The Multi-year Through-the-cycle and Point-in-time Probability of Default

Jessica Hallblad
Abstract

This thesis examines how the through-the-cycle probability of default (TTC PD) and point-in-time probability of default (PIT PD) relate to each other in the multi-year horizon. In a first step to analyze this issue, the Nelson-Siegel function is used to estimate the term structure of TTC PD based on historical average default rates reported by Moody’s. The one-factor Merton model is applied to transform the term structure of TTC PDs into a term structure of PIT PDs by incorporating the economic cycle effect. The economic cycle is estimated by applying the Hodrick-Prescott filter to the X-12-ARIMA trend-cycle component. This latent component is initially extracted from the real GDP in Sweden. The state of the economy in a particular time period is predicted by an autoregressive moving average model. It is observed that the marginal PIT PD is larger than the marginal TTC PD in periods of recession, while it is smaller than the marginal TTC PD in expansions. Furthermore, the empirical results indicate that the cumulative TTC PD tends to be higher than the cumulative PIT PD. It is also observed that the distinction between the TTC and PIT PD decreases with the time horizon.

Key words: Credit risk, through-the-cycle probability of default, point-in-time probability of default, one-factor Merton model, Nelson-Siegel function, X-12-ARIMA program, Hodrick-Prescott filter, ARIMA model
# Contents

1 Introduction .................................................. 1

2 Literature review ............................................. 3

3 Probability of default ........................................ 5
   3.1 TTC and PIT PD ......................................... 5
   3.2 The term structure of PD ................................. 7

4 Economic state factor ....................................... 9
   4.1 Extracting the economic state factor ..................... 9
   4.2 Model the economic state factor ......................... 12

5 Data ................................................................... 13

6 Empirical results .............................................. 14
   6.1 Estimated Nelson-Siegel functions ....................... 14
   6.2 Estimated model for the economic state factor ......... 16
   6.3 Estimated term structure of TTC and PIT PDs .......... 17

7 Conclusions ..................................................... 19

8 References ...................................................... 21

A Test statistics .................................................. 24

B The algorithm of X-12-ARIMA ............................. 26

C The Henderson filter ......................................... 29
1 Introduction

The probability of default (PD) of an obligor is a fundamental risk parameter in credit risk analysis. The important role of credit risk in historical banking failures, especially in the latest sub-prime mortgage crisis, has resulted in an increasing interest in credit risk modeling. In brief, credit risk refers to the risk of losses due to some credit event as, for example, the default of a counterparty (Wilson, 1997a,b). The PD of an obligor depends on firm specific risk factors as well as macroeconomic risk factors. Papers focusing on the latter risk factor show that the PD tends to increase in recessions while it tends to decrease in periods of expansion (e.g., Wilson, 1997a,b; Nickell, Perraudin and Varotto, 2000; Bangia et al., 2002). The impact of macroeconomic conditions on the PD has resulted in two different PD estimates, through-the-cycle (TTC) and the point-in-time (PIT) PD. A TTC PD estimate is largely unaffected by the economic state, while a PIT PD estimate varies with the economic cycle.

In the Basel consultative document on the proposed internal rating based (IRB) approach, the Basel committee on banking supervision (BCBS) stated the first formal distinction between TTC and PIT PD estimates (BCBS, 2000). The BCBS provides international recommendations, known as the Basel accords, on banking regulations in order to avoid and mitigate the impact of financial instability. The Basel I accord only captures credit risk, while the Basel II and III accords incorporate other types of risks. In context of credit risk, the Basel accords provide guidelines for calculating the capital banks need to hold to cover losses due to obligor defaults. The Basel II accord allows banks to base their capital requirements on internal as well as external rating system. The first step when banks develop their internal PD models is to determine whether the PD should exhibit TTC or PIT behavior. This decision is of key importance as it affects for example loan pricing, profitability, shareholder value, early warning of defaults, level and volatility of capital requirements. The calculation of regulatory capital is based on the concept of expected and unexpected losses. Expected losses are losses the banks expect to occur and are referred to the cost of doing business. These losses are covered through pricing of credit exposures and provisioning. Unexpected losses, on the other hand, are considered unforeseeable and it is these losses the capital requirements aim to cover. To estimate the unexpected losses (UL) three input factors are used in the Basel II formula, exposure at default (EAD), loss given default (LGD) and PD

\[
UL = q_{0.999}(loss) - E(loss) \approx EAD \cdot LGD \cdot (f(PD) - PD),
\]

where \( q_{0.999} \) is the 99.9% quantile of the loss variable, \( E \) denotes expected value and \( f \) is a strictly increasing function based on the one-factor Merton model, which is discussed further in section 3. If the PD used in equation (1) varies with the economic cycle, i.e. the PD exhibit PIT behavior, it results in cyclical movements in the capital requirements. This is the background to the discussion regarding pro-cyclicality in Basel II. The pro-cyclicality issue concerns the possibility that the Basel framework might cause an amplification of cyclical movements in the economy. For example, in recessions, when the PD tends to increase, capital requirements of banking institutions should also be
increased. This might eventually lead to an increase in capital costs and reduction in credit supply. These effects may further amplify the economic downturn. The opposite effect might occur in periods of economic expansion (e.g., Bangia et al., 2002; Klaassen, Koopman and Lucas, 2005). One proposed method for solving this pro-cyclical problem is to force banks to use TTC ratings. However, this approach is difficult to implement in practice since banks may not have access to the required default data to estimate TTC PD, or they prefer to use more risk-sensitive PIT PD (Carlehed and Petrov, 2012).

Since the PIT PD is sensitive to variations in the macroeconomic environment, banks tend to prefer to use PIT PDs in credit risk pricing and management. The TTC PD, on the other hand, is stable over time and therefore often used to calculate the required capital. The different nature of the TTC and PIT PD is derived from the information they contain. The PIT PD is based on all current information regarding obligors credit worthiness, including both firm specific and macroeconomic risk factors. The TTC PD, by contrast, is a measure of borrowers’ long-run credit quality, where the effects of the economic cycle average close to zero. The different nature of the inputs used to estimate the TTC and PIT PD results in different behavior of these two PD estimates. The TTC PD tends to move slowly and persistent due to the low frequency of the shocks affecting the firm’s long-run credit worthiness. The high frequency and transient nature of macroeconomic shocks cause the PIT PD to move fast. This results in larger variability in the PIT PD than in the TTC PD.

A number of earlier papers have examined the distinction between the TTC and PIT PD empirically (e.g., Nickell, Perraudin and Vorotto, 2000; Bangia et al. 2002). The results indicate that the PIT PD is significant higher than the TTC PD in recessions, while the PIT PD in periods of expansion tends to be lower than the TTC PD. These studies primarily focus on the one-year horizon. Credit exposures, however, often have maturities greater than one year and proper credit risk pricing and management involves analyses over longer periods.

To my knowledge no previous studies have investigated the distinction between the TTC and PIT PD for time periods longer than one year. This observation constitutes the content of this paper where the purpose is to analyze how the TTC and PIT PD relate to each other in the multi-year horizon. To examine this issue the term structure of TTC and PIT PDs are estimated and compared. First the Nelson-Siegel function is used to estimate the term structure of TTC PDs from historical average default rates reported by the rating agency Moody’s. The next step is to transform the term structure of TTC PDs into a term structure of PIT PDs. In this thesis the so-called one-factor Merton model is used to accomplish this. As it will be shown in section 3, to convert TTC PDs into PIT PDs, a variable that indicates the economic state is required. This economic state variable is modeled by a latent economic cycle component extracted from the quarterly real GDP in Sweden. Specifically, the latent economic cycle component is estimated by applying the Hodrick-Prescott (HP) filter to the unobserved trend-cycle component obtained from the X-12-ARIMA procedure. Thereafter, an autoregressive integrated moving average (ARIMA) is used to predict the state of the economy.

The outline of this paper is as follows. Section 2 involves a review of different
models developed in the literature for estimating and moving between TTC and PIT PD. Following this section, the one-factor Merton model used to convert TTC PDs into PIT PDs is presented. Later in this section it is described how the term structure of the two PDs is obtained. In section 4 it is described how the economic state variable is modeled and estimated. Followed by a description of the default data and the GDP variable. Thereafter the empirical results are presented and the last section consists of conclusions.

2 Literature review

In Wilson (1997a,b) one of the first credit risk models that link macroeconomic conditions to PD is developed. Empirical application of this model indicates that the PD contains cyclical movements, especially in periods of recession when the PD increases dramatically. Thereafter, several models have been proposed in the literature to incorporate the impact of macroeconomic environment on the PD (e.g., Belkin, Forest and Suchower, 1998; Kim, 1999, Nickell, Perraudin and Varotto, 2000; Bangia et al., 2002). Credit risk models are usually divided into two categories, structural models based on the framework in Merton (1974) and reduced form models pioneered by Jarrow and Turnbull (1995), Jarrow, Lando and Turnbull (1997) and Duffie and Singleton (1999). The structural models aim to provide an explicit relationship between default events and capital structure, while the reduced form approach model defaults as exogenous events driven by a stochastic process. These two models use different methods to incorporate the effect of macroeconomic environment on PD. Structural models measure the cyclical impact on the PD by incorporating systematic risk factors into the specification of the process driving the variation in the firms’ asset values. On the other hand, in reduced form models default events are usually modeled as a Poisson process with time-varying default intensity, where the intensity function partly consists of macroeconomic variables.

Wilson (1997a,b) present a reduce form model to explicitly link the impact of the economic state to rating transitions. A rating transition matrix consists of probabilities to move between different rating classes, where the last column contain PDs. Specifically, he use a logit model to estimate the PD conditional on the economic state. A constructed macroeconomic index that is defined as an autoregressive process is used as a proxy variable for the systematic risk factor. In particular, this index is built on several macroeconomic variables.

Belkin, Forest and Suchower (1998) employ the Merton framework to derive a standard normally distributed cycle index and propose a method to estimate transition probabilities based on this index. In the literature, their model is usually referred to as the one-factor Merton model. The cycle index is constructed to be negative in periods of contraction and positive in expansions. Specifically, the cycle index is positive when the default rate is lower than average default rate and the ratio of upgrades to downgrades is higher than average. In periods of contraction the revers is true. The credit cycle index is estimated based on default rates provided by the rating agencies Moody’s and Standard&Poor’s.
Inspired by Belkin, Forest and Suchower (1998) and Wilson (1997a,b), Kim (1999) applies an ordered probit model to estimate migration probabilities conditional on the economic state. Similar to Wilson (1997a,b), he creates a credit cycle index based on macroeconomic variables such as interest rates, real GDP growth and unemployment rates.

Nickell, Perraudin and Varotto (2000) extend the ordered probit model in Kim (1999), by permitting the transition probabilities to be conditional on not only the economic cycle, but also on the industry and the country domicile. In contrast to Kim (1999), the economic state variable is modeled as a discrete variable instead of a continuous macroeconomic index. They use the GDP growth to divide the economic state into three states, high, medium and low growth periods. The transition matrices are then estimated separately for these three stages of the economy. The migration probabilities estimated conditional on the medium state are interpreted as the unconditional or TTC probabilities. The empirical results imply that the economic cycle variable is the most important factor to explain the variation in transition probabilities, especially the movements in the PD. They also observe that lower rating classes are more affected by the economic cycle than higher rating classes.

A similar study is performed by Bangia et al. (2002), which also propose a method to estimate transition matrices conditional on different discrete economic states. However, unlike Nickell, Perraudin and Varotto (2000), Bangia et al. (2002) employ a structural approach to derive the systematic shifts in transition probabilities. The economic cycle is divided into two stages, recession and expansion, by following the classification provided by The National Bureau of Economic Research of US’s economic state. First they estimate the transition matrix unconditional of the economic state from rating statistics provided by Standard&Poor’s and thereafter separately for the two economic stages. In order to examine whether the recession and expansion matrices differ, they compare these two transition matrices to the unconditional migration matrix. The empirical analysis indicates a distinct difference between the recession and expansion transition matrices. The most striking distinction between these matrices is the PDs, which increase significantly in periods of recession.

Another interesting approach to model the systematic risk factor used in for example Koopman, Lucas and Klassen (2005), is to employ macroeconomic tools such as the HP-filter to extract a latent economic cycle component. Koopman, Lucas and Klaassen (2005) decompose U.S business failure rates into an autoregressive component and a time-varying cycle component by the HP-filter. These components are then interpreted as systematic risk factors. The purpose of their study is primary to analyze the existence of cyclical movements in default data and whether these cycles correspond to economic cycles. The empirical results indicate that the default rate consists of long-run movements and cyclical pattern. Furthermore, the extracted credit cycle has similar length as the economic cycles.

As a consequence of the Basel II framework more recently studies have focused on the PD instead of the whole migration matrix. Aguais et al. (2008) derive a method to move between the TTC and PIT PD by using a latent credit cycle index in line with
the one-factor Merton model. They use the concept of distance-to-default to transform a TTC PD into a PIT PD. The distance-to-default is referred to the distance between the expected value of the assets and the default point, which is derived by the Merton model. The PD is defined as the standard normal cumulative distribution function of the negative value of distance-to-default. Following Belkin, Forest and Schuwer, Aguais et al. (2008) derive the credit cycle index form default rates, which is positive when economic condition is better than historical average and negative when the economic condition is worse than historical average.

Carlehed and Petrov (2012) propose a model similar to Aguais et al. (2008). In contrast to Aguais et al. (2008), they do not use the concept of distance-to-default to move between TTC and PIT PDs. A detailed description of the methodology derived in Carlehed and Petrov (2012) to convert TTC PDs into PIT PDs is given in the next section.

### 3 Probability of default

#### 3.1 TTC and PIT PD

The methodology derived in this thesis for moving between the TTC and PIT PD is based on a one-factor Merton model framework. According to Merton (1974), a borrower defaults if the value of its assets falls below the amount borrowed. Hence, the PD is the probability that the asset value is below the debt value at some given point in time. Most credit risk models, including the one-factor Merton model, are built on asset returns rather than the asset value itself. In the one-factor Merton model the asset return, $r_i$, of borrower $i$ is assumed to be driven by a single systematic factor, $Z$, and an idiosyncratic risk factor, $\varepsilon_i$, as

$$r_i = \sqrt{\rho_i}Z + \sqrt{1-\rho_i}\varepsilon_i,$$

where the parameter $\rho_i$ is the correlation between the asset return and systematic risk factor. Both risk factors, $Z$ and $\varepsilon_i$, are assumed to be independent standard normal distributed variables. This assumption means that the asset return also must be a standard Gaussian distributed variable. It is important to note that these assumptions are rarely fulfilled in practice but accepted in credit risk modeling. Assuming that the return is normally distributed, means that the PD can be expressed as the probability that a standard normal variable, in this case $r_i$, falls below a certain threshold, $d_i$. Thus, the unconditional or TTC PD, $PD_i$, is

$$PD_i = P(r_i < d_i) = \Phi(d_i),$$

where $\Phi$ is the cumulative distribution function of the standard Gaussian distribution. Later in this section the TTC PD is discussed further. By inverting equation (3), the threshold can be expressed as a function of the TTC PD

$$d_i = \Phi^{-1}(PD_i),$$

where $\Phi^{-1}$ denotes the inverse cumulative standard normal distribution function.
Now, consider a particular outcome \( z \) for the economic state factor. Given \( z \) the conditional or PIT PD, \( PD_i(z) \), can be calculated by

\[
PD_i(z) = P(r_i < d_i | Z = z) \\
= P(\sqrt{\rho_i}Z + \sqrt{1-\rho_i}\varepsilon_i < d_i | Z = z) \\
= P\left( \varepsilon_i < \frac{d_i - \sqrt{\rho_i}z}{\sqrt{1-\rho_i}} \right) \\
= \Phi\left( \frac{d_i - \sqrt{\rho_i}z}{\sqrt{1-\rho_i}} \right) \\
= \Phi\left( \phi^{-1}(PD_i) - \sqrt{\rho_i}z \right). \tag{5}
\]

In words, the PIT PD is a function of the TTC PD, systematic risk factor and correlation parameter. The value of the systematic risk or economic state factor reflects the position in the economic cycle at some stated time period. Since the economic state factor is assumed to be standard normally distributed, the value of this variable is negative when the economic condition is worse than the equilibrium or “normal” state, and positive when the economic environment is better than the normal state (Aguais et al., 2008). In section 4 it is described how the value \( z \) is estimated. Another crucial issue in estimating the PIT PD is to determine the value of the correlation parameter. In the initial version of the Basel II accord the correlation parameter or the factor weight was fixed to 0.2 for all obligors (BCBS, 2001). However, empirical results show that the correlation parameter varies with the PD, therefore the factor weight is expressed as a function of the PD in the current Basel framework (BCBS, 2005). For simplicity, \( \rho_i \) is assumed to be constant over time and fixed to 0.2 for all issuers in this thesis.

Equation (5) can be used to transform TTC PDs into PIT PDs for given values of the economic state factor. In order to use this formula the TTC PD must be estimated. This is a challenging task since there is no fixed definition of the TTC PD. In the literature several interpretations of the TTC PD exists. One view is that the TTC PD should be a long run average, while other states that it should be a worst case scenario (Aguais et al., 2008). Following for example Carlehed and Petrov (2012), the TTC PD is defined as the average of PIT PD, where the average is over all economic states

\[
PD_i = E_Z(PD_i(z)) \\
= \int_{-\infty}^{\infty} \Phi\left( \frac{d_i - \sqrt{\rho_i}z}{\sqrt{1-\rho_i}} \right) \phi(z)dz \\
= \Phi(d_i), \tag{6}
\]

where \( \phi \) is the density function of the standard normal distribution. Note that the expression of the TTC PD stated in equation (6) is the same as the classical definition of TTC PD given in equation (2). Empirically, the TTC PD is estimated by calculating the long run average of historical PIT PDs. Here, long-run should be taken as at least sufficiently long to cover a full economic cycle. Such average PD estimates are reported
by external rating agencies. Specifically, average default frequencies reported by Moody’s are used as TTC estimates in this thesis. In the next section these TTC PD estimates are discussed in detail.

3.2 The term structure of PD

The term structure of PD describes the relationship between PD and time to maturity. The first step in multi-period PD analysis is to determine the length of one time period. Following for example Bangia et al. (2002) and Carlehed and Petrov (2012), a time period corresponds to a quarter. This choice is primarily based on the frequency at which new information about the economic state is provided. In principle, there are three different approaches used in the literature to obtain the term structure of PD, from historical default data, as implied probabilities from market spreads of defaultable bonds or through Merton’s option theoretic approach (see Bluhm, Overbeck and Wagner, 2003, for an overview). As discussed above, in this study historical default rates reported by Moody’s are used to estimate the term structure of TTC PD.

Moody’s apply a dynamic pool approach to calculate multi-year default rates. A pool consists of all issuers holding the same credit rating at a given formation date. Moody’s employ the letter designation, Aaa, Aa, A, Baa, Ba, B, Caa, Ca and C, to indicate the credit rating, where Aaa is the highest rating grade and C is the lowest rating grade. The default or survival status of the issuers in each pool is followed over some stated time horizon \( T \). This time horizon is divided into evenly spaced time intervals of length \( t \). For annual pool spacing default data, which is used in this thesis, a time interval corresponds to a year. In each time interval, some fraction of the pool that has survived up to that time may default. In multi-period PD analysis, the distinction between marginal and cumulative PD becomes important. The marginal PD is the probability that an issuer that has survived in the pool until the beginning of the time period \( t \) defaults during this particular time interval. More specific, the marginal PD in time interval \( t \) for a pool of issuers formed on date \( y \) holding rating \( r \) is defined as the number of defaults, \( x(t) \), that occur in the pool during the time interval \( t \) divided by the effective size of the pool, \( n(t) \), at the start of time \( t \). The cumulative PD is the probability that an obligor defaults at any time between now and time period \( T \). Average cumulative PD, \( Q^r(t) \), for a given historical time period is estimated by first calculating weighted average marginal PDs, \( q^r(t) \), from all the available pool marginal default rates in the historical data set \( Y \)

\[
q^r(t) = \frac{\sum_{y \in Y} x^r_y(t)}{\sum_{y \in Y} n^r_y(t)}
\] (7)

and then estimate cumulative average PDs by

\[
Q^r(T) = 1 - \prod_{t=1}^{T}(1 - q^r(t)).
\] (8)

Note that if \( T = 1 \) the cumulative PD and marginal PD reconcile. The cumulative PD typically increases with maturity for all rating classes as an indicator of that the
Credit risk increases with the time horizon. The marginal PD, on the other hand, either increases or decreases with time depending on the borrower's initial credit rating (Bluhm, Overbeck and Wagner, 2003). Ratings can broadly be divided into investment grade ratings and speculative grade ratings. If Moody's letter designation is used, investment grade ratings are referred to the rating classes, Aaa, Aa, A and Baa, and the speculative grade ratings are related to the remaining rating classes. The default risk for large and solid investment grade firms is low on the near term, while their credit worthiness for the longer term is less certain. In other words, the marginal PD for investment grade issuers increase with maturity. The ability for speculative grade borrowers, for example small growing or highly leverage firms, to fulfill their near term debt obligation is highly uncertain. In the long term, when these firms have surmounted obstacles and survived without a default, their risk of default decrease. Thus, the marginal PD for speculative grade obligors decrease with time (Fons, 1994).

As mentioned above, the cumulative average default rates calculated by Moody’s are reported on yearly basis. These yearly TTC PDs can be converted into quarterly TTC PDs by estimating the continuous term structure of TTC PDs. An interesting approach for estimating these continuous term structures is to take advantage of the similarities between modeling default rates and interest rates. Note that the relationship between the cumulative and marginal PD stated in equation (7) is analogous to the link between the spot and forward interest rate. This similarity implies that the curve fitting problem faced in this thesis can be equated with the modeling of spot yield curves. After the continuous term structure of TTC PDs has been estimated, the marginal quarterly TTC PDs can be calculated. Therefore, by applying equation (5) these quarterly marginal TTC PDs can be transformed into marginal PIT PDs.

The Nelson-Siegel function is widely used to model the term structure of interest rates. This function is able to produce a wide range of curve shapes, including U-shape, S-shape, monotonic or inverted U-shape. Nelson and Siegel (1987) suggest to model the spot rate as

\[
Q(t, \beta) = \beta_0 + \beta_1 \frac{1 - \exp\left(-\frac{t}{\tau}\right)}{\tau} + \beta_2 \left[\frac{1 - \exp\left(-\frac{t}{2\tau}\right)}{\tau} - \exp\left(-\frac{t}{\tau}\right)\right],
\]

(9)

where \( t \) is time to maturity, \( \beta_0, \beta_1 \) and \( \beta_2 \) are interpreted respectively as the level factor, the slope factor and curvature factor and \( \tau \) is the rate of exponential decay. The curve is downward sloping if \( \beta_1 \) is positive, while it is upward sloping if \( \beta_1 \) is negative. Furthermore, the curve is hump shaped if \( \beta_2 \) is larger than zero. On the other hand, if

\(^1\)Nelson and Siegel (1987) derive the spot rate by integrating the instantaneous forward rate, \( h(t) \), where the instantaneous forward rates is expresses as a solution to a second-order differential equation with real and equal roots, \( h(t)\beta = \beta_0 + \beta_1 \exp\left(-\frac{t}{\tau}\right) + \beta_2 \left[\exp\left(-\frac{t}{\tau}\right) \exp\left(\frac{t}{\tau}\right)\right] \). In PD context, the instantaneous forward rate or the hazard rate is the rate at which default events occurs. The marginal PD and hazard rates are closely related, more specific the marginal PD between \( t \) and \( t + \Delta t \) is equal to \( h(t)\Delta t \), assuming that hazard rate is continuous. This implies if \( \Delta t = 1 \), which is the case in this thesis, the hazard rate and marginal PD reconcile. See Bluhm, Overbeck and Wagner for a comprehensive explanation of the hazard rate.
$\beta_2$ is smaller than zero the curve is U-shaped. The shape parameter $\tau$ determines both the steepness of the slope and the location of the maximum or the minimum.

It may be pointed out that there are several extensions of the Nelson-Siegel function proposed in the literature. For example Svensson (1994), extended the Nelson-Siegel function by an additional polynomial term to achieve more flexibility and Diebold and Li (2006) introduced a dynamic version of the Nelson-Siegel model. De Poorter (2007) provides an overview of these variants. The estimation procedures related to these extended functions are usually developed for the data available to model yield curves and therefore not suitable to estimate the term structure of PDs, where the available data is more limited. Furthermore, a graphical analysis of the empirical data indicate that the standard Nelson-Siegel function is able to capture the shape of the term structure of TTC PD.

The estimation procedure of the Nelson-Siegel model is complicated by that function is nonlinear in the shape parameter $\tau$. There are mainly two estimation procedures used in the literature. One approach is to estimate all the parameters simultaneously by nonlinear regression techniques. Cairns and Pritchard (2001) among others, however, show that the nonlinear estimators are extremely sensitive to the starting values used. Moreover, these authors remark that the probability of getting local optima is high. In order to avoid nonlinear estimation procedures, Nelson and Siegel (1987) linearize their model by fixing $\tau$ to a pre-determined value. The linearized model is thereafter estimated by ordinary least square for a whole grid of $\tau$ values, where the model with highest $R^2$ is selected. These estimated parameters have been reported to behave erratically over time and to have relatively large variances (De Pooter, 2007). Despite these drawbacks, the linearized estimation approach is the most frequently used method in the literature and also in this thesis. In order to estimate the parameters efficiently and consistently, no autocorrelation or heteroscedasticity are allowed in the residuals. The Ljung-Box test statistic is used to test the null hypothesis that the residuals are independent. Furthermore, the possible existence of conditional heteroscedasticity is tested by the Breusch-Pagan test. See Appendix A for a detailed description of these test statistics.

4 Economic state factor

4.1 Extracting the economic state factor

As mentioned in section 2, there are several different methodologies used in the literature to model the systematic risk factor. One common approach is to construct an economic cycle index based on macroeconomic variables (e.g. Wilson, 1997a,b; Kim, 1998; Nickell, Perraudin and Vorotto, 2000). In this thesis a latent cycle component extracted from the real GDP in Sweden is used as a proxy variable for the economic state factor. Macroeconomic time series are usually assumed to consist of multiple components corresponding to a long-term trend, a medium-term economic cycle, seasonal movements and short-term irregular variation. The HP-filter derived in Hodrick and Prescott (1997) is widely used in the macroeconomic literature for estimating the economic cycle. Since
the time series used as input to this filter is assumed to have no seasonal variations, seasonal adjusted data is commonly applied to the HP-filter. Even though, the HP-filter is frequently used, it is also criticized. The major caveats associated with the HP-filter are that it provides poor estimates at the beginning and end of the sample period, the value of the smoothness parameter is to some extent chosen arbitrary, business cycle can be generated even if none is present in the original data and noise in the cyclical signal significant disturbs its interpretation (Kraiser and Marawall, 1999). Kaiser and Marawall (1999) among others, however, show that the performance of the filter can be improved by using the latent trend-cycle component instead of seasonal adjusted series. Examples of these trend-cycle components are the ones produced by the X-12-ARIMA procedure or the ones obtained in the model-based decomposition of a series, as in programs STAMP or SEATS. In this study the X-12-ARIMA approach is used to estimate the trend-cycle component.

The X-12-ARIMA method decompose the original time series, $y_t$, into a trend-cycle component, $tc_t$, a seasonal component, $s_t$, and an irregular component, $i_t$. The first step in the X-12-ARIMA procedure is to determine the mode of decomposition, for example multiplicative or additive decomposition. The appropriate decomposition method depends on the relationship between the three time series components. The variability in most macroeconomic variables increase with the level of the series, indicating that the components are multiplicative related (Findley, 1998). To correct for this pattern, macroeconomic series are usually transformed into natural logarithms. This transformation changes the multiplicative association between the components into an additive relationship. In the additive model the original time series is the sum of the three components

$$y_t = tc_t + s_t + i_t.$$  \hfill (10)

In the X-12-ARIMA procedure a set of moving averages is used to decompose the original into the three time series components. A moving average, $m_t$, of order $g + f + 1$ and weights $w_j$ can be written as

$$m_t = \sum_{j=-g}^{f} w_j x_{t+j},$$ \hfill (11)

where $x_{t+j}$ is the series the moving average is applied to, $g$ and $f$ is the number of data points before and after the target point, respectively. The moving average is said to be centered when $f$ is equal to $g$. Furthermore, if $w_j$ is equal to $w_{-j}$ the moving average is symmetric. Observe that it is impossible to smooth the first $g$ observations and the last $f$ observations of the series. This issue is known as the endpoint problem. The X-12-ARIMA procedure use ARIMA modeling techniques to extend the original time series with backcasts and forecasts to correct for this issue. The original time series is also adjusted for calendar effects, outliers and trading day effects. It is outside the scoop of this thesis to describe all these pre-adjustments. For a detailed review see Findley et al. (1998).
The X-12-ARIMA decomposition procedure consists of three steps. The process is described briefly here and in Appendix B a more detailed review is provided. In the first step a crude trend-cycle component is estimated by applying a centered moving average to the original time series. This estimated trend-cycle component is thereafter subtracted from the original series. After subtracting this component, the residual series consists of seasonal and irregular variations. In order to estimate the preliminary seasonal component a seasonal moving average is applied to the seasonal-irregular component for each quarter separately. These seasonal factors, however, are biased because their yearly averages may not be equal to their mean. Unbiased estimates are obtained by subtracting each seasonal factor by a centered moving average of the preliminary seasonal factors. Thereafter, the initial seasonal adjusted series, $A_t$, is estimated by removing the unbiased seasonal component from the original series. This process described above is then repeated in step two with a few modifications. The trend-cycle component often has too much curvature to be adequately represented by a simple moving average. The X-12-ARIMA program corrects for this by using Henderson moving averages to estimate the trend-cycle component in the second and third step. The Henderson moving average can only be applied to non-seasonal series and can therefore not be used directly in step one. The Henderson weights are designed, so that the long-term trend estimates are as smooth as possible and reproduce a wide range of curvatures that may include peaks and troughs. There are several different methodologies developed in the literature to estimate the Henderson weights. The widely used technique derived by Gray and Thomson (1996) is presented in Appendix C. In the third step of the X-12-ARIMA procedure, the final trend-cycle and irregular components are estimated.

The economic cycle component is then obtained by removing the trend from the X-12-ARIMA trend-cycle component. As already stated above, in this thesis the well-known HP-filter is used to separate the trend and cycle component. The HP-filter is an algorithm that smooth the original time series, $tc_t$, to estimate its trend component, $t_t$. The cycle component, $c_t$, is obtained by taking the difference between the original series and the trend. Hence, the original time series is assumed to be the sum of the cycle component and the trend component

$$tc_t = t_t + c_t. \quad (12)$$

The trend is estimated by solving the following minimization problem

$$\min_{t_t} \left\{ \sum_{t=1}^{N} (tc_t - t_t)^2 + \lambda \sum_{t=2}^{N}[(t_t - t_{t-1}) - (t_{t-1} - t_{t-2})]^2 \right\}, \quad (13)$$

where $\lambda$ is a nonnegative smoothing parameter. The first part of equation (13), the squared sum of the difference between the original time series and the trend, is a measurement of the goodness of fit. The second part, which consists of the sum of squared second differences in the trend, is a smoothness measure. The technical specification of the HP-filter imposes a trade-off between the goodness of fit and the smoothness of the trend. In particular, the smoother the trend, the poorer is the fit of the trend to the
original series and vice versa. This trade-off is governed by the smoothing parameter. A lower value of $\lambda$ produces a trend that closely follows the original series, while a higher value of $\lambda$ gives a more smoother trend and thereby allows for more variation in the cycle component. Note that for $\lambda$ equal to zero the trend component is equivalent to the original series, and as $\lambda$ approaches infinity the solution of equation (13) goes towards the least square fit of a linear trend model. Hodrick and Prescott (1997) recommended setting the smoothing parameter to 1600 for quarterly data. This suggestion is usually followed in applied work and also in this thesis. As a final step, the cycle component is standardized. Henceforth this estimated standardized economic state factor is denoted $Z_t$.

4.2 Model the economic state factor

The purpose of extracting the economic cycle is to predict the state of the economy at future times in order to transform the term structure of TTC PDs into a term structure of PIT PDs. Box-Jenkins’ ARIMA $\langle p, d, q \rangle$ models are frequently used to forecast the values for a particular time series at future times. The letters, $p$, $d$, and $q$, indicate the order of the autoregressive (AR), integrated (I) respective moving average (MA) parts. By using lag operators, $LZ_t = Z_{t-1}$, the general ARIMA $\langle p, d, q \rangle$ can be written as

$$\phi_p(L)(1 - L)^dZ_t = \alpha + \theta_q(L)a_t,$$

(14)

where $a_t$ is a white noise variable, the polynomials $\phi(L) = 1 - \phi_1L - ... - \phi_pL^p$ and $\theta_q(L) = 1 - \theta_1L - ... - \theta_qL^q$ represent the AR respective MA parts, $(1 - L)^d$ is the integrated component and $\alpha$ is a constant.

The first step in identifying the ARIMA model specification is to determine the order of integration, which is referred to the number of differences required to obtain a stationary series. The HP-filter is constructed so that the cycle component is stationary (Hodrick and Prescott, 1997). In other terms, the order of integration for the estimated economic state series is equal to zero.

When the time series is stationary, the order of the AR and MA terms can be determined. Each ARMA model is associated with specific pattern in the autocorrelation function (ACF) and the partial autocorrelation function (PACF). The orders of $p$ and $q$ are determined by comparing the estimated ACF and PACF of the observed time series with the theoretical functions of ARMA processes. The ACF and PACF provide effective tools for identifying pure AR($p$) or MA($q$) models. The theoretical ACF and PACF of models with both AR and MA terms have infinitely many nonzero values, which makes it difficult to identify mixed models from the sample ACF and PACF. Tsay and Tiao (1984) implemented the extended autocorrelation function (EACF) for the purpose of identifying mixed models. It is, however, usually difficult to determine the most accurate model based on these estimated autocorrelation functions. The Akaike information criterion (AIC) is commonly used to select the final model specification by comparing the different models considered in the autocorrelation analysis described above. The selection rule for this criterion is to choose the model with lowest value of the AIC.
To examine whether the residuals from the ARIMA model are independent the Ljung-Box test statistic is used. Furthermore, the test statistic developed in Engle (1983) is applied to test the null hypothesis that no autoregressive conditional heteroscedasticity (ARCH) effects exist in the residuals. For a comprehensive description of these test statistics see Appendix A.

The selected model is thereafter used to forecast the economic state several periods ahead. Let $\hat{Z}_h(l)$ denote the forecast of $Z_{h+l}$, where $h$ is the forecast origin and $l \geq 1$ is the forecast horizon. Several criterions are available for determining the optimal forecast. Usually the objective is to obtain small prediction errors. A predictor is therefore often selected by minimizing the expected squared prediction error

$$E\{[Z_{h+l} - \hat{Z}_h(l)]^2 | F_h\},$$

(15)

where $F_h$ is all available information at time $h$ and $E\{·|F_h\}$ denotes the conditional expectation given the information set $F_h$. It can relatively easy be derived that the optimal predictor, $\hat{Z}_h(l)$, given the information in $h$, is the conditional expectation of $Z_{h+l}$ given the information $F_h$. This results in the following expression for the $l$-step ahead forecast of a general ARMA ($p,q$) model

$$\hat{Z}_h(l) = E(Z_{h+l}|F_h) = \alpha + \sum_{i=1}^{p} \phi_i \hat{Z}_h(l-i) - \sum_{i=1}^{q} \theta_i a_h(l-i),$$

(16)

where, as stated above, $\alpha$ is a white noise variable, $\phi$ and $\theta$ are the parameters related to the AR respective MA terms and $\alpha$ is a constant. Furthermore, $\hat{Z}_h(l-i) = Z_{h+l-i}$ if $l-i \leq 0$, $a_h(l-i) = 0$ if $l-i > 0$ and $a_h(l-i) = a_{h+l-i}$ if $l-i \leq 0$. Note that the one-step ahead forecast is based on only observed values, while the multi-steps ahead forecasts are based on both observed and previous forecasted values. The confidence intervals for the predictions typically increase with the forecast horizon due to the multi-steps ahead forecasts are based on several previous predicted values.

5 Data

Since the default data already has been described in detail, only a brief presentation of the data is provided here. The cumulative average default rates from one to ten years used in this thesis are collected from Moody’s ”annual default study: Corporate default and recovery rates, 1920-2011” for the seven rating classes given in section 3. This broad division of rating classes is preferred to more detailed classification since more observations are included in each rating class. These average default frequencies are calculated based on historical default statistics between the years 1970 and 2011. Furthermore, the data is corrected for withdrawals, meaning that it is assumed that issuers whose ratings are withdrawn would have faced the same risk of default as other similarly rated issuers if they had stayed in the sample.

Quarterly real GDP of Sweden is used to extract the economic cycle component. The time series of GDP extends from the first quarter of year 1993 to the fourth quarter
of year 2013. This data is obtained from the Statistics of Sweden (SCB). The length of the time series is limited by the new measurement method introduced in year 1993, which complicates the comparison between new and old data. In Figure 1 the real GDP of Sweden is illustrated. As shown in Figure 1, the real GDP series consists of quarterly seasonal movements. It is also observed that the seasonal fluctuations around the trend do not vary with the level of the time series. This observation indicates that an additive decomposition approach is appropriate to employ for extracting the trend-cycle component. In Table 1 descriptive statistics of the real GDP series is presented.

![Figure 1: Real GDP of Sweden, 1993:1 - 2013:4](image)

Table 1: Descriptive statistics

<table>
<thead>
<tr>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Max</th>
<th>Min</th>
<th>Kurtosis</th>
<th>Skewness</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>739 809.1</td>
<td>122 569.2</td>
<td>962 229.0</td>
<td>499 085.0</td>
<td>1.868</td>
<td>-0.109</td>
<td>84</td>
</tr>
</tbody>
</table>

Note: Real GDP is measured in million SEK.

6 Empirical results

6.1 Estimated Nelson-Siegel functions

The first step in the empirical analysis is to estimate the term structure of TTC PDs for each rating class based on average default rates reported by Moody’s. The Nelson-Siegel models are first linearized by fixing τ to a pre-determined value and thereafter estimated with ordinary least squares. This procedure is repeated for a whole grid of τ values ranging from 1 to 10. The models with the highest adjusted R-square is chosen. In Table 2 the estimated Nelson-Siegel parameters and the related p-values are presented. Most coefficients are significant at 5 percent significance level. The results from the Ljung-Box and Breusch-Pagan tests indicate no presence of either autocorrelation or conditional heteroscedasticity in the residuals. In particular, the Ljung-Box test statistic is performed for the lags one to four. The estimated value of β₀ represents the level of
Table 2: Nelson-Siegel functions

<table>
<thead>
<tr>
<th>Rating</th>
<th>$\beta_0$</th>
<th>Coef</th>
<th>P-value</th>
<th>$\beta_1$</th>
<th>Coef</th>
<th>P-value</th>
<th>$\beta_2$</th>
<th>Coef</th>
<th>P-value</th>
<th>$\tau$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aaa</td>
<td>0.025</td>
<td>0.038</td>
<td>-0.024</td>
<td>0.042</td>
<td>-0.034</td>
<td>0.018</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aa</td>
<td>0.025</td>
<td>0.000</td>
<td>-0.026</td>
<td>0.057</td>
<td>-0.026</td>
<td>0.000</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>0.176</td>
<td>0.021</td>
<td>-0.176</td>
<td>0.007</td>
<td>-0.165</td>
<td>0.012</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baa</td>
<td>0.212</td>
<td>0.002</td>
<td>-0.215</td>
<td>0.067</td>
<td>-0.201</td>
<td>0.006</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ba</td>
<td>0.459</td>
<td>0.000</td>
<td>-0.469</td>
<td>0.001</td>
<td>0.303</td>
<td>0.000</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>0.822</td>
<td>0.000</td>
<td>-0.843</td>
<td>0.026</td>
<td>0.363</td>
<td>0.000</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caa-C</td>
<td>0.960</td>
<td>0.039</td>
<td>-0.912</td>
<td>0.000</td>
<td>0.513</td>
<td>0.000</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The curves. As expected the value of this coefficient is larger for lower ratings than for higher ratings. This indicates that the PD decreases with the level of rating. As a consequence of that $\beta_1$ is negative for all ratings, the term structure of cumulative TTC PD is upward sloping for each rating class. The coefficient $\beta_2$ is negative for the rating classes of investment grade, which implies that the curves related to these rating classes are U-shaped. In contrast, for the speculative grade ratings the values of $\beta_2$ are positive. Hence, the curves related to the investment grade ratings are hump shaped. The interpretation of the estimated values of $\tau$ is more difficult. In brief, higher values indicate flatter curves. In Figure 2 the estimated Nelson-Siegel curves for the rating classes of investment grade are illustrated.

![Figure 2: Cumulative TTC PD for the rating classes Aaa to Baa](image)

The cumulative TTC PDs for investment grade ratings increase at an exponential rate. Among the rating classes illustrated in Figure 2, lower ratings have a flatter curve than higher ratings. This means that the PD increases at a higher rate for higher ratings than for lower ratings. Unfortunately this pattern is difficult to observe in the figure because of the level differences in PD between the rating classes. In Figure 3 the estimated curves for the rating classes related to speculative grade are illustrated.
For speculative grade ratings, the cumulative TTC PDs increase with a decreasing rate over time. The lower the rating the more hump shaped the Nelson-Siegel curve is. This implies, the cumulative PDs related to the poorest rating classes decrease at a faster rate than the cumulative PDs for the higher ratings illustrated in Figure 2.

6.2 Estimated model for the economic state factor

The output from the X-12-ARIMA program is extensive and difficult to summarize here. The test statistics performed to investigate the existence of seasonality, a F-test and nonparametric test,\(^2\) indicate significance seasonal movements at 1 percent significance level. The final trend-cycle component is estimated by a 7-term Henderson filter. The standardized economic cycle component estimated by applying the HP-filter to the trend-cycle component is illustrated in Figure 4.

\(^2\)See Findley et al. (1998) for an explanation of these tests.
Based on the estimated ACF, PACF and EACF several model specifications are considered appropriate for modeling the economic state factor. For all these models the AIC is estimated, where the ARIMA \((2, 0, 1)\) model obtain the lowest AIC value. The estimated ARIMA model is presented in Table 4 below.

### Table 3: The ARIMA model

<table>
<thead>
<tr>
<th>Coef</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.036</td>
</tr>
<tr>
<td>AR1</td>
<td>1.445</td>
</tr>
<tr>
<td>AR2</td>
<td>-0.632</td>
</tr>
<tr>
<td>MA1</td>
<td>0.756</td>
</tr>
</tbody>
</table>

The adjusted R-squared for the estimated ARIMA \((2, 0, 1)\) model is equal to 0.885. The results from the Ljung-Box and Engel’s tests indicate no presence of either autocorrelation or ARCH effects in the residuals. The Ljung-Box test statistic is performed up to four lags. The estimated ARIMA model is thereafter used to forecast the economic state six years ahead by the method described in section 4.2. In other terms, the forecast horizon is equal to 24 since the time series is on quarterly basis. The fourth quarter of year 2013 is the last observation in the real GDP time series, which is used as the forecasting origin. Thus, the first forecast corresponds to the predicted value of \(Z\) related to the first quarter of year 2014.

#### 6.3 Estimated term structure of TTC and PIT PDs

Based on the estimated Nelson-Siegel functions, the marginal TTC PDs are calculated for each quarter six years ahead. However, usually yearly PDs are more interesting than quarterly PDs. To transform quarterly marginal TTC PDs into yearly marginal PDs equation (7) is applied. The estimated yearly marginal TTC PDs are presented in Table 4.
Table 4: Marginal TTC PD (%)

<table>
<thead>
<tr>
<th>Rating</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aaa</td>
<td>0.000</td>
<td>0.000</td>
<td>0.013</td>
<td>0.032</td>
<td>0.051</td>
<td>0.073</td>
</tr>
<tr>
<td>Aa</td>
<td>0.010</td>
<td>0.058</td>
<td>0.082</td>
<td>0.097</td>
<td>0.104</td>
<td>0.108</td>
</tr>
<tr>
<td>A</td>
<td>0.067</td>
<td>0.129</td>
<td>0.171</td>
<td>0.207</td>
<td>0.237</td>
<td>0.260</td>
</tr>
<tr>
<td>Baa</td>
<td>0.202</td>
<td>0.298</td>
<td>0.394</td>
<td>0.462</td>
<td>0.509</td>
<td>0.534</td>
</tr>
<tr>
<td>Ba</td>
<td>2.423</td>
<td>2.361</td>
<td>2.359</td>
<td>2.343</td>
<td>2.315</td>
<td>2.277</td>
</tr>
<tr>
<td>B</td>
<td>4.097</td>
<td>6.239</td>
<td>6.076</td>
<td>5.905</td>
<td>5.729</td>
<td>5.548</td>
</tr>
</tbody>
</table>

For investment grade ratings, the marginal TTC PDs increase between each year. In contrast, the marginal TTC PDs for speculative grade ratings decrease between each year. These patterns are directly related to the shape of the estimated Nelson-Siegel curves. The estimated quarterly marginal TTC PDs and the forecasted values of the economic state are used in equation (5) to calculate marginal PIT PDs for each quarter until year 2019. By using the same approach as for TTC PDs, the quarterly marginal PIT PDs are transformed into yearly marginal PIT PDs. These estimates are represented in Table 5.

Table 5: Marginal PIT PD (%)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Aaa</td>
<td>0.000</td>
<td>0.001</td>
<td>0.029</td>
<td>0.019</td>
<td>0.041</td>
<td>0.075</td>
</tr>
<tr>
<td>Aa</td>
<td>0.006</td>
<td>0.086</td>
<td>0.116</td>
<td>0.081</td>
<td>0.096</td>
<td>0.111</td>
</tr>
<tr>
<td>A</td>
<td>0.034</td>
<td>0.199</td>
<td>0.242</td>
<td>0.156</td>
<td>0.211</td>
<td>0.268</td>
</tr>
<tr>
<td>Baa</td>
<td>0.096</td>
<td>0.623</td>
<td>0.537</td>
<td>0.328</td>
<td>0.453</td>
<td>0.546</td>
</tr>
<tr>
<td>Ba</td>
<td>0.664</td>
<td>3.065</td>
<td>2.781</td>
<td>1.839</td>
<td>2.219</td>
<td>2.398</td>
</tr>
<tr>
<td>B</td>
<td>2.024</td>
<td>7.989</td>
<td>8.196</td>
<td>4.369</td>
<td>5.419</td>
<td>5.759</td>
</tr>
<tr>
<td>Caa-C</td>
<td>11.824</td>
<td>16.344</td>
<td>15.761</td>
<td>8.179</td>
<td>11.069</td>
<td>11.493</td>
</tr>
</tbody>
</table>

In contrast to the marginal TTC PDs, PIT PDs for investment grade ratings either increase or decrease from one year to another depending on the value of the economic state factor. Similar patterns are observed in PIT PDs related to speculative grade rating classes. It is observed that the marginal PIT PD varies around the TTC PD. In the analysis of multi-year PDs it is also interesting to examine the cumulative PDs. Again, equation (7) is applied, in this case to estimate yearly cumulative PDs based on the yearly marginal PDs. These calculations are performed for both TTC and PIT PDs. In Table 6 the cumulative TTC PDs are presented.
Table 6: Cumulative TTC PD (%)

<table>
<thead>
<tr>
<th>Rating</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aaa</td>
<td>0.000</td>
<td>0.000</td>
<td>0.013</td>
<td>0.045</td>
<td>0.097</td>
<td>0.161</td>
</tr>
<tr>
<td>Aa</td>
<td>0.010</td>
<td>0.067</td>
<td>0.149</td>
<td>0.249</td>
<td>0.353</td>
<td>0.458</td>
</tr>
<tr>
<td>A</td>
<td>0.067</td>
<td>0.198</td>
<td>0.367</td>
<td>0.573</td>
<td>0.808</td>
<td>1.069</td>
</tr>
<tr>
<td>Baa</td>
<td>0.202</td>
<td>0.499</td>
<td>0.891</td>
<td>1.349</td>
<td>1.851</td>
<td>2.382</td>
</tr>
<tr>
<td>B</td>
<td>4.097</td>
<td>10.080</td>
<td>15.544</td>
<td>20.531</td>
<td>25.083</td>
<td>29.242</td>
</tr>
<tr>
<td>Caa-C</td>
<td>17.628</td>
<td>28.348</td>
<td>37.353</td>
<td>44.932</td>
<td>51.322</td>
<td>56.722</td>
</tr>
</tbody>
</table>

The interpretation of the patterns in the cumulative TTC PDs has already been given in section 6.1. The estimated cumulative PIT PDs are presented in Table 7.

Table 7: Cumulative PIT PD (%)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Aaa</td>
<td>0.000</td>
<td>0.001</td>
<td>0.020</td>
<td>0.033</td>
<td>0.074</td>
<td>0.153</td>
</tr>
<tr>
<td>Aa</td>
<td>0.006</td>
<td>0.082</td>
<td>0.168</td>
<td>0.210</td>
<td>0.295</td>
<td>0.407</td>
</tr>
<tr>
<td>A</td>
<td>0.034</td>
<td>0.203</td>
<td>0.444</td>
<td>0.540</td>
<td>0.740</td>
<td>1.004</td>
</tr>
<tr>
<td>Baa</td>
<td>0.096</td>
<td>0.419</td>
<td>0.854</td>
<td>1.079</td>
<td>1.518</td>
<td>2.046</td>
</tr>
<tr>
<td>Ba</td>
<td>0.664</td>
<td>3.709</td>
<td>6.098</td>
<td>7.355</td>
<td>9.318</td>
<td>11.438</td>
</tr>
<tr>
<td>Caa-C</td>
<td>11.824</td>
<td>26.235</td>
<td>35.649</td>
<td>40.912</td>
<td>47.571</td>
<td>53.596</td>
</tr>
</tbody>
</table>

Similar as the cumulative TTC PDs, the cumulative PIT PDs increase over time for all rating classes. The cumulative TTC PDs are larger than cumulative PIT PDs for each year. It is also observed that the difference between the TTC and PIT PDs tends decrease over time.

7 Conclusions

This thesis examines how the TTC and PIT PD relate to each other in the multi-year horizon. In the first step to analyze this issue empirically the Nelson-Siegel function is used to estimate the term structure of TTC PDs based on average default rates reported by Moody’s. The term structure of TTC PDs is thereafter converted into a structure of PIT PDs by using a Merton model framework to incorporate the economic cycle effect. As already mentioned, a drawback of this study is that the asset correlation parameter is assumed to be fixed at 0.2 for all rating classes. Furthermore, the economic state factor is estimated from the real GDP of Sweden by applying the HP-filter to the trend-cycle component produced by the X-12-ARIMA procedure. An ARIMA model is thereafter used to predict the state of the economy for particular time periods. The forecasted values of the cycle component associated with the one-year horizon are positive, which indicates a period of expansion. Consistent with previous studies, for example Bangia et al. (2002), the PIT PD is lower than the TTC PD for all rating classes in the one-year horizon.
horizon when the economy is in expansion.

The Nelson-Siegel curves are upward sloping for all rating classes indicating that the default risk increases with time. Similar to the empirically results in Fons (1994), the cumulative TTC PD for investment grade ratings increase at an exponential rate with maturity, while the cumulative TTC PD for speculative grade ratings increase with a decreasing rate as the horizon grows. As a direct consequence of the shapes of the Nelson-Siegel curves, the term structures of marginal TTC PD for investment and speculative grade ratings are upward respectively downward sloping. Similar trends are observed in the term structure of the marginal PIT PD. However, unlike the TTC PD, the PIT PD vary with the economic cycle.

The empirical results indicate that the cyclical movements in the marginal PIT PD decreases with the horizon. This means that the distinction between the marginal TTC and PIT PD diminish as the horizon grows. The decreasing difference between the term structure of TTC and PIT PDs is even more clear in the analysis of cumulative PDs. The cumulative TTC PD is higher than the cumulative PIT PD for all time periods. The multi-period PD is affected by the economic state in the current period as well as all previous periods. In the first periods when the impact of the economic cycle is as largest, the economic state factor takes positive values. This could be the explanation to why the TTC PD is higher than the PIT PD for the whole period considered. Furthermore, since the TTC PD usually is interpreted as the long-run PD, these patterns indicate that the PIT PD tends to revert to its long-run trend. However, in order to investigate whether the TTC PD and PIT PD convergence eventually even longer horizons need to be analyzed. As a consequence of the high uncertainty in long-run forecasts, it is difficult to examine this question empirically. The empirical results show that the PIT PD varies around the TTC PD for the whole term structure and that the PIT PD tends to move against the TTC PD.

An improvement of this study would be to develop the method used for moving between TTC and PIT PD to better capture the dependence in PD over time. In order to further improve the analysis the correlation parameter should be estimated based on the PD. Another question for future studies is to examine the term structures of TTC and PIT PD given different initial states of the economy. It would also be interesting to investigate whether the term structures of TTC and PIT PD eventually converges by using more sophisticated time series analysis to model the economic cycle.
8 References


A Test statistics

In this appendix the test statistics applied to test for autocorrelation and heteroscedasticity are presented. First, the Ljung-Box test statistic is used to examine if the residuals from the Nelson-Siegel and ARIMA models are independent. The null hypothesis of this test statistic is $H_0: \rho_1 = \ldots = \rho_m = 0$ and the alternative hypothesis is $H_a: \rho_i \neq 0$ for some $i \in \{1, \ldots, m\}$, where $\rho_i$ are the autocorrelation coefficients. The test statistic is defined as

$$Q(m) = T(T + 2) \sum_{l=1}^{m} \frac{\hat{\rho}_l^2}{T-l}$$

where $T$ is the sample size, $\hat{\rho}_l$ is the estimated autocorrelation coefficients at lag $l$ for the residuals and $m$ is the number of lags being tested. Under the null hypothesis this test statistic is asymptotically distributed as a chi-squared distribution with $m$ degrees of freedom. The decision rule is to reject the null hypothesis if $Q(m) > \chi^2(\alpha)(m)$, where $\chi^2(\alpha)(m)$ denotes the $100(1-\alpha)$th percentile of a chi-squared distribution with $m$ degrees of freedom. The empirical results from this test implies that the null cannot be rejected for either the nonlinear model or the ARIMA model.

To test the presence of conditional heteroskedacity in the Nelson-Siegel models the Breuch-Pagan test is used. The basis in the Breuch-Pagan test is to perform the auxiliary regression

$$a_t^2 = \beta_0 + \beta_1 x_1 + \ldots + \beta_m x_m + e_t$$

where $a_t^2$ is the squared residuals from the Nelson-Siegel model, $x_i$ is the explanatory variables and $e_t$ is the error term. The null hypothesis is $H_0: \beta_1 = \ldots = \beta_m = 0$. Furthermore, the test statistic is defined as

$$LM = nR^2$$

where $n$ is the sample size and $R^2$ is the coefficient of determination of the auxiliary regression. This test statistic is asymptotically distributed as $\chi^2(m-1)$ under the null hypothesis of homoscedasticity. The empirical results indicate that the null hypothesis cannot be rejected.

Engel’s (1982) Lagrange multiplier test is applied to examine the existence of autoregressive conditional heteroscedasticity effects in the residuals from the ARIMA model. This test is equivalent to the standard F statistic. Engel’s test statistic is performed by estimating

$$a_t^2 = \beta_0 + \beta_1 a_{t-1}^2 + \ldots + \beta_m a_{t-m}^2 + e_t, t = m+1, \ldots, T,$$

where $T$ is the sample size, $a_t^2$ is the squared residuals from the ARIMA model, $m$ is a pre specified integer, $e_t$ is the error term. The null hypothesis is $H_0: \beta_1 = \ldots = \beta_m = 0$. Let $SSR_0 = \sum_{t=m+1}^{T} (a_t^2 - \bar{\omega})^2$, where $\bar{\omega} = (1/T) \sum_{t=1}^{T} a_t^2$ is the sample mean of $a_t^2$ and $SSR_1 = \sum_{t=m+1}^{T} \hat{e}_t^2$, where $\hat{e}_t$ is the estimated residuals from equation (16). The test statistic can be written as

$$F = \frac{(SSR_0 - SSR_1)/m}{SSR_1/(T-2m-1)}$$

24
This test statistic is asymptotically distributed as a chi-squared distribution with $m$ degrees of freedom under the null hypothesis. The decision rule is to reject the null hypothesis if $F > \chi^2_{\alpha}(m)$, where $\chi^2_{\alpha}(m)$ denotes the $100(1 - \alpha)$th percentile of a chi-squared distribution with $m$ degrees of freedom. Based on the empirical results from this test the null hypothesis cannot be rejected.
The algorithm of X-12-ARIMA

The purpose of this review is to describe the estimation procedure of the trend-cycle component, not to provide an exhaustive description of the X-12-ARIMA program. For a detailed explanation of this seasonal adjustment program see Findley et al. (1998). The component decomposition is performed in three steps following the procedure of X-11.

In the X-12-ARIMA program the trend-cycle component, $TC_t$, is estimated by moving averages. The simplest form of moving averages is the symmetric moving average. First consider the following time series:

$$x_1, x_2, x_3, ..., x_{t-1}, x_t, x_{t+1}, ..., x_{t+n}$$

(22)

where the index is referred to the time period. The symmetric moving average include an equal number of data points on both sides of the target data point and each data point have the same weights. For example, assume that the span of data for the moving average is equal to three, the moving average at time period $t$, $m_t$, is then calculated as follows:

$$m_t = \frac{x_{t-1} + x_t + x_{t+1}}{3} = \frac{1}{3} x_{t-1} + \frac{1}{3} x_t + \frac{1}{3} x_{t+1}$$

(23)

The moving average defined in equation (23) is denoted as a 3x1 moving average. The calculation of a moving average in a particular time periods is more complicated when the span of data for the moving average is equal to an even number. Assume the span equals four, the following moving average

$$x_{t-2} + x_{t-1} + x_t + x_{t+1}$$

(24)

is the average between time period $t - 1$ and $t$. Furthermore, a moving average for the next four data points

$$x_{t-1} + x_t + x_{t+1} + x_{t+2}$$

(25)

is the average between time period $t$ and $t + 1$. The moving averages stated in equation (24) and (25) are called a 4x1 moving average. An average of the two 4x1 moving averages defined above would be centered at time period $t$.

$$m_t = \frac{1}{2} \left( \left( \frac{x_{t-2} + x_{t-1} + x_t + x_{t+1}}{4} \right) + \left( \frac{x_{t-1} + x_t + x_{t+1} + x_{t+2}}{4} \right) \right)$$

$$= \frac{1}{8} x_{t-2} + \frac{1}{4} x_{t-1} + \frac{1}{4} x_t + \frac{1}{4} x_{t+1} + \frac{1}{8} x_{t+2}$$

(26)

The average defined in equation (26) is called a 4x2 moving average. This sort of average is applied to estimate the initial trend-cycle. Following the notation in section 4.1, where $Y_t$ is the original time series, the initial trend-cycle is estimated as follows:

$$TC_t^{(1)} = \frac{1}{8} Y_{t-2} + \frac{1}{4} Y_{t-1} + \frac{1}{4} Y_t + \frac{1}{4} Y_{t+1} + \frac{1}{8} Y_{t+2}$$

(27)
The trend-cycle is removed from the original time series by subtracting the estimated trend-cycle component from the original time series.

\[ SI_t^{(1)} = Y_t - TC_t^{(1)} \]  

(28)

where \( SI_t \) is a series consisting of seasonal and irregular movements. The seasonal component, \( S_t \), is estimated by a seasonal moving average. A seasonal moving average is a combination of equally weighted moving averages. In this step a 3x3 moving average is plied to the SI series for each quarter separately to estimate the initial preliminary seasonal component. A 3x3 filter is the average of three consecutive 3x1 moving average. This method implicitly assumes that the irregular component behaves as a stochastic variable that is symmetrically distributed around its expected value and therefore can be eliminated by averaging.

\[ \hat{S}_t^{(1)} = \frac{1}{9} SI_{t-8}^{(1)} + \frac{2}{9} SI_{t-4}^{(1)} + \frac{3}{9} SI_t^{(1)} + \frac{2}{9} SI_{t+4}^{(1)} + \frac{1}{9} SI_{t+8}^{(1)} \]  

(29)

The initial seasonal component is estimated by normalizing the preliminary initial seasonal factor.

\[ S_t^{(1)} = \hat{S}_t^{(1)} - \left( \frac{\hat{S}_{t-2}}{8} + \frac{\hat{S}_{t-1}}{4} + \frac{\hat{S}_t}{4} + \frac{\hat{S}_{t+1}}{4} + \frac{\hat{S}_{t+2}}{8} \right) \]  

(30)

The initial seasonal adjustment series, \( A_t \), is estimated by subtracting the seasonal component defined in equation (30) from the original time series.

\[ A_t^{(1)} = Y_t - S_t^{(1)} \]  

(31)

The calculations described above represent the first step. This procedure is then repeated in step two with a few modifications. In step one the components are estimated by equally weighted moving averages or combination of them. Although, the structure of these filters are often two simple to accurate model the trend-cycle component. It exists other moving averages with more complex patterns of unequal weights. In step two, the trend-cycle component is estimated by the Henderson average. The Henderson filter cannot be applied to series consisting of seasonal movements, therefore this filter is applied to the seasonal adjusted series produced in step one.

\[ TC_t^{(2)} = \sum_{j=-H}^{H} h_j^{2H+1} A_{t+j}^{(1)} \]  

(32)

where \( 2H+1 \) is the length of the filter and \( h_j \) is the Henderson weight. The determination of these weights are explained in appendix C. The length of the filter is determined by the size of the irregular component. In particular, the larger the irregular component the longer span of data is required to estimate the trend-cycle component. Similar as in step one, the trend-cycle component is then subtracted from the original time series.

\[ SI_t^{(2)} = Y_t - TC_t^{(2)} \]  

(33)
In the X-11 the preliminary seasonal component is estimated by a 3x5 seasonal moving average. Although, in the X-12-ARIMA several filters are available to estimate this component. The specification of the moving average is selected based on the criterion provided by Lothian (1984). In this presentation a 3x5 moving average is applied.

\[
\hat{S}_t^{(2)} = \frac{1}{15} S_{t-12}^{(2)} + \frac{2}{15} S_{t-8}^{(1)} + \frac{3}{15} S_{t-4}^{(2)} + \frac{3}{15} S_{t}^{(2)} + \frac{3}{15} S_{t+4}^{(2)} + \frac{2}{15} S_{t+8}^{(2)} + \frac{1}{15} S_{t+12}^{(2)} \tag{34}
\]

In equation (36) below the seasonal component is defined.

\[
S_t^{(2)} = \hat{S}_t^{(2)} - \left( \frac{\hat{S}_{t-2}}{8} + \frac{\hat{S}_{t-1}}{4} + \frac{\hat{S}_{t-1}}{4} + \frac{\hat{S}_{t+2}}{4} + \frac{\hat{S}_{t+2}}{8} \right) \tag{35}
\]

Similar as in the first step one, the seasonal adjustment series is estimated by subtracting the seasonal component defined in equation (35) from the original time series.

\[
A_t^{(2)} = Y_t - S_t^{(2)} \tag{36}
\]

Equations (32) to (36) represent the second step. To further improve the estimation of the trend-cycle component, the Henderson filter is applied to seasonal adjusted series produced in step two. The final trend-cycle is defined as

\[
TC_t^{(3)} = \sum_{j=-H}^{H} h_j^{2H+1} A_{t+j}^{(2)} \tag{37}
\]

where the value of H not necessarily is the same as in step two. The final irregular component is estimated by subtracting the final trend-cycle component from the seasonal adjusted series defined in equation (36).

\[
I_t^{(3)} = A_t^{(2)} - TC_t^{(3)} \tag{38}
\]

The trend-cycle component defined in equation (37) is applied to the HP-filter.
C The Henderson filter

The derivation of the Henderson weights below follows the summary in Findley et al. (1998) of the approach developed by Gray and Thomson (1996). First, assume that the seasonal adjusted series, \( A_{t+j} \), for \(-H + 3 \leq j \leq H\) is specified as

\[
A_{t+j} = \alpha + \beta(t + j) + \gamma(t + j)^2 + I_{t+j}
\]  

(39)

where \( I_{t+j} \) are Gaussian variates with mean 0 and variance \( \sigma^2 \), which, for different \( j \), are independent. In this review only filters that estimate unbiased values of the quadratic trend at time \( t \) are considered. This implies when \( I_{t+j} \) in equation (39) are zero for all \( j \), this must hold

\[
\sum_{j=-H}^{H} h_j A_{t+j} = \alpha + \beta t + \gamma t^2
\]

(40)

for all values of \( \alpha, \beta \) and \( \gamma \). Let \( \Delta \) denote the differencing operator, this imply \( \Delta A_t = A_t - A_{t-1} \) and \( \Delta h_j = h_j - h_{j-1} \). Furthermore let \( E \) denote expectation. Among the filter satisfying equation (40) the Henderson filter is the minimizer of the following smoothness measure

\[
E \left( \Delta^3 \sum_{j=-H}^{H} h_j A_{t+j} \right)^2
\]

(41)

If \( h_j = 0 \) for \( j = \pm(H + 1), \pm(H + 2), \pm(H + 3) \). The smoothness measure in equation (41) can be reduced to a smoother measure of filter coefficients as

\[
E \left( \Delta^3 \sum_{j=-H}^{H} h_j A_{t+j} \right)^2 = \sigma^2 \sum_{j=-H}^{H+3} (\Delta^3 h_j)^2
\]

(42)

With

\[
q_j(H) = \{(H + 1)^2 - j^2\}{(H + 2)^2} - j^2\}{(H + 3)^2} - j^2\}
\]

(43)

and \( a \) and \( b \) being determined by

\[
a \sum_{j=-H}^{H} q_j(H) + b \sum_{j=-H}^{H} q_j(H) j^2 = 1
\]

(44)

\[
a \sum_{j=-H}^{H} q_j(H) j^2 + b \sum_{j=-H}^{H} q_j(H) j^4 = 0
\]

(45)

the Henderson coefficients are given by

\[
h_j^{2H+1} = q_j(H)(a + bj^2), -H \leq j \leq H
\]

(46)

These weights are used in equation (11) to estimate the trend-cycle component.