of the diversity in default rates in our sample compared to the period Fons examines. He used a sample period prior to 1982, where default rates were significantly below the rates observed in the 90’es. Therefore it is more sensible to compare the results with Duffee (1999). He used a sample from 1985-95 on individual bond prices, and found that the spreads of AA to BBB rated firms increased about 1.5 to 5 bp per each additional year to maturity. The slope of the curves estimated here, might then appear high, but still, we cannot conclude anything as the sample period is different and Duffee examines individual bond prices. Nevertheless, these issues emphasize the importance of the presence of a positive constant to dampen the steepness of the credit curves. If the AAA curve was estimated without the constant the slope (15Y) would approximately increase from 34 bp to 64 bp.

Turning our attention to the middle and lower chart, we see that the positive slopes are ensured by the negative $\kappa + \lambda$. By comparing the two charts, it is seen that the slope (15Y) of the AA curve does not increase with increasing instantaneous default risk, as the curves begin to flatten at the long end. The medium term slope (8Y), however, increases for all ratings with increasing instantaneous default rates (the slope of the BBB curve decreases, due to the higher level of the hazard rate process). Compared to Duffee (1999) we do not find that the higher slopes for lower rated bonds are caused by lower $\kappa + \lambda$, but instead find that the differences primarily owes to differences in the implied hazard rates. Despite the higher recovery rate, it is also seen that the AAA and AA curves are almost identical when default rates are high, but when default rates are low the AAA curve is essentially flat across maturities. Of course, this is explained by the low level of $\kappa$ but is primarily a consequence of different underlying distributions of the hazard rates. The AA-hazard rate has a higher mean and lower 2.5 percentile compared to AAA.

In this relation Figure 3-4 depicts the implied default probabilities. First of all, it is seen that the hazard rates are always positive. This is a very comfortable result, as it is a necessary condition for excluding arbitrage from the economy. Furthermore, the historical evidence of credit risk discussed earlier is well reflected in the implied default probabilities. We see that up to 1999, low levels and low volatility characterize the series, indicating strong market conditions. After 1999 default intensities as well as volatilities rise considerably. The large fluctuations ultimo 1998 are
Part III: Empirical Investigation

caused by the liquidity crisis, and therefore not credit risk, and the large peak in 2001 is the September 11th effect.

Figure 3-4: Estimated Implied Instantaneous Default Rates, 1994-2002

The figure illustrates the times series of the estimated instantaneous default rates, of AAA, AA, A and BBB-rated corporate bond indices, implied by the translated one-factor hazard rate models, 1994-2002.

Not surprisingly, default probabilities exhibit strong correlation, which decrease with lower ratings. The default rates, furthermore, show significant time variations and ARCH-like effects. Table 3-7 reports the volatilities of $h_i$ and reveals a general pattern of higher default volatility of lower rated firms.

Table 3-7: Descriptive Statistics and Correlation Structure Between Default Probabilities

<table>
<thead>
<tr>
<th>Rating</th>
<th>Descriptive Statistics (bp)</th>
<th>Correlation Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\hat{E}(\hat{h}_i)$</td>
<td>$\text{Std}(\hat{h}_i)$</td>
</tr>
<tr>
<td>AAA</td>
<td>51</td>
<td>16</td>
</tr>
<tr>
<td>AA</td>
<td>54</td>
<td>18</td>
</tr>
<tr>
<td>A</td>
<td>66</td>
<td>28</td>
</tr>
<tr>
<td>BBB</td>
<td>87</td>
<td>29</td>
</tr>
</tbody>
</table>

The table displays the mean of the fitted instantaneous default probabilities and their respective standard errors, 1994-2002. Furthermore, the table reveals the correlation between the estimated instantaneous default probabilities, of AAA, AA, A and BBB-rated corporate bond indices, implied by the translated one-factor hazard rate models.

Considering the mean fitted $\hat{h}_i$, it appears that the intensities reported in Figure 3-4, compared to the empirically observed default rates are very high. Even for AAA firms, for which actually observed default rates are very rare, the mean fitted $\hat{h}_i$ is still high. Intuitively, this can be justified
Part III: Empirical Investigation

over shorter periods, but how can we explain that implied default probabilities vary systematically above actual probabilities?

The aspect that $\mathbb{P}(z < T) \ll \mathbb{Q}(z < T)$ is well-known within the framework of models of instantaneous default probabilities (See Lando 2000, Duffee 1996, 1999), and there are several ways to explain the difference. Firstly, it is plausible to imagine that in times of recession defaults arrive systematically. Combined with the logic that the marginal utility of additional wealth in such periods is high, the required compensation for the non-diversifiable risk will be high (formalized by large negative $\lambda$). The risk adjusted default probabilities used for pricing, i.e. the probabilities under the equivalent Martingale measure, will therefore differ significantly from the actual default probabilities. This aspect is very important to notice, but it also, as stressed by Duffee (1996), has the depressive effect that empirical default probabilities cannot straightforward be used for pricing.

Another important aspect to consider is that $h$ is inverted from credit spreads containing components non-related to credit risk. We have already addressed this aspect in 1.3.1, so we do not go into detail about this. The important thing to notice here is that components such as tax-differences between Treasury and Corporate bonds, liquidity, and aspects concerning incomplete information will translate into premiums over Treasury rates. Without a formal way to identify these components they will as a result be subsumed into $h$. In addition, in our specific analysis another problem emerges. We do not have information on the credit spreads in the short end, which is problematic as we model the instantaneous default probabilities. If the data contained this information the precisions of the estimates would inevitable improve, especially in series that increase significantly in the short end.

As for the two-factor model we conclude this section with an evaluation of the prediction accuracy. Table 3-8 presents the measurement errors, and as seen, there is no clear tendency across ratings, but between maturity segments it appears that the measurement errors at the long end are about twice the size of the errors of the short maturity, indicating that the model performs relatively better when fitting prices of shorter bonds.
Table 3-8: Estimates of the Measurement Error Standard Deviation, Weekly Data, 1994 to 2002

<table>
<thead>
<tr>
<th>Rating</th>
<th>$\Omega_{\text{five-seven}}$</th>
<th>$\Omega_{\text{ten-fifteen}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>0.0075</td>
<td>0.0225</td>
</tr>
<tr>
<td>AA</td>
<td>0.0091</td>
<td>0.0186</td>
</tr>
<tr>
<td>A</td>
<td>0.0095</td>
<td>0.0203</td>
</tr>
<tr>
<td>BBB</td>
<td>0.0089</td>
<td>0.0138</td>
</tr>
</tbody>
</table>

The table illustrates the fitted measurement errors of the four ratings, implied by the one-factor hazard-rate Model, 1994-2002.

Turning more specifically to the accuracy measures it is obvious from Table 3-9 that the model predicts the shorter yields more precisely than the longer yields. This does not surprise as the inherent structure of the ML-model entails a more precise estimation of the shorter yields (the short 1-3 years spread is implemented as the true yield, and the long 10-15 years spread is only used as a calibration point).

Table 3-9: The One-Factor Model’s Time-Series Predictive Ability separated by Ratings, 1994 to 2002

<table>
<thead>
<tr>
<th>Maturity</th>
<th>Rating</th>
<th>Correlation</th>
<th>Max Dev.</th>
<th>Mean Dev.</th>
<th>RMSE</th>
<th>Theils U</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-5 years</td>
<td>AAA</td>
<td>0.872</td>
<td>51.4 bp</td>
<td>15.4 bp</td>
<td>19.7 bp</td>
<td>28.0%</td>
</tr>
<tr>
<td></td>
<td>AA</td>
<td>0.871</td>
<td>44.3 bp</td>
<td>11.1 bp</td>
<td>14.8 bp</td>
<td>18.7%</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>0.947</td>
<td>36.0 bp</td>
<td>9.8 bp</td>
<td>13.2 bp</td>
<td>14.4%</td>
</tr>
<tr>
<td></td>
<td>BBB</td>
<td>0.933</td>
<td>89.0 bp</td>
<td>22.8 bp</td>
<td>26.2 bp</td>
<td>17.0%</td>
</tr>
<tr>
<td>5-7 years</td>
<td>AAA</td>
<td>0.850</td>
<td>61.1 bp</td>
<td>15.5 bp</td>
<td>19.7 bp</td>
<td>25.8%</td>
</tr>
<tr>
<td></td>
<td>AA</td>
<td>0.917</td>
<td>35.9 bp</td>
<td>10.2 bp</td>
<td>13.2 bp</td>
<td>15.6%</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>0.961</td>
<td>34.4 bp</td>
<td>8.1 bp</td>
<td>10.7 bp</td>
<td>10.6%</td>
</tr>
<tr>
<td></td>
<td>BBB</td>
<td>0.921</td>
<td>101.6 bp</td>
<td>29.4 bp</td>
<td>33.5 bp</td>
<td>20.1%</td>
</tr>
<tr>
<td>7-10 years</td>
<td>AAA</td>
<td>0.846</td>
<td>59.5 bp</td>
<td>13.3 bp</td>
<td>17.8 bp</td>
<td>24.8%</td>
</tr>
<tr>
<td></td>
<td>AA</td>
<td>0.868</td>
<td>55.8 bp</td>
<td>24.4 bp</td>
<td>26.4 bp</td>
<td>30.1%</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>0.951</td>
<td>67.1 bp</td>
<td>24.6 bp</td>
<td>27.4 bp</td>
<td>26.9%</td>
</tr>
<tr>
<td></td>
<td>BBB</td>
<td>0.920</td>
<td>92.5 bp</td>
<td>27.8 bp</td>
<td>32.2 bp</td>
<td>19.2%</td>
</tr>
<tr>
<td>10-15 years</td>
<td>AAA</td>
<td>0.758</td>
<td>80.5 bp</td>
<td>18.9 bp</td>
<td>24.6 bp</td>
<td>31.3%</td>
</tr>
<tr>
<td></td>
<td>AA</td>
<td>0.835</td>
<td>94.6 bp</td>
<td>35.5 bp</td>
<td>39.9 bp</td>
<td>42.4%</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>0.951</td>
<td>83.5 bp</td>
<td>42.8 bp</td>
<td>44.9 bp</td>
<td>42.2%</td>
</tr>
<tr>
<td></td>
<td>BBB</td>
<td>0.880</td>
<td>95.1 bp</td>
<td>33.8 bp</td>
<td>39.3 bp</td>
<td>22.6%</td>
</tr>
<tr>
<td>15+ years</td>
<td>AAA</td>
<td>0.516</td>
<td>94.6 bp</td>
<td>29.9 bp</td>
<td>38.0 bp</td>
<td>39.4%</td>
</tr>
<tr>
<td></td>
<td>AA</td>
<td>0.779</td>
<td>135.3 bp</td>
<td>67.1 bp</td>
<td>71.5 bp</td>
<td>66.5%</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>0.894</td>
<td>178.3 bp</td>
<td>84.5 bp</td>
<td>88.4 bp</td>
<td>73.7%</td>
</tr>
<tr>
<td></td>
<td>BBB</td>
<td>0.838</td>
<td>107.6 bp</td>
<td>40.2 bp</td>
<td>46.8 bp</td>
<td>25.4%</td>
</tr>
</tbody>
</table>

The table illustrates the pricing accuracy of the translated one-factor hazard-rate model on AAA, AA, A and BBB-rated corporate bond indices, 1994-2002. In relation to the utilized credit spreads, the table presents a general view of the model’s performance through the correlation, maximum and mean deviation and the more specific measures RMSE (root mean squared error) and Theil $U$ statistic.
The best fit is found in the short (3-5 years) and intermediate (5-7 years) yields with RMSE (Theil $U$) generally below 20 bp (20%). This is an acceptable level but still far higher than the predictions of the Treasury term structure. As a benchmark the same accuracy measures are calculated for a naive Martingale model. The results are not shown here but are to some extent depressing as the one-factor model only performs marginally better when focusing on the RMSE’s or the $U$ statistics. Moreover, except for the BBB-rating the table reveals a pattern in the shorter yields as all the accuracy measures improve with lower ratings. When focusing on the longer yields it seems as this tendency reverses but the measures for the BBB-rating disclose that the model performs rather well in predicting bonds of longer maturities of this rating. Taking the correlations into account unveil that the problem with the AA- and A-rating is that the model is not capable of capturing the level of these longer yields. All things equal, we cannot expect the same level of performance as in the two-factor model, so in summery we find comfort in the model’s ability to replicate yields of shorter bonds (less that 7 years).

### 3.3 Robustness and Specification Tests

In this chapter we evaluate the consistency of the estimated parameters and explain possible mis-specifications. First of all, as the estimates proved sensitive towards the chosen error structure we investigate the consistency of the parameters with regards to alternative error structures. Secondly, for reasons outlined in section 3.3.2 the inference is questionable in finite samples. So, in order to derive more confident inference we estimate the empirical distribution function (EDF) for the parameters, and hereby obtain new critical values. Thirdly, if our estimated parameters are reliable and valid we expect the normalized innovations to be white noise (Duffee 1996, 1999) i.e. the estimated processes are well-specified. These issues unveil key information concerning both the estimation methodology and model limitations, and are therefore very important when evaluating the robustness of the model and eventually its reliability.

#### 3.3.1 Alternative Error Structures

We stated earlier that the estimates would be independent of the chosen error structure if the models are well-specified. As already seen, the model is not necessarily well-specified, and the parameters may therefore alter to changes in test designs (the measurement error structure). To address this question and to investigate the robustness of the empirical findings revealed in the previous section, this section presents and evaluates different model specifications. The section also serves as an examination of whenever it is possible to improve the performance of the model on the troublesome long maturities by changes in test design.
Part III: Empirical Investigation

For the Treasury term-structure model the following alternatives are evaluated

**Alternative I:**
\[
\log P(t,T_i) = -Y(t,T_i) \tau_1 + \varepsilon_{1,i} \\
\log P(t,T_2) = -Y(t,T_2) \tau_2 - \varepsilon_{1,i} \\
\log P(t,T_3) = -Y(t,T_3) \tau_3 + \varepsilon_{2,i} \\
\log P(t,T_4) = -Y(t,T_4) \tau_4 + \varepsilon_{3,i} \\
\log P(t,T_5) = -Y(t,T_5) \tau_5 + \varepsilon_{4,i} \\
\log P(t,T_{10}) = -Y(t,T_{10}) \tau_{10} + \varepsilon_{5,i} \\
\log P(t,T_{15}) = -Y(t,T_{15}) \tau_{15} - (\varepsilon_{1,i} + \varepsilon_{2,i})
\]

**Alternative II:**
\[
\log P(t,T_i) = -Y(t,T_i) \tau_1 \\
\log P(t,T_2) = -Y(t,T_2) \tau_2 + \varepsilon_{1,i} \\
\log P(t,T_3) = -Y(t,T_3) \tau_3 + \varepsilon_{2,i} \\
\log P(t,T_4) = -Y(t,T_4) \tau_4 + \varepsilon_{3,i} \\
\log P(t,T_5) = -Y(t,T_5) \tau_5 + \varepsilon_{4,i} \\
\log P(t,T_{10}) = -Y(t,T_{10}) \tau_{10} \\
\log P(t,T_{15}) = -Y(t,T_{15}) \tau_{15} + \varepsilon_{4,i}
\]

where the definition of \( Y(t,T_k) \tau_k \) is analogous to equation (3-15). In the first alternative the aim is to define the true yields as a mix of short-term and long-term yields and only include the intermediate 4-year yield as a calibration point. The second alternative fits the 1-year and 10-year yields exactly and places measurement errors on all other utilized yields.

For the credit-spread term structure we propose the following two alternative test designs

**Alternative I:**
\[
\log \tilde{P}_{obs}(t,T_1,C,\phi) = \log \tilde{P}_{calc}(t,T_1,C,\phi) \\
\log \tilde{P}_{obs}(t,T_2,C,\phi) = \log \tilde{P}_{calc}(t,T_2,C,\phi) + \varepsilon_{1,i} \\
\log \tilde{P}_{obs}(t,T_3,C,\phi) = \log \tilde{P}_{calc}(t,T_3,C,\phi) + \varepsilon_{2,i} \\
\log \tilde{P}_{obs}(t,T_4,C,\phi) = \log \tilde{P}_{calc}(t,T_4,C,\phi) + \varepsilon_{3,i} \\
\log \tilde{P}_{obs}(t,T_5,C,\phi) = \log \tilde{P}_{calc}(t,T_5,C,\phi) + \varepsilon_{4,i}
\]

**Alternative II:**
\[
\log \tilde{P}_{obs}(t,T_1,C,\phi) = \log \tilde{P}_{calc}(t,T_1,C,\phi) \\
\log \tilde{P}_{obs}(t,T_2,C,\phi) = \log \tilde{P}_{calc}(t,T_2,C,\phi) + \varepsilon_{1,i} \\
\log \tilde{P}_{obs}(t,T_3,C,\phi) = \log \tilde{P}_{calc}(t,T_3,C,\phi) + \varepsilon_{2,i} \\
\log \tilde{P}_{obs}(t,T_4,C,\phi) = \log \tilde{P}_{calc}(t,T_4,C,\phi) + \varepsilon_{3,i} \\
\log \tilde{P}_{obs}(t,T_5,C,\phi) = \log \tilde{P}_{calc}(t,T_5,C,\phi) + \varepsilon_{4,i}
\]

In the first alternative the aim is to place more weight on the short end, whereas the second alternative seeks to incorporate the information contained in the medium and long maturities. In order to reduce computations and complexities to a minimum we model the shortest maturity segment as the true yield in both alternatives.

The maximum likelihood estimates are presented in Table 3-10 along with the estimates of the original model.
Part III: Empirical Investigation


<table>
<thead>
<tr>
<th>Parameter</th>
<th>$\kappa_i$</th>
<th>$\theta_i$</th>
<th>$\sigma_i^2$</th>
<th>$\lambda_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: Parameters of $x_1$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Original model</td>
<td>0.6055 (0.0248)</td>
<td>0.0337 (0.0013)</td>
<td>0.0069 (0.0003)</td>
<td>-0.0922 (0.0223)</td>
</tr>
<tr>
<td>Alternative I</td>
<td>0.3249 (0.0076)</td>
<td>0.04713 (0.0026)</td>
<td>0.0083 (0.0011)</td>
<td>-0.0011 (0.0005)</td>
</tr>
<tr>
<td>Alternative II</td>
<td>0.3869 (0.0090)</td>
<td>0.0161 (0.0017)</td>
<td>0.0024 (0.0001)</td>
<td>-0.0016 (0.0004)</td>
</tr>
<tr>
<td>Panel B: Parameters of $x_2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Original model</td>
<td>0.2245 (0.0866)</td>
<td>0.0048 (0.0019)</td>
<td>0.0112 (0.0002)</td>
<td>-0.2178 (0.0867)</td>
</tr>
<tr>
<td>Alternative I</td>
<td>0.1561 (0.0051)</td>
<td>0.0025 (0.0001)</td>
<td>0.0109 (0.0049)</td>
<td>-0.2850 (0.0052)</td>
</tr>
<tr>
<td>Alternative II</td>
<td>0.0002 (0.0016)</td>
<td>0.0033 (0.0015)</td>
<td>0.0009 (0.0017)</td>
<td>-0.0082 (0.0008)</td>
</tr>
<tr>
<td>Panel C: Parameters of $h$ for single-A ratings</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Original model</td>
<td>0.1731 (0.0234)</td>
<td>0.0086 (0.0017)</td>
<td>0.0124 (0.0039)</td>
<td>-0.2401 (0.0313)</td>
</tr>
<tr>
<td>Alternative I</td>
<td>0.1871 (0.0796)</td>
<td>0.0080 (0.0018)</td>
<td>0.0136 (0.0041)</td>
<td>-0.2201 (0.0907)</td>
</tr>
<tr>
<td>Alternative II</td>
<td>0.1031 (0.0516)</td>
<td>0.0122 (0.00621)</td>
<td>0.0194 (0.0119)</td>
<td>-0.2599 (0.0107)</td>
</tr>
</tbody>
</table>

The table displays the original and alternative parameter estimates of the translated two-factor Treasury model and the translated one-factor hazard-rate model for A-rated credit spreads 1994-2002. The values in parentheses are standard errors based on the Hessian. The mean log-likelihood values are 21.4, 20.2 and 15.3 respectively for the two-factor Treasury model and 12, 11.3 and 8.3 for the one-factor hazard-rate model.

As seen Table 3-10 reveals noticeable differences in the estimated parameters, so our hypothesis that the parameters are sensitive with regards to the chosen test design is generally well-founded. Compared to the original model, except for the mean-reversion parameter of the first state variable, the parameters of alternative I (treasury or hazard) are however less different than the parameters under alternative II. Intuitively this result does not surprise, as the only difference between the specification of the original Treasury model and alternative I is the interchange of the 4-year and 15-year yields. In addition, the parameters of the hazard-rate model are also roughly unchanged under the first alternative, i.e. when focusing on the shorter yield-segments. Primarily this can be explained by the fact that we still use the first yield-segment as a true yield, and interchange the second and third, and fourth and fifth yield-segments in the error structure, which to a high extend are expected to coincide.

The real dissimilarities are introduced when considering the two other alternatives. In the case of the two-factor model, the second state variable is predicted to be a martingale with very small
Part III: Empirical Investigation

volatility, and except from $\kappa$, all the parameter values are small compared to the two other models. Especially, the predicted long-run mean of 1.53% seems very implausible. Turning to the hazard-rate model the changes in the estimated parameters are again seen in the perspective of using the first yield segment as a true yield. As the measurement errors are placed on the longer yield segments the results are not expected to coincide, as the properties of the short and long yields are different i.e. when using longer matured bonds we expect a lower $\kappa$ and a higher $\sigma^2$ and $\lambda$.

The ideal approach, in formally examining if there is a significant difference between the three designs, would be to test whether the individual parameters are different. Because the structure in the alternative model formulations are very similar and as the parameter solutions are functions of the same underlying data this is not statistically possible. As an alternative we employ the likelihood ratio test for non-nested hypotheses developed by Voung (1989) where the null hypothesis is

$$E\left\{ \ln \frac{f_1(\psi_1)}{f_2(\psi_2)} \right\} = 0$$

for two models with density functions $f_i$ where $\psi_i$ represents the parameter vector. In light of the differences in the parameter estimates we at least hope to test the second alternative model significantly different from the original i.e. that the test statistic is significantly larger than zero, as it indicate that the original test design captures the properties of the data more adequately. The conclusions are however twofold, since a rejection of the null also indicates a possible misspecification of the derived model framework.

Voung showed that the test statistic

$$n^{1/2} \left\{ \frac{\ln L_1(\psi_1) - \ln L_2(\psi_2)}{S_n} \right\}$$

where

$$S_n^2 = E\left[ (\ln L_1(\psi_1) - \ln L_2(\psi_2))^2 \right] - E\left[ (\ln L_1(\psi_1) - \ln L_2(\psi_2))^2 \right]$$

$$= n^{-1} \sum_{i=1}^n \left( \ln L_{1i}(\psi_1) - \ln L_{2i}(\psi_2) \right)^2 - \left( n^{-1} \sum_{i=1}^n \left( \ln L_{1i}(\psi_1) - \ln L_{2i}(\psi_2) \right) \right)^2$$

is asymptotic a standard normal, where $L_i(\psi)$ represents the respective likelihood functions. The results of the tests are shown in Table 3-11.
Part III: Empirical Investigation

Table 3-11: Model Selection Test Statistics, Weekly Data, 1994 to 2002

<table>
<thead>
<tr>
<th></th>
<th>Two-Factor Treasury Model</th>
<th>One-Factor Credit-spread Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>t-statistic</td>
<td>t-statistic</td>
</tr>
<tr>
<td>Original Model versus Alternative I</td>
<td>0.363</td>
<td>1.182</td>
</tr>
<tr>
<td>Original Model versus Alternative II</td>
<td>5.152</td>
<td>2.632</td>
</tr>
</tbody>
</table>

The table shows the results of applying Young’s likelihood ratio test for non-nested hypotheses on the time series of log-likelihood values for the alternative treasury and hazard-rate models (A-rated bonds), 1994-2002.

As expected, we see that the likelihood ratio test rejects both second alternatives while not rejecting the first alternatives. Yet, both results are in a sense problematic, since we saw from Table 3-10 that the parameters were different and individually significant. In principle, the estimated parameters ought to be within a small range of each other, as if not, it indicates that the model is misspecified. As the test cannot satisfactorily reveal the differences we have to examine the forecasting abilities of each test design, which is not presented here but illustrated in Appendix G. In summary we find that both alternative treasury designs fit the short end to a wide extent but suffers ruthlessly at the longer maturities. Although the alternative models follow the evolution in the observed yields depicted by the high partial correlation estimates they cannot fix the level for which reason they always underestimate the observed yields as exemplified through the high accuracy measures. As a consequence the model with the highest forecasting power is the original model. The same conclusions apply for the credit-spread models, and, in line with the findings in the previous chapter, the tests therefore tells us that the models exhibit problems in describing the long term yields even if we change the test design to cope more explicitly with this yield segment.

After reviewing the results of this section, we find it problematic that our empirical findings are sensitive to the chosen error structure, and we interpret this as providing evidence of misspecification. Essentially this is also why we conduct an explicit specification analysis later in this chapter, but the results revealed here might equally well, be a consequence of inadequate estimation methodology. The results unveil that our framework does not adequately capture the inherent dynamics of the whole term structure i.e. it does not manage to incorporate all the information found in the yields. In light of Duffee (1996) this postulate seems plausible as he encountered similar problems with the MLE-inversion approach and employed a Kalman Filter instead (Duffee, 1999). However this hypothesis is not investigated here.

3.3.2 Finite-sample Consistency

The analyses in chapter 3.2 relied on first-order asymptotic theory when evaluating the significances of the parameter estimates. This, however, is critical for two reasons. First of all, the assumption of asymptotic normal parameter estimates may very well be violated. Recall that asymp-
Part III: Empirical Investigation

totic distribution theory often is applied in approximating the cumulative distribution function (CDF) of the investigated statistics. The asymptotic distributions of many econometric statistics are standard normal or chi-square, possibly after centering and normalization, regardless of the distribution from which the data were sampled. Such statistics are called asymptotically pivotal, meaning that their asymptotic distributions do not depend on unknown population parameters. If the statistic is asymptotically pivotal and the sample is sufficiently large the asymptotic distribution can be applied without loss of generality. However, as this thesis only utilizes a small sample of the total population of corporate bonds, i.e. 7 years, we have a possible finite sample bias, and consequently the finite-sample distributions may certainly differ from the asymptotic. Secondly, we saw that the default-free state variables are close at being non-stationary under the pricing measure, and that the hazard rates are explosive. From Dickey and Fuller (1981) we know that the presence of a unit root implies that the estimators are superconsistent, whereby the assumption of asymptotic normality is invalid (Davidson and MacKinnon, 1993).

As a consequence we investigate the empirical distributions of the estimators through the bootstrap methodology. It amounts to treating the data as if they were the population for the purpose of evaluating the distribution of interest, which hereby renders more efficient and reliable inference. As stated by Horowitz (1999), under mild regularity conditions, the bootstrap yields an approximation to the distribution of an estimator or test statistic that is at least as accurate as the approximation obtained from first-order asymptotic theory. In fact, the bootstrap is often more accurate in finite samples. Thus, it can provide a practical method for improving upon first-order approximations, as it is not unusual for an asymptotically unbiased estimator to have a large finite-sample bias. As a result, the nominal probability that a test based on an asymptotic critical value rejects a true null hypothesis can be very different from the true rejection probability (the probability of a Type I error).

To see this Horowitz (1999) define $F_n(\psi)$ as the estimated finite-sample CDF for the investigated parameter vector, $\psi$. Employing a first-order Taylor expansion around $\tilde{\psi} = \psi_0$, yields

$$F_n(\psi) = F_n(\psi_0) + \left[ \frac{\partial F_n(\psi_0)}{\partial \psi} \right] (\psi - \psi_0) + O(n^{-2})$$

(3-23)

and from the Lindberg-Levy’s central limit theorem we know that

---

62 For a thorough examination see Horowitz (1999).
63 This improvement is labelled asymptotic refinement.
Part III: Empirical Investigation

\[ n^{\frac{1}{2}} \left[ F_n (\psi) - F_n (\psi_0) \right] \xrightarrow{d} N(0, V) \]

thus, the asymptotic CDF of the test statistic is the normal distribution, with an error of approximation of \( O(n^{1/2}) \). In evaluating the accuracy of the bootstrap method Horowitz (1999) develop higher-order asymptotic approximations of the CDF, which yield an approximation error of \( O(n^{-1}) \) if the investigated statistics are asymptotic pivotal. Thus, if we expect that the investigated statistics in this thesis are pivotal then the bootstrap method provides asymptotic refinements.\(^{64}\)

3.3.2.1 The CMLBoot in GAUSS

The bootstrap methodology is implemented in GAUSS through the CMLBoot procedure, which only requires marginal changes in the original CML program. The CMLBoot saves the parameter estimates for every trial in a new data set. This set is afterwards centered around the original maximum likelihood estimates \( \psi_0 \), which gives the bootstrapped covariance matrix

\[
\text{Cov} \left[ \tilde{\psi}_b \right] = \frac{1}{B} \sum_{b=1}^{B} \left[ \tilde{\psi}_b - \psi_0 \right] \left[ \tilde{\psi}_b - \psi_0 \right]'
\]

where the number of samples to be drawn are set to \( B = 1000 \). Thus, the EDF is then determined by

\[ T^*_n = n^{1/2} \frac{\tilde{\psi}_b - \psi_0}{\text{Std} (\tilde{\psi}_b)} \]

for every trial, \( b \). Sorting the values of \( T^*_n \), the new critical values are obtained in the usual way i.e. selecting the 950’th observation from the sorted data set (or just use the CMLBlimits procedure).

3.3.2.2 Revision of the critical values

During the Bootstrap analysis we encountered several problems. Firstly, the EDF of the coefficients of correlation with the Treasury term structure provided nonsense results. Therefore we do not report the new critical values for \( \beta_1 \) and \( \beta_2 \). Secondly, for reasons outlined in section 3.1.4, the

\(^{64}\) Albeit the advantage of the bootstrap method, it should be emphasized that it still is an approximation and therefore not exact. Even when theory indicates that bootstrapping provides asymptotic refinements the numerical performance may be poor. In some cases the numerical accuracy of bootstrap approximations may be even worse than the accuracy of first-order approximations, which is most likely to happen when the asymptotic covariance matrix of the estimators is nearly singular. Thus, the bootstrap should not be used uncritically. See Horowitz (1999) pp. 3.
Part III: Empirical Investigation

majority of the runs of the hazard model for the BBB-rating did not converge providing us with insufficient information about the EDF. The results are therefore excluded. Finally, as the ML-model was unable to pin down the original estimates of the constants ($\alpha_r$ and $\alpha_h$) they were fixed in the bootstrap analysis and therefore not shown in Table 3-12.

As illustrated in the table the first-order critical values are substantially below the bootstrapped critical values. This fact depicts the obvious finite-sample bias and that the empirical distribution functions display fat tails. As a consequence the estimated parameters of the models become less significant. Besides the market price of risk parameter of the first state variable, this does not change the conclusions of the parameters in the two-factor Treasury term-structure model. In contrast the fat tails of the EDF’s imply that many of the parameter estimates of the hazard rates, except for the A-rating, become insignificant. This, of course, indicates that the one-factor model experiences problems in capturing the dynamics of the credit spread, concretized in high parameter variance.

Table 3-12: Bootstrapped Critical Values and original t-statistics

<table>
<thead>
<tr>
<th>State variables</th>
<th>Parameters</th>
<th>$\kappa_i$</th>
<th>$\theta_i$</th>
<th>$\sigma_i^2$</th>
<th>$\lambda_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1$</td>
<td>t-stat</td>
<td>24.42</td>
<td>26.13</td>
<td>25.27</td>
<td>-4.13</td>
</tr>
<tr>
<td></td>
<td>critical</td>
<td>20.44</td>
<td>19.71</td>
<td>22.08</td>
<td>-4.92</td>
</tr>
<tr>
<td>$x_2$</td>
<td>t-stat</td>
<td>2.59</td>
<td>2.56</td>
<td>53.72</td>
<td>-3.67</td>
</tr>
<tr>
<td></td>
<td>critical</td>
<td>2.06</td>
<td>1.67</td>
<td>36.74</td>
<td>-3.00</td>
</tr>
<tr>
<td>$h_{AAA}$</td>
<td>t-stat</td>
<td>2.91</td>
<td>3.10</td>
<td>3.56</td>
<td>-5.37</td>
</tr>
<tr>
<td></td>
<td>critical</td>
<td>10.59</td>
<td>10.76</td>
<td>4.11</td>
<td>-6.87</td>
</tr>
<tr>
<td>$h_{AA}$</td>
<td>t-stat</td>
<td>5.73</td>
<td>5.09</td>
<td>7.01</td>
<td>-6.37</td>
</tr>
<tr>
<td></td>
<td>critical</td>
<td>7.99</td>
<td>11.18</td>
<td>5.89</td>
<td>-5.43</td>
</tr>
<tr>
<td>$h_A$</td>
<td>t-stat</td>
<td>7.39</td>
<td>5.01</td>
<td>6.91</td>
<td>-7.67</td>
</tr>
<tr>
<td></td>
<td>critical</td>
<td>4.41</td>
<td>4.99</td>
<td>6.21</td>
<td>-5.99</td>
</tr>
</tbody>
</table>

This table shows the bootstrapped and asymptotic critical values for the parameters in the two-factor Treasury term-structure model and the one-factor hazard rate model (A-rated), 1994-2002. The table only depicts the CIR processes.

Even though we encountered some problems when investigating the excess yields of Treasuries and credit spreads of AAA-rated bonds, our results indicated that the market price of risk parameters were all significant when using the asymptotic critical values. It is obvious from the table that this conclusion to some extend is revised as we find the $\lambda$ parameter of the first state variable and the hazard rate of AAA-rated bonds to be insignificant. This corresponds to our initial expectations. Generally, we are still capable of validating the commonly observed upward sloped term structures as the risk premiums in the majority are significant (the $\lambda$ in $x_1$ is significant at a 10%...
Part III: Empirical Investigation

level). Furthermore, although the preliminary analysis of the diffusion function in section 2.4.2.2 revealed possible misspecification, Table 3-12 still discloses significant volatility estimates except for the AAA-rating. This implies that the results of the hazard rates are verified, i.e. the higher uncertainty associated with bonds of lower ratings makes investors require a higher risk premium or equivalently a lower price.

The real dissimilarities when comparing the inference based on the asymptotic and empirical distribution functions are the implied insignificances of the $\kappa$ and $\theta$ estimates. Recall that the credit-spread term structures in the period 1994-1999 were typically upward sloped with low volatility but from 2000 high volatility and level changes predominated the series. Such shifts are anticipated to influence the efficiency of the estimated parameters. Especially, when the level of the term structures fluctuate the $\theta$ estimates are difficult to measure precisely. This is revealed through the insignificant estimates when utilizing the empirical distribution function. Furthermore, the difficulties in efficiently pinning down the $\theta$ parameters influence the models ability in capturing the mean reversion in the data. As a consequence the $\kappa$ and $\theta$ estimates of the AAA- and AA-ratings are insignificant, which implies that the equilibrium default probabilities are very small and perhaps not even larger than the incremental estimate of 30 bp. Nevertheless, we find that the A-rating still involves a significantly higher equilibrium default probability, which reconciles the empirically higher spread for the A-rating compared to AAA and AA. We additionally believe that this aspect of insignificant $\kappa$ and $\theta$ estimates is a consequence of the investigation in section 2.4.2.1, which revealed possible misspecification of the drift function. The probable non-linearities might influence the models ability in estimating the drift parameters. The bootstrap analysis confirms this element and, however, implies that capturing the properties of the data sufficiently requires a different model framework than the one proposed in this thesis.

3.3.3 Specification Tests

In the above chapters we outlined some empirical problems with our model framework. To narrow the source of these problems we examine the core of the framework i.e. the inherent CIR processes. As we several times have been reminded that the drift and volatility structures of the Treasury and hazard rate formulations might be misspecified this chapter formally investigates the potential problem. In doing so we follow Duffee (1996) and define the estimated innovations $e_{i,t}$ as $[\hat{\psi}_{i,t} - E_{r,t-1}(\psi_{i,t})]/\sqrt{Var_{r,t-1}(\psi_{i,t})}$, where $\psi_i$ is either the interest or default rate.\textsuperscript{65} The intuition is

\textsuperscript{65}The definitions of $E_{r,t-1}(\bullet)$ and $Var_{r,t-1}(\bullet)$ are the conditional moments in the CIR model cf. (3-23).
Part III: Empirical Investigation

that if the processes are correctly specified, prior values of $\varepsilon$ should not help forecast $\varepsilon$ – or stated otherwise the normalized innovations should resemble white noise.

Like Duffee (1996) we investigate the drift function by fitting $\varepsilon_{i,t}$ to an AR(3)

$$
\varepsilon_{i,t} = \alpha + \sum_{j=1}^{3} \phi_j \varepsilon_{i,t-j} + \eta_{i,t}
$$

If the model is correctly specified the AR coefficients should equal zero. The results of the analysis are illustrated in the table below.

**Table 3-13: Specification Test of the Predictability of the Estimated Innovations to the Instantaneous Interest Rate and Default Probabilities**

<table>
<thead>
<tr>
<th></th>
<th>Intercept</th>
<th>Lag 1</th>
<th>Lag 2</th>
<th>Lag 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treasury</td>
<td>-0.023</td>
<td>-0.015</td>
<td>-0.110</td>
<td>-0.111</td>
</tr>
<tr>
<td></td>
<td>(-0.56)</td>
<td>(0.28)</td>
<td>(-2.10)</td>
<td>(-2.07)</td>
</tr>
<tr>
<td>AAA</td>
<td>0.061</td>
<td>-0.229</td>
<td>-0.042</td>
<td>-0.030</td>
</tr>
<tr>
<td></td>
<td>(0.83)</td>
<td>(-4.60)</td>
<td>(-0.82)</td>
<td>(-0.59)</td>
</tr>
<tr>
<td>AA</td>
<td>0.050</td>
<td>-0.237</td>
<td>-0.044</td>
<td>-0.013</td>
</tr>
<tr>
<td></td>
<td>(0.93)</td>
<td>(-4.76)</td>
<td>(-0.84)</td>
<td>(-0.25)</td>
</tr>
<tr>
<td>A</td>
<td>0.036</td>
<td>-0.254</td>
<td>-0.041</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(1.22)</td>
<td>(-5.13)</td>
<td>(-0.80)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>BBB</td>
<td>0.034</td>
<td>-0.094</td>
<td>0.021</td>
<td>0.066</td>
</tr>
<tr>
<td></td>
<td>(1.51)</td>
<td>(-1.89)</td>
<td>(-0.43)</td>
<td>(1.31)</td>
</tr>
</tbody>
</table>

The table shows the autoregression analysis of the normalized innovations to the instantaneous rates (Treasury and AAA, AA, A and BBB-rated), 1994-2002. T-statistics are in parentheses.

From Table 3-13 it is evident that specification error is present. For the Treasury and AAA to A-rated bonds the normalized shocks $\eta$ exhibit mean reversion, indicating that estimated values of $\kappa_i$ do not adequately capture the extend of mean reversion in $\psi_i$. The combined point estimates generally indicate that 30 percent of a given shock is reversed within three weeks. Surprisingly we find that the estimated mean reversion for the BBB-rated indices is rich enough to capture the mean reversion in the instantaneous default rates, as all t-statistics are insignificant implying the wanted white noise process.

Comparing these results with the preliminary analysis of the drift function in section 2.4.2.1 gives the indication that the problems with the estimated $\kappa_i$ may arise from possible misspecification of the drift function in the CIR model rather than problems with the derived ML-model. Still, the analysis conducted above is an examination of the models capabilities in replicating the properties of the underlying Treasury yields and credit spreads, which to some extend are less encouraging.
If we take a closer look at the models ability to capture the volatility in the observed yields, it is very important that the square-root processes capture the true relation between the state variables and innovations herein. The intuition is that when the derived diffusion functions are correctly specified there are no ARCH-effects in the innovations, i.e. the models capture the implied volatility structure in the instantaneous rates. We investigate this very important aspect by estimating

$$\varepsilon_{it}^2 = \alpha + \sum_{j=1}^{3} \phi_j \varepsilon_{i,j-1}^2 + \psi \varepsilon_{i,j-1} + \eta_{it}$$

where the coefficients should be insignificantly different from zero if the diffusion functions are correctly specified. Davidian and Carroll (1987) conclude that when distributions are characterized by fat tails, it is often more efficient to estimate volatility functions using absolute values of residuals instead of squares. Therefore we also estimate the above equation substituting $\varepsilon_{it}^2$ with $|\varepsilon_{it}|$ as we expect fat tails in our distributions cf. the bootstrap analysis. The results of the investigation are illustrated in the table below where the first sets of numbers are the results when using the squared residuals.

### Table 3-14: Specification Test of the Predictability of the Volatility of Estimated Shocks to the Instantaneous Interest Rate and Default Probabilities.

<table>
<thead>
<tr>
<th>Rating</th>
<th>Intercept</th>
<th>Lag 1</th>
<th>Lag 2</th>
<th>Lag 3</th>
<th>$\psi_{i,t-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treasury</td>
<td>2.809</td>
<td>-0.020a</td>
<td>0.055a</td>
<td>0.052a</td>
<td>4.493</td>
</tr>
<tr>
<td></td>
<td>1.275</td>
<td>0.065a</td>
<td>0.089a</td>
<td>0.135</td>
<td>1.579</td>
</tr>
<tr>
<td>AAA</td>
<td>4.651</td>
<td>0.138</td>
<td>0.044a</td>
<td>0.054a</td>
<td>4.077</td>
</tr>
<tr>
<td></td>
<td>0.684</td>
<td>0.597</td>
<td>-0.022a</td>
<td>0.025a</td>
<td>0.620</td>
</tr>
<tr>
<td>AA</td>
<td>0.326</td>
<td>0.171</td>
<td>0.044a</td>
<td>0.063a</td>
<td>2.156</td>
</tr>
<tr>
<td></td>
<td>0.855</td>
<td>0.731</td>
<td>-0.031a</td>
<td>0.021a</td>
<td>1.661</td>
</tr>
<tr>
<td>A</td>
<td>0.510</td>
<td>0.205</td>
<td>-0.030a</td>
<td>0.018a</td>
<td>1.055a</td>
</tr>
<tr>
<td></td>
<td>0.556</td>
<td>0.377</td>
<td>-0.015a</td>
<td>0.054a</td>
<td>0.576a</td>
</tr>
<tr>
<td>BBB</td>
<td>0.682</td>
<td>0.243</td>
<td>-0.072a</td>
<td>0.037a</td>
<td>0.873a</td>
</tr>
<tr>
<td></td>
<td>0.899</td>
<td>0.326</td>
<td>-0.042a</td>
<td>0.095a</td>
<td>1.718</td>
</tr>
</tbody>
</table>

This table shows the ARCH analysis of the normalized innovations to the instantaneous interest and default rates (AAA, AA, A and BBB-rating), 1994-2002. The estimates with the superscript a are not significant at the 5% level.

When viewing the results of the Treasury model the table only indicates a very weak ARCH-effect as only the third lag of the absolute value is significant at the 5% level. Turning to the credit-spread models both sets of results show that there are ARCH-effects in $\varepsilon$ irrespectively of the chosen rating. The evidence of ARCH-effects is stronger when volatility is measured in abso-
Part III: Empirical Investigation

lute values as the AR(1) coefficients are larger. Independent of the chosen method we find that the volatility is positively associated with prior values of $\psi$ but this relationship is only clearly significant when analyzing the instantaneous interest rate and default rates of AAA and AA-ratings. Based on this evidence, the assumption that the volatility of instantaneous rates rises with the square root appears to understate the true sensitivity of the volatility. The conclusion, that the tendencies are most prevalent in the higher ratings, entails that the derived model performs better when reproducing the empirical properties of the lower rating classes. This is consistent with the accuracy measures in section 3.2.2. The results are in accordance with the preliminary diffusion analysis conducted in section 2.4.2.2, where the analogous standardized residuals exposed ARCH-effects revealing some difficulties with the specified volatility model.

Like the bootstrap analysis, the results of this section indicate that the CIR model neglects to incorporate certain empirical features of the data. This postulate may revise the usefulness of the derived model framework, and especially the applied affine-factor model. The results suggest a model capable of incorporating stochastic levels and perhaps higher volatility elasticity cf. CKLS (1992). However, we see the drift function as the most problematic. Indeed, we believe if a stochastic long run level is incorporated like in the double decay model, the improvements in refining the volatility structure may be minimum.

3.4 Pricing Accuracy and Model Evaluation

In light of the above-mentioned limitations, we now shift our attention to the overall fit of the full three-factor model. We do not report a thorough examination of the overall sample fit as the conclusions are trivially in light of the examination of the separate factors in chapter 3.2. Instead we report the overall conclusions and discuss the usefulness of the model.

The accuracy of the model is reported in Appendix F. In summary, the RMSE (Theil $U$) range from 10.7 (1.6%) to 33.5 (4.7%) across all ratings in the short and intermediate maturities, and in the long end the accuracy is 17.8 (2.75%) to 88.4 (12.0%). This tendency is graphically illustrated in Figure 3-5, and verifies the model’s inability in describing longer maturities. In addition, we find diversity in the accuracy of the predictions for the long and short maturities when focusing on the different credit ratings. For shorter maturities the fit generally improves with lower rating, whereas the opposite is observed for longer maturities. More explicitly we find the model’s replicating ability the highest with RMSE (Theil $U$) as low as 10.7 bp – 13.2 bp (1.60% – 2.03%) for the shorter A-rated indices and RMSE (Theil $U$) of 17.8 bp – 38.0 bp (2.75% - 5.35%) for the longer AAA-rated indices. This tendency makes sense in the long end, as it is interpreted in line
Part III: Empirical Investigation

with the observed volatilities of the credit spreads, where the lower rated bonds exhibit higher volatility cf. Appendix A. The tendency in the short spreads is in contrast to the observed volatilities, but explained by the model’s inability to pin down the estimate of the constant term.

Figure 3-5: The Observed and Predicted Term Structure of A-Rated Corporate Bonds Indices

The figure illustrates the observed and predicted time series, of the maturity segments of A-rated corporate bond indices, implied by the three-factor corporate bond model, 1994-2002.

To set the model’s usefulness in perspective, and investigate whether it is a functional tool we would like to conduct an out of sample analysis. But, as mentioned in chapter 3.1 the derived log-likelihood models did not exclude the possibility of non-stationary state variables. This choice should, of course, be seen in light of the empirical evidence when using the CIR model as the process that governs the state variables. Chen and Scott (1993) argue that at least one of the state variables describing the default-free term structure must have a low $\kappa$ (i.e. is close to bearing non-stationary) if the properties of long-term yields are to be sufficiently captured by the model. Furthermore, we saw that capturing the empirical behavior of the credit spreads requires not only non-stationary but explosive intensities under the pricing measure. This, however, reveals a serious obstacle in conducting an out of sample analysis because we need to simulate the state variables. When a process is non-stationary or explosive such a simulation is bound to render nonsense, as the path taken by the process is highly spurious with a trend that cannot be alleviated.
Therefore we cannot use the derived three-factor model to forecast the evolution of corporate bond indices, and in application this means that we cannot price credit derivatives unless very simple assumptions are employed.

Instead, as the model fits the prices of shorter matured bonds well, we can calibrate or estimate the model to fit the observed term structures on a daily basis. This enables us to investigate whether the market prices the bonds consistently i.e. the model reveals which bonds are priced too high or low. But as Duffee (1996, 1999) ultimately states with his analysis on individual bond prices, the model is by no means a success. The real strengths of the model are therefore not in practical applications, it is more in a theoretical sense of understanding the underlying dimensions of corporate bond prices. In particular, to determine which features of the model that are empirically well-supported and which are not.
References


Moody’s Investors Service (2002) ”Special Comments” Moody’s.


Appendix

Appendix A: The Term Structures

Appendix B: Explaining Credit Spreads – Regression and Cointegration

Appendix C: Excess Yields in the CIR Model

Appendix D: Investigation of the SDE

Appendix E: Pricing Accuracy of the Treasury Term-structure Model

Appendix F: Pricing Accuracy of the Three-Factor Model

Appendix G: Alternative Error Structures

Appendix H: GAUSS program for the two-factor Treasury Model

Appendix I: GAUSS program for the one-factor Credit-Spread Model

Appendix J: Deriving the Ricatti Equations

Appendix K: Deriving the Affine Solutions in the CIR Model
Appendix A: The Term Structures

The Historical US Treasury Yields, 1991 - 2002

This figure presents the historical movement of yields to maturity of ten selected zero coupon yields. The sample contains 133 monthly observations in the period January 1991 to January 2002.


<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>STD</td>
<td>Mean</td>
</tr>
<tr>
<td>1-year</td>
<td>5.05</td>
<td>1.09</td>
<td>4.90</td>
</tr>
<tr>
<td>2-year</td>
<td>5.41</td>
<td>1.03</td>
<td>5.48</td>
</tr>
<tr>
<td>3-year</td>
<td>5.68</td>
<td>0.96</td>
<td>5.90</td>
</tr>
<tr>
<td>4-year</td>
<td>5.88</td>
<td>0.94</td>
<td>6.29</td>
</tr>
<tr>
<td>5-year</td>
<td>6.01</td>
<td>0.91</td>
<td>6.45</td>
</tr>
<tr>
<td>7-year</td>
<td>6.27</td>
<td>0.91</td>
<td>6.82</td>
</tr>
<tr>
<td>10-year</td>
<td>6.54</td>
<td>0.91</td>
<td>7.20</td>
</tr>
<tr>
<td>15-year</td>
<td>6.85</td>
<td>0.89</td>
<td>7.58</td>
</tr>
<tr>
<td>20-year</td>
<td>6.96</td>
<td>0.89</td>
<td>7.74</td>
</tr>
</tbody>
</table>

This table presents the sample means and standard deviations (STD) of the zero-coupon yields. The sample contains 558 weekly observations in the period between January 1991 and January 2002. The two subperiods are January 1991 to December 1995 and January 1996 to January 2002, with 254 and 304 observations, respectively.
AAA-Rated Bond Indices:

This figure presents the historical movement of yields to maturity of five selected credit spreads for AAA-rated corporate bonds indices. The sample contains 133 monthly observations in the period January 1991 to January 2002.

Summary Statistics for Credit Spreads on AAA-rated Bond Indices, 1991 – 2002

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>STD</td>
<td>Mean</td>
</tr>
<tr>
<td>2-year</td>
<td>0.73</td>
<td>0.45</td>
<td>1.35</td>
</tr>
<tr>
<td>4-year</td>
<td>0.72</td>
<td>0.26</td>
<td>0.95</td>
</tr>
<tr>
<td>6-year</td>
<td>0.71</td>
<td>0.25</td>
<td>0.82</td>
</tr>
<tr>
<td>8-year</td>
<td>0.67</td>
<td>0.27</td>
<td>0.59</td>
</tr>
<tr>
<td>10-year</td>
<td>0.83</td>
<td>0.26</td>
<td>0.79</td>
</tr>
</tbody>
</table>

This table presents the sample means and standard deviations (STD) of the credit spread on AAA-rated corporate bond indices. The sample contains 558 weekly observations in the period between January 1991 and January 2002. The two subperiods are January 1991 to December 1993 and January 1994 to January 2002, with 146 and 412 observations, respectively.
AA-Rated Bond Indices:

This figure presents the historical movement of yields to maturity of five selected credit spreads for AA-rated corporate bonds indices. The sample contains 133 monthly observations in the period January 1991 to January 2002.

Summary Statistics for Credit Spreads on AA-rated Bond Indices, 1991 – 2002

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>STD</td>
<td>Mean</td>
</tr>
<tr>
<td>2-year</td>
<td>0.71</td>
<td>0.25</td>
<td>0.95</td>
</tr>
<tr>
<td>4-year</td>
<td>0.83</td>
<td>0.27</td>
<td>1.06</td>
</tr>
<tr>
<td>6-year</td>
<td>0.83</td>
<td>0.26</td>
<td>0.87</td>
</tr>
<tr>
<td>8-year</td>
<td>0.84</td>
<td>0.29</td>
<td>0.74</td>
</tr>
<tr>
<td>10-year</td>
<td>0.96</td>
<td>0.30</td>
<td>0.88</td>
</tr>
</tbody>
</table>

This table presents the sample means and standard deviations (STD) of the credit spread on AA-rated corporate bond indices. The sample contains 558 weekly observations in the period between January 1991 and January 2002. The two subperiods are January 1991 to December 1993 and January 1994 to January 2002, with 146 and 412 observations, respectively.
A-Rated Bond Indices:

This figure presents the historical movement of yields to maturity of five selected credit spreads for A-rated corporate bonds indices. The sample contains 133 monthly observations in the period January 1991 to January 2002.


<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>STD</td>
<td>Mean</td>
</tr>
<tr>
<td>2-year</td>
<td>0.90</td>
<td>0.37</td>
<td>1.24</td>
</tr>
<tr>
<td>4-year</td>
<td>1.05</td>
<td>0.36</td>
<td>1.34</td>
</tr>
<tr>
<td>6-year</td>
<td>1.07</td>
<td>0.34</td>
<td>1.17</td>
</tr>
<tr>
<td>8-year</td>
<td>1.06</td>
<td>0.38</td>
<td>0.99</td>
</tr>
<tr>
<td>10-year</td>
<td>1.20</td>
<td>0.37</td>
<td>1.17</td>
</tr>
</tbody>
</table>

This table presents the sample means and standard deviations (STD) of the credit spread on A-rated corporate bond indices. The sample contains 558 weekly observations in the period between January 1991 and January 2002. The two subperiods are January 1991 to December 1993 and January 1994 to January 2002, with 146 and 412 observations, respectively.
This figure presents the historical movement of yields to maturity of five selected credit spreads for BBB-rated corporate bonds indices. The sample contains 133 monthly observations in the period January 1991 to January 2002.

### Summary Statistics for Credit Spreads on BBB-rated Bond Indices, 1991 – 2002

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2-year</td>
<td>Mean 1.43, STD 0.69</td>
<td>Mean 1.96, STD 0.53</td>
<td>Mean 1.20, STD 0.62</td>
<td></td>
</tr>
<tr>
<td>4-year</td>
<td>Mean 1.55, STD 0.58</td>
<td>Mean 2.02, STD 0.33</td>
<td>Mean 1.35, STD 0.53</td>
<td></td>
</tr>
<tr>
<td>6-year</td>
<td>Mean 1.55, STD 0.51</td>
<td>Mean 1.74, STD 0.27</td>
<td>Mean 1.46, STD 0.53</td>
<td></td>
</tr>
<tr>
<td>8-year</td>
<td>Mean 1.53, STD 0.45</td>
<td>Mean 1.42, STD 0.21</td>
<td>Mean 1.55, STD 0.49</td>
<td></td>
</tr>
<tr>
<td>10-year</td>
<td>Mean 1.69, STD 0.42</td>
<td>Mean 1.74, STD 0.16</td>
<td>Mean 1.65, STD 0.48</td>
<td></td>
</tr>
</tbody>
</table>

This table presents the sample means and standard deviations (STD) of the credit spread on BBB-rated corporate bond indices. The sample contains 558 weekly observations in the period between January 1991 and January 2002. The two subperiods are January 1991 to December 1993 and January 1994 to January 2002, with 146 and 412 observations, respectively.
Appendix B: Explaining Credit Spreads – Regression and Cointegration

January 1991-2002

Phillips-Perron Tests

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>t-Stat</td>
<td>-2.73</td>
</tr>
<tr>
<td>Critical value</td>
<td>-3.42</td>
</tr>
</tbody>
</table>

This table presents the Phillips-Perron Unit-Root test based on weekly observations in the period January 1991-2002. The critical values are obtained at the 5% significance level.

Unit-Roots Tests AAA – BBB Rated Bonds

<table>
<thead>
<tr>
<th>ADF-test Spread to Treasury</th>
<th>01/1991-01/2002</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short Index</td>
<td></td>
</tr>
<tr>
<td>AAA</td>
<td>-1.70</td>
</tr>
<tr>
<td>AA</td>
<td>-2.78</td>
</tr>
<tr>
<td>A</td>
<td>-2.12</td>
</tr>
<tr>
<td>BBB</td>
<td>-2.50</td>
</tr>
<tr>
<td>Intermediate Index</td>
<td></td>
</tr>
<tr>
<td>AAA</td>
<td>-1.87</td>
</tr>
<tr>
<td>AA</td>
<td>-2.09</td>
</tr>
<tr>
<td>A</td>
<td>-1.66</td>
</tr>
<tr>
<td>BBB</td>
<td>-2.38</td>
</tr>
<tr>
<td>Long Index</td>
<td></td>
</tr>
<tr>
<td>AAA</td>
<td>-1.75</td>
</tr>
<tr>
<td>AA</td>
<td>-3.03</td>
</tr>
<tr>
<td>A</td>
<td>-1.76</td>
</tr>
<tr>
<td>BBB</td>
<td>-2.36</td>
</tr>
</tbody>
</table>

This table presents the ADF-Unit-Root test based on weekly observations in the period January 1991-2002. The critical values are obtained at the 5% significance level. The credit spreads are defined as the difference between the corporate bond yields and the corresponding zero-coupon yields. E.g. the short index denotes the spread based one a duration of two years of the corporate and Treasury bond, respectively.
## January 1994-2002
### Unit-Roots Tests AAA – BBB Rated Bonds

<table>
<thead>
<tr>
<th>ADF-test</th>
<th>Spread to Treasury</th>
<th>01/1994-01/2002</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>t-Stat</td>
</tr>
<tr>
<td>Short Index</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AAA</td>
<td></td>
<td>-3.12</td>
</tr>
<tr>
<td>AA</td>
<td></td>
<td>-2.93</td>
</tr>
<tr>
<td>A</td>
<td></td>
<td>-2.75</td>
</tr>
<tr>
<td>BBB</td>
<td></td>
<td>-2.60</td>
</tr>
</tbody>
</table>

| Intermediate Index |       |       |               |                     |
| AAA                | -2.56 | -3.42 | I(1)         |
| AA                 | -2.67 | -3.42 | I(1)         |
| A                  | -2.46 | -3.42 | I(1)         |
| BBB                | -2.37 | -3.42 | I(1)         |

| Long Index |       |       |               |                     |
| AAA       | -2.26 | -3.42 | I(1)         |
| AA        | -2.89 | -3.42 | I(1)         |
| A         | -2.29 | -3.42 | I(1)         |
| BBB       | -2.06 | -3.42 | I(1)         |

This table presents the ADF-Unit-Root test based on weekly observations in the period January 1994-2002. The critical values are obtained at the 5% significance level. The credit spreads are defined as the difference between the corporate bond yields and the corresponding zero-coupon yields. E.g. the short index denotes the spread based on a duration of two years of the corporate and Treasury bond, respectively.

<table>
<thead>
<tr>
<th>Spread to Treasury</th>
<th>Constant</th>
<th>ΔLevel ST</th>
<th>ΔLevel LT</th>
<th>ΔSlope</th>
<th>Non-LE</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Short Index</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AAA</td>
<td>-0.001</td>
<td>-0.113</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.10</td>
</tr>
<tr>
<td>(0.51)</td>
<td>(-4.68)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>-0.003</td>
<td>-0.121</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.099</td>
</tr>
<tr>
<td>(-1.35)</td>
<td>(-5.12)</td>
<td></td>
<td></td>
<td>-</td>
<td>(1.71)</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>-0.002</td>
<td>-0.125</td>
<td>-</td>
<td>-</td>
<td>0.148</td>
<td>0.12</td>
</tr>
<tr>
<td>(0.83)</td>
<td>(-4.56)</td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BBB</td>
<td>0.002</td>
<td>-0.118</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.10</td>
</tr>
<tr>
<td>(0.41)</td>
<td>(-2.54)</td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| **Intermediate Index** |          |          |           |        |        |    |
| AAA                | 0.002    | -0.072   | -         | -0.043 | -0.120 | 0.09 |
| (1.12)             | (-3.96)  |          |           | (-1.87)| (-2.14)|    |
| AA                 | 0.000    | -0.104   | -0.068    | -      | -      | 0.12 |
| (0.25)             | (-5.20)  |          | (-2.85)   | -      |        |    |
| A                  | 0.001    | -0.086   | -         | -      | -      | 0.07 |
| (0.41)             | (-4.73)  |          |           | -      |        |    |
| BBB                | 0.002    | -0.133   | -0.067    | -      | -      | 0.10 |
| (0.79)             | (-6.36)  |          | (-2.12)   | -      |        |    |

| **Long Index**     |          |          |           |        |        |    |
| AAA                | 0.000    | -0.186   | -         | -      | -      | 0.26 |
| (0.25)             | (-9.02)  |          |           | -      |        |    |
| AA                 | 0.000    | -0.170   | -         | -      | -      | 0.21 |
| (0.08)             | (-8.74)  |          |           | -      |        |    |
| A                  | 0.000    | -0.151   | -         | -      | -      | 0.21 |
| (0.30)             | (-11.56) |          |           | -      |        |    |
| BBB                | 0.000    | -0.235   | -         | -      | -      | 0.32 |
| (0.29)             | (-12.93) |          |           | -      |        |    |

The t-statistics are in parentheses. The standard errors used in calculating the t-statistics are adjusted for heteroscedasticity and autocorrelation using the Newey-West estimator.
### Johansen Maximum Likelihood Cointegration Results: AAA – BBB rated, January 1994-2002

<table>
<thead>
<tr>
<th>Index Type</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Short Index</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda$ (eigenvalue)</td>
<td>0.041</td>
<td>0.041</td>
<td>0.040</td>
<td>0.040</td>
</tr>
<tr>
<td>L.R. Value Critical Value</td>
<td>21.87$^*\ $</td>
<td>21.87$^*\ $</td>
<td>21.56$^*\ $</td>
<td>22.27$^*\ $</td>
</tr>
<tr>
<td>Conclusion</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td><strong>Intermediate Index</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda$ (eigenvalue)</td>
<td>0.045</td>
<td>0.050</td>
<td>0.050</td>
<td>0.045</td>
</tr>
<tr>
<td>L.R. Value Critical Value</td>
<td>26.83</td>
<td>28.81</td>
<td>27.36</td>
<td>25.06</td>
</tr>
<tr>
<td>Conclusion</td>
<td>cointegration</td>
<td>cointegration</td>
<td>cointegration</td>
<td>cointegration</td>
</tr>
<tr>
<td><strong>Long Index</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda$ (eigenvalue)</td>
<td>0.065</td>
<td>0.051</td>
<td>0.052</td>
<td>0.058</td>
</tr>
<tr>
<td>L.R. Value Critical Value</td>
<td>33.17</td>
<td>27.93</td>
<td>27.32</td>
<td>30.99</td>
</tr>
<tr>
<td>Conclusion</td>
<td>cointegration</td>
<td>cointegration</td>
<td>cointegration</td>
<td>cointegration</td>
</tr>
</tbody>
</table>

The critical values are obtained at the 5% significance level ($a$ refers to a significant result at the 10% level). The null hypothesis is that there is no cointegration vector (rank = 0) against the alternative hypothesis that there is one cointegrating vector (rank = 1).
Normalized Cointegration Coefficients – One Cointegrating Equation

Short Index

AAA: $CS = -0.101 \text{Level ST} - 0.069 \text{Slope}$
AA: $CS = -0.117 \text{Level ST} - 0.145 \text{Slope}$
A: $CS = -0.123 \text{Level ST} - 0.178 \text{Slope}$
BBB: $CS = -0.178 \text{Level ST} - 0.292 \text{Slope}$

Intermediate Index

AAA: $CS = -0.117 \text{Level ST} - 0.055 \text{Slope}$
AA: $CS = -0.129 \text{Level ST} - 0.115 \text{Slope}$
A: $CS = -0.162 \text{Level ST} - 0.165 \text{Slope}$
BBB: $CS = -0.221 \text{Level ST} - 0.276 \text{Slope}$

Long Index

AAA: $CS = -0.122 \text{Level LT} - 0.031 \text{Slope}$
AA: $CS = -0.179 \text{Level LT} - 0.062 \text{Slope}$
A: $CS = -0.192 \text{Level LT} - 0.062 \text{Slope}$
BBB: $CS = -0.250 \text{Level LT} - 0.009 \text{Slope}$
Appendix C: Excess Yields in the CIR Model

Descriptives of Excess Yields, AAA-BBB Rated Bond Indices, 1994-2002

<table>
<thead>
<tr>
<th>Excess Yield</th>
<th>Mean</th>
<th>STD</th>
<th>Mean/STD</th>
<th>Calculations based on mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E(2)</td>
<td>0.112</td>
<td>0.091</td>
<td>1.23</td>
<td>E(CS) 0.5052</td>
</tr>
<tr>
<td>E(4)</td>
<td>0.133</td>
<td>0.098</td>
<td>1.36</td>
<td>STD 0.1876</td>
</tr>
<tr>
<td>E(6)</td>
<td>0.159</td>
<td>0.127</td>
<td>1.25</td>
<td>Range CS [0.143;1.181]</td>
</tr>
<tr>
<td>E(8)</td>
<td>0.198</td>
<td>0.167</td>
<td>1.20</td>
<td>θ 0.244 (0.1876-1.30)</td>
</tr>
<tr>
<td>E(10)</td>
<td>0.314</td>
<td>0.218</td>
<td>1.44</td>
<td>αh 0.261 (0.5052 – 0.244)</td>
</tr>
<tr>
<td>Mean</td>
<td>0.183</td>
<td>0.148</td>
<td>1.30</td>
<td>Range ĥ [-0.118;0.920]</td>
</tr>
<tr>
<td>AA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E(2)</td>
<td>0.113</td>
<td>0.090</td>
<td>1.26</td>
<td>E(CS) 0.6252</td>
</tr>
<tr>
<td>E(4)</td>
<td>0.117</td>
<td>0.096</td>
<td>1.22</td>
<td>STD 0.2137</td>
</tr>
<tr>
<td>E(6)</td>
<td>0.188</td>
<td>0.131</td>
<td>1.43</td>
<td>Range CS [0.2816;1.293]</td>
</tr>
<tr>
<td>E(8)</td>
<td>0.246</td>
<td>0.173</td>
<td>1.42</td>
<td>θ 0.293 (0.2137-1.37)</td>
</tr>
<tr>
<td>E(10)</td>
<td>0.347</td>
<td>0.231</td>
<td>1.50</td>
<td>αh 0.332 (0.6252 – 0.293)</td>
</tr>
<tr>
<td>Mean</td>
<td>0.202</td>
<td>0.153</td>
<td>1.37</td>
<td>Range ĥ [-0.051;0.960]</td>
</tr>
<tr>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E(2)</td>
<td>0.149</td>
<td>0.111</td>
<td>1.34</td>
<td>E(CS) 0.7806</td>
</tr>
<tr>
<td>E(4)</td>
<td>0.169</td>
<td>0.121</td>
<td>1.40</td>
<td>STD 0.3304</td>
</tr>
<tr>
<td>E(6)</td>
<td>0.257</td>
<td>0.156</td>
<td>1.64</td>
<td>Range CS [0.3367;1.7586]</td>
</tr>
<tr>
<td>E(8)</td>
<td>0.309</td>
<td>0.189</td>
<td>1.63</td>
<td>θ 0.502 (0.3304-1.52)</td>
</tr>
<tr>
<td>E(10)</td>
<td>0.410</td>
<td>0.259</td>
<td>1.58</td>
<td>αh 0.279 (0.7806 – 0.502)</td>
</tr>
<tr>
<td>Mean</td>
<td>0.259</td>
<td>0.176</td>
<td>1.52</td>
<td>Range ĥ [0.058;1.480]</td>
</tr>
<tr>
<td>BBB</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E(2)</td>
<td>0.135</td>
<td>0.119</td>
<td>1.13</td>
<td>E(CS) 1.2385</td>
</tr>
<tr>
<td>E(4)</td>
<td>0.145</td>
<td>0.136</td>
<td>1.07</td>
<td>STD 0.6477</td>
</tr>
<tr>
<td>E(6)</td>
<td>0.249</td>
<td>0.203</td>
<td>1.22</td>
<td>Range CS [0.4670;3.4850]</td>
</tr>
<tr>
<td>E(8)</td>
<td>0.355</td>
<td>0.255</td>
<td>1.31</td>
<td>θ 0.821 (0.0677-1.27)</td>
</tr>
<tr>
<td>E(10)</td>
<td>0.432</td>
<td>0.269</td>
<td>1.61</td>
<td>αh 0.417 (1.2385 – 0.821)</td>
</tr>
<tr>
<td>Mean</td>
<td>0.263</td>
<td>0.206</td>
<td>1.27</td>
<td>Range ĥ [-0.050;0.960]</td>
</tr>
</tbody>
</table>

The table shows the descriptives for the credit spreads, their implied estimates of the state-variable processes and the inherent ranges of the state variables.
Appendix D: Investigation of the SDE

Treasury


AAA-rated corporate bonds


AA-rated corporate bonds

A-rated corporate bonds


BBB-rated corporate bonds


Autocorrelations for the squared scaled residuals and GARCH-analysis of the scaled residuals

<table>
<thead>
<tr>
<th>Volatility model</th>
<th>ARCH</th>
<th>GARCH</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coef.</td>
<td>STD</td>
</tr>
<tr>
<td>AAA</td>
<td>0.073</td>
<td>0.033</td>
</tr>
<tr>
<td>AA</td>
<td>0.157</td>
<td>0.054</td>
</tr>
<tr>
<td>A</td>
<td>0.216</td>
<td>0.043</td>
</tr>
<tr>
<td>BBB</td>
<td>0.137</td>
<td>0.032</td>
</tr>
</tbody>
</table>

The table shows the volatility analysis of the time series for the scaled residuals of AAA-BBB rated bonds. The p-values are based on the z-statistics.
Appendix E: Pricing Accuracy of the Treasury Term-structure Model

The Two-Factor and the Naïve Martingale Model’s Pricing Accuracy, 1994 - 2002

<table>
<thead>
<tr>
<th>Maturity</th>
<th>Correlation</th>
<th>Correlation (Δ)</th>
<th>Max. Deviation</th>
<th>Mean Dev.</th>
<th>RMSE</th>
<th>Theil U</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-year</td>
<td>0.995</td>
<td>0.977</td>
<td>59.6 bp</td>
<td>10.5 bp</td>
<td>13.1 bp</td>
<td>2.44%</td>
</tr>
<tr>
<td></td>
<td>(0.989)</td>
<td>(0.276)</td>
<td>(66.7)</td>
<td>(10.4)</td>
<td>(14.2)</td>
<td>(2.65)</td>
</tr>
<tr>
<td>2-year</td>
<td>0.999</td>
<td>0.991</td>
<td>47.6 bp</td>
<td>5.7 bp</td>
<td>7.4 bp</td>
<td>1.32%</td>
</tr>
<tr>
<td></td>
<td>(0.987)</td>
<td>(0.419)</td>
<td>(63.4)</td>
<td>(11.4)</td>
<td>(15.3)</td>
<td>(2.73)</td>
</tr>
<tr>
<td>3-year</td>
<td>0.998</td>
<td>0.992</td>
<td>28.2 bp</td>
<td>9.7 bp</td>
<td>12.0 bp</td>
<td>2.07%</td>
</tr>
<tr>
<td></td>
<td>(0.985)</td>
<td>(0.545)</td>
<td>(62.5)</td>
<td>(11.7)</td>
<td>(15.5)</td>
<td>(2.67)</td>
</tr>
<tr>
<td>4-year</td>
<td>0.998</td>
<td>0.994</td>
<td>23.9 bp</td>
<td>8.4 bp</td>
<td>10.2 bp</td>
<td>1.74%</td>
</tr>
<tr>
<td></td>
<td>(0.984)</td>
<td>(0.596)</td>
<td>(60.1)</td>
<td>(11.8)</td>
<td>(15.4)</td>
<td>(2.62)</td>
</tr>
<tr>
<td>5-year</td>
<td>0.998</td>
<td>0.995</td>
<td>21.6 bp</td>
<td>4.7 bp</td>
<td>6.2 bp</td>
<td>1.05%</td>
</tr>
<tr>
<td></td>
<td>(0.983)</td>
<td>(0.647)</td>
<td>(59.6)</td>
<td>(11.6)</td>
<td>(15.1)</td>
<td>(2.54)</td>
</tr>
<tr>
<td>6-year</td>
<td>0.999</td>
<td>0.998</td>
<td>21.7 bp</td>
<td>2.9 bp</td>
<td>4.0 bp</td>
<td>0.66%</td>
</tr>
<tr>
<td></td>
<td>(0.983)</td>
<td>(0.675)</td>
<td>(59.8)</td>
<td>(11.5)</td>
<td>(14.9)</td>
<td>(2.47)</td>
</tr>
<tr>
<td>7-year</td>
<td>0.999</td>
<td>0.998</td>
<td>20.0 bp</td>
<td>3.2 bp</td>
<td>4.2 bp</td>
<td>0.68%</td>
</tr>
<tr>
<td></td>
<td>(0.982)</td>
<td>(0.703)</td>
<td>(58.6)</td>
<td>(11.4)</td>
<td>(14.7)</td>
<td>(2.41)</td>
</tr>
<tr>
<td>8-year</td>
<td>0.999</td>
<td>0.999</td>
<td>20.8 bp</td>
<td>4.0 bp</td>
<td>5.1 bp</td>
<td>0.82%</td>
</tr>
<tr>
<td></td>
<td>(0.983)</td>
<td>(0.726)</td>
<td>(57.2)</td>
<td>(11.3)</td>
<td>(14.6)</td>
<td>(2.35)</td>
</tr>
<tr>
<td>9-year</td>
<td>0.999</td>
<td>0.998</td>
<td>23.2 bp</td>
<td>5.7 bp</td>
<td>7.0 bp</td>
<td>1.03%</td>
</tr>
<tr>
<td></td>
<td>(0.983)</td>
<td>(0.743)</td>
<td>(57.6)</td>
<td>(11.1)</td>
<td>(14.3)</td>
<td>(2.29)</td>
</tr>
<tr>
<td>10-year</td>
<td>0.999</td>
<td>0.999</td>
<td>25.0 bp</td>
<td>7.2 bp</td>
<td>8.7 bp</td>
<td>1.20%</td>
</tr>
<tr>
<td></td>
<td>(0.983)</td>
<td>(0.758)</td>
<td>(58.6)</td>
<td>(10.9)</td>
<td>(14.1)</td>
<td>(2.22)</td>
</tr>
<tr>
<td>15-year</td>
<td>0.984</td>
<td>0.994</td>
<td>48.4 bp</td>
<td>10.0 bp</td>
<td>13.8 bp</td>
<td>2.10%</td>
</tr>
<tr>
<td></td>
<td>(0.984)</td>
<td>(0.823)</td>
<td>(49.5)</td>
<td>(9.8)</td>
<td>(12.7)</td>
<td>(1.93)</td>
</tr>
<tr>
<td>20-year</td>
<td>0.961</td>
<td>0.990</td>
<td>66.5 bp</td>
<td>15.6 bp</td>
<td>20.8 bp</td>
<td>3.11%</td>
</tr>
<tr>
<td></td>
<td>(0.986)</td>
<td>(0.844)</td>
<td>(40.5)</td>
<td>(9.4)</td>
<td>(12.1)</td>
<td>(1.81)</td>
</tr>
</tbody>
</table>

The table shows the accuracy measures for the two-factor model on all yields. The numbers in parentheses are the corresponding benchmark values of the martingale model.
Appendix F: Pricing Accuracy of the Three-Factor Model

The Three-Factor Model's Pricing accuracy separated by Ratings, 1994 - 2002

<table>
<thead>
<tr>
<th>Maturity</th>
<th>Rating</th>
<th>Correlation</th>
<th>Theil U</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-5 years</td>
<td>AAA</td>
<td>0.983</td>
<td>3.16%</td>
</tr>
<tr>
<td></td>
<td>AA</td>
<td>0.981</td>
<td>2.34%</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>0.985</td>
<td>2.03%</td>
</tr>
<tr>
<td></td>
<td>BBB</td>
<td>0.962</td>
<td>3.81%</td>
</tr>
<tr>
<td></td>
<td>AAA</td>
<td>0.977</td>
<td>3.07%</td>
</tr>
<tr>
<td></td>
<td>AA</td>
<td>0.987</td>
<td>2.04%</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>0.988</td>
<td>1.60%</td>
</tr>
<tr>
<td></td>
<td>BBB</td>
<td>0.951</td>
<td>4.72%</td>
</tr>
<tr>
<td>5-7 years</td>
<td>AAA</td>
<td>0.974</td>
<td>2.75%</td>
</tr>
<tr>
<td></td>
<td>AA</td>
<td>0.973</td>
<td>3.97%</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>0.978</td>
<td>4.02%</td>
</tr>
<tr>
<td></td>
<td>BBB</td>
<td>0.958</td>
<td>4.45%</td>
</tr>
<tr>
<td>7-10 years</td>
<td>AAA</td>
<td>0.954</td>
<td>3.70%</td>
</tr>
<tr>
<td></td>
<td>AA</td>
<td>0.979</td>
<td>5.83%</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>0.976</td>
<td>6.42%</td>
</tr>
<tr>
<td></td>
<td>BBB</td>
<td>0.938</td>
<td>5.30%</td>
</tr>
<tr>
<td>10-15 years</td>
<td>AAA</td>
<td>0.859</td>
<td>5.35%</td>
</tr>
<tr>
<td></td>
<td>AA</td>
<td>0.960</td>
<td>9.90%</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>0.894</td>
<td>11.97%</td>
</tr>
<tr>
<td></td>
<td>BBB</td>
<td>0.920</td>
<td>6.04%</td>
</tr>
</tbody>
</table>

The table shows the accuracy measures for the full three-factor model, 1994-2002, divided into maturity segments and ratings.
Appendix G: Alternative Error Structures


<table>
<thead>
<tr>
<th>Maturity</th>
<th>Correlation</th>
<th>Max Dev.</th>
<th>Mean Dev.</th>
<th>RMSE</th>
<th>Theil U</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Orig. model</td>
<td>Alt. I</td>
<td>Alt. II</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-year</td>
<td>0.977</td>
<td>0.984</td>
<td>-</td>
<td>59.3 bp</td>
<td>10.5 bp</td>
</tr>
<tr>
<td>4-year</td>
<td>0.994</td>
<td>0.993</td>
<td>0.983</td>
<td>23.9 bp</td>
<td>33.5 bp</td>
</tr>
<tr>
<td>6-year</td>
<td>0.998</td>
<td>0.996</td>
<td>0.987</td>
<td>21.7 bp</td>
<td>33.7 bp</td>
</tr>
<tr>
<td>10-year</td>
<td>0.999</td>
<td>0.997</td>
<td>-</td>
<td>25.0 bp</td>
<td>110.7 bp</td>
</tr>
<tr>
<td>15-year</td>
<td>0.994</td>
<td>0.989</td>
<td>0.981</td>
<td>48.4 bp</td>
<td>302.2 bp</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Maturity</th>
<th>Correlation</th>
<th>Max Dev.</th>
<th>Mean Dev.</th>
<th>RMSE</th>
<th>Theil U</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-5 year</td>
<td>Orig. model</td>
<td>0.947</td>
<td>36.0 bp</td>
<td>9.8 bp</td>
<td>13.2 bp</td>
</tr>
<tr>
<td></td>
<td>Alt. I</td>
<td>0.944</td>
<td>42.8 bp</td>
<td>10.9 bp</td>
<td>14.9 bp</td>
</tr>
<tr>
<td></td>
<td>Alt. II</td>
<td>0.825</td>
<td>61.7 bp</td>
<td>18.7 bp</td>
<td>22.9 bp</td>
</tr>
<tr>
<td>5-7 year</td>
<td>Orig. model</td>
<td>0.961</td>
<td>34.4 bp</td>
<td>8.1 bp</td>
<td>10.7 bp</td>
</tr>
<tr>
<td></td>
<td>Alt. I</td>
<td>0.955</td>
<td>41.6 bp</td>
<td>10.3 bp</td>
<td>13.5 bp</td>
</tr>
<tr>
<td></td>
<td>Alt. II</td>
<td>0.803</td>
<td>71.5 bp</td>
<td>21.1 bp</td>
<td>23.7 bp</td>
</tr>
<tr>
<td>7-10 year</td>
<td>Orig. model</td>
<td>0.951</td>
<td>67.1 bp</td>
<td>24.6 bp</td>
<td>27.4 bp</td>
</tr>
<tr>
<td></td>
<td>Alt. I</td>
<td>0.906</td>
<td>72.7 bp</td>
<td>30.1 bp</td>
<td>33.7 bp</td>
</tr>
<tr>
<td></td>
<td>Alt. II</td>
<td>0.830</td>
<td>90.6 bp</td>
<td>47.4 bp</td>
<td>54.4 bp</td>
</tr>
<tr>
<td>10-15 year</td>
<td>Orig. model</td>
<td>0.951</td>
<td>83.5 bp</td>
<td>42.8 bp</td>
<td>44.9 bp</td>
</tr>
<tr>
<td></td>
<td>Alt. I</td>
<td>0.867</td>
<td>90.9 bp</td>
<td>53.0 bp</td>
<td>58.0 bp</td>
</tr>
<tr>
<td></td>
<td>Alt. II</td>
<td>0.890</td>
<td>113.1 bp</td>
<td>69.5 bp</td>
<td>72.2 bp</td>
</tr>
<tr>
<td>15+ year</td>
<td>Orig. model</td>
<td>0.894</td>
<td>178.3 bp</td>
<td>84.5 bp</td>
<td>88.4 bp</td>
</tr>
<tr>
<td></td>
<td>Alt. I</td>
<td>0.856</td>
<td>189.2 bp</td>
<td>88.5 bp</td>
<td>91.1 bp</td>
</tr>
<tr>
<td></td>
<td>Alt. II</td>
<td>0.898</td>
<td>221.6 bp</td>
<td>111.2 bp</td>
<td>119.5 bp</td>
</tr>
</tbody>
</table>

The table shows selected accuracy measures for the three alternative Treasury two-factor models.

The table shows selected accuracy measures for the three alternative one-factor hazard rate models on A-rated corporate bonds, 1994-2002.
Appendix H: GAUSS program for the Two-Factor Treasury Model

/* Gauss' - Constrained Maximum Likelihood */

new;
closeall;
library cm1;
#include cm1.ext;
cmlset;
cls;

load data[558,17]=P:\gauss\PHF\treasury.txt;
data=data[147:558,1 2 4 6 10 15];
_n=rows(data);
__title="Estimation of the Two-Factor Treasury Term Structure Model";

start={6.0550,3.3661,-0.0922,0.69170,2.2451,0.4825,-2.1781,1.1213,1.4053,4.2226,0.4842,2.1491,-0.0001};

/* Descent and line Search Specification */
_cml_Algorithm=3; /*BFGS=1 - default=3*/
_cml_LineSearch=3; /*BHHH=5 - default=2*/
_cml_cmlTry=100;
_cml_GradMethod=1;
/*_cml_GradOrder=3;*/ /*antal punkter i numeriske gradienter - default=2*/
_cml_TrustRegion=0.1;
/*Det er en begrænsning på retningen i iterationen således at parameterestimaterne
ikke bliver urealistiske*/
_cml_GridSearch=1;
/*Hvis linesearch fejler vil Gauss prøve i ny retning*/
_cml_Switch=(1 3, 0.001 0.001, 10 10, 0.0001 0.0001);
/*Switcher mellem BFGS og Newton når nedenstående kriterier er opfyldt*/
_cml_DirTol=1E-7;
/*Optimeringen slutter hvis parametrene ikke ændrer sig på 7 dec.*/
_cml_DFTol=1E-13;
_cml_GradCheckTol=1E-3;
_cml_CovPar=1;
_cml_FinalHess;
_row=0;
_cml_IterData;
_cml_Diagnostic=0;
_cml_Active={1,1,1,1,1,1,1,1,1,1,1,1};
_cml_ParNames={"Kappa1","Theta1","Lambda1","SigmaSQ1","Kappa2","Theta2",
"Lambda2","SigmaSQ2",
"Omega1","Omega2","Omega3","Omega4","Constant"};

{k,f,g,cov,retcode}=cml(data,0,&loglh,start);
Call cmlprt(k,f,g,cov,retcode);

_XAll = zeros(_n,2);
_Loglh = zeros(_n,1);
x1=_XAll[,,1];
x2=_XAll[,,2];
/* Maximum Likelihood Procedure */

Proc LogLh(b, data);
Local
K1, T1, Sq1, K2, T2, Sq2, O1, O2, O3, O4, c1, A2, B2, e1, e2, e3, e4, lnFe1, lnFe2, lnFe3, lnFe4, x1, x2, m1, m2, v1, v2, Jac, detijac, G1, G2, tau, delta, i, j, Y, XAll, ASum, Ax, Bx, Cx, Qx, inFx1, inFx2, orderx, alpha, mb, mb1, mb2, mbo1, mbo2, cnt1, cnt2, LogLh;

K1 = b[1]/10; T1 = b[2]/100; Sq1 = (b[4]/10)^2; K2 = b[5]/10; T2 = b[6]/100; L2 = b[7]/10; Sq2 = (b[8]/10)^2; O1 = b[9]/1000; O2 = b[10]/1000; O3 = b[11]/100; O4 = b[12]/100; c1 = b[13];

delta = 1/52;
Tau = 4|6|10|15;

/* Dimensioning the variables */
A1 = zeros(6,1);
A2 = zeros(6,1);
B1 = zeros(6,1);
B2 = zeros(6,1);
Jac = zeros(2,2); /* The Jacobian */
detijac = zeros(1,1); /* The Inverse Jacobian */
Y = zeros(2,1); /* Yields */
ASum = zeros(2,1); /* True Yields */
X = zeros(2,1); /* State Variables */
XAll = zeros(rows(Data), 2); /* All State Variables */
Ax = zeros(2,1); /* Mellemregninger til Bessel-funktionen*/
Cx = zeros(2,1);
Bx = zeros(2,1);
Qx = zeros(2,1);
mb = zeros(2,6);
orderx = zeros(2,1);
alpha = zeros(2,1);

/* General constraints*/
_cml_bounds = zeros(13,2);
_cml_bounds =
{0 1e10, 0 1e10, -1e10 0, 0 1e10, 0 1e10, -1e10 0, 0 1e10, 1e10 0, 1e10 0, -1e10 1e10, 1e10 -1e10};
i = 1;
G1 = (((k1+11)^2)+2*Sq1)^(0.5);
G2 = (((k2+12)^2)+2*Sq2)^(0.5);

do while i<=6;
  A1[i] = ((2*G1*exp(0.5*tau[i]*(k1+11+G1)))/
(2*G1+(k1+11+G1)*exp(G1*tau[i])-1))^(2*K1*T1/Sq1);
  A2[i] = ((2*G2*exp(0.5*tau[i]*(k2+12+G2)))/
(2*G2+(k2+12+G2)+exp(G2*tau[i])-1))^(2*K2*T2/Sq2);
  B1[i] = 2*(exp(G1*tau[i])-1)/(2*G1+(k1+11+G1)*exp(G1*tau[i])-1));
  B2[i] = 2*(exp(G2*tau[i])-1)/(2*G2+(k2+12+G2)*exp(G2*tau[i])-1));
  i = i + 1;
Endo;

/* Jacobimatrikser */
if rank(Jac)==2 and det(Jac)>0;
    detijac=det(inv(Jac));
else;
    detijac=0.001;
endif;

/* De sande afkast defineres */
Y[1]=(data[1,1]/100)*tau[1]+(data[1,2]/100)*tau[2];
Y[2]=(data[1,3]/100)*tau[3]+(data[1,4]/100)*tau[4]+(data[1,5]/100)*tau[5];


/* Udtryk for statevariablene opstilles */
if rank(jac)==2;
    X=inv(Jac)*(Y-ASum);
    if X[1]<0; X[1]=0.001; endif;
    if X[2]<0; X[2]=0.001; endif;
else;
    X[1]=0.001;
    X[2]=0.001;
endif;

XAll[1,1]=X[1];
XAll[1,2]=X[2];

_loglh[1]=loglh;
j=2;

Do while j <= _n; /* antal observationer fra jan 1994 til jan 2002 */
Y[1]=(data[j,1]/100)*tau[1]+(data[j,2]/100)*tau[2];


/* Udtryk for statevariablene opstilles */
if rank(jac)==2;
    X=inv(Jac)*(Y-ASum);
    if X[1]<0; X[1]=0.001; endif;
    if X[2]<0; X[2]=0.001; endif;
else;
    X[1]=0.001;
\[ \text{X}[2]=0.001; \]
\end{verbatim}

\[ \text{XAll}[j,1]=\text{X}[1]; \]
\[ \text{XAll}[j,2]=\text{X}[2]; \]

/* QML - Anvendelse af momenter fremfor modificeret besselfunction*/

\[ m_1=\text{XAll}[j-1,1]\times \exp(-k_1 \delta)+t_1\times(1-\exp(-k_1 \delta)); \]
\[ v_1=\text{XAll}[j-1,1]\times(\text{sq1}/k_1)\times(\exp(-k_1 \delta)-\exp(-2k_1 \delta))+t_1\times(\text{sq1}/2k_1)\times(1-\exp(-k_1 \delta))^2; \]
\[ m_2=\text{XAll}[j-1,2]\times \exp(-k_2 \delta)+t_2\times(1-\exp(-k_2 \delta)); \]
\[ v_2=\text{XAll}[j-1,2]\times(\text{sq2}/k_2)\times(\exp(-k_2 \delta)-\exp(-2k_2 \delta))+t_2\times(\text{sq2}/2k_2)\times(1-\exp(-k_2 \delta))^2; \]

/* ML - Anvendelse af den modificerede besselfunction*/

\[ C_x[1]=2K_1/((\text{sq1})\times(1-\exp(-K_1 \delta)))); \]
\[ C_x[2]=2K_2/((\text{sq2})\times(1-\exp(-K_2 \delta)))); \]
\[ A_x[1]=C_x[1]\times \text{X}[1]; \]
\[ A_x[2]=C_x[2]\times \text{X}[2]; \]
\[ B_x[1]=C_x[1]\times \text{XAll}[j-1,1]\times \exp(-K_1 \delta); \]
\[ B_x[2]=C_x[2]\times \text{XAll}[j-1,2]\times \exp(-K_2 \delta); \]
\[ Q_x[1]=(2K_1T_1/\text{sq1})-1; \]
\[ Q_x[2]=(2K_2T_2/\text{sq2})-1; \]
\[ \text{orderx}[1]=\text{floor}(Q_x[1]); \]
\[ \text{orderx}[2]=\text{floor}(Q_x[2]); \]
\[ \text{Alpha}[1]=Q_x[1]-\text{floor}(Q_x[1]); \]
\[ \text{Alpha}[2]=Q_x[2]-\text{floor}(Q_x[2]); \]

if (\text{orderx}[1]>=1) and ((2*(\text{Ax}[1]*\text{Bx}[1])^0.5)<5);
\[ \text{mbol}=\text{mbesseli}(2*(\text{Ax}[1]*\text{Bx}[1])^0.5,\text{orderx}[1]+1,\text{alpha}[1]); \]
\[ \text{mb1}=\text{mbol}[1,\text{orderx}[1]+1]; \]
else;
\[ \text{if 2*(Ax[1]*Bx[1])^0.5 < 5; \}
\[ \text{mb1}=\text{mbesseli1}(2*(\text{Ax}[1]*\text{Bx}[1])^0.5); \]
\[ \text{else; \}
\[ \text{mb1}=0.001; \]
\[ \text{endif; \}
\[ \text{endif; \}

if (\text{orderx}[2]>=1) and ((2*(\text{Ax}[2]*\text{Bx}[2])^0.5)<2);
\[ \text{mb02}=\text{mbesseli}(2*(\text{Ax}[2]*\text{Bx}[2])^0.5,\text{orderx}[2]+1,\text{alpha}[2]); \]
\[ \text{mb2}=\text{mb02}[1,\text{orderx}[2]+1]; \]
else;
\[ \text{if 2*(Ax[2]*Bx[2])^0.5 < 2; \}
\[ \text{mb2}=\text{mbesseli0}(2*(\text{Ax}[2]*\text{Bx}[2])^0.5); \]
\[ \text{else; \}
\[ \text{mb2}=0.001; \]
\[ \text{endif; \}
\[ \text{endif; \}

if ((\text{Ax}[1]>0) and (\text{Bx}[1]>0) and (\text{Cx}[1]>0) and (\text{mb1}>0) and (\text{mb1}/=0.001));
\[ \text{lnFx1}=\text{ln}(\text{Cx}[1])-\text{Ax}[1]-\text{Bx}[1]+0.5*Q_x[1]\times(\text{ln}(\text{Ax}[1])-\text{ln}(\text{Bx}[1]))+\text{ln}(\text{mb1}); \]
```plaintext
else;
    lnFx1=-0.5*ln(2*pi*v1)-0.5*((X[1]-m1)^2)/v1;
    cnt1=cnt1+1;
endif;

if ((Ax[2]>0) and (Bx[2]>0) and (Cx[2]>0) and (mb2>0) and (mb2/=0.001));
    lnFx2=ln(Cx[2])-Ax[2]-Bx[2]+0.5*Qx[2]*(ln(Ax[2])-ln(Bx[2]))+ln(mb2);
else;
    lnFx2=-0.5*ln(2*pi*v2)-0.5*((X[2]-m2)^2)/v2;
    cnt2=cnt2+1;
endif;
/* Measurement erroers */
lnFe1=-0.5*ln(2*pi*O1^2)-0.5*(e1/O1)^2;
lnFe2=-0.5*ln(2*pi*O2^2)-0.5*(e2/O2)^2;
lnFe3=-0.5*ln(2*pi*O3^2)-0.5*(e3/O3)^2;
lnFe4=-0.5*ln(2*pi*O4^2)-0.5*(e4/O4)^2;

/* Log-Likelihood funktionen*/
LogLh=LogLh+ln(detijac)+lnFx1+lnFx2+lnFe1+lnFe2+lnFe3+lnFe4;

_XAll[j,1]=XAll[j,1];
_XAll[j,2]=XAll[j,2];
_loglh[j]=loglh;

j=j+1;
Endo;

Retp(LogLh);
EndP;

/*Tæthedsfunktioner*/;
library cml, pgraph;
cmlset;
_cml_kernel=1;
__output=1;
{px, py, smth}=cmlDensity(x1,0);
Density=px-py;
```
Appendix I: GAUSS program for the One-Factor Credit-Spread Model

```gauss
new;closeall;
library cml;
#include cml.ext;
cmlset;

/* *** TREASURY PARAMETERS *** */
/* kappa1   theta1    lambda1   sigmasq1
kappa2   theta2    lambda2   sigmasq2  constant */
let par= 0.64055  0.031924  -0.11514  0.006574
    0.27132  0.003842  -0.35744  0.001098  -0.0001;

/* *** CORPORATE PARAMETERS *** */
/* Rating  1 = AAA
2 = AA
3 = A
4 = BBB */
rating=4;

/*Kappa Theta Lambda Sigmasq Constant Beta1 Beta2 Omega1 Omega2 Omega3 */
if rating eq 1;
    load data[558,14]=P:\gauss\PHF\corp3A.txt;
    _rec=0.70;
    _interest = 0.05;
    let start=0.004901 0.026354 -0.220541 0.006259 0.001341 0.023248 0.09468 0.01493 0
    0.180357;
elseif rating eq 2;
    load data[558,14]=P:\gauss\PHF\corp2A.txt;
    _rec=0.65;
    _interest = 0.05;
    let start=0.08672 0.018443 -0.342364  0.005231 0.002341 0.026108 0.189685
    0.001115 0 0.010480;
elseif rating eq 3;
    load data[558,14]=P:\gauss\PHF\corp1A.txt;
    _rec=0.60;
    _interest = 0.06;
    let start=0.08672 0.018443 -0.342364  0.005231 0.002341 0.026108 0.189685
    0.001115 0 0.010480;
elseif rating eq 4;
    load data[558,14]=P:\gauss\PHF\corp3B.txt;
    _rec=0.55;
    _interest = 0.07;
    let start=0.1873  0.00801  -0.2012    0.0113   0.0059   0.01308  -0.0747  0.0115 0
    0.005480;
endif;

_Tau=data[147:558,7 8 9 11];
_Data=data[147:558,1 2 3 5]./100;

load state[412,2]=P:\gauss\PHF\state.txt;
_State=State[1:412,1:2];

load trea[558,17]=P:\gauss\PHF\treasury.txt;
_Trea=trea[147:558,1 2 4 6 10 15]);
```
/* BRUGERINDTASTNING */
_delta=1/52;
_couponFreq = 2;
_obsStart = 1;
_obsEnd = 412; /*rows(_Data);*/

/* VARIABLES */
_k=cols(_Tau);
_n=rows(_Data);
_h=zeros(_n,1);
_g=zeros(_n,1);
_loglh=zeros(_n,1);
_Tyield=zeros(_n,1);
_Cyield=zeros(_n,1);
_spread=zeros(_n,1);

/* *** START COMPUTATIONS *** */

/* Measurement error 1 */
{timep1,nopayments1,cf1,redebt1,kmax1}=cashflows(_tau[.,4],_interest,_couponfreq);
_timep1=timep1; _nopayments1=nopayments1; _cf1=cf1; _kmax1=kmax1;
clear timep1,nopayments1,cf1,redebt1,kmax1;

/* Measurement error 2 */
{timep3,nopayments3,cf3,redebt3,kmax3}=cashflows(_tau[.,3],_interest,_couponfreq);
_timep3=timep3; _nopayments3=nopayments3; _cf3=cf3; _kmax3=kmax3;
clear timep3,nopayments3,cf3,redebt3,kmax3;

/* True yield */
{timep4,nopayments4,cf4,redebt4,kmax4}=cashflows(_tau[.,1],_interest,_couponfreq);
_timep4=timep4; _nopayments4=nopayments4; _cf4=cf4; _kmax4=kmax4;
clear timep4,nopayments4,cf4,redebt4,kmax4;

/* *** ML Iteration *** */
__title="Estimation of the One-Factor Credit Spread Model";
_cml_Algorithm=4;
_cml_LineSearch=2;
_cml_cmlTry=20;
_cml_GradMethod=1;
_cml_GridSearch=1;
_cml_Diagnostic=0;
_cml_CovPar=2;
/*_cml_Switch={1 3, 0.001 0.001, 10 10, 0.0001 0.0001};*/
_cml_DirTol=1E-4;
_cml_TrustRegion=0.1;
_cml_GradCheckTol=1E-5;
_cml_DFTol=1E-10;
_row=10;
_cml_Active={1,1,1,1,1,1,1,1,1,1};
_cml_ParamNames="Kappa","Theta","Lambda","SigmaSQ","Constant","Beta1","Beta2","Omega1"
,"Omega2","Omega3";
_cml_bounds=zeros(10,2);
_cml_bounds={0 1e10,0 1e10,-1e10 0,0 1e10,-1e10 1e10,-1e10 1e10,0 1e10,0 1e10,0 1e10};

{k,f,g,cov,retcode}=cml(_data,0,&loglh,start);
call cmlprt(k,f,g,cov,retcode);
Proc LogLh(b, _data);

Local
k1,t1,l1,sq1,k2,t2,l2,sq2,k3,t3,l3,sq3,o1,o2,o3,c1,c2,A1,A2,A3,B1,B2,B3,B3d,w1,w2,w3,
g1,g2,g3,q1,q2,q3,beta1,beta2,dfcallT,dfcall1,dfcall2,dfcall3,dfobs,dfobs1,dfobs2,dfobs3,
pvcall,pvcall1,pvcall2,pvcall3,pvcallf,pvobs,pvobs1,pvobs2,pvobs3,pvcallc,diffT,diff,
parM,ytm,ytm1,ytm3,low,high,D,e1,e2,e3,lne1,lne2,lne3,iter1,m1,v1,Cg,Ag,Bg,Qg,orderbyg,A
lpha,mb0,mb1,lnFg,i,j,p,l,z,LogLh;

dfobs1=zeros(_k1max+1,1);
dfcall1=zeros(_k1max+1,1);
pvobs1=zeros(_n,1);
pvcall1=zeros(_n,1);
e1=zeros(_n,1);
lne1=zeros(_n,1);

dfobs3=zeros(_k3max+1,1);
dfcall3=zeros(_k3max+1,1);
pvobs3=zeros(_n,1);
pvcall3=zeros(_n,1);
e3=zeros(_n,1);
lne3=zeros(_n,1);

dfcallT=zeros(_k4max+1,1);
dfobs=zeros(_k4max+1,1);
dfcallC=zeros(_k4max+1,1);
pvcall= zeros(_n,1);
pvcallT=zeros(_n,1);
pvobs= zeros(_n,1);
diffT=zeros(_n,1);
diff=zeros(_n,1);

lnFg=zeros(_n,1);
lm=zeros(_n,1);
vl=zeros(_n,1);
Ag=zeros(_n,1);
Cg=zeros(_n,1);
Bg=zeros(_n,1);
Qg=zeros(_n,1);
mb0=zeros(_n,3);
mb1=zeros(_n,3);
orderbyg=zeros(_n,1);
alpha=zeros(_n,1);

z=_state[.,1].*(1+b[6]*(1-_rec))-_state[.,2].*(1+b[7]*(1-_rec));
k3=b[1];t3=(1-_rec)*b[2];l3=b[3];sq3=(1-_rec)*b[4];c2=b[5];beta1=b[6];
beta2=b[7];o1=b[8];o2=b[9];o3=b[10];

parM=par;
parM[2]=1+beta1*(1-_rec))*par[2];
parM[4]=1+beta1*(1-_rec))*par[4];
parM[6]=1+beta2*(1-_rec))*par[6];
parM[8]=1+beta2*(1-_rec))*par[8];
c1=par[9];
g3=((k3+l3)^2)+2*sq3)^0.5;
q3=k3+l3-g3;
D=zeros(_n,1);
LogLh=zeros(_n,1);
/* *** FØRSTE OBSERVATION *** */

i=1;

_g[i]=0.02;
high=1; low=0; iter1=1;
do while iter1 le 30;

do while j le _NoPayments4[i]+1;
    \{A1, A2, B1, B2\}={CIR2_AB(parM, _timep4[i, j])};
    w3=1-exp(-g3*_timep4[i, j]);
    A3=(2*k3*t3/sq3)*ln(2*g3*exp(q3*0.5*_timep4[i, j])/(2*g3+q3*w3));
    B3=-2*w3/(2*g3+q3*w3);
    dfcalT[j]=exp(A1+A2+B1*z[i,1]+B2*z[i,2]+_timep4[i,j]*c1);
    dfcalC[j]=exp(A1+A2+A3+B1*z[i,1]+B2*z[i,2]+B3*_g[i]+_timep4[i,j]*(c1+c2));  j=j+1;
endo;

PVCal[i]=_cf4[i, .]*dfcalC;
{ytm}=yldtomat(PvCal[i], _timep4[i, .], _cf4[i, .]);
diff[i]=(_data[i, 1]-ytm);
if diff[i] gt 0;
    high=_g[i];
    _g[i]=(_g[i]+low)/2;
else;
    low=_g[i];
    _g[i]=(_g[i]+high)/2;
endif;
iter1=iter1+1;
endo;

if _g[i] le 0.0001;
    _g[i]=0.0001;
endif;
pvcalT[i]=_cf4[i, .]*dfcalT;
{ytm}=yldtomat(PvCal[i], _timep4[i, .], _cf4[i, .]);
_CYield[i]=ytm;
_spread[i]=(_data[i, 1]-ytm);

/* *** GENNEMKØRER ØVRIGE OBSERVATIONER *** */

i=2;
Do while i le _ObsEnd;

_g[i]=0.02;
high=1; low=0; iter1=1;
do while iter1 le 30;

do while j le _NoPayments4[i]+1;
    \{A1, A2, B1, B2\}={CIR2_AB(parM, _timep4[i, j])};
    w3=1-exp(-g3*_timep4[i, j]);
    A3=(2*k3*t3/sq3)*ln(2*g3*exp(q3*0.5*_timep4[i, j])/(2*g3+q3*w3));
    B3=-2*w3/(2*g3+q3*w3);
    dfcalT[j]=exp(A1+A2+B1*z[i,1]+B2*z[i,2]+_timep4[i,j]*c1);
\[
dfcalC[j] = \exp(A1 + A2 + A3 + B1z[i,1] + B2z[i,2] + B3g[i] + _timep4[i,j] * (c1 + c2));
\]
\[
j = j + 1;
\]
\[
end;
\]
\[
PVCal[i] = _cf4[i,.] * dfcalC;
\]
\[
\{ytM3\} = yldtomat(PVCal[i], _timep4[i, .], _cf4[i, .]);
\]
\[
diff[i] = _tau[i,1] * (_data[i,1] - ytM3);
\]
\[
if diff[i] gt 0;
\]
\[
high = _g[i];
\]
\[
_g[i] = (_g[i] + low) / 2;
\]
\[
else;
\]
\[
low = _g[i];
\]
\[
_g[i] = (_g[i] + high) / 2;
\]
\[
endif;
\]
\[
iter1 = iter1 + 1;
\]
\[
end;
\]
\[
if _g[i] le 0.00001;
\]
\[
_g[i] = 0.00001;
\]
\[
endif;
\]
\[
pvcalT[i] = _cf4[i, .] * dfcalT;
\]
\[
\{ytM3\} = yldtomat(PVCal[i], _timep4[i, .], _cf4[i, .]);
\]
\[
_CYield[i] = ytM3;
\]
\[
_spread[i] = (_data[i,1] - ytM3);
\]
\[
/* *** THE JACOBIAN *** */
\]
\[
B3d = zeros(_k4max+1,1);
\]
\[
j = 1;
\]
\[
do while j le _NoPayments4[i] + 1;
\]
\[
w3 = 1 - exp(-g3 * _timep4[i, j]);
\]
\[
B3d[j] = -2 * w3 / (2 * g3 + q3 * w3);
\]
\[
j = j + 1;
\]
\[
end;
\]
\[
D[i] = -(1/PVCal[i]) * _cf4[i, .] * (B3d * dfcalC);
\]
\[
/* *** MEASUREMENT ERRORS *** */
\]
\[
j = 1;
\]
\[
do while j le _nopayments1[i] + 1;
\]
\[
(A1, A2, B1, B2) = CIR2_AB(ParM, _timep1[i, j]);
\]
\[
w3 = 1 - exp(-g3 * _timep1[i, j]);
\]
\[
A3 = (2 * k3 * t3 / sq3) * ln(2 * g3 * exp(q3 * 0.5 * _timep1[i, j]) / (2 * g3 + q3 * w3));
\]
\[
B3 = -2 * w3 / (2 * g3 + q3 * w3);
\]
\[
dfcal1[j] = \exp(A1 + A2 + A3 + B1z[i,1] + B2z[i,2] + B3g[i] + _timep1[i, j] * (c1 + c2)) * _cf1[i, j];
\]
\[
j = j + 1;
\]
\[
end;
\]
\[
j = 1;
\]
\[
do while j le _nopayments3[i] + 1;
\]
\[
(A1, A2, B1, B2) = CIR2_AB(ParM, _timep3[i, j]);
\]
\[
w3 = 1 - exp(-g3 * _timep3[i, j]);
\]
\[
A3 = (2 * k3 * t3 / sq3) * ln(2 * g3 * exp(q3 * 0.5 * _timep3[i, j]) / (2 * g3 + q3 * w3));
\]
\[
B3 = -2 * w3 / (2 * g3 + q3 * w3);
\]
\[
dfcal3[j] = \exp(A1 + A2 + A3 + B1z[i,1] + B2z[i,2] + B3g[i] + _timep3[i, j] * (c1 + c2)) * _cf3[i, j];
\]
\[ j = j + 1; \]
\[ \text{end}; \]
\[ \{ytm\} = \text{yldtomat(sumc(dfca1l),_timep1[i,..],_cf1[i,..]);} \]
\[ ytm1 = ytm; \]
\[ \{ytm\} = \text{yldtomat(sumc(dfca13),_timep3[i,..],_cf3[i,..]);} \]
\[ ytm3 = ytm; \]
\[ e1[i] = (_data[i,4] - ytm1); \]
\[ e3[i] = (_data[i,3] - ytm3); \]
\[ \text{if } o1 > 0; \]
\[ lne1[i] = -0.5 * \ln(2 * \pi * O1^2) - 0.5 * (e1[i] / O1)^2; \]
\[ \text{else; } \]
\[ lne1[i] = 0; \]
\[ \text{endif; } \]
\[ \text{if } o3 > 0; \]
\[ lne3[i] = -0.5 * \ln(2 * \pi * O3^2) - 0.5 * (e3[i] / O3)^2; \]
\[ \text{else; } \]
\[ lne3[i] = 0; \]
\[ \text{endif; } \]
\[ /* *** QML APPROXIMATION *** */ \]
\[ m1[i] = _g[i-1,1] * \exp(-k3*_delta) + t3*(1-\exp(-k3*_delta)); \]
\[ v1[i] = _g[i-1,1]*(sq3/k3)*(\exp(-k3*_delta)-\exp(-2*k3*_delta)) \]
\[ + t3*(sq3/2*k3)*(1-\exp(-k3*_delta))^2; \]
\[ /* *** ML *** */ \]
\[ Cg[i] = 2*K3/((Sq3)*(1-\exp(-K3*_delta))); \]
\[ Ag[i] = Cg[i]*_g[i]; \]
\[ Bg[i] = Cg[i]*_g[i-1]*\exp(-K3*_delta); \]
\[ Qg[i] = (2*K3*T3/Sq3)-1; \]
\[ orderg[i] = \text{floor}(Qg[i]); \]
\[ Alpha[i] = Qg[i] - \text{floor}(Qg[i]); \]
\[ \text{if } (\text{orderg}[i] > 1) \text{ and } (2*(Ag[i]*Bg[i])^0.5 < 5); \]
\[ mb0 = \text{mbessel1}(2*(Ag[i]*Bg[i])^0.5,orderg[i]+1,alpha[i]); \]
\[ mb1 = \text{mb0[1,orderg[i]+1]}; \]
\[ \text{else; } \]
\[ \text{if } 2*(Ag[i]*Bg[i])^0.5 < 5; \]
\[ mb1 = \text{mbessel1}(2*(Ag[i]*Bg[i])^0.5); \]
\[ \text{else; } \]
\[ mb1 = 0.001; \]
\[ \text{endif; } \]
\[ \text{endif; } \]
\[ \text{if } (_g[i-1] \text{ gt } 0.001) \text{ and } (\text{Ag}[i]>0) \text{ and } (\text{Bg}[i]>0) \text{ and } (\text{Cg}[i]>0) \text{ and } (\text{mb1}>0) \text{ and } \text{mb1/=0.001}); \]
\[ \text{lnFg}[i] = \ln(Cg[i]) - Ag[i] - Bg[i] + 0.5 * Qg[i] * \ln(Ag[i]) - \ln(Bg[i])) + \ln(mb1); \]
\[ \text{else; } \]
\[ \text{if } _g[i-1] \text{ gt } 0.001; \]
\[ \text{lnFg}[i] = -0.5 * \ln(2 * \pi * v1[i]) - 0.5 * ((_g[i] - m1[i])^2) / v1[i]; \]
\[ \text{else; } \]
\[ \text{lnFg}[i] = 0; \]
\[ \text{endif; } \]
\[ \text{endif; } \]
\[ \text{Loglh}[i] = \ln(D[i]) + \lnFg[i] + lne1[i] + lne3[i]; \]
i=i+1;
endp;

Retp(LogLh);
EndP;

{vnam,mean,var,std,min,max,valid,mis}=dstat(0,_g);

/*Tæthedsfunktioner;*/
library cml, pgraph;
cmlset;
state=_state[.,1];
_cml_kernel=1;
__output=1;
var="State2";
{px, py, smth}=cmlDensity(_g,var);
Density=px~py;

/ * ********* Procedures ********* */
proc(5)=cashflows(maturity,interest,freq);
Local NoPayments,n_,i_,kmax,p_,rest,repay;

n_=rows(maturity);
kmax=Floor(maxc(maturity)*freq);
CF=zeros(n_,kmax+1);
TimeP=zeros(n_,kmax+1);
NoPayments=zeros(n_,1);
ReDebt=(100*ones(n_,1))-zeros(n_,kmax+1);
NoPayments=Floor(maturity.*freq);

i_=1;
do while i_<=n_;  
RePay=100/(NoPayments[i_]+1);
p_=0;
do while p_<=NoPayments[i_];
  timep[i_,p_+1]=maturity[i_]-[(NoPayments[i_] - p_)*1/freq];
  CF[i_,p_+1]=RePay+ReDebt[i_,p_+1].*(interest/freq);
  ReDebt[i_,p_+2]=ReDebt[i_,p_+1]-RePay;
p_+=p_+1;
endo;
i_+=i_+1;
endo;
clear n_ i_ p_;
retp(timep,nopayments,cf,redebt,kmax);
endp;

proc(1)=yldtomat(price,tmat,cashflow);
local n,ytm,iter,dur,moddur,presval,p1;
/* Compute the Yield to Maturity - given price, cashflows and time of payments */
n=rows(price);
if rows(tmat)==1;
tmat=tmat.*ones(n,cols(tmat));
endif;
if (n ne rows(tmat)) or (rows(tmat) ne rows(cashflow));
  print "dimension error";
  stop;
else;
    if (cols(price) ne 1) or (cols(tmat) ne cols(cashflow));
        print "dimension error";
        stop;
    endif;
endif;
/* calculate yields to maturities */
ytm=(0.05.*ones(n,1));
iter=1;
do while iter le 20;
presval=sumc((exp(-ytm.*tmat).*cashflow'));
p1=sumc((-tmat.*cashflow.*exp(-ytm.*tmat))');
ytm=ytm+(price-presval)./p1;
iter=iter+1;
endo;
retp(ytm);
endp;

proc(4)=CIR2_AB(parM,tau);
local A1,A2,B1,B2,k1,t1,l1,sq1,k2,t2,l2,sq2,g1,g2,q1,q2,w1,w2;
/* Computes the functions A and B in the two-factor CIR model */
parM: The modified parameter values for the treasury term structure
    { kappa1 theta1 lambda1 sigma1   kappa2 theta2 lambda2 sigma2 }
tau: time to maturity (in years)
*/
k1=parM[1];t1=parM[2];l1=parM[3];sq1=parM[4];k2=parM[5];t2=parM[6];l2=parM[7];sq2=parM[8];
g1=((k1+l1)^2)+2*sq1)^(0.5);
g2=((k2+l2)^2)+2*sq2)^(0.5);
q1=k1+l1-g1;
q2=k2+l2-g2;
w1=1-exp(-g1*tau);
w2=1-exp(-g2*tau);
A1=(2*k1*t1/sq1)*ln(2*g1*exp(q1*0.5*tau)
    /(2*g1+q1*w1));
A2=(2*k2*t2/sq2)*ln(2*g2*exp(q2*0.5*tau)
    /(2*g2+q2*w2));
B1=-2*w1/(2*g1+q1*w1));
B2=-2*w2/(2*g2+q2*w2);
retp(A1,A2,B1,B2);
endp;
Appendix J: Deriving the Ricatti Equations

To derive the general ordinary differential equations we adopt the notation of Dai & Singleton (1997). The derivation draws on Cochrane (2001). For ease of exposition we drop the time notation in the majority of the formulae. The model setup in the affine case is therefore

\[
\begin{align*}
\dot{x} &= K(\theta - x)dt + \Sigma \sqrt{S} dw \\
r &= \delta_n + \delta' x \\
S_i &= \alpha_i + \beta_i' x \\
\frac{d\pi}{\pi} &= -\Lambda dt - \Lambda' dw, \quad \Lambda = \Sigma^{-1} \sqrt{S} \lambda
\end{align*}
\]

(J.1) (J.2) (J.3) (J.4)

where \( S \) is a diagonal matrix and (J.4) is the equation that the stochastic discount factor satisfies. The existence of this discount factor or state price deflator, \( \pi(t) \), ensures that \( \pi(t)x(t) \) is a martingale under the probability measure \( \mathbb{P} \); i.e. for any time \( s > t \),

\[
E_x \left[ \pi(s) x(s) \right] = E_x \left[ \pi(t) x(t) \right].
\]

(J.5)

By Girsanov’s Theorem this price deflator is equivalent to the accumulation of the money market account i.e. the time discount factor under the equivalent martingale measure. Employing Ito’s lemma to (J.5), recalling that \( \sigma(t)\Lambda(t) = u(t) - r(t)x(t) \), gives the stated result in (J.4).

From Cochrane (2001) and section 2.1.2 we know that the basic bond pricing equation can be written as

\[
E_x \left( \frac{dp}{p} \right) = -E_x \left( \frac{dp}{p} \frac{d\pi}{\pi} \right)
\]

(J.6)

where the left-hand side is the expected excess return. If we specified bonds by their maturity date this equation could be applied directly. However if we are looking for a bond pricing equation that fixes maturity rather than the date (J.6) must be revised. In this case we need the instantaneous holding period return of a bond with fixed maturity, where the effect of the fact, that you sell younger bonds than you buy is incorporated\footnote{The general derivation of the holding period return can be found in Cochrane (2001) equation (19.2) p. 349.}

\[
hpr = \frac{dp}{p} - \frac{1}{p} \frac{\partial p}{\partial \tau} dt.
\]

(J.7)

Thus, the fundamental pricing equation, applied to the price of bonds of given maturity is
\[ E_i \left( \frac{dp}{p} \right) - \left( \frac{1}{p} \frac{\partial p}{\partial \tau} + r \right) dt = -E_i \left( \frac{dp}{p} \frac{d\pi}{\pi} \right). \] (J.8)

The further derivation relies on finding the expression of \( \frac{dp}{p} \). Using Ito’s lemma does this

\[ \frac{dp}{p} = \frac{1}{p} \frac{\partial p}{\partial x} dx + \frac{1}{2} \frac{1}{p} \frac{\partial^2 p}{\partial x^2} dx \cdot \] (J.9)

Recall the general partial derivatives

\[ \frac{1}{P(x, \tau)} \frac{\partial P(x, \tau)}{\partial \tau} = -\frac{1}{P(x, \tau)} \frac{\partial P(x, \tau)}{\partial t} = \left\{ -\frac{\partial A(\tau)}{\partial \tau} - \frac{\partial B(\tau)}{\partial \tau} \right\} x \] (J.10a)

\[ \frac{1}{P(x, \tau)} \frac{\partial P(x, \tau)}{\partial x} = B(\tau) \] (J.10b)

\[ \frac{1}{P(x, \tau)} \frac{\partial^2 P(x, \tau)}{\partial x \partial x'} = B(\tau)B' \] (J.10c)

the first term in (J.8) is

\[ E_i \left( \frac{dp}{p} \right) = B(\tau) E_i \left[ K (\theta - x) dt + \Sigma \sqrt{S} \; dw \right] + \]

\[ \frac{1}{2} E_i \left[ K (\theta - x) dt + \Sigma \sqrt{S} \; dw \right] B(\tau) B'(\tau) \left[ K (\theta - x) dt + \Sigma \sqrt{S} \; dw \right] \] (J.11)

\[ = B(\tau) K (\theta - x) dt + \frac{1}{2} E_i \left[ \Sigma \sqrt{S} \; dw \right] B(\tau) B'(\tau) \left[ \Sigma \sqrt{S} \; dw \right] \]

where the simplification of the last term relies on the fact that \( E_i \left( dtdt \right) = E_i \left( dwdw \right) = 0 \).

If we use the notation \([i] \), to denote the \( i \)th element of the \( n \) dimensional vector of state variables, and recalling that \( E_i \left( dw_i dw_j \right) = dt \) and \( E_i \left( dw_i dw_j \right) = 0 \), the last term can be rewritten as

\[ \frac{1}{2} E_i \left[ \Sigma \sqrt{S} \; dw \right] B(\tau) B'(\tau) \left[ \Sigma \sqrt{S} \; dw \right] = \frac{1}{2} \sum_i \left[ \Sigma' B(\tau) \right] \left[ \alpha_i + \beta_i \right] dt. \] (J.12)

Substituting (J.12) into (J.11) yields:

\[ E_i \left( \frac{dp}{p} \right) = B(\tau) K (\theta - x) dt + \frac{1}{2} \sum_i \left[ \Sigma' B(\tau) \right] \left[ \alpha_i + \beta_i \right] dt. \] (J.13)
The right-hand side of (J.8) is

\[- E_i \left( \frac{dp}{p} \frac{d\tau}{\tau} \right) = E_i \left[ \left( B(\tau)^{'} \right) \left[ K(\theta - x)dt + \Sigma \sqrt{S}dw \right] + \right. \]

\[\left. \frac{1}{2} \left[ K(\theta - x)dt + \Sigma \sqrt{S}dw \right] B(\tau)B(\tau)^{'} \left[ K(\theta - x)dt + \Sigma \sqrt{S}dw \right] \right] (rdt + \Lambda dw) \]

\[= B(\tau)^{'} \Sigma \sqrt{S}dw \left( \Sigma^{-1} \sqrt{S} \lambda \right)^{'} dw = \Sigma \sqrt{S} dw d\lambda \Sigma^{-1} \sqrt{S} \lambda \]

\[= \sum_i B(\tau), \lambda_i \left( \alpha_i + \beta_i^{'} x \right) dt \]

where

\[E_i \left( B(\tau)^{'} \left[ K(\theta - x)dt + \Sigma \sqrt{S}dw \right] rdt + \Lambda dw \right)^{'} = B(\tau)^{'} \Sigma \sqrt{S} dw d\lambda \Sigma^{-1} \sqrt{S} \lambda \]

and

\[E_i \left[ K(\theta - x)dt + \Sigma \sqrt{S}dw \right] B(\tau)B(\tau)^{'} \left[ K(\theta - x)dt + \Sigma \sqrt{S}dw \right] rdt + \Lambda dw \right) = 0 \]

Substituting (J.13) and (J.14) into (J.8) baring expression (J.10a) and (J.2) in mind, and eliminating \(dt\), we get

\[E_i \left( \frac{dp}{p} \right) = - \frac{1}{2} \frac{\partial \rho}{\partial \tau} + r \right) dt = - E_i \left( \frac{dp}{p} \right) \]

\[\Leftrightarrow \]

\[B(\tau)^{'} K(\theta - x) + \frac{1}{2} \sum_i \left[ \Sigma B(\tau)^{'} \right] \left( \alpha_i + \beta_i^{'} x \right) - \left( \frac{\partial A(\tau)}{\partial \tau} + \frac{\partial B(\tau)^{'}}{\partial \tau} x + \delta_o + \delta^{'} x \right) \]

\[= \sum_i B(\tau), \lambda_i \left( \alpha_i + \beta_i^{'} x \right). \]

Rearranging the above in terms of the constant and the state variables, recalling that the expressions must be identically zero we get

\[\frac{\partial A(\tau)}{\partial \tau} = B(\tau)^{'} K \theta + \frac{1}{2} \sum_i \left[ \Sigma B(\tau)^{'} \right] \alpha_i - \sum_i B(\tau), \lambda_i \alpha_i - \delta_o \]

\[= B(\tau)^{'} K \theta + \sum_i \left( \frac{1}{2} \left[ \Sigma B(\tau)^{'} \right] - B(\tau), \lambda_i \right) \alpha_i - \delta_o \]

\[\frac{\partial B(\tau)^{'}}{\partial \tau} = -K' B(\tau) + \frac{1}{2} \sum_i \left[ \Sigma B(\tau)^{'} \right] \beta_i - \sum_i B(\tau), \lambda_i \beta_i - \delta \]
Appendix K: Deriving the Affine Solutions in the CIR Model

In the following we use the affine expression \( P(x, \tau) = \exp(A(\tau) + B(\tau)'x) \). Recall, that if the state variable is governed by a one-factor CIR-process, the ordinary differential equations are

\[
\begin{align*}
\frac{\partial A(\tau)}{\partial \tau} &= \kappa \theta B(\tau) \\
\frac{\partial B(\tau)}{\partial \tau} &= -B(\tau)(\kappa + \lambda) + \frac{1}{2} \sigma^2 B(\tau)^2 - 1.
\end{align*}
\]

(K.1a) 

(K.1b)

Recognizing that (K.1b) is a separable differential equation it can be rewritten as

\[
\int \frac{dB}{\frac{1}{2} \sigma^2 B^2 - (\kappa + \lambda)B - 1} = \int d\tau
\]

(K.2)

where the right-hand side equals \( \tau + c \), where \( c \) is a constant. This term cannot explicitly be defined before the left-hand side of equation (K.2) is solved. To do so we manipulate and rewrite the left-hand side as

\[
\int \frac{dB}{\frac{1}{2} \sigma^2 B^2 - (\kappa + \lambda)B - 1} = \frac{2}{\sigma^2} \int \frac{dB}{(2/\sigma^2)(\frac{1}{2} \sigma^2 B^2 - (\kappa + \lambda)B - 1)}
\]

(K.3)

The trick is now to define \(-2/\sigma^2\) as \([\kappa + \lambda]^2 - (\kappa + \lambda)^2 - 2\sigma^2\) and for convenience set \( \gamma \) equal \( \sqrt{(\kappa + \lambda)^2 + 2\sigma^2} \). Thus, equation (K.3) can now be further manipulated to give

\[
\frac{2}{\sigma^2} \int \frac{dB}{B^2 - (2(\kappa + \lambda)B)/\sigma^2 + ((\kappa + \lambda)^2 - \gamma^2)/\sigma^2}.
\]

(K.4)

Bearing in mind, that \((\kappa + \lambda)^2 - \gamma^2\) can be written as \((\kappa + \lambda - \gamma)(\kappa + \lambda + \gamma)\), it is relatively easy to express (K.4) as

\[
\frac{2}{\sigma^2} \int \frac{dB}{(B - (\kappa + \lambda - \gamma)/\sigma^2)(B - (\kappa + \lambda + \gamma)/\sigma^2)}.
\]

(K.5)

which can be separated into two integrals simplifying the integration process \(^2\)

\(^2\) Collecting the two fractions on the same fraction line and reducing can verify this.
\[
\int \frac{-1/\gamma}{B-(\kappa-\lambda-\gamma)/\sigma^2}dB + \int \frac{1/\gamma}{B-(\kappa+\lambda+\gamma)/\sigma^2}dB
\]
\[
= \frac{-1}{\gamma} \log \left( \frac{B-(\kappa+\lambda-\gamma)/\sigma^2}{B-(\kappa+\lambda+\gamma)/\sigma^2} \right) + \frac{1}{\gamma} \log \left( \frac{B-(\kappa+\lambda+\gamma)/\sigma^2}{B-(\kappa+\lambda-\gamma)/\sigma^2} \right)
\]
\[
= \frac{-1}{\gamma} \log \left( \frac{B-(\kappa+\lambda-\gamma)/\sigma^2}{B-(\kappa+\lambda+\gamma)/\sigma^2} \right)
\]

(K.6)

From the separable differential equation it follows, that (K.6) must equal \( \tau + c \), so

\[
\frac{-1}{\gamma} \log \left( \frac{B-(\kappa+\lambda-\gamma)/\sigma^2}{B-(\kappa+\lambda+\gamma)/\sigma^2} \right) = \tau + c \iff \frac{B-(\kappa+\lambda-\gamma)/\sigma^2}{B-(\kappa+\lambda+\gamma)/\sigma^2} = ke^{-\gamma \tau}
\]

(K.7)

where \( k = e^{-\gamma \tau} \). It is now straightforward to rearrange in terms of \( B \), and after some manipulation we get

\[
B = \frac{\kappa+\lambda-\gamma}{\sigma^2} - \frac{\kappa+\lambda+\gamma}{\sigma^2} ke^{-\gamma \tau}
\]

\[
1 - ke^{-\gamma \tau}
\]

(K.8)

The expression for the constant \( k \) is now derived from the boundary condition \( B(0) = 0 \), which simplifies to equalling the numerator in (K.8) to zero

\[
\frac{\kappa+\lambda-\gamma}{\sigma^2} - \frac{\kappa+\lambda+\gamma}{\sigma^2} k = 0 \iff \frac{\kappa+\lambda-\gamma}{\sigma^2} = \frac{\kappa+\lambda+\gamma}{\sigma^2} k \iff k = \frac{\kappa+\lambda-\gamma}{\kappa+\lambda+\gamma}
\]

(K.9)

By substituting (K.9) into (K.8), bearing the identity \(-2\sigma^2 = (\kappa + \lambda)^2 - \gamma^2\) in mind, gives us the final expression for \( B(\tau) \)
This is the first part of equation 23 in Cox, Ingersoll & Ross (1985). There is an obvious difference between (K.10) and equation 23 in CIR, which can be traced back to the affine form of the bond price. In this dissertation we follow $P(x, \tau) = \exp(A(\tau) + B(\tau)^2x)$ where CIR in contrast apply a minus in front of $B$ in the bond price equation.

In the derivation of $A(\tau)$ it is convenient to define the denominator of (K.10) as $h(\tau)$ and rewriting the equation as:

\[
B = \frac{\kappa + \lambda - \gamma}{\sigma^2} - \frac{\kappa + \lambda - \gamma}{\sigma^2} e^{-\tau} = \frac{\kappa + \lambda - \gamma}{\sigma^2} \left(1 - e^{-\tau}\right)
\]

\[
= \frac{(\kappa + \lambda + \gamma)(\kappa + \lambda - \gamma)}{\sigma^2} \left(1 - e^{-\tau}\right) = \frac{(\kappa + \lambda)^2 - \gamma^2}{\sigma^2} \left(1 - e^{-\tau}\right)
\]

\[
= \frac{2\gamma + (\kappa + \lambda - \gamma)e^{-\tau}}{2\gamma + (\kappa + \lambda + \gamma)(e^{-\tau} - 1)} .
\]

\text{(K.10)}

Now, the evaluation of the integral defining the solution for $A(\tau)$ is straightforward

\[
A(\tau) = \kappa \theta \int B(\tau) d\tau = \kappa \theta \left[\frac{\kappa + \lambda - \gamma}{\sigma^2} - \frac{2}{\sigma^2} \log(h(\tau)) + c\right] .
\]

\text{(K.12)}

Employing the boundary condition $A(0) = 0$, where $h(0) = 2\gamma$, the constant $c$ must equal $2\log(2\gamma)/\sigma^2$, and the final expression for $A(\tau)$ is
\[ A(\tau) = \frac{\kappa \theta}{\sigma^2} \left[ (\kappa + \lambda - \gamma) \tau - 2 \log \frac{h(\tau)}{2\gamma} \right] = \frac{2\kappa \theta}{\sigma^2} \left[ \log \left( e^{(\kappa + \lambda - \gamma)\tau/2} \right) - \log \frac{h(\tau)}{2\gamma} \right] \]

\[ = \frac{2\kappa \theta}{\sigma^2} \log \left[ \frac{2e^{\gamma (\kappa + \lambda + \gamma) \tau}}{h(\tau)} \right] = \frac{2\kappa \theta}{\sigma^2} \log \left[ \frac{2e^{\gamma (\kappa + \lambda + \gamma) \tau}}{2\gamma + (\kappa + \lambda + \gamma)(e^{\tau \gamma} - 1)} \right] \quad (K.13) \]

This is the last part of equation 23 in Cox, Ingersoll & Ross (1985b) and the solution is exactly the same. The reason why is once again traced back to the affine form of the bond price. If we follow CIR in the definition of the form of the affine bond price, we would get a minus in front of the ODE in (K.1a), and (K.1b) is simultaneously subject to changes. The new solution to \( B(\tau) \) would be the negative of the one stated in (K.10). When substituting this result into (K.12), utilizing the new differential equation for \( A(\tau) \), the two minuses cancel out, yielding the same solution to \( A(\tau) \) as stated in (K.13).

Thus, the full expression for the price of a pure discount bond, depending solely on the two independent state variables, governed by CIR-processes is

\[ P(x, \tau) = \exp \left( \sum_{i=1}^{2} A_i(\tau) + \sum_{i=1}^{2} B_i(\tau) X_i \right) = \prod_{i=1}^{2} \exp[A_i(\tau)] \exp \left[ \sum_{i=1}^{2} B_i(\tau) X_i \right] \]

\[ = \left[ \frac{2\gamma e^{\gamma (\kappa_1 + \lambda_1 + \gamma_1) \tau}}{2\gamma_1 + (\kappa_1 + \lambda_1 + \gamma_1)(e^{\tau \gamma_1} - 1)} \right]^{\frac{2\kappa \theta}{\sigma^2}} \left[ \frac{2\gamma_2 e^{\gamma (\kappa_2 + \lambda_2 + \gamma_2) \tau}}{2\gamma_2 + (\kappa_2 + \lambda_2 + \gamma_2)(e^{\tau \gamma_2} - 1)} \right]^{\frac{2\kappa \theta}{\sigma^2}} \exp \left[ \frac{-2(e^{\tau \gamma_1} - 1)}{2\gamma_1 + (\kappa_1 + \lambda_1 + \gamma_1)(e^{\tau \gamma_1} - 1)} X_1 + \frac{-2(e^{\tau \gamma_2} - 1)}{2\gamma_2 + (\kappa_2 + \lambda_2 + \gamma_2)(e^{\tau \gamma_2} - 1)} X_2 \right] \quad (K.14) \]