UNIVERSITY OF CALGARY

The Duration of Bull and Bear Markets in the Dow Jones Industrial Average

by

Kin Pan Benny Leung

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF ARTS

DEPARTMENT OF ECONOMICS

CALGARY, ALBERTA

APRIL 2006

UNIVERSITY OF CALGARY

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled "The Duration of Bull and Bear Markets in the Dow Jones Industrial Average" submitted by Kin Pan Benny Leung in partial fulfilment of the requirements of the degree of Master of Arts.

____
Supervisor, Dr. Apostolos Serletis, Department of Economics

____
Dr. Frank J. Atkins, Department of Economics

____
Dr. L. Bos, Department of Mathematics

April 12, 2006
Date

Abstract

In this thesis, I used a special algorithm to separate the Dow Jones Industrial Average into bull and bear markets. Once the bull and bear markets are gathered, I examined whether they are positively duration dependent using discrete and continuous time tests as well as dynamic hazard functions. We find that bull and bear markets tend exhibit positive duration dependence suggesting mean reversion in stock prices. I also studied the impact of business cycles on the duration of bull and bear markets. In particular, I found that economic expansions tend to encourage bull markets but discourage bear markets. On the other hand, economic recessions tend to encourage bear markets but discourage bull markets.

Acknowledgements

I would like to take this opportunity to express my gratitude and indebtness to my supervisor, Dr. Apostolos Serletis, for providing guidance and encouragement throughout this challenging task. His generosity will never be forgotten. I would also like to thank Dr. Frank Atkins, Dr. John Rowse, Dr. Michael Scott Taylor and Dr. Daniel Gordon for giving me awesome instructions throughout my university career. I would have never made it this far without their teachings.

Great thanks are also given to my good friends and classmates: Arvind Magesan, Graham Thomson and Jason Vaccaro. I will never forget those long nights when we hacked through tedious economic problems. Again, I would have never made it this far without their company.

I also feel a great sense of gratitude towards my family, whom provided endless mental and financial support. I will never forget how you have stood by me regardless of the situation.

Finally, I would like to dedicate this work to God, whom, by grace has forgiven all my wrong doings and allowed me to encounter all the great people mentioned in the above.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

The behavior of stock prices is a popular research area among economic and finance researchers. A common question that is asked by the research community is the validation of the random walk model of stock prices. If the random walk model holds, then predictability of stock prices is impossible. On the other hand, if the random walk model is incorrect, then stock prices may exhibit phenomena such as mean reversion. Although there has been some contributions that test for mean reversion in financial data, many researchers still find it difficult to reject the random walk model using traditional time series methodologies. As such, a purpose of this thesis is to check for mean reversion in stock prices using non-traditional techniques by examining the duration aspects of bull and bear markets in the Dow Jones Industrial Average. In particular, we test for the presence of positive duration dependence (an analog of mean reversion) in the bull and bear markets by using a continuous time test and several discrete time tests. Then, we use a dynamic duration model to model the hazard functions of bull and bear markets from which we can qualitatively uncover the presence of positive duration dependence.

Another common question among the research community is whether economic innovations have an impact on the behavior of stock prices. In fact, many studies have shown that there is strong evidence for switching behavior in stock prices around business cycle turning points. Specifically, stock returns tend to switch from negative to positive around the troughs of business cycles; the opposite switching behavior is observed around the peaks of business cycles. Although switching is traditionally a time series concept, an analog can be developed in terms of duration models. As such, we seek to identify switching behavior in stock prices around business cycle turning points by examining the effects of changes in the economy on the hazards of bull and bear markets.

This thesis is organized as follows. Chapter 2 describes a method for selecting bull and bear markets from a time series of stock prices. Chapter 3 develops continuous time and discrete time duration dependence tests and presents the test results for bull and bear markets. Chapter 4 describes methods for modeling hazard functions in discrete time and presents qualitative results concerning the hazard functions for bull and bear markets estimated by Bayesian means. Finally, chapter 5 concludes the thesis.

Chapter 2

A Simple Framework For Selecting Bull and Bear Markets

2.1 Selection Algorithm

The idea of bear (bull) markets correspond to an extended period of generally decreasing (increasing) stock prices (Chauvet and Petter, 2000). In other words, bear (bull) markets correspond to an extended period at which stock returns are negative (positive). More commonly, the finance community defines a bear (bull) market as a minimum of 20% decrease (increase) in stock prices over an extended period of time. In relation to macroeconomics, bear and bull markets are analogous to busts and booms in business cycles - see King and Plosser (1994), Watson (1994) and Harding and Pagan (2002). Given this analogy, bear markets are simply defined as the movement of stock prices from a local peak to a local trough. Similarly, bull markets are referred to as the movement of stock prices from a local trough to a local peak. Using these definitions, we discuss an algorithm that systematically selects the local peaks and troughs of a time series of stock prices.

As mentioned in the above, bear and bull markets are analogous to busts and booms in business cycles. In fact, Bry and Boschan (BB) (1971) have devised an algorithm for dissecting a monthly GDP time series[1] into appropriate business cycles

[1]It should be noted that the data is smoothed before it is applied to the BB algorithm to avoid problems resulting from outliers.

of different time lengths. In a nutshell, the BB algorithm seeks out local peaks and troughs according to some user defined rules. Specifically, the algorithm splits the series into frames such that each frame contains a full business cycle. After that, the algorithm seeks out the highest (peaks) and lowest (troughs) points within each frame. As simple as the BB algorithm sounds, the algorithm is actually quite difficult to implement in practice.

The first difficulty arises from selecting full cycles by slicing the time series into frames. For the purpose of this thesis, the basic definition of a full cycle is given by a sequence of trough-peak-trough[2]. However, this definition is difficult to implement in real data. First, it is not uncommon in time series data that a general trend is composed of several smaller cycles as depicted in Panel a of Figure 2.1. The problem here is to decide whether the extreme point sequence a-b-c should be treated as an individual cycle or as part of an uptrend defined by a-d. To overcome this issue, a minimum length for a cycle must be defined. In the business cycle context, the BB algorithm sets the minimum cycle length to 15 months. Thus, in reference to Panel a, a-b-c should be regarded as a cycle if it has a duration of at least 15 months. Otherwise, the sequence should be incorporated into the uptrend defined by a-d. Another problem associated with framing is the issue of skipped cycles. This issue arises when a large trend crosses over several frames of 15 months as depicted by Panel b in Figure 2.1. A simple solution to this is to enlarge the frame until a full cycle is fitted. Thus, frame size in the BB algorithm has a lower bound of 15 months but no upper bound.

[2]A full cycle may also be defined by a peak-trough-peak sequence. However, the programming of the algorithm will require minor modification if this definition is used.

In the meantime, the BB algorithm simultaneously determines the local peaks and troughs of the series. In the simplest sense, peaks and troughs are defined to be the highest and lowest points. In addition, an extreme point cannot be classified as a peak unless it is sandwiched between two troughs. Likewise, an extreme point cannot be a trough unless it is sandwiched between two peaks. Again, these definitions are difficult to implement in practice. For example, there are many instances where the time series hits the same high value after a period of fluctuations as depicted by Panel c of Figure 2.1[3]. Such double turns are a problem because whether we select the first or second peak as the local peak has great impact in timing and duration measures. The prescription to this problem depends on how the series behaves between the two peaks. If the movement of the series is mainly descending between the two peaks, then the first peak should be selected as the local peak. On the other hand, if the movement of the series is mainly ascending between the two peaks, then the second peak should be chosen as the local peak. In addition to double turns, step patterns of the time series also present a problem for the identification of peaks and troughs. As an example, consider Panel d of Figure 2.1 where the series maintains a peak and a trough level for several time periods. Notice that the duration of the down trend is dependent on whether R_1 or R_2 is chosen as the local peak and whether V_1 or V_2 is chosen as the local trough. As a rule of thumb the BB algorithm suggests to use the last of the equal values as the turning point such that R_2 and V_2 are selected as the local peak and trough respectively. It follows that the time gap between turning points is also crucial to the selection procedure. If the time length is too short

[3]The high points in panel C are often referred to as the resistance in technical analysis (Murphy, 1999). Similarly, a support is formed if the series continually returns to a same low value.

then the selection algorithm will lead to spurious results due to the erratic nature of economic time series. In the BB algorithm, the minimum time gap between turning points is six months.

While the BB algorithm is designed for dissecting macroeconomic data, the same principle can also be used for financial time series as suggested by Pagan and Sossounov (PS) (2003). However, given the nature of financial time series, some modifications to the BB algorithm are required; we refer to the modified BB algorithm as the PS algorithm. First, since much attention has been given to the behavior of $\Delta \ln P_t$, the PS algorithm examines the natural log of the series rather than the series itself. The second difference between the BB algorithm and the PS algorithm is that the PS algorithm does not require the time series to be smoothed. The reason for this is because data smoothing removes extreme movements (i.e. outliers) that are actually of interest when studying the behavior stock prices. Third, the framing specification of the PS algorithm is slightly different than that of the BB algorithm. In particular, Pagan and Sossounov (2003) suggest to use a full cycle of at least 16 months. In addition, the minimum time gap between two turning points is set to four months according to the Dow Theory (Hamilton, 1919). Of course, the difference between the peak price and the trough price must be at least 20% to satisfy the definition of bear and bull markets. As a note, although the PS algorithm is applied on monthly data in the work by Pagan and Sossounov (2003), it can also be used in data with different frequencies. Since the algorithm is designed to capture broad market movements (Gonzales *et al*, 2005), we do not advise on using data with frequency higher than weekly data.

2.2 Bear and Bull Market Characteristics

Once the turning points of the time series are defined by the PS algorithm, several characteristics of bear and bull markets may be calculated. To compute these characteristics, it is necessary to devise a counting mechanism to separate bear and bull phases in the data. Pagan and Sossunouv (2003) suggest to use a binary variable S_t that is equal to unity when the series is in the bull phase and zero otherwise at time t. With this binary variable, the average duration of bull markets is calculated. In particular, the total time spent in the bull state is given by $\sum_{t=1}^{T} S_t$ and the number of peaks is given by $NTP = \sum_{t=1}^{T-1} (1 - S_{t+1}) S_t$. As such, the average duration of bull markets is simply

$$\widehat{D} = \frac{1}{NTP} \sum_{t=1}^{T} S_t \tag{2.1}$$

Similarly, the average amplitude (i.e. the magnitude of price change) of bull markets is given by

$$\widehat{A} = \frac{1}{NTP} \sum_{t=1}^{T} S_t \Delta \ln P_t \tag{2.2}$$

A more interesting characteristic maybe the cumulative movement of stock prices over the entire bull market. To calculate this, let Z_t be the cumulative sum of $\Delta \ln P_t$ during a bull market such that Z_t resets to zero whenever the market switches to the bear state. Specifically, Z_t is given by

$$Z_t = S_t Z_{t-1} + S_t \Delta \ln P_t \tag{2.3}$$

with $Z_0 = 0$. As such the average cumulative movement over the bull markets is simply

$$\widehat{C} = \frac{TC}{NTP} \tag{2.4}$$

where TC is the total cumulative movement given by

$$TC = \sum_{t=1}^{T} Z_t \tag{2.5}$$

In accordance with Pagan and Harding (2002), it is useful to think of a bull market selected by the PS algorithm as a triangle as depicted by Figure 2.2. With point A being the trough and point C the peak, the linear path $\overline{AB}$ is the cumulative movement of the bull market given by equation (2.3). In reality, $\overline{AB}$ is just an approximation such that it may deviate from the actual path. To measure how well the triangle approximates the actual path, Pagan and Harding (2002) calculate an excess index given by

$$EX_i = \frac{C_{Ti} - C_i + 0.5A_i}{D_i} \tag{2.6}$$

where C_{Ti} is the total cumulative movement of the actual path. If $\overline{AB}$ approximates the actual path well, then the index is close to one. Otherwise, EX_i is close to zero. The average excess index is simply the average of all the EX_i's. A similar logic is used for calculating the average excess index for bear markets.

Many bull markets increase more than the 20% benchmark. To account for these 'strong' bull markets, consider another binary variable $I[a]$ where a is a boolean that defines the 'strong' market. In particular, a is $(1 - S_{t+1}) S_t Z_t > 0.20$ for strong bull markets such that $I[a]$ is unity when a is true and zero otherwise. Consequently, the fraction of strong bull markets is just

$$B^+ = \frac{1}{NTP} \sum_{t=1}^{T-1} I\left[(1 - S_{t+1}) S_t Z_t > 0.20\right] \tag{2.7}$$

Finally, similar statistics are calculated for bear markets by setting S_t to unity during a bear phase and zero otherwise. The fraction of 'strong' bear markets is calculated using the boolean $(1 - S_{t+1}) S_t Z_t < -0.20$ such that

$$B^{-} = \frac{1}{NTP} \sum_{t=1}^{T-1} I\left[(1 - S_{t+1}) S_t Z_t < -0.20\right] \tag{2.8}$$

2.3 Empirical Results: Bull and Bear Markets in the Dow Jones Industrial Average

We use a weekly series of the logarithm of the Dow Jones Industrial Average from January 1928 to May 2005 for analysis. Using the PS algorithm, we dissect the series into its bull and bear components. In particular, we capture absolute changes in the logarithm of stock price of 20% or higher. In addition, the minimum length of any phase must be at least four weeks in length rather than four months. To parallel analysis from previous work such as Cochran and DeFina (1995a), Cochran and DeFina (1995b) and Ohn, Taylor and Pagan (2004), we also separate the series into pre WWII and post WWII subsamples.

Tables 2.1 records the simple summary statistics of bear and bull markets for the three samples. Under the full sample, we identify 59 bear markets and 59 bull markets. Under the post WWII sample, we identify 43 bear markets and 43 bull markets. The pre WWII sample, however, has weak sample size with only 10 bear markets and 9 bull markets. Generally speaking, the mean durations of bear and bull markets are larger than their corresponding standard deviations suggesting the possibility of positive duration dependence under the discrete time duration dependence test tests. On the other hand, the Pre WWII bull markets are suspected of exhibiting negative duration dependence under the discrete time tests because the average duration is less than the standard deviation. Finally, it is observed that bull markets are, on average, longer than bear markets.

A more detail examination of the bull and bear markets will require the examination of the other characteristics described in section 2.2. Table 2.2 and 2.3 record such characteristics for the selected bull and bear markets respectively. First, we observe that not only are bull markets longer than bear markets but they are also stronger as suggested by the mean amplitude and mean cumulated movements. Second, we observe that both bull markets and bear markets deviate quite far away from the triangle approximation as suggested by the mean excess movements. Finally, there is a much larger fraction of the bull markets that exceed the 20% price movement benchmark than bear markets.

2.4 Conclusion

We describe a systematic method for separating a weekly time series of the Dow Jones Industrial Average into bull and bear market components. On average, we find that bull markets are stronger and lengthier than bear markets. In addition, we also find that actual asset price movements deviate quite a bit from the triangular approximated that is used in the selection algorithm. Finally, there is a larger fraction of bull markets with price movements that exceed the 20% benchmark than bear markets.

Figure 2.1

Problems associated with turning point selection.

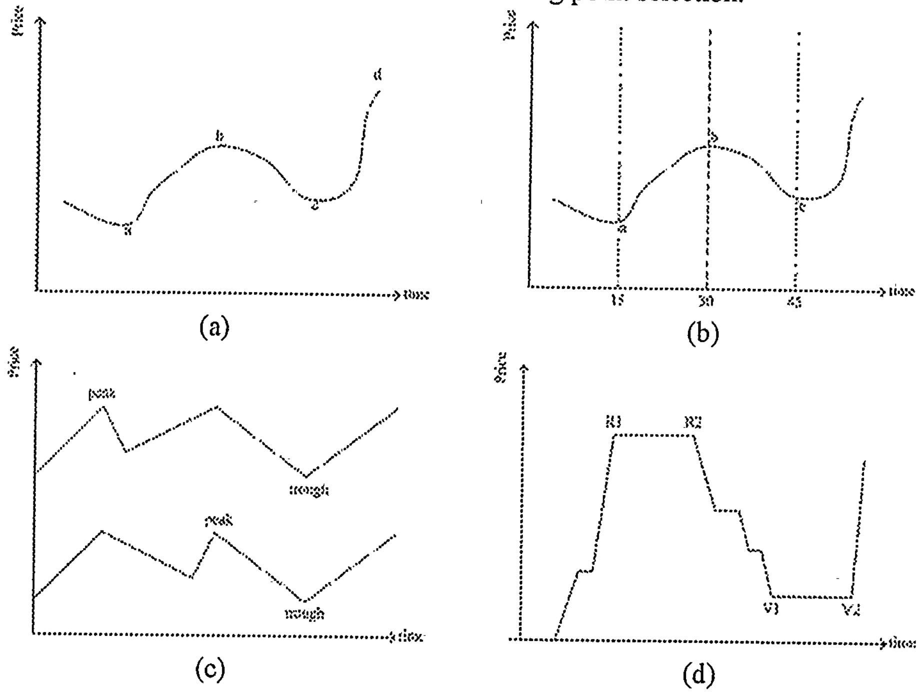
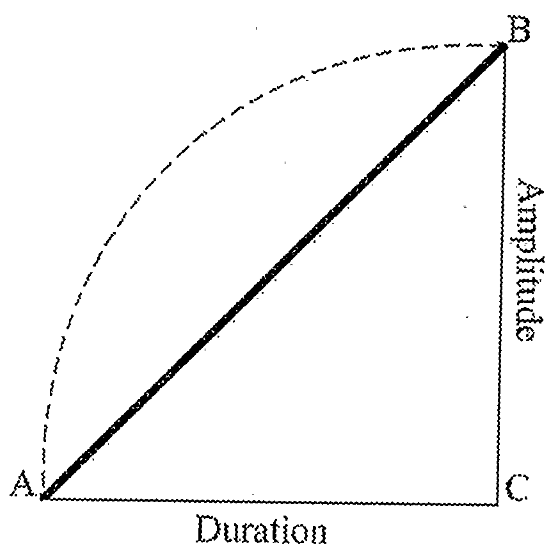


Figure 2.2

An illustration of the triangle approximation for a bull market.



Note: The actual path of stock price movement is indicated by the arch whereas the estimated path is indicated by the solid line.

Table 2.1

Summary Statistics of Bull and Bear Market Durations

Bull Markets	Full Sample	Pre WWII Sample	Post WWII Sample
Mean Duration	28.98	23.67	31.07
Standard Deviation	22.08	25.53	22.65
Max Duration	104	89	104
Min Duration	4	6	4
Sample Size	59	9	43
Bear Markets	**Full Sample**	**Pre WWII Sample**	**Post WWII Sample**
Mean Duration	15.71	17.90	15.14
Standard Deviation	11.32	12.65	11.50
Max Duration	104	41	52
Min Duration	4	4	4
Sample Size	59	10	43

Note: Duration is measured in weeks. Sample size indicates the number of bull markets in upper panel and the number of bear markets in the lower panel

Table 2.2

Characteristics of Bull Markets

	Full Sample	Pre WWII Sample	Post WWII Sample
Mean Duration	28.98	23.67	31.07
Mean Amplitude	0.26	0.37	0.24
Mean Cumulated Movement	5.71	6.48	5.96
Mean Excess Movement	0.012	0.012	0.015
B^+	0.58	0.89	0.53

Table 2.3

Characteristics of Bear Markets

	Full Sample	Pre WWII Sample	Post WWII Sample
Mean Duration	15.71	17.90	15.14
Mean Amplitude	-0.19	-0.39	-0.15
Mean Cumulated Movement	-1.84	-4.23	-1.35
Mean Excess Movement	0.012	0.016	0.008
B^-	0.25	0.60	0.16

Chapter 3

Some Statistical Tests for Duration Dependence

3.1 Motivation

Duration analysis has received much attention in the study of economics. Formally, duration analysis involves the study of the time that an economic agent takes to leave a specific state[1]; such a state is commonly referred to as a spell. For example, duration analysis has been applied to labor economics to study the durations unemployment in different individuals and labor disputes - see Nickel (1979), Lancaster (1979), Kennan (1985) and Kiefer (1988). Other economic applications of duration analysis include consumer choice and marketing (Vilcassim and Jain, 1991), industrial organization (Pakes and Schankerman, 1984) and political economy (King *et al.*, 1990). More recently, duration analysis has found its way into macroeconomics and finance - see Engle and Russell (1998), Partington and Stevenson (2001) and Zuehlke (2003).

An important concept in duration analysis is the idea of duration dependence. A spell is said to be duration dependent if its hazard probability depends on the duration of the spell itself. Duration dependence can be further subdivided into positive and negative duration dependence. Positive duration dependence implies that the hazard probability of a spell increases with the duration of that spell. In other words, mature spells are more likely to end than mature spells under the

[1] In the context of this thesis, the 'agent' is the stock market with bear and bull markets being the states.

influence of positive duration dependence. Similarly, negative duration dependence implies that the hazard probability tends to fall with the duration of the spell such that younger spells are more likely to end than older spells.

The notion of duration dependence is of importance in finance. For example, McQueen and Thorley (1994) argue that speculative bubbles are evident in the stock market if a run of abnormally high (or low) returns is negatively duration dependent. More importantly, duration dependence has strong implications on the behavior of stock returns. Early work by Samuelson (1965) and Leroy (1973) suggest that stock prices follow a random walk such they are unpredictable. However, this view has been challenged by Lo and MacKinlay (1988), Fama and French (1988), Poterba and Summers (1988), Shiller (1989) and Boudohk and Richardson (1994) as they identify mean reverting behavior in stock prices. This in turn suggests stock prices do not necessarily follow the random walk model as prices regularly return to their mean after a period of positive or negative deviations. In the context of duration dependence, mean reversion simply implies positive duration dependence in a run of positive or negative returns (Cochran and Defina, 1995a). The reason for this is simple. In the absence of cyclical behavior, there will be no evidence for duration dependence because trends tend to be consistent. On the other hand, a cyclical component along with the presence of positive duration dependence implies that stock prices will eventually return to their long run trend level, hence mean reverting.

The purpose of this chapter is to identify the presence of mean reversion in stock prices using some statistical tests for duration dependence. In conjunction with the simulation results from chapter 2, the presence of positive duration dependence in bull and bear markets duration will support the rejection of the random walk model. In the next two sections, we present some statistical tests for testing for duration dependence. The first section concerns a continuous time test that is used by Cochran and Defina (1995a) and Cochran and Defina (1995b). The latter section concerns a set of discrete time duration dependence tests that are used in Ohn *et al.* (2004).

3.2 A Continuous Time Duration Dependence Test

A popular assumption in duration analysis is to treat duration time as a continuous random variable. Here, we outline a parametric test for duration dependence in the continuous time framework using the ideas presented in Kiefer (1988) and Davidson and Mackinnon (2004). Let x be a positive continuous random variable that accounts for the duration of a spell (bear or bull markets). Further let $PDF(x)$ and $CDF(x)$ be the probability distribution and cumulative distribution of x respectively. The probability distribution $PDF(x)$ is associated with the survivor function $S(x)$, which justifies the probability that a spell is still present at time x. Mathematically, the survivor function is given by

$$S(x) = 1 - PDF(x) \tag{3.1}$$

Together with $S(x)$, the cumulative distribution $CDF(x)$ defines the continuous time hazard function $h(x)$ where

$$h(x) = \frac{CDF(x)}{S(x)} \tag{3.2}$$

The continuous time hazard function has a specific conditional probabilistic meaning. In particular, it represents the probability that a spell will end given that it has already lasted up to time x. Duration independence is evident when $h'(x) = 0$ whereas duration dependence is evident when $h'(x) \neq 0$. In particular, we have positive duration dependence when $h'(x) > 0$ and negative duration dependence when $h'(x) < 0$.

The distributional assumption on x is crucial to the behavior of the hazard function $h(x)$. Suppose that x is exponentially distributed such that

$$CDF(x) = 1 - e^{-\theta x} \tag{3.3}$$

where $\theta > 0$. Using equation (3.2) the hazard function must be

$$h(x) = \theta \tag{3.4}$$

Taking the first derivative of the above, it is clear that the spell is duration independent because $h'(x) = 0$ regardless of the value of x. The exponential distribution is of crucial importance because it serves as a null hypothesis of no duration dependence for the parametric test.

A simple and versatile alternative to the exponential distribution is the Weibull distribution. The Weibull distribution takes form of

$$CDF(x) = 1 - e^{-(\theta x)^{\alpha}} \tag{3.5}$$

such that the hazard function must be

$$h(x) = \alpha\theta^{\alpha}x^{\alpha-1} \tag{3.6}$$

where $\alpha > 0$. Taking the first derivative of the above yields

$$h'(x) = (\alpha - 1)\,\alpha\theta^{\alpha}x^{\alpha-2} \tag{3.7}$$

When $\alpha = 1$, equation (3.6) collapses to equation (3.4) such that duration independence is evident as the weibull model is identical to the exponential distribution under this condition. On the other hand, negative duration dependence is evident when $\alpha < 1$ because equation (3.7) becomes negative for all x. Similarly, positive duration dependence is evident when $\alpha > 1$ because equation (3.7) becomes positive for all x.

At this point, it is clear that the null hypothesis for the continuous time test is $H_0 : \alpha = 1$ where as the alternative hypotheses are $H_1 : \alpha > 1$ for positive duration dependence and $H_1 : \alpha < 1$ for negative duration dependence. The estimation of the parameter α involves the method of maximum likelihood. Specifically, the loglikelihood function for the Weibull distribution is given by $LL(x, \alpha, \theta) = \sum_{i=1}^{n} \ln h(x, \alpha, \theta) + \sum_{i=1}^{n} \ln S(x, \alpha, \theta)$. Using equation (3.2), the loglikelihood can be further expressed as

$$LL(x, \alpha, \theta) = n \ln \alpha + n \ln \theta + (\alpha - 1) \sum_{i=1}^{n} \ln x_i + \theta^{\alpha} \sum_{i=1}^{n} x_i^{\alpha} \tag{3.8}$$

such that estimates for α and θ can be obtained by maximizing the above in the usual way[2]. Moreover, $\sqrt{n}\left(\widehat{\alpha} - \alpha\right)$ is asymptotically normally distributed with mean zero and variance equal to the second derivative of equation (3.8). As such hypothesis testing and confidence interval construction can be performed in the usual way: $\tau = \left(\widehat{\alpha} - \alpha\right) / se\left(\widehat{\alpha}\right)$ and $\alpha = \widehat{\alpha} \pm \tau_c se\left(\widehat{\alpha}\right)$ where τ is the test statistic and τ_c is the critical value.

[2]The usual way implies finding values of α and θ that maximizes the observation of the data.

3.3 Discrete Time Duration Dependence Tests

3.3.1 Some Preliminaries

The assumption that x is a positive continuous random variable may not be realistic for our purpose. As a matter of fact, duration data collected from a time series should be considered as discrete data because each time period in a time series is measured in fixed intervals (i.e. day, week, month, etc.). Given this, we now define x as a positive discrete random variable such that $x \in \{0, 1, 2, ...\}$.

To construct the discrete time duration dependence tests, we need to define a base case (i.e. a null hypothesis) that always bring about duration independence. A commonly used probability distribution for modeling discrete duration data is the geometric distribution. The geometric distribution takes the form of

$$P(x) = p(1-p)^x \tag{3.9}$$

with moment generating function

$$G(x) = \frac{p}{1-(1-p)e^x} \tag{3.10}$$

The first derivative of $G(x)$ evaluated at $x = 0$ gives the first moment of the geometric distribution such that the population mean is given by

$$\mu_x = \frac{1-p}{p} \tag{3.11}$$

Similarly, the second moment is given by

$$G''(x=0) = 2\left(\frac{1-p}{p}\right)^2 + \left(\frac{1-p}{p}\right) \tag{3.12}$$

such that the corresponding population variance is

$$\sigma_x^2 = \frac{1-p}{p^2} \tag{3.13}$$

Using the work by Ohn *et al.* (2004), the hazard function can be written as

$$h(x) = \frac{P(x = X)}{P(x \leq X)} \tag{3.14}$$

If x is geometrically distributed, then equation (3.11), equation (3.13) and equation (3.14) implies that $h(x) = p$ such that the spell of interest is duration independent. As such, the null hypothesis for the discrete time duration dependence tests is that x is geometrically distributed. On the other hand, deviation from the geometric distribution constitutes the alternative hypothesis of duration dependence. In particular, if the average of x is greater than its standard deviation, then positive duration dependence should be suspected. On the other hand, if the average of x is less than its standard deviation, then negative duration dependence should be suspected (Lancaster, 1990).

With the null and alternative hypotheses defined, we can go about constructing tests for duration dependence in the discrete time framework. We begin our discussion by discussing four weak form tests. These tests are termed 'weak' because they only compare the first two moments of the data that of the geometric distribution. Next, we move on to a strong form test that compares the estimated of the density of the data with the geometric distribution.

3.3.2 Some Weak Form Tests

The Zero Plim Test

The zero plim tests is a discrete time duration dependence test developed by Mudambi and Taylor (1991). First, assume that x is geometrically distributed with mean and variance defined by equation (3.11) and equation (3.13) respectively. For small values of p, it is evident that $\mu_x^2 \approx \sigma_x^2$ such that the population is not expected too over- or underdispersed. As such, for a sufficiently large sample of x we can write

$$\text{plim}\left[\frac{\overline{x}}{s_x} - 1\right] = 0 \tag{3.15}$$

if x is actually geometrically distributed. In other words, we are suggesting that a sufficiently large sample of x will converge (at least in the first two moments) to the geometric distribution if x truly follows the geometric distribution. On the other hand, if x is not geometrically distributed, then equation (3.15) will not hold. With this in mind, we write the test statistic as

$$z_1 = \sqrt{n}\left(\frac{\overline{x}}{s_x} - 1\right) \tag{3.16}$$

where n is the number of observations, $\overline{x}$ is the sample mean and s_x the sample standard deviation. Using the central limit theorem, the distribution of z_1 must converge to the standard normal as n gets infinitely large. With finite samples, however, Mudambi and Taylor (1991) found that the distribution of z_1 is very skewed so that it is necessary to bootstrap the critical values.

Although the zero plim test is classified as a weak form test, it is not weak in terms of statistical power. Rather, the test is expected to be powerful because the population mean and standard deviation will generally be different when duration independence is violated (Lancaster ,1990). For example, we expect negative (positive) duration dependence if z_1 is significantly negative (positive). Similarly, duration independence is identified if z_1 is statistically insignificant. However, as noted by Mudambi and Taylor (1991), the zero plim test is extremely sensitive to hazard functions with erratic behavior as well as those that are monotonically increasing or decreasing. As such, the sign of z_1 must be interpreted with caution.

The Method of Moments Test

An analog to the zero plim test is the method of moment test (Mudambi and Taylor, 1991). According to the method of moments methodology, the k^{th} moment of the population is equivalent to the k^{th} moment of the sample. As such, large deviations between the hypothesized population moments and sample moments suggest that the sample violates the hypothesized assumptions on the population.

Using the sample mean, it is easy to see that a maximum likelihood estimator for p is given by

$$\widehat{p} = \frac{1}{1+\overline{x}} \tag{3.17}$$

Similarly, a maximum likelihood estimator for p can also be derived from the sample variance such that

$$\widehat{p} = \frac{-1+\sqrt{1+4s_x^2}}{2s_x^2} \tag{3.18}$$

Since $\bar{x}$ and s_x^2 are considered consistent estimators for the population mean and variance, equation (3.17) and equation (3.18) will both converge in probability to p under the null hypothesis of duration independence. To construct the method of moments test, we rearrange equation (3.17) so that $(1-\widehat{p})/\widehat{p}^2$ is an unbiased and consistent estimator for $\overline{x}$. Under the null hypothesis of the geometric distribution, $(1-\widehat{p})/\widehat{p}^2$ is a consistent estimator for the population mean μ_x for a sufficiently large sample of x. Then, conditional on s_x, the test statistic can be written as

$$z_2 = \frac{\sqrt{n}\left[\overline{x} - \left(\frac{1-\widehat{p}}{\widehat{p}^2}\right)\right]}{s_x} \tag{3.19}$$

which, according to the central limit theorem, is approximately normal under the null hypothesis. However, z_2 is highly skewed under finite samples so the critical values should be bootstrapped. The method of moment test is expected to be powerful because deviation from the null hypothesis will overthrow the consistency of $\widehat{p}$. Finally, it should be noted that z_1 and z_2 will converge numerically to each other for s_y that are much larger than one. As such, it is expected that the two test statistics should have similar behavior.

The Generalized Method of Moments Test

The generalized method of moments (GMM) test is another analog to the zero plim and method of moments tests (Mudambi and Taylor, 1995). This test is assumed to be superior to the zero plim and method of moments tests because it operates unconditionally on the sample variance.

To derive the GMM test, we employ the GMM method proposed by Tauchen (1985). First, we compute $\widehat{p}$ via maximum likelihood and use it to calculate the score function $\widehat{\eta}_i = \eta(x_i, \widehat{p})$. In accordance with Mudambi and Taylor (1992), the loglikelihood function for the geometric distribution is

$$LL = n\ln(p) + \ln(1-p)\sum_{i=1}^{n} x_i \tag{3.20}$$

while the score function is

$$\widehat{\eta}_i = \frac{1}{\widehat{p}} - \frac{x_i}{1-\widehat{p}} = (\overline{x} - x_i)\left(1 + \frac{1}{\overline{x}}\right) \tag{3.21}$$

Next we select an auxiliary criterion function $c(x,p)$ such that a large absolute value of $(1/n)\sum c(x_i,\widehat{p})$ will lead to speculation on whether the underlying loglikelihood model is correct. In particular, the auxiliary criterion function takes the form of

$$c(x,p) = \left[x - \frac{1-p}{p}\right]^2 - \frac{1-p}{p^2} \tag{3.22}$$

At this point, the general idea behind the GMM test is clear. Under the geometric distribution, equation (3.22) has an expected value of zero so that we are implicitly checking whether the sample variance is the same as the population variance implied by the geometric distribution. To check this, we regress

$$\widehat{c}_i = b_0 + b_1\widehat{\eta}_i + disturbance \tag{3.23}$$

and confirm whether the intercept term is significant. The estimate for b_0 is just

$$\widehat{b}_0 = \frac{1}{n}\sum_{i=1}^{n} c(y_i,\widehat{p}) = \left[\left(\frac{1}{n}\right)\sum_{i=1}^{n}(y_i - \overline{y})^2\right] - \overline{y}^2 - \overline{y}$$

so that the GMM test statistic is

$$z_3 = \left[\left(\frac{1}{n}\right)\sum_{i=1}^{n}(y_i - \overline{y})^2\right] - \overline{y}^2 - \overline{y} \tag{3.24}$$

Again, the central limit theorem suggests that z_3 converges in distribution to the standard normal but the critical values should be bootstrapped when dealing with finite samples.

The SB Test

An alternative to the duration dependence tests in the above is the SB regression test proposed by Ohn *et al* (2004). The test is termed SB because it is constructed using binary variables that describe the state. Lets begin by analyzing bull markets by considering the binary variable S_t that accounts for the state of the stock market at time t. In particular, S_t assumes unity when during a bull market and zero during a bear market[3]. Under the assumption of a constant hazard function, S_t must follow a Markov process such that it can be written as an AR(1) as suggested by Hamilton (1989). Specifically, the AR(1) process is given by

$$S_t = c_0 + c_1 S_{t-1} + disturbance \tag{3.25}$$

where $c_0 = p_{1|0} = 1 - p_{0|1}$ and $c_1 = p_{1|1} + p_{0|0} - 1$[4]. Under duration dependence, the market state at t will depend on the market state at $t-1$ as well as the duration of that state. To investigate this, we modify the above to

$$S_t = c_0 + c_1 S_{t-1} + c_2 S_{t-1} d_{t-1} + disturbance \tag{3.26}$$

where d_t is the number of consecutive periods spent in a particular state[5]. As such duration dependence implies that c_2 is significantly different from zero. That is, testing the null hypothesis $H_0 : c_2 = 0$ is equivalent to testing duration independence.

However, the SB test must be constructed with caution. In particular, the test cannot be constructed using a sample consisting of bear and bull markets because

[3]Duration dependence in bear markets may be explored by setting S_t to one when the state is a bear market and zero otherwise.

[4]The notation $p_{a|b}$ denotes the conditional probability of the stock market switching from state a to state b.

[5]Durland and McCurdy (1994) suggest that the relationship between S_t and d_{t-1} is nonlinear. However, our linear model is still suitable becasue we are just testing whether or not there is a relationship between the two variables.

doing so will introduce conditional hetroskedasticity in the disturbance terms[6]. To overcome this, bear and bull markets must be treated separately. Consider, for example, the sequence $S_t = \{1, 1, 1, 0, 0, 1, 1, 1, 0, 0, 1, 1, 0\}$ such that periods 1-3, 6-8 and 11-12 are bull markets. A sample for S_t, S_{t-1} and d_{t-1} consisting of bear and bull markets is

S_t	S_{t-1}	d_{t-1}
1	-	0
1	1	1
1	1	2
0	1	3
0	0	0
1	0	0
1	1	1
1	1	2
0	1	3
0	0	0
1	0	0
1	1	1
0	1	2

[6]It follows that $V\left(disturbance|S_{t-1} = 0\right) = p_{1|0}\left(1 - p_{1|0}\right)$ and $V\left(disturbance|S_{t-1} = 1\right) = p_{1|1}\left(1 - p_{1|1}\right)$.

whereas a sample for bull markets only is given by

S_t	S_{t-1}	d_{t-1}
1	1	1
1	1	2
0	1	3
1	1	1
1	1	2
0	1	3
1	1	1
0	1	2

Under the bull-markets-only sample, we see that S_{t-1} is a column of ones such that equation (3.26) can be rewritten as

$$S_t = c_0 + c_2 d_{t-1} \tag{3.27}$$

Here the test statistic is just the standard t test statistic for c_2 so that positive duration dependence in bull markets is evident when c_2 is significantly positive whereas negative duration dependence in bull markets is implied if c_2 is significantly negative. As a reminder, the relationship between S_t and d_{t-1} may be nonlinear as suggested by Durland and McCurdy (1994) so that the power of the SB test should be questioned. Finally, it should be noted that the SB test is asymptotically equivalent to the GMM test. An elegant proof for this claim is found in Ohn *et al* (2004).

3.3.3 A Strong Form Test

As discussed previously, a strong form test compares the estimated density of the data with the density hypothesized distribution rather than comparing only the first two moments. A good candidate is the chi-square goodness of fit test (chi-square test for short) that is used by Diebold and Rudebusch (1991) to test for duration dependence in business cycles. The general idea behind the chi-square test is to divide the sample into K bins and compare the observed frequencies with the expected frequencies generated by the hypothesized distribution for each bin. If the hypothesized distribution is true, then the observed and expected frequencies will be very close to each other. Of course, the hypothesized distribution for our purpose is the geometric distribution. Formally, the test statistic is given by

$$\chi^2 = \sum_{i=1}^{K} \frac{(O_i - E_i)^2}{E_i}$$

where O_i and E_i are the observed and expected frequencies of the i^{th} bin.

Bin selection is important in order to obtain satisfactory results from the chi-square test. Hoel (1954) suggests bins should be selected such that the expected frequency should be at least five for each bins. To be on the safe side, we follow the suggestion by Ohn *et al* (2004) and set expected frequency to at least six. E_i is organized such that E_1 corresponds to the bin with the lowest values for the range of realizations of x whereas E_K corresponds to the bin with the highest values for the range of realizations of x. Further, the last bin must be defined with care. The bins corresponding to $E_1, E_2, ..., E_{K-1}$ are constructed with expected frequencies that are closest to six from the right. Since, extreme durations are scarce, it is often the case that E_K is less than five. In such an event, we combine bins K and $K-1$ together to construct the last bin for the chi square test.

3.4 Empirical Results: Evidence for Duration Dependence in Bull and Bear Markets

Using the selected bear and bull markets from chapter 2, we test whether they exhibit duration dependence by applying continuous time and discrete time tests. Since there is some uncertainty surrounding the exact timing of the turning points, we adjust the duration data as suggested by McCulloch (1975) and Diebold and Rudebusch (1990). Namely, we remove different positive values of τ_0, which is at most the historical minimum, from the duration data obtained from the selection algorithm.

We begin the analysis with the continuous time parametric Weibull test on bull and bear markets. Under this test, we check for positive duration dependence, negative duration dependence and duration independence by checking whether the test statistic is greater than one, less than one or equal to one respectively. Table 3.1 summarizes the test results for the three samples of bear markets. It is clear that there is some evidence concerning positive duration dependence in bear markets for the full and post WWII samples as some of the test statistics are statistically greater than unity. However, the evidence is only present for τ_0 less than two and then disappears thereafter. The bear markets in the pre WWII sample, however, does not seem to exhibit duration dependence for all values of τ_0. This finding is quite different from Cochran and DeFina (1995a) who uncovered some evidence for positive duration dependence in the pre WWII sample. Table 3.2 presents the continuous time parametric Weibull test for bull markets. Here, we find strong evidence concerning positive duration dependence in bull markets for the full sample as well as the post WWII sample; the test statistics are greater than one for all values of τ_0. On the other hand, the bull markets for the pre WWII sample does not show any evidence concerning duration dependence. Again our finding is different from Cochrane and DeFina (1995a) who failed to identify positive duration dependence in bull markets except in the pre WWII sample. An explanation for the discrepancies is that our data set is different from Cochran and DeFina's (1995a) study. In particular, Cochran and Defina's pre WWII sample is larger than our pre WWII sample. Furthermore, a different selection algorithm[7] is used in Cochran and DeFina (1995a), which is another possible source for discrepancy. It should also be noted that our test results for bear and bull markets in the pre WWII sample should be questioned. Specifically, the small sample size in the pre WWII sample decreases the power of the continuous time parametric Weibull test.

Next, we venture into the discrete time realm to test for duration dependence in bear and bull markets. Given the summary statistics in table 1 and table 2, we specifically test for negative duration dependence in bull markets for the pre WWII sample and positive duration dependence for the other bull market samples. First, consider the weak form tests. Table 3.3 summarizes the zero plim test results for bear markets. Under the full sample, we find evidence for positive duration dependence for $\tau_0 = \{1, 2\}$. However, the statistics show no evidence for positive duration dependence for the pre WWII and post WWII subsamples. Table 3.4

[7]Cochran and DeFina (1995b) used the method proposed by Cohen *et al.* (1987) to select bear and bull markets. It is hard to say whether the PS algorithm is superior to the Cohen *et al.* (1987) method; an indepth comparison between the two methods are required for a conclusive answer.

summarizes the zero plim test results for bull markets. This test shows positive duration dependence for the full sample and post WWII sample when $\tau_0 = \{1, 2\}$. The bull markets in pre WWII sample, however, does not appear to be duration dependent as suggested by the zero plim test.

The method of moment test results for bear markets are recorded in table 3.5. Like the zero plim test results for bear markets, we fail to uncover evidence for duration dependence in the pre WWII and post WWII subsamples. However, we do observe some presence of positive duration dependence in the full sample for $\tau_0 = \{1, 2\}$. The method of moment test results for bull markets are recorded in table 3.6 from which we see some evidence of duration dependence. In particular, the method of moment test identified positive duration dependence in the full sample when $\tau_0 = 1$. The test has also detected evidence of positive duration dependence in the post WWII sample for $\tau_0 = \{1, 2\}$. Further, the bull markets in the pre WWII sample seems to exhibit negative duration dependence for $\tau_0 = \{5, 6\}$.

The generalized method of moments test results for bear and bull markets are recorded in table 3.7 and table 3.8 respectively. Like the results from the zero plim test and method of moment for bear markets, we fail to uncover evidence for duration dependence in the pre WWII and post WWII subsamples. However, there is evidence for positive duration dependence in the full sample when τ_0 is equal to one. There is also evidence for positive duration dependence in the post WWII sample when τ_0 is one as well as in the full sample for $\tau_0 = \{1, 2\}$. The pre WWII sample, on the other hand, seems to exhibit negative duration dependence when τ_0 is six.

Table 3.9 summarizes the SB test results for bear markets. There is no evidence for positive duration dependence in bear markets except in the full sample bear markets when $\tau_0 = 1$. Table 3.10 summarizes the SB test results for bull markets. In the full sample and post WWII sample, we uncover evidence for positive duration dependence in bull markets for $\tau_0 = \{1, 2\}$. The pre WWII sample, however, does not appear to show any duration dependence under the SB test.

To complete the analysis using discrete time duration dependence tests, we now move on to the strong form Chi square goodness of fit test. Regrettably, the pre WWII sample is too small to construct meaningful bins for the testing procedure. As such, we discard the pre WWII sample for the Chi square goodness of fit test. Table 3.11 summarizes the Chi square goodness of fit test results for the full sample bear market data. By inspection, we see considerable clustering of observed frequencies for $\tau_0 = \{1, 2, 3\}$, which in turn suggests deviation from the geometric distribution. Consider the case where τ_0 is equal to one. We see very low observed frequencies (i.e. 0 and 2) in bins $[0, 1]$ and $[2, 3]$ whereas high observed frequencies (i.e. 15, 12 and 11) are evident in bins $[4, 6]$, $[7, 9]$ and $[19, 25]$. In comparison, these observed frequencies are quite different from their corresponding expected frequencies. Since the expected frequencies are generated from the geometric distribution, the deviations makes it clear that the full sample bear markets are duration dependent when τ_0 is equal to one. A similar behavior is also observed in cases where $\tau_0 = 2$ and $\tau_0 = 3$. In fact, the Chi square test statistics convey that the null hypothesis of duration independence is rejected such that positive duration dependence is evidence for $\tau_0 = \{1, 2, 3\}$. On the contrary, for the case where τ_0 is four, the observed frequencies seems to be very close the expected frequencies suggesting that positive duration dependence is not very likely. In fact, as the Chi square test statistic suggests, there is no statistical evidence for positive duration dependence in full sample bear markets when τ_0 is four. Using a similar argument, we observe, from Table 3.12, positive duration dependence in full sample bull markets when τ_0 is set at 1 and 2 but not when τ_0 is set at 3 and 4. Table 3.13 and Table 3.14 concern the post WWII bear and bull markets respectively. There is evidence for positive duration dependence in the post WWII bear markets for all values of τ_0 under consideration. On the other hand, we have evidence for positive duration dependence for post WWII bull markets only when τ_0 is equal three.

In comparison to previous work, our results are somewhat different from the findings of Ohn, Taylor and Pagan (2004), who, generally speaking, found strong evidence for positive duration dependence in bull and bear markets. However, the evidence duration dependence for pre WWII bull markets is still weak in their results. A logical explanation for the discrepancies is that different data are used in the analysis. Namely, the Ohn, Taylor and Pagan (2004) study used a monthly S&P500 data set whereas the weekly Dow Jones Industrial Average is used in this study. As a note of caution, the danger of the small pre WWII sample remains problematic in the discrete tests since the small sample may decrease statistical power of our tests.

3.5 Conclusion

In summary, we find some evidence of positive duration dependence in bear and bull markets for the full sample as well as the post WWII sample using continuous and discrete time tests. However, there is no evidence for duration dependence in the pre WWII bear markets whereas there is some evidence for negative duration dependence in the pre WWII bull markets. As a remark, the results for the pre WWII samples should be questioned due to their relatively sample size. In addition the results are sensitive to τ_0, which is related to the uncertainty of turning point identification. Further research is required to find out why that is the case.

Table 3.1

Continuous Time Weibull Parametric Test for Bear Markets

Minimum phase	Full sample	Pre WWII sample	Post WWII sample
1	1.41814**	1.49323	1.36164**
2	1.30373**	1.38065	1.24966*
3	1.17355	1.24442	1.12285
4	1.08924	1.42294	1.01033

Note: **p-value* < *0.10*; ** *p-value* < *0.05*; ****p-value* < *0.01*

Table 3.2

Continuous Time Weibull Parametric Test for Bull Markets

Minimum phase	Full sample	Pre WWII sample	Post WWII sample
1	1.37627***	1.13875	1.41279**
2	1.31262**	1.08697	1.34618**
3	1.24009**	1.03099	1.26838**
4	1.20828*	0.96813	1.25765*
5	-	0.89061	-
6	-	0.95500	-

Note: **p-value* < *0.10*; ** *p-value* < *0.05*; ****p-value* < *0.01*

Table 3.3

Discrete Time Zero Probability Limit Test for Bear Markets

Minimum phase	Full sample	Pre WWII sample	Post WWII sample
1	2.3570***	1.0614	1.5061
2	1.6268**	0.8115	0.9358
3	0.9480	0.5616	0.3655
4	0.2692	0.3116	-0.2048

Note: *p-*value* < 0.10; ** p-*value* < 0.05; ***p-*value* < 0.01

Table 3.4

Discrete Time Zero Probability Limit Test for Bull Markets

Minimum phase	Full sample	Pre WWII sample	Post WWII sample
1	2.0545**	-0.3364	2.1485**
2	1.7066*	-0.4539	1.8591*
3	1.3587	-0.5714	1.5695
4	1.1080	-0.6889	1.2800
5	-	-0.8065	-
6	-	-0.9240*	-

Note: *p-*value* < 0.10; ** p-*value* < 0.05; ***p-*value* < 0.01

Table 3.5

Discrete Time Method of Moment Test for Bear Markets

Minimum phase	Full sample	Pre WWII sample	Post WWII sample
1	2.6367**	1.1839	1.7850
2	1.9588*	0.9340	1.2147
3	0.2799	0.6841	0.6444
4	0.6011	0.4341	0.0742

Note: *p-value < 0.10; ** p-value < 0.05; ***p-value < 0.01

Table 3.6

Discrete Time Method of Moment Test for Bull Markets

Minimum phase	Full sample	Pre WWII sample	Post WWII sample
1	2.2265**	-0.2782	2.2917*
2	1.8786	-0.3957	2.0022*
3	1.5307	-0.5132	1.7127
4	1.1827	-0.6308	1.4231
5	-	-0.7483*	-
6	-	-0.8658*	-

Note: *p-value < 0.10; ** p-value < 0.05; ***p-value < 0.01

Table 3.7

Discrete Time Generalized Method of Moment Test for Bear Markets

Minimum phase	Full sample	Pre WWII sample	Post WWII sample
1	-5.2130**	-3.9682	-3.3563
2	-3.7562	-3.1216	-2.2387
3	-2.3938	-2.3250	-1.2001
4	-1.1395	-1.5786	-0.2405

Note: **p-value* < 0.10; ** *p-value* < 0.05; ****p-value* < 0.01

Table 3.8

Discrete Time Generalized Method of Moment test for Bull Markets

Minimum phase	Full sample	Pre WWII sample	Post WWII sample
1	-3.9987*	0.2490	-4.2059*
2	-3.3243*	0.5121	-3.6251
3	-2.6741	0.7637	-3.0602
4	-2.0479	1.0036	-2.5147
5	-	1.2320	-
6	-	1.4487*	-

Note: **p-value* < 0.10; ** *p-value* < 0.05; ****p-value* < 0.01

Table 3.9

Discrete Time SB Test for Bear Markets

Minimum phase	Full sample	Pre WWII sample	Post WWII sample
1	0.00166**	0.00196	0.00127
2	0.00125	0.0017	0.00080
3	0.00070	0.00135	0.00019
4	0.00033	0.00178	0.00038

Note: **p-value* $<$ *0.10;* ** *p-value* $<$ *0.05;* ****p-value* $<$ *0.01*

Table 3.10

Discrete Time SB Test for Bull Markets

Minimum phase	Full sample	Pre WWII sample	Post WWII sample
1	0.00042**	0.00014	0.00050**
2	0.00037*	0.00023	-0.00045*
3	0.00031	0.00034	-0.00040
4	0.00027	0.00045	-0.00040
5	-	0.00059	-
6	-	0.00049	-

Note: **p-value* $<$ *0.10;* ** *p-value* $<$ *0.05;* ****p-value* $<$ *0.01*

Table 3.11

Goodness of Fit Test for Full Sample Bear Markets.

$\tau_0 = 1$			$\tau_0 = 2$		
Interval	E	O	Interval	E	O
$[0, 1]$	7.27124	0	$[0, 1]$	7.74815	0
$[2, 3]$	6.37512	2	$[2, 3]$	6.73062	8
$[4, 6]$	8.12028	15	$[4, 6]$	8.47553	12
$[7, 9]$	6.66639	12	$[7, 9]$	6.86204	10
$[10, 13]$	7.06996	5	$[10, 13]$	7.16176	5
$[14, 18]$	6.58432	7	$[14, 18]$	6.53378	7
$[19, 25]$	6.23953	11	$[19, 25]$	6.02575	11
> 25	10.67313	7	> 25	9.46235	6
	$\chi^2(6) = 25.89817$	(0.00023)		$\chi^2(6) = 16.94706$	(0.00948)
$\tau_0 = 3$			$\tau_0 = 4$		
Interval	E	O	Interval	E	O
$[0, 1]$	8.29188	2	$[0, 1]$	8.91755	8
$[2, 3]$	7.12654	8	$[2, 3]$	7.56971	9
$[4, 5]$	6.12497	10	$[4, 5]$	6.42559	7
$[6, 8]$	7.61195	10	$[6, 8]$	7.86417	7
$[9, 11]$	6.06504	4	$[9, 11]$	6.15040	4
$[12, 15]$	6.21431	5	$[12, 15]$	6.16808	6
> 15	16.28423	20	$[16, 21]$	6.17657	10
	-	-	> 21	9.72786	8
	$\chi^2(5) = 9.87037$	(0.07899)		$\chi^2(6) = 3.94110$	(0.68465)

Table 3.12

Goodness of Fit Test for Full Sample Bull Markets

$\tau_0 = 1$			$\tau_0 = 2$		
Interval	E	O	Interval	E	O
$[0, 3]$	7.73088	1	$[0, 2]$	6.10191	1
$[4, 7]$	6.71789	5	$[3, 6]$	7.16569	5
$[8, 12]$	7.17336	6	$[7, 10]$	6.19501	5
$[13, 17]$	6.01835	12	$[11, 15]$	6.57733	13
$[18, 24]$	6.83355	11	$[16, 21]$	6.46506	9
$[25, 32]$	6.00627	5	$[22, 29]$	6.69196	7
$[33, 44]$	6.36776	9	$[30, 39]$	6.04071	7
$[45, 64]$	6.13099	5	$[40, 55]$	6.07407	6
> 64	6.02083	5	> 55	7.68813	6
	$\chi^2(7) = 16.61533$	(0.02005)		$\chi^2(7) = 12.95453$	(0.07322)

$\tau_0 = 3$			$\tau_0 = 4$		
Interval	E	O	Interval	E	O
$[0, 2]$	6.31957	3	$[0, 2]$	6.55323	4
$[3, 6]$	7.38591	5	$[3, 6]$	7.61972	6
$[7, 10]$	6.35039	4	$[7, 10]$	6.51269	5
$[11, 15]$	6.70098	12	$[11, 15]$	6.82684	10
$[16, 21]$	6.53730	9	$[16, 21]$	6.60635	10
$[22, 29]$	6.70274	7	$[22, 29]$	6.70452	5
$[30, 40]$	6.45973	7	$[30, 40]$	6.37280	8
$[41, 58]$	6.18717	6	> 40	11.80367	11
> 58	6.35606	6		-	-
	$\chi^2(7) = 8.58646$	(0.28373)		$\chi^2(6) = 5.81223$	(0.44455)

Table 3.13

Goodness of Fit Test for Post War Bear Markets

$\tau_0 = 1$			$\tau_0 = 2$		
Interval	E	O	Interval	E	O
$[0, 2]$	7.97032	0	$[0, 2]$	8.49333	1
$[3, 5]$	6.49297	9	$[3, 5]$	6.81573	13
$[6, 9]$	6.82499	15	$[6, 9]$	7.04108	10
$[10, 14]$	6.28381	2	$[10, 14]$	6.33980	4
$[15, 22]$	6.49716	9	$[15, 22]$	6.35212	9
> 22	8.93075	8	> 22	7.95784	6
	$\chi^2(4) = 22.71191$	(0.00015)		$\chi^2(4) = 22.71191$	(0.00313)
$\tau_0 = 3$			$\tau_0 = 4$		
Interval	E	O	Interval	E	O
$[0, 1]$	6.29607	1	$[0, 1]$	8.06341	7
$[2, 4]$	7.75859	13	$[2, 3]$	6.55135	7
$[5, 7]$	6.11855	10	$[4, 6]$	7.59734	10
$[8, 11]$	6.19523	2	$[7, 10]$	7.06535	2
$[12, 17]$	6.28811	5	$[11, 16]$	6.36267	5
> 17	10.34338	12	> 16	7.35986	12
	$\chi^2(4) = 13.82817$	(0.00786)		$\chi^2(4) = 7.77959$	(0.10000)

Table 3.14

Goodness of Fit Test for Post War Bull Markets

$\tau_0 = 1$			$\tau_0 = 2$		
Interval	E	O	Interval	E	O
$[0, 4]$	6.48857	3	$[0, 4]$	6.69003	3
$[5, 10]$	6.50728	3	$[5, 10]$	6.66884	4
$[11, 17]$	6.14111	10	$[11, 17]$	6.24871	10
$[18, 26]$	6.08621	8	$[18, 26]$	6.13887	7
$[27, 39]$	6.15830	8	$[27, 39]$	6.13808	9
< 39	11.61844	11	< 39	11.11531	10
	$\chi^2(4) = 7.37627$	(0.11729)		$\chi^2(4) = 6.92248$	(0.14004)
$t_0 = 3$			$\tau_0 = 4$		
Interval	E	O	Interval	E	O
$[0, 4]$	6.90436	4	$[0, 4]$	7.13284	4
$[5, 10]$	6.83807	3	$[5, 10]$	7.01547	5
$[11, 17]$	6.35844	11	$[11, 17]$	6.47008	10
$[18, 26]$	6.18844	6	$[18, 26]$	6.23420	6
$[27, 39]$	6.10940	9	$[27, 39]$	6.07120	8
> 39	10.60120	10	> 39	10.07608	10
	$\chi^2(4) = 8.17172$	(0.08549)		$\chi^2(4) = 4.50299$	(0.34219)

Chapter 4

Modelling the Hazard Function in the Discrete Time Framework

4.1 The Basic Discrete Time Hazard Model

As a complement to the duration dependence tests, it is useful to model the hazard functions of bear and bull markets to get a visual confirmation on how the hazard probabilities behave over time. Furthermore, modeling the hazard functions will also enable us to determine how hazard probabilities are affected by exogenous variables. Research on duration modeling is performed using a continuous time framework where time is assumed to be observed continuously - see Lawless (1982), Kalbfleisch and Prentice (1980) and Lancaster (1990) for detailed outlines on such modeling techniques. However, as noted previously, the continuous time assumption is not valid for our application where time is measured is discrete intervals such as days, weeks, months, etc. In this chapter, we will discuss conventional discrete time techniques as outlined by Allison (1982) and Singer and Willet (2003).

Consider $t = 1, 2, 3, \ldots$ time periods and $i = 1, 2, \ldots, n$ observed bull markets beginning at $t = 1$. Further, let t_i be the time period when the i^{th} bull market terminates. In addition, let x_i be a vector of k covariates that influence the duration of the i^{th} bull market. Now the hazard probability of the i^{th} bull market is the conditional probability that it will not be observed at $t+1$ given that it has already lasted t periods. A set of these hazard probabilities as a function of time is defined as the hazard function. Formally, it is written as

$$h_{it} = P(T_i = t | T_i \geq t, x_i) \tag{4.1}$$

such that the hazard function is a mapping of real values into the probability space. The behavior of this map is of crucial importance. For example, if the hazard function increases with time, then the risk of a bull market terminating rises with respect to time suggesting positive duration dependence. If, on the other hand, the hazard function is flat with respect to time then there is evidence for zero duration dependence. Lastly, there is evidence for negative duration dependence if the hazard function is a decreasing function of time, which in turn suggests that longer bull markets are at lower risks of termination.

At first glance, one may be tempted to use a linear probability model to model the hazard function. Unfortunately, h_{it} is not bounded in the linear probability model. Using the linear probability model will lead to logically impossible results such as $h_{it} > 1$ and $h_{it} < 0$. However, this is not to say that a linear setup should be abandoned. Rather, we need to transform h_{it} such that a linear model of the transformed h_{it} will lead to plausible results. One such transform is the logit transform - see Cox (1972), Brown (1975) and Thompson (1977)[1]. Mathematically, the logit transform is written as

$$\text{logit} h_{it} = \ln\left(\frac{h_{it}}{1 - h_{it}}\right) \tag{4.2}$$

[1]Cox (1972) used the logit link in the continuous time context. Brown (1975) and Thompson (1977) used the logit link for discrete time modeling. Nevertheless, the discrete time model converges to its continuous conterpart for sufficiently small time intervals (Thompson, 1977).

where $\frac{h_{it}}{1-h_{it}}$ is the odds ratio. To model logith_{it} as a function of time and covariates, we investigate the logit model

$$\text{logit}h_{it} = \gamma_{0,t} + \gamma_1' x_i \tag{4.3}$$

from which we can obtain the hazard function using the logit link

$$(\text{logit}h_{it}) = h_{it} = \frac{1}{1 + \exp\left(-\gamma_{0,t} - \gamma_1' x_i\right)} \tag{4.4}$$

The advantage of using the logit link is that it results in plausible hazard probabilities even when logith_{it} is not bounded. For extremely negative logith_{it} values, equation (4.4) still leads to positive probabilities because $\lim_{\text{logit} \to -\infty} h_{it} = 0$. On the other hand, extremely large values of logith_{it} will lead to hazard probabilities that are very close to one because $\lim_{\text{logit} \to +\infty} h_{it} = 1$.

The model specified by equation (4.3) can be broken down into two components. First, $\gamma_{0,t}$ is a sequence of parameters that defines the baseline logit hazard function. In fact, each $\gamma_{0,t}$ represents the logit hazard at each time period t. For example, $\gamma_{0,1}$ is the logit hazard at period 1, $\gamma_{0,2}$ is that logit hazard at period 2 and so forth. Note that $\gamma_{0,t}$ can take on any real values such that the functional form of the baseline logit hazard function is not explicitly defined under this specification. The purpose of this setup is to add flexibility to the model by allowing the data to "speak for itself" such that the true shape of the hazard function can be revealed.

To better illustrate this, we adopt the expanded form of the model that is described in Singer and Willet (2003) where the subscript i is dropped for simplicity. Let D_t be the time indicator dummy variable that takes on unity in the time period it represents and zero elsewhere. Further, let t range from one to J such that we can write the expanded hazard as

$$\text{logit}h_t = \left[\gamma_{0.1}D_1 + \gamma_{0,2}D_2 + ... + \gamma_{0,J}D_J\right] + \gamma_1'x \tag{4.5}$$

where the terms in $[\cdot]$ define the baseline group. We ignore the effects of the covariates x (i.e. $x = 0$) such that each $\gamma_{0,t}$ acts as multiple intercepts for each time period. For example, only D_1 is unity in period 1 such that $\text{logit}h_1 = \gamma_{0,1}$, only D_2 is unity in period 2 such that $logith_2 = \gamma_{0,2}$, and so forth. Clearly, the sequence $\{\gamma_{0,1}, \gamma_{0,2}, ..., \gamma_{0,J}\}$, when taken together, represents the values of flexible baseline logit function at each period. It is also possible to extract the direction of duration dependence from the sequence of γ_0's. In particular, there is evidence for definite negative duration dependence when $\gamma_{0,1} > \gamma_{0,2} > ... > \gamma_{0,J}$. On the other hand, positive duration dependence is evident when $\gamma_{0,1} < \gamma_{0,2} < ... < \gamma_{0,J}$ is observed. More importantly the logit hazard need not be well behaved. For example, the hazard function may be U-shaped if $\gamma_{0,1} > \gamma_{0,2} > ... > \gamma_{0,m-1} < \gamma_{0,m} < \gamma_{0,m+1} < ... < \gamma_{0,J}$. Finally, there is zero duration dependence if $\gamma_{0,1} = \gamma_{0,2} = ... = \gamma_{0,J}$. To be complete, the hazard function in terms of probabilities may be obtained by calculating $1/\left(1 + \exp\left(-\gamma_{0,t}\right)\right)$ for each period.

The second component of equation (4.3), γ_1', defines the impact of the covariates on the logit hazard function. Let there be one dummy covariate, x_1, such that we can write

$$\text{logit}h_t = \left[\gamma_{0,1}D_1 + \gamma_{0,2}D_2 + ... + \gamma_{0,J}D_J\right] + \gamma_1 x_1 \tag{4.6}$$

Using this specification, the covariate effect is the same across all time periods such that the covariate's effect is proportional. In particular, the logit hazard, when $x_1 = 1$, is $\gamma_{0,1} + \gamma_1$ in period 1, $\gamma_{0,2} + \gamma_1$ in period 2, $\gamma_{0,3} + \gamma_1$ in period 3 and so forth. Now, the sign of γ_1 is of crucial importance. Specifically, the logit hazard is lowered when γ_1 is negative suggesting that the probabilities of termination of a bull market are lowered as a result of the presence of x_1 for every point in time. Conversely, the hazard probabilities are raised for every point in time when γ_1 is positive. However, the effect of x_1 need not be restricted to the proportionality assumption but may vary over time. To accommodate this we can rewrite equation (4.3) as

$$\text{logit}h_{it} = \gamma_{0,t} + \gamma'_{1,t}x_i \tag{4.7}$$

so that the simplified expanded model becomes

$$\text{logit}h_t = \left[\gamma_{0,1}D_1 + \gamma_{0,2}D_2 + ... + \gamma_{0,J}D_J\right] + \left[\gamma_{1,1}x_1D_1 + \gamma_{1,2}x_1D_2 + ... + \gamma_{1,J}x_1D_J\right] \tag{4.8}$$

Under the new specification in equation (4.8), each of the $\gamma_{1.t}$ represents the impact of x_1 in each period t. As such the logit hazard, with the presence of x_1, is $\gamma_{0,1}+\gamma_{1,1}$ in period 1, $\gamma_{0,2} + \gamma_{1,2}$ in period 2, $\gamma_{0,3} + \gamma_{1,3}$ in period 3 and so forth. Like the $\gamma_{0,t}$'s, the sequence of $\gamma_{1,t}$'s may take on any real values. In particular, the effect of the covariate decreases over time when we observe $\gamma_{1,1} > \gamma_{1,2} > ... > \gamma_{1,J}$. Likewise, the covariate's effect may be increasing over time when $\gamma_{1,1} < \gamma_{1.2} < ... < \gamma_{1,J}$. If $\gamma_{1,1} = \gamma_{1,2} = ... = \gamma_{1,J}$ then the proportionality assumption on x_1 is valid. Basically, the flexible assumption allows us to capture the erratic behavior of the $\gamma_{1,t}$'s that would never be captured if the proportionality assumption is used right off the back.

The estimation of equation (4.4) utilizes the maximum likelihood method where we seek a set of model parameters that maximizes the model's likelihood of observing the data. Assuming that the standard maximum likelihood assumptions hold (see

Arjas and Haara, 1987), the corresponding likelihood function for the problem is given by

$$L = \prod_{i=1}^{n} P\left(T_i = t_i\right)^{y_i} P\left(T_i > t_i\right)^{1-y_i} \tag{4.9}$$

where

$$P\left(T_i = t\right) = h_{it} \prod_{j=1}^{t-1} \left(1 - h_{ij}\right) \tag{4.10}$$

$$P\left(T_i > t\right) = \prod_{j=1}^{t} \left(1 - h_{it}\right) \tag{4.11}$$

and y_i is a binary variable that is unity if the i^{th} bull market terminates at t and zero otherwise. We call the variable y_i the failure indicator. For example, a vector of duration data $\begin{bmatrix} 2 & 3 & 2 \end{bmatrix}$ in terms of a vector of failure indicators is $\begin{bmatrix} 0 & 1 & 0 & 0 & 1 & 0 & 1 \end{bmatrix}$.

The mathematics behind equation (4.9) is quite messy. However, we can make the mathematics more tractable by using the loglikelihood function. The corresponding loglikelihood function is

$$LL = \sum_{i=1}^{n} y_i \ln\left(\frac{h_{it}}{1 - h_{it}}\right) + \sum_{i=1}^{n} \sum_{j=1}^{t_i} \ln\left(1 - h_{it}\right) \tag{4.12}$$

which is obtained by substituting equations (4.10) and (4.11) into the natural log of equation (4.9). Since the hazard probability may be expressed as a function of parameters as in equation (4.4), we can obtain estimates for the model parameters by maximizing equation (4.12) with respect to $\gamma_{0,t}$ and $\gamma'_{1,t}$. This operation can be performed by the numerical procedures preprogrammed in standard statistical software that are capable of logit regressions.

However, it is worth mentioning that the basic discrete time duration model is not flawless. For data sets with large number of intervals, the number of parameters in the model for modeling the time effects alone becomes very large. This becomes problematic because it may lead to maximum likelihood estimates[2] that do not converge. Also, data can often be sparse at large t's because extreme events are usually rare such that maximum likelihood estimates may not exist for those time periods. Several authors such as Mantel & Hankey (1978), Efron (1988) and Yamaguchi (1993) have suggested the use of piecewise polynomials to overcome these hurdles. However, we do not advise on their approach because it may overlook the abrupt changes in the true underlying hazard function. Rather, a more sophisticated nonparametric approach is required. We will discuss such an approach in the next section.

4.2 The Dynamic Discrete Time Hazard Model

We concluded the previous section by stating some obvious problems of the basic discrete time duration model and that a new approach is required for modeling discrete time duration data. Here, we introduce a dynamic discrete time duration model (dynamic model for short) that applies state space techniques to duration data - see Fahrmeir (1994) and Fahrmeir and Wagenpfeil (1996). This approach is preferred for two reasons. First, unlike the Mantel and Hankey (1978), Efron (1988) and Yamaguchi (1993) type models, the dynamic model is nonparametric in the sense that there is no predefined functional form for the hazard function. This flexibility allows us to capture important abrupt changes in the hazard function. Second, the dynamic model allows estimation and smoothing of the hazard function

[2] The problem becomes worse if we are to model the time effects as well as the time varying effects of covariates.

and covariate effects simultaneously, which can be done through semi-Bayesian or fully Bayesian methods. This characteristic allows us to overcome the maximum estimation problem as a result of over parameterization in the basic model.

To setup the dynamic model, let r_{it}, with i and $t \geq 1$, be risk indicators defined by

$$r_{it} = \left\{ \begin{array}{l} 1 \text{ if the } i^{th} \text{ run is at risk of termination at } t \\ 0 \text{ otherwise} \end{array} \right\}$$

such that the vector $r_t = (r_{it}, i \geq 1)$ is the risk vector and the risk set R_t contains all runs that are at risk of termination at t. Furthermore, let x_t and y_t be vectors of all x_{it} and y_{it} that belong to the risk set respectively. Formally, we can write $x_t = (x_{it}, i \in R_t)$ and $y_t = (y_{it}, i \in R_t)$. The histories of covariates, failure and risk indicators up to time period t are simply given by $x_t^* = (x_1, ..., x_t)$, $y_t^* = (y_1, ..., y_t)$ and $r_t^* = (r_1, ..., r_t)$.

The dynamic nature of the model comes from the fact that the conditional probability of failure is based on historical information. Namely, we seek to model $P\left(y_t | y_{t-1}^*, x_t^*, r_t^*, \gamma_{0,t}^*, \gamma_1^*, Q\right)$ where Q is the covariance matrix of the transition equation. In relation to state space nomenclature, we define equation (4.4) as the measurement equation of the system[3]. To make the notation more compact, we can rewrite the measurement equation as

$$h_{it} = F\left(z_{it}' \gamma_t\right) \tag{4.13}$$

where F is the logit link function, z_{it}' is the design matrix and γ_t is the state vector that contains the baseline parameter $\gamma_{0,t}$, as well as the time varying effects of co-

[3]The measurement equation relates observation to the underlying state of the system - see Chatfield (2004) for further details regarding state space terminologies.

variates $\gamma_{1,t}$. Now, to model the stochastic variation of the baseline group and time varying effects of the covariate, we need to make assumptions on how the parameters evolve through time. The simplest assumption is that the dynamics of the state vector follows a first order random walk process such that the transition equation may be written as

$$\alpha_t = \Phi\alpha_{t-1} + \xi_t \tag{4.14}$$

where Φ is an identity matrix and ξ_t is white noise that is normally distributed with mean 0 and variance Q.

To illustrate this setup, consider the case where there is only one covariate x_{1it}. Thus, the design matrix is simply $z'_{it} = \begin{bmatrix} 1 & x_{1it} \end{bmatrix}$ and the state vector is just $\gamma'_t = \begin{bmatrix} \gamma_{0,t} & \gamma_{1,t} \end{bmatrix}$. The first order random walk transition model for this case in its expanded form can simply be expressed as

$$\begin{bmatrix} \gamma_{0,t} \\ \gamma_{1,t} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \gamma_{0,t-1} \\ \gamma_{1,t-1} \end{bmatrix} + \begin{bmatrix} \xi_{0,t} \\ \xi_{1,t} \end{bmatrix} \tag{4.15}$$

with $\xi_{0,t}$ and $\xi_{1,t}$ assumed to be white noise. It is also assumed that $\xi_{0,t} \sim N(0, \sigma_0^2)$ and $\xi_{1,t} \sim N(0, \sigma_1^2)$. In case we want to impose the proportionality assumption, we can pin down the time varying effect of x_1 by forcing σ_1^2 to zero. However, the proportionality assumption will not be considered in this thesis.

Before describing an estimation procedure, it is necessary to specify some assumptions concerning conditional independence so that we can specify the model in terms of joint densities. First, given x_{it} and γ_t, the current individual failure indicators y_{it} do not depend on γ^*_{t-1} and Q such that

$$P\left(y_{it} | y^*_{t-1}, x^*_t, r^*_t, \gamma^*_{0,t}, \gamma^*_{1,t}, Q\right) = P\left(y_{it} | y^*_{t-1}, x^*_t, r^*_t, \gamma^*_t, Q\right) = P\left(y_{it} | \gamma_t, x_{it}\right)$$

This suggests that the conditional information of γ_t^* on y_t is already reflected in γ_t alone. Second, given y_{t-1}^*, x_t^*, r_t^* and Q, the failure individual indicators y_{it} that belong to the risk set R_i are conditionally independent such that

$$P\left(y_t|y_{t-1}^*, x_t^*, r_t^*, \gamma_t^*, Q\right) = \prod_{i\in R_t} P\left(y_{it}|y_{t-1}^*, x_t^*, r_t^*; \gamma_t\right)$$

or

$$h_t = \prod_{i\in R_t} h_{it}$$

This assumption is just a weaker form of Assumption 2 found in Arjas and Haara (1987), which is unconditional in nature and is likely to hold if a common cause of failure is included in the covariate process Thirdly, we assume that the model parameters follow a Markov process such that

$$P\left(\gamma_t|\gamma_{t-1}^*, y_{t-1}^*, x_t^*, r_t^*, Q\right) = P\left(\gamma_t|\gamma_{t-1}\right)$$

Forth, conditional on y_{t-1}^*, $x_{t-1}^*$and r_{t-1}^*, the covariates x_t and risk indicators r_t do not depend on γ_{t-1}^* and Q. This assumption is identical to Assumption 1 in Arjas and Haara (1987) and is likely to hold for external covariates as well as time independent covariates. Finally, it is assumed that the white noise sequence ξ_t is independent of the initial states of $\gamma_1, ..., \gamma_J, x_1$ and r_1. Together, the assumptions imply that

$$P\left(y_t|y_{t-1}^*, x_t^*, r_t^*, \gamma_t^*, Q\right) = \prod_{i\in R_t} P\left(y_{it}|x_{it}, \gamma_t\right) \tag{4.16}$$

In order to estimate the model parameters, namely γ_t, let the g_0, Q_0 and Q be known and fixed for the moment. A full Bayesian estimation for γ_t would require the computation of the posterior density $P\left(\gamma|y_J^*, x_J^*, r_J^*\right)$, which is not observed. For low dimensional problems, traditional Monte Carlo techniques will suffice. However, for high dimensional problems, the Monte Carlo method becomes infeasible and the Markov Chain Monte Carlo (MCMC) scheme is required. The MCMC will be examined in further detail in the next chapter.

Prior to the development of the MCMC, Fahrmeir (1992), Fahrmeir and Wagenpfeil (1996) and Timmermann and Lunde (2004) use a semi-Bayesian approach to estimate the model by focusing on the posterior mode of the parameters rather than the posterior density[4]. In order to apply this, one must repeatedly apply Bayes' theorem to the posterior density to get

$$P\left(\gamma|y_J^*, x_J^*, r_J^*\right) \propto \prod_{t=1}^{J} \prod_{i\in R_t} P\left(y_{it}|x_{it}, \gamma_t\right) \times \prod_{t=1}^{J} P\left(\gamma_t|\alpha_{t-1}\right) P\left(\gamma_0\right) \tag{4.17}$$

Given this proportionality, maximizing the posterior density is analogous to maximizing the right hand side of equation (4.17).

Taking the natural log of the right hand side of equation (4.17) and applying the measurement and transition equations results in the penalized log-likelihood function, we can obtain

$$PL\left(\gamma\right) = \sum_{t=1}^{J}\sum_{i\in R_t} l_{it}\left(\gamma_t\right) - \frac{1}{2}\left(\gamma_0 - g_0\right)' Q_0^{-1}\left(\gamma_0 - g_0\right) - \frac{1}{2}\sum_{t=1}^{J}\left(\gamma_t - \gamma_{t-1}\right)' Q^{-1}\left(\gamma_t - \gamma_{t-1}\right) \tag{4.18}$$

where

$$l_{it}\left(\gamma_t\right) = y_{it}\ln F\left(z_{it}'\gamma_t\right) + \left(1 - y_{it}\right)\ln\left(1 - F\left(z_{it}'\gamma_t\right)\right) \tag{4.19}$$

is the individual log-likelihood contribution. The maximum likelihood solution for equation (4.17) is simply the conditions at which the posterior density is maximized.

[4] With $\alpha = \alpha_J^* = \left(\alpha_0', \alpha_1', ..., \alpha_J'\right)'$, the posterior mode estimate is just $a = \left(a_{0|J}', a_{1|J}', ..., a_{J|J}'\right)'$, which is identical to $\arg\max_\alpha \left\{p\left(\alpha|y_J^*, x_J^*, r_J^*\right)\right\}$ or the maximum likelihood for the posterior density.

However, one can move away from the Bayesian context by focusing directly on the penalized log-likelihood function. The first term of equation (4.18) accounts for the goodness of fit between the state vector and the data. In general, the higher its value, the better the fit. But, as always, goodness of fit may come with a penalty. For example, there may be a set of candidate state vectors, say $\{\widetilde{\gamma}_t\}$, that fits the data very well but the deviation between successive state vectors are too large. In such an event, the resulting estimates of state vectors may be too jagged and erratic such that useful statistical inference cannot be obtained. In order to prevent this, the last two terms of equation (4.18) penalize large deviations such that $\{\widetilde{\gamma}_t\}$ will lead to a low penalized log-likelihood value. In other words, the goal of the penalized log-likelihood is to find an optimal balance between goodness of fit and smoothness such that the value of equation (4.18) is maximized.

A numerical solution for equation (4.18) requires sophisticated algorithms. For instance, Fahrmeir (1992), Fahrmeir and Wagenpfeil (1996) and Timmermann and Lunde (2004) use an iteratively weighted Kalman filter and smoother. This method is based on the recursive fisher scoring scheme (Fahrmeir and Kaufmann, 1991). Althought, in practice, the hyperparameters g_0, Q_0 and Q need not be known. In fact, the unknown hyperparameters may be estimated via a EM-type scheme by setting initial conditions on g_0, Q_0 and Q (Fahrmeir and Goss, 1992). However, they are treated as being deterministic in this semi-Bayesian framework. Allowing the hyperparameters to be stochastic will bring the problem from a semi-Bayesian setting to a fully Bayesian setting where the MCMC is required. Finally, the prior specification in equation (4.18) need not be a first order random walk as it is possible to impose any AR(p) process (Fahrmeir and Lang, 2001) while random walk models with a local trend may also be used.

4.3 Markov Chain Monte Carlo

Estimation of the dynamic model via posterior mode is merely a semi-Bayesian approach. For a method to qualify as a fully Bayesian approach, it must be able to estimate at least the first and second moments of the posterior density such that the hyperparameters are considered to be stochastic random variables. As mentioned in the previous chapter, fully Bayesian inference requires the use of the MCMC scheme. The purpose of this chapter is to provide a brief discussion concerning the MCMC and its application to the dynamic model as outlined by Fahrmeir and Knorr-Held (1997), Fahrmeir and Tutz (2001) and Tsay (2002). For in depth discussions on MCMC and related topics, we advise the reader to consult Casella and George (1992), Chib and Greenberg (1995) and Gilks, Richardson and Spiegelhalter (1996).

Bayesian inference utilizes Bayes' rule to obtain a solution for the moments of the posterior distribution of interest. To illustrate this, let θ be an arbitrary vector of parameters that belongs to the space Θ. According to Bayes' theorem, given data D, the posterior distribution $P(\theta|D)$ assumes the proportionality

$$P(\theta|D) = \frac{P(D|\theta)\,P(\theta)}{\int P(D|\theta)\,P(\theta)\,d\theta} \propto P(D|\theta)\,f(\theta) \tag{4.20}$$

with $P(D|\theta)$ being the likelihood of observing D given θ. Given this, the first moment of the posterior distribution is

$$E(\theta|D) = \frac{\int \theta P(D|\theta)\,P(\theta)\,d\theta}{\int P(D|\theta)\,P(\theta)\,d\theta} \tag{4.21}$$

such that its solution requires numerical integration[5]. As stated previously, if θ is of high dimensions (i.e. above 20), the integrals become complex such that conventional methods such as the Monte Carlo is infeasible.

An alternative to the Monte Carlo is the MCMC. The general idea of the MCMC is to simulate a Markov chain on Θ that eventually converges to $P(\theta|D)$ upon a sufficient number of repetitions. The nature of $P(\theta|D)$ is extremely complicated in high dimensions but we can make computation more tractable by examining its full conditional counterparts by subdividing θ into components.

For illustrative purposes, let θ be a three component vector such that $\theta = \begin{bmatrix} \theta_1 & \theta_2 & \theta_3 \end{bmatrix}$. The full conditionals for this case are

$$P_1(\theta_1|\theta_2, \theta_3, D), \; P_2(\theta_2|\theta_1, \theta_3, D), \; P_3(\theta_3|\theta_1, \theta_2, D) \tag{4.22}$$

Using the Gibbs sampler (Geman and Geman, 1984), we make updates on each of the full conditionals by repeatedly sampling the sub vectors from their corresponding full conditionals. First, we choose some initial values $\left(\theta_1^{(0)}, \theta_2^{(0)}, \theta_3^{(0)}\right)$ and draw the first set of subvectors $\left(\theta_1^{(1)}, \theta_2^{(1)}, \theta_3^{(1)}\right)$. Next, using $\left(\theta_1^{(1)}, \theta_2^{(1)}, \theta_3^{(1)}\right)$ we draw another set of subvectors $\left(\theta_1^{(2)}, \theta_2^{(2)}, \theta_3^{(2)}\right)$. By repeating this process k times, a sequence

$$\left(\theta_1^{(1)}, \theta_2^{(1)}, \theta_3^{(1)}\right), \left(\theta_1^{(2)}, \theta_2^{(2)}, \theta_3^{(2)}\right), \ldots, \left(\theta_1^{(x)}, \theta_2^{(x)}, \theta_3^{(x)}\right)$$

is obtained. By setting k to a sufficiently large number, the Markov chain theory predicts that $\left(\theta_1^{(k)}, \theta_2^{(k)}, \theta_3^{(k)}\right)$ will eventually converge in distribution to $(\theta_1, \theta_2, \theta_3)$. Of course, the example can be expanded to an n-dimensional θ such that we sample

[5]The second moment is built on the first moment via a moment generating function. Thus, a numerical solution for the second moment of the posterior distribution also requires numerical integration. The same is true for higher moments of the posterior distribution.

from $\left(\theta_1^{(k)}, \theta_2^{(k)}, ..., \theta_n^{(k)}\right)$. It should be noted that the Gibbs sampler is implemented only when the full conditionals are a well defined. In reality, it is often the case that there are no closed form solutions for the full conditionals such that other algorithms are needed.

A popular alternative to the Gibbs Sampler is the Metropolis-Hasting (MH) algorithm. The difference between the MH algorithm and the Gibbs Sampler is that not all proposed update $\left(\theta_1^{(k)}, \theta_2^{(k)}, ..., \theta_n^{(k)}\right)$ from the MH algorithms are accepted. Instead, the update is accepted if the acceptance probability of it lies in a particular range, say 30% to 100%[6].

To demonstrate this formally, consider the full conditional of the component vector θ_1. On the k^{th} updating step, a proposal θ_1' is proposed by the transition kernel[7] $P(\theta_1' \to \theta_1; \theta_2\theta_3)$. Now, the acceptance probability, δ_1, is just

$$\delta_1 = \min\left\{1, \frac{P(\theta_1'|\theta_2\theta_3)\, P(\theta_1' \to \theta_1; \theta_2\theta_3)}{P(\theta_1|\theta_2\theta_3)\, P(\theta_1 \to \theta_1'; \theta_2\theta_3)}\right\} \tag{4.23}$$

If δ_1 lies in range (i.e. the second argument is between 30% to 100%), then the proposed update will be accepted. Otherwise, another proposal must be generated until δ_1 is within the proper acceptance probability range. Finally, it is easy to see that the Gibbs Sampler is just a special case of the MH algorithm. In fact, the Gibbs Sampler specifies

$$f(\theta_1 \to \theta_1'; \theta_2\theta_3) = f(\theta_1'|\theta_2, \theta_3)$$

such that equation (4.23) collapses to unity suggesting that all proposals are ac-

[6]This range is dependent on the programming of the software. The BayesX software has an average acceptance probability between 30% and 70%.

[7]The transition kernel is actually defined by the transition model.

cepted.

In the context of the dynamic duration model with first order random walk transition, the parameter vector θ consists of γ and Q. Since the hyperparameters are now considered to be stochastic, the posterior distribution in equation (4.17) must be rewritten as .

$$P\left(\gamma|y_J^*, x_J^*, r_J^*\right) \propto \prod_{t=1}^{J} \prod_{i\in R_t} P\left(y_{it}|x_{it}, \gamma_t\right) \times \prod_{t=1}^{J} P\left(\gamma_t|\gamma_{t-1}, Q\right) P\left(Q\right) \qquad (4.24)$$

where

$$\prod_{t=1}^{J} \prod_{i\in R_t} P\left(y_{it}|x_{it}, \gamma_t\right) \qquad (4.25)$$

is equivalent to $P\left(D|\theta\right)$ and

$$\prod_{t=1}^{J} P\left(\gamma_t|\gamma_{t-1}, Q\right) P\left(Q\right) \qquad (4.26)$$

is the prior distribution for the unknown parameters. With the posterior distribution redefined, the full conditional for the time varying parameters γ results in the proportionality

$$P\left(\gamma_t|\gamma_{s\neq t}, Q, y_J^*, x_J^*, c_J^*\right) \propto \prod_{i\in R_t} P\left(y_{it}|x_{it}, \gamma_t\right) \times P\left(\gamma_t|\gamma_{s\neq t}, Q\right) \qquad (4.27)$$

where $\prod_{i\in R_t} P\left(y_{it}|x_{it}, \gamma_t\right)$ is defined by the measurement equation and $P\left(\gamma_t|\gamma_{s\neq t}, Q\right)$ defined by the transition equation. In accordance with the first order random walk transition model, $P\left(\gamma_t|\gamma_{s\neq t}, Q\right)$ is identical to $N\left(\mu_t, \Sigma_t\right)$ where

$$N\left(\mu_t, \Sigma_t\right) = \left\{ \begin{array}{cc} N\left(\gamma_{t+1}, Q\right) & (t=1) \\ N\left(\frac{1}{2}\gamma_{t-1} + \frac{1}{2}\gamma_{t+1}, \frac{1}{2}Q\right) & (t=2,...,J-1) \\ N\left(\gamma_{t-1}, Q\right) & (t=J) \end{array} \right\} \qquad (4.28)$$

In light of Knorr-Held (1996), γ_t is updated by drawing on the conditional distribution $P(\alpha_t|\alpha_{s\neq t}, Q)$ via the MH algorithm. The acceptance probability in this case is just

$$\delta_\alpha = \min\left\{1, \frac{P(y_t|\alpha_t')}{P(y_t|\alpha_t)}\right\} \tag{4.29}$$

where $P(y_t|\gamma_t)$ is given by $\prod_{i\in R_t} P(y_{it}|x_{it}, \gamma_t)$. Again, the δ_γ must be greater than 30% for the proposed γ_t to be accepted as an update.

Now, recalling that Q is not defined in the measurement equation, it is necessary to make some assumptions on Q. The full conditional for Q is simply

$$P(Q|\gamma, y_J^*, x_J^*, c_J^*) \sim P(Q|\gamma) \tag{4.30}$$

such that the updating of Q is relatively simple if Q is distributed according to an inverse gamma prior $IG(a, b)$. Specifically, when Q is distributed according to $IG(a, b)$, we have

$$Q \propto Q^{-a-1} \exp\left(-\frac{b}{Q}\right) \tag{4.31}$$

However, there is still no consensus concerning the hyperparameterization for the above but many researchers find that small values for a and b work well. Since the results from the MCMC may be sensitive to the selection of the hyperparameters, it is necessary to try out various combinations of them to see whether the results are significantly different across the different sets of hyperparameters.

Finally, it is important to ensure good mixing properties and convergence in the simulation of the Markov chain. The convergence component is ensured by the BayesX software because the program will not stop resampling until convergence is evident. Normally, however, this property is checked by examining the time series plots of the parameters. Mixing properties can be examined by checking the autocorrelation functions of the parameters. Specifically, low values of the autocorrelation function indicate good mixing properties. It must be noted that each of the parameters must be checked. For a case where we are only studying the time effect, it is necessary to check all J autocorrelation functions (one for each time parameter). Once these two characteristics are ensured, the simulation results for γ and Q is equivalent to the marginal distributions $P(\gamma|x, y, c)$ and $P(Q|x, y, c)$. These marginal distribution are used to estimate the posterior distribution from which the hazard function and covariate effects are calculated.

4.4 Empirical Results: Hazard Functions of Bull and Bear Markets

4.4.1 Modeling the Time Effect

To complement the findings from the duration dependence tests, we seek to model the underlying hazard functions of the bear and bull markets using the dynamic model with first order random walk transition. In particular, if there is positive (negative) duration dependence, then we will observe generally upward (downward) sloping hazard functions. Otherwise, if the processes of interest are duration independent, then the hazard functions will appear relatively flat. Specifically, we are investigating the hazard specification

$$h_{it} = F(\gamma_{0t}) \tag{4.32}$$

where $F(\cdot)$ is the logit link function and γ_{0t} are the time parameters.

Using the BayesX software, we estimate and plot the hazard functions with their ± 1 standard error bands on the logit scale for bear and bull markets for all the samples using the MCMC technique. Since the results may be dependent on the choice of hyperparameters, we select three sets of hyperparameters - $a = 0.001$ $b = 0.001$, $a = 1$ $b = 0.005$ and $a = 0.0001$ $b = 0.0001$. Then, we apply them to each data set in order to see whether the results differ across hyperparameters. To check for mixing properties, we also plot the autocorrelation functions for the results.

However, it must be noted that checking all autocorrelation functions for each data set can be a daunting task! For example, if the number of estimated parameters for one data set is 40, then it will be necessary to examine all 40 autocorrelation functions. To avoid confusion, the mean autocorrelation function and the maximum[8] autocorrelation function are examined instead. Finally, the figures are arranged in sets for three unless stated otherwise. For example, Figures 4.1 through 4.3 plot the results for the pre WWII bull market data set while Figures 4.4 through 4.6 plot the results for the post WWII bull market data set. In particular, the first of the three figures depicts results for $a = 0.001$ $b = 0.001$, the second depicts results for $a = 1$ $b = 0.005$ and the third depicts results for $a = 0.0001$ $b = 0.0001$.

Figures 4.1 through 4.3 plot the results for pre WWII bull markets. From panel a of the figures, it is quite clear that the general shape of the hazard functions are not sensitive to the choice of hyperparameters. On the other hand, the mixing properties are quite sensitive to the choice of hyperparameters as indicated by Panel b of the figures. For example, Figure 4.3 has mean and maximum autocorrelation functions that are higher in value than Figures 4.1 and 4.2 suggesting inferior mixing properties.

[8]The mean autocorrelation function is the average of all the autocorrelation functions. The purpose for this is to get an idea of what the average mixing property is like. The maximum autocorrelation function is the one that has the largest values. The purpose for it is to get an idea of what the worst mixing property looks like.

Alternatively, Figure 4.2 is inferior to Figure 4.1 such that the hyperparameter set $a = 1$ $b = 0.005$ has superior mixing properties.

Figures 4.4 through 4.6 plot the results for post WWII bull markets. Again, the general shape of the hazard functions are not too different across the different sets of hyperparameters but the mixing properties behave otherwise. For this sample, it is clear that Figure 4.4 is superior over Figures 4.5 and 4.6. In particular, the maximum autocorrelation function in Figure 4.5 is very unsatisfactory as the function seem quite persistent. In sum, the optimal hyperparameter set for post WWII bull markets is $a = 0.001$ $b = 0.001$.

Figures 4.7 through 4.9 plot the results for full sample bull markets. Once again, the general shape of the hazard functions themselves are insensitive to the choice of hyperparameters while mixing properties remain sensitive. In particular, Figures 4.7 and 4.9 are inferior to Figure 4.8 as their maximum autocorrelation functions are slow in their decay. As such the optimal hyperparameter set for this case is $a = 1$ $b = 0.005$.

Figures 4.10 through 4.12 plot the results for pre WWII bear markets. The general shape of the hazard functions are more or less the same across the different sets of hyperparameters. As for mixing properties, it can be seen that Figure 4.11 is superior to Figures 4.10 and 4.12. In particular, Figure 4.12 displays the worst mixing since the mean as well as the maximum autocorrelation functions are very slow in their decay. In sum, the hyperparameter set $a = 1$ $b = 0.005$ has superior mixing properties.

Figures 4.13 through 4.15 portrait the results for post WWII bear markets. Again, the general shape of the hazard functions across the different hyperparameter sets while the mixing properties seem different. Judging by the maximum autocorrelation functions alone, it is difficult to draw conclusions on the optimal hyperparameter set since all maximum autocorrelation functions are slow in their decay. As such, the decision rests on the mean autocorrelation functions. Specifically, as indicated by the figures, the optimal hyperparameter set is $a = 1$ $b = 0.005$ because it produces the mean autocorrelation function with the lowest values.

Figures 4.16 trough 4.18 plot the results for full sample bear markets. Once again, the general shape of the hazard functions and the maximum autocorrelation functions are quite similar across the different hyperparameter sets. Thus, the decision rests on the mean autocorrelation functions. From Panel b of the figures, it is evident that the optimal hyperparameter set is $a = 1$ $b = 0.005$ because it produces the mean autocorrelation function with the lowest values.

To complement the duration dependence test results and to make the results more interpretable, we compare the bull and bear hazard functions across the different samples on the probability scale. Figure 4.19 plots the mean bull market hazard functions for the pre WWII, post WWII and full samples on the probability scale using optimal hyperparameters. Notice that positive duration dependence for the post WWII sample is evident as its hazard function is almost always increasing. In particular, the chance for bull market termination starts at about 36% in week 1 and eventually rises to about 58% by week 104. This finding is somewhat consistent with the duration dependence test.

Although not obvious, the full sample bull hazard function also exhibit positive duration dependence. In particular, the function starts at a hazard probability of 47% in week 1 and then rises to a probability to about 50% in week 20. After that, the hazard probability slowly decreases to 47% by week 53 and then rises to an all time high probability of about 52% by week 104. In general, the full sample bull hazard function is an increasing function, which is consistent with some of the test results. The pre WWII bull hazard function, on the other hand, does not seem like an increasing function. Starting at a hazard probability of about 50% at week 1, the function rises to about 55% at week 25 and then falls to a hazard probability of about 45% by week 70 or so. Beyond week 70, the hazard function returns to the probability of about 50% at the terminal. Since the function fails to reach a high (or a low) at its terminal, we argue that it fails to exhibit positive duration dependence as suggested by weak evidence from the statistical tests.

Figure 4.20 plots the bear market hazard functions for the pre WWII, post WWII and full samples on the probability scale. The post WWII and full sample bear hazard functions behave erratically with time. However, both of them manage to attain all time highs at the functions' terminus. As such, we argue that both functions exhibit positive duration dependence, which is consistent with some of the statistical tests. On the other hand, the pre WWII bear hazard function consistently increases with time. For example, starting at a hazard probability of about 44% in week 1, the hazard probability rises steadily to about 55% by week 41. Thus, the pre WWII bear hazard function suggests positive duration dependence, which is different from the statistical tests. We argue that the discrepancy is a result of poor sample size of pre WWII bear markets.

4.4.2 The Impact of the Business Cycle on Bull and Bear Markets Hazards

There is very little theoretical guidance to the choice of covariate variables for modeling bear and bull markets hazard functions. However, it is commonly believed that stock price movements are sensitive to the underlying economic environment. For instance, Chen *et al.* (1986) find a positive relationship between industrial production and stock market returns. Furthermore, some empirical studies have identified switching behavior in the conditional means and higher moments of stock returns as a result of changes in the underlying economic state - see Schwert (1989), Turner *et al.* (1989) and Schaller and Van Norden (1997). Further, McQueen and Roley (1993) have identified asymmetric responses of stock returns to macroeconomic innovations. In particular, Perex-Quiros and Timmermann (1998) and DeStefano (2004) have found asymmetric switching about business cycle turning points. In particular, stock returns tend to be positive in the first half of an economic boom and then diminishes to zero (or near zero) in the second half of the boom. On the other hand, stock returns become negative in the first half of an economic recession and then slowly move back to the zero (or near zero) range in the second half of the recession.

In this section, we wish to revisit such switching properties around business cycle turning points by using duration analysis. In particular, if there is switching behavior in stock prices about business cycle turning points, then the logit hazard of bull and bear markets must also be responsive to changes in the underlying economic condition.

To account for the underlying economic state, we use the business cycle dates that are published by the National Bureau of Economic Research (NBER). It should be noted that the definition of business cycles in the NBER is different than that of conventional economics textbooks. In particular, the NBER definition of a recession is defined as a significant decrease in economic activity (i.e. real GDP, employment, real income, industrial production and wholesale-retail sales) over a period of more than a few months. An opposite definition holds for a economic boom according to the NBER. Unfortunately, the latest NBER business cycle date is November 2001. As such, we can only use a sample starting from January 1928 to November 2001 rather than the full sample.

Using the business cycle dates, we generate two dummy variables *expansion* and *contraction* as covariates. The variable *expansion* assumes the value of unity when the economy is in an economic boom and zero otherwise. Similarly, the variable *contraction* is equal to one when the economy is enduring a recession and zero otherwise. Since the economy can change its state during the course of a bear or bull market, the variables *expansion* and *contraction* are actually time varying covariates.

What we are interested in is how the hazard probabilities of bear and bull market respond to changes to the underlying economic environment. In particular, the question of interest is how the hazard functions of bear and bull markets change as a result of the economy changing from an expansion to a recession and vice versa. The dynamic duration models of interest are

$$h_{it} = F\left(\gamma_{0t} + \gamma_{1t} expansion\right) \tag{4.33}$$

and

$$h_{it} = F\left(\gamma_{0t} + \gamma_{1t} contraction\right) \tag{4.34}$$

The baseline hazard functions for each of the models are defined by the sequence $\{\gamma_{0t}\}$ that corresponds to each of the models.

These baseline functions must be interpreted with care. Consider, for example, the model defined by equation (4.33). Further, suppose the event of interest is bull markets. When *expansion* is equal to zero, the baseline hazard function for the bull market is given by the sequence $\{\gamma_{0t}\}$ defined by equation (4.33). Since the economy can either be in a boom or a recession (and nothing else), this baseline hazard function is actually the hazard function for a bull market where the underlying economic state is a recession. When the economy switches from a recession to an expansion, however, the hazard function begins to deviate from the baseline according to the sequence of $\{\gamma_{1t}\}$. That is, the sequence $\{\gamma_{1t}\}$ captures the dynamic effects of changing from an economic recession to an expansion on the baseline bull market hazard function. Similarly, the baseline hazard function for equation (4.34) is actually the hazard function for a bull market during an economic boom. When the economy changes from a boom to a recession, the variable *contraction* assumes the value of unity. Given this change, the sequence $\{\gamma_{1t}\}$ captures the impact of the economy changing from a contraction state to a boom state on the baseline bull market hazard function. In sum, the baseline functions defined by each model have different meanings.

It should be noted that results from our models are not directly comparable to previous studies that used time series techniques although some analogies do exist. Perex-Quiros and Timmermann (1998) and DeStefano (2004), as a whole, find switching behavior in stock returns around business cycle turning points. In particular, they find that stock returns tend change from positive to negative around the transition point at which the economy changes from boom to recession. On the other hand, stock returns tend to change from negative to positive around the transition point at which the economy changes from recession to boom. Our models, however, is not capable of directly predicting this. What our models are capable of predicting is how a change in the state of the economy impact the baseline hazard of a run of negative or positive returns (i.e. bear or bull markets).

Analogous to the time series studies, the switching behaviors are similar to changes in the hazard function of the bull and bear as a result of changes in the economic environment. For example, if stock returns tend to switch from negative to positive around business cycle troughs, then it must be true that bull (bear) markets become less (more) likely to terminate when the economy changes from contraction to expansion. Similarly, if stock returns tend to switch from positive to negative around business cycle troughs, then it must be true that bull (bear) markets become more (less) likely to terminate when the economy changes from boom to recession.

As in the previous section, we wish to conduct some sensitivity analysis concerning the choice of hyperparameters as a well as checks for proper mixing. The difference here is that it is necessary to examine the main time effect (i.e. the hazard function alone) as well as the covariate effect. Again, the figures are grouped in sets of threes such that the first of the three depicts results for $a = 0.001$ $b = 0.001$, the second depicts results for $a = 1$ $b = 0.005$ and the third depicts results for $a = 0.0001$ $b = 0.0001$. Also, each figure is divided into four panels: Panel a depicts the hazard function, Panel b depicts the mean and maximum autocorrelation functions for the time parameters, Panel c depicts the time varying effects of the covariate and Panel d depicts the mean and maximum autocorrelation functions for the covariate effects.

Figures 4.21 through 4.23 plot the results according to equation (4.33) for bull markets using the three hyperparameter sets. Judging from the positive sloping baseline hazard functions, it seems that bull markets under economic contractions exhibit positive duration dependence. In other words, older bull markets are more likely to terminate than younger ones. In addition, it is also true that the general shapes of the baseline hazard functions are insensitive to the choice of hyperparameters; the same can be said concerning the time varying effect of *expansion*. On the other hand, the mixing properties are somewhat sensitive to the choice of a and b. Next, consider the autocorrelation functions for the time parameters depicted in the figures. The mean autocorrelation functions look very similar but the maximum autocorrelation function in Figure 4.22 is superior in comparison to the other two figures. Further, the mean and maximum autocorrelation functions for the covariate effects in Figure 4.22 are also superior. Thus, we conclude that the optimal hyperparameter set for this case is $a = 1$ $b = 0.005$. With this in mind, we turn to Panel c of Figure 4.22 and examine the time varying effects of *expansion* more closely. Judging by the mean effects of *expansion*, it is clear that the change from an economic contraction to economic expansion has negative effects on the baseline hazard. In other words, such a change in the economic environment tends to promote bull markets by decreasing their termination probabilities. However, the magnitude of the effects are not the same throughout the entire time horizon. In particular, the effect has a stronger negative impact on younger bull markets than on older ones. This time varying impact is even more obvious if the confidence intervals are taken into account. Specifically, the effects are negative up to week 33 and then become insignificant thereafter. Simply put, the statistically significant results are consistent with the switching behavior suggested by previous time series studies.

Figures 4.24 through 4.26 plot the results according to equation (4.34) for bull markets using the three hyperparameter sets. The positively sloped baseline hazard functions from Panel a of the figures suggest positive duration dependence in bull markets under economic expansions. Once again, the shapes of the functions are insensitive to the choice of a and b. On the other hand, the same cannot be said for the effects of *contraction*. In particular, the plot for the effects of *contraction* in Figure 4.25 appears flatter than ones in Figures 4.24 and 4.26. As for mixing properties, the autocorrelation functions for the time effects are very well behaved for all three sets of hyperparameters. However, the autocorrelation functions are superior in Figure 4.26 suggesting that the optimal choice for a and b are 1 and 0.005 respectively. With this in mind, we study Panel b of Figure 4.25 more carefully. The effects of *contraction* are significantly positive throughout the entire time horizon suggesting that the change from economic expansion to contraction tends to discourage bull markets. More importantly, the effects seem to be constant throughout the entire time horizon. This in turn implies that the effects from such change in the economy have the same impact on young bull markets as the older bull markets. Again, these statistically significant finding is consistent with the switching behavior suggested by previous time series studies.

Figures 4.27 through 4.29 plot the results according to equation (4.33) for bear markets using the three hyperparameter sets. As with bull markets during economic recession, bear markets during economic recession seem to exhibit positive duration dependence as indicated by the positively sloped baseline hazard functions. As well, the general shapes of the baseline hazard functions are not very different across the three hyperparameter sets. This observation is also evident for the time varying effects of *expansion*. The mixing properties for each hyperparameter sets are given by the autocorrelation functions. It is quite clear that the mean and maximum autocorrelation functions for the time effect in Figure 4.27 are unsatisfactory.

On the other hand, the autocorrelation functions for the time effect in Figure 4.29 are inferior to that in Figure 4.28. As such, the optimal hyperparameter choice for this case is $a = 1$ $b = 0.005$. With this in mind, we study Panel c of Figure 4.28 more closely. By inspection, the mean effects of *expansion* are positive and time varying throughout the entire time horizon. This suggests that, on average, the change from economic contraction to expansion tend to discourage bear markets. However, once the confidence intervals are accounted for, the time varying effects of the change in economic activity is statistically insignificant throughout the entire time horizon. In other words, we fail to identify significant switching behavior.

Figures 4.30 through 4.32 plot the results according to equation (4.34) for bear markets using the three hyperparameter sets. From the figures, it is clear that the shapes of the time varying effects of *contraction* are almost identical across the three hyperparameter sets. However, the shapes of the baseline hazard functions are sensitive to the choice of hyperparameters. While the generally upward sloping baseline hazard functions in Figures 4.30 and 4.32 appear similar, the baseline hazard function in Figure 4.31 look quite different. In particular, the baseline hazard function in Figure 4.31 is much more erratic than the other two. However, since the erratic function still manages to reach a high at its terminal, we argue that it exhibit positive duration dependence. In other words, all three baseline hazard functions suggest that older bear markets are more likely to terminate than younger ones.

To check the mixing properties of the results, it is necessary to examine the autocorrelation functions. The mean and maximum autocorrelation functions for the main time effect are very similar across Figures 4.30 to 4.32. On the other hand, the mean and maximum autocorrelation functions concerning *contraction* look quite different. In particular the maximum autocorrelation functions for *contraction* in Figures 4.30 and 4.32 are unsatisfactory while the one in Figure 4.31 is very well behaved. Furthermore, although the mean autocorrelation functions for *contraction* all look satisfactory, it is quite clear that the one in Figure 4.31 is superior as it has the lowest values. Thus, the optimal hyperparameter set for this case is $a = 1$ $b = 0.005$. With this in mind, we study Panel c of Figure 4.31 with a little bit more care.

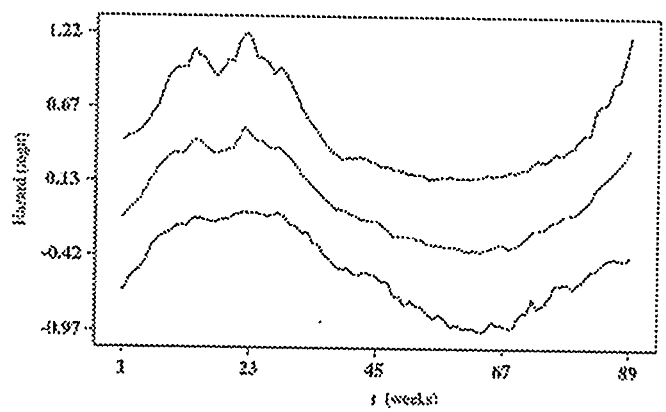
Starting from week 1, the mean effects of *contraction* are negative up week 32. In other words, within this time horizon, the effect of the economy changing from expansion to contraction tends to promote bear markets. Beyond week 32, however, we discover an unexpected finding: the mean effects become positive such that such change in the economy actually discourages bear markets. A possible explanation for this unexpected finding has to do with synchronicity. In particular, it is very possible that the timing of the turning points selected by the selection algorithm is not exact. Thus, the timing of bear markets and the NBER business cycle dates may be out-of-sync leading to the unexpected results. Likewise, the unexpected results maybe a result of improper business cycle selection by the NBER. Finally, once the confidence intervals are taken into account, we find that the effects of *contraction* are statistically insignificant. In sum, we fail to identify significant switching behavior

4.5 Conclusion

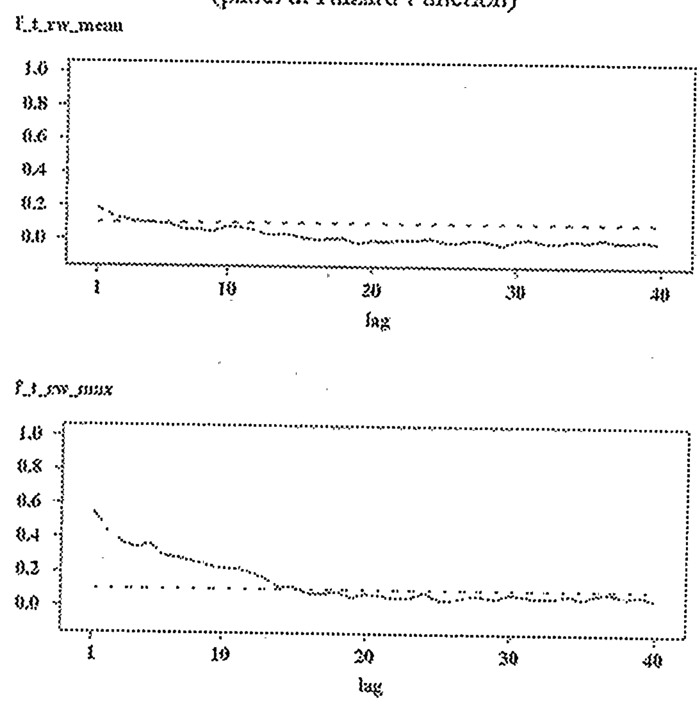
We provide a dynamic framework for estimating discrete time hazard functions for bull and bear market. We find that most of the hazard functions generated by the dynamic model using the MCMC approach seem to exhibit positive duration dependence with the exception for the pre WWII hazard function. Nevertheless, the results are generally consistent with the findings from the statistical test from the previous chapters. Finally, the shape of the hazard functions within the same sample are not sensitive to the choice of hyperparameters. However, the mixing properties are quite sensitive to the choice of a and b.

Next, we seek to identify switching behavior of stock returns around business cycle turning points using dynamic duration models. In particular, we find that changes in the economic environment have statistically significant impact on the hazard functions of bull market hazards. For example, the change from economic expansion to contraction tends to increase the hazard of bull markets. On the other hand, the change from economic contraction to expansion tends to decrease the hazard of bull markets. These observations are consistent with the switching behavior identified by previous studies using conventional time series techniques. However, we find that changes in the economic environment have statistically insignificant impact on the hazard functions of bull markets. Once statistical significance is discarded, we find that the mean effect of the variable *expansion* is positive on the hazard of bear markets. Likewise, the mean effect of the variable *contraction* is generally negative on the hazard of bear markets. Thus, we fail to identify switching behavior of stock prices around business cycle turning points using bear market data. Finally, we observe that mixing properties of the results are sensitive to the selection of a and b.

Figure 4.1
Pre WWII Bull Market Hazard Function with a=0.001 and b=0.001.



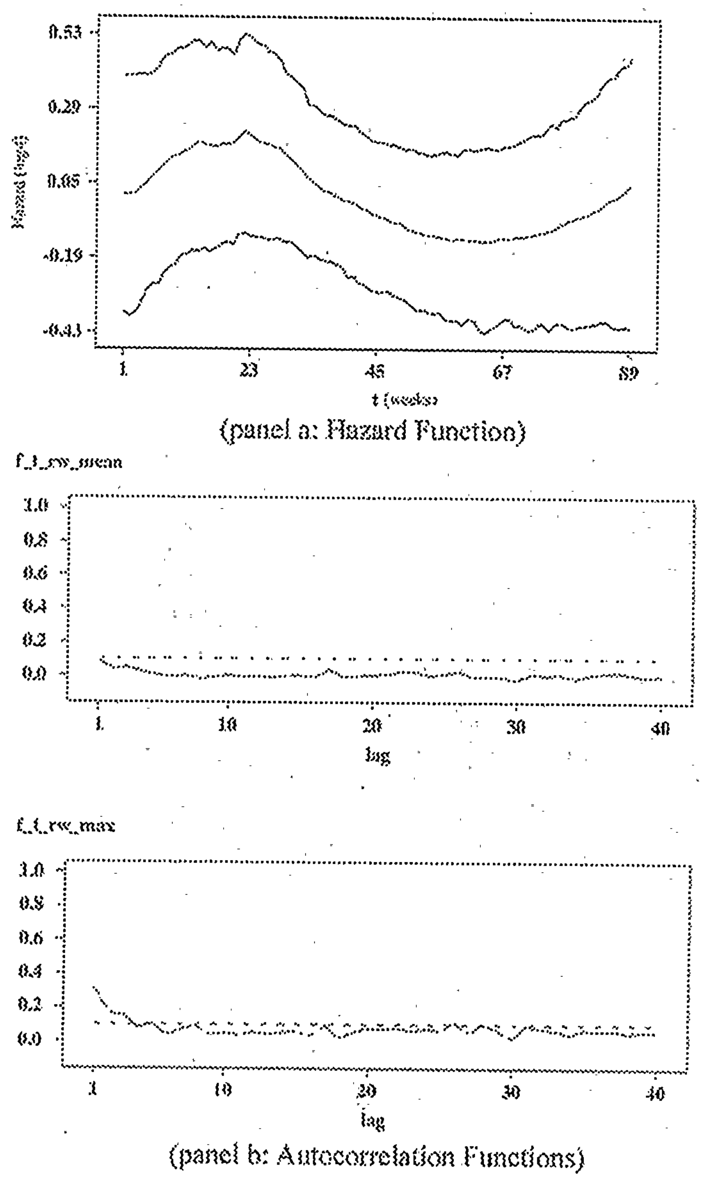
(panel a: Hazard Function)



(panel b: Autocorrelation Functions)

Note: Panel b contains two autocorrelation functions for the time parameters. The upper one is the mean autocorrelation function while the lower one is the max autocorrelation function.

Figure 4.2
Pre WWII Bull Market Hazard Function with $a=1$ and $b=0.005$.



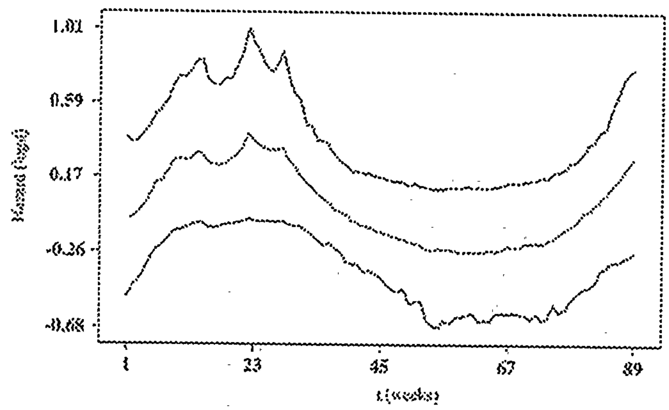
(panel a: Hazard Function)

(panel b: Autocorrelation Functions)

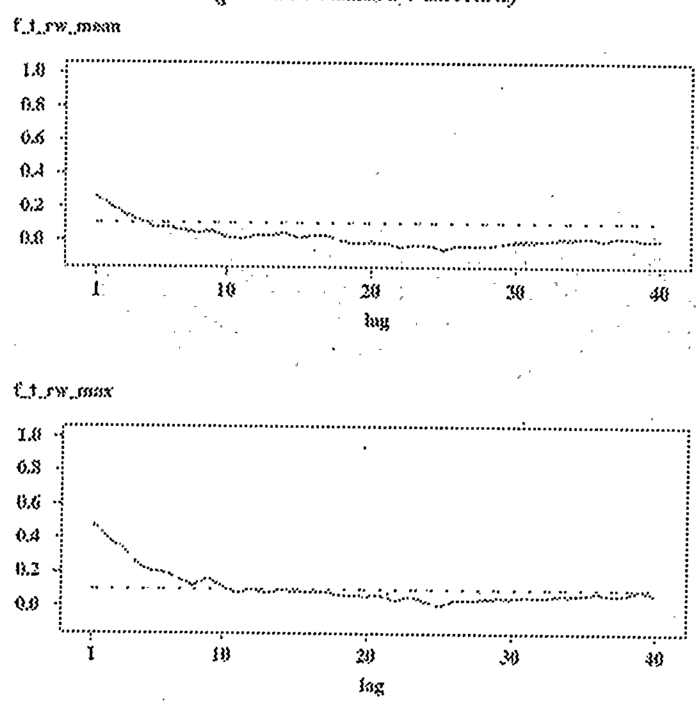
Note: Panel b contains two autocorrelation functions for the time parameters. The upper one is the mean autocorrelation function while the lower one is the max autocorrelation function.

Figure 4.3

Pre WWII Bull Market Hazard Function with a=0.0001 and b=0.0001.



(panel a: Hazard Function)

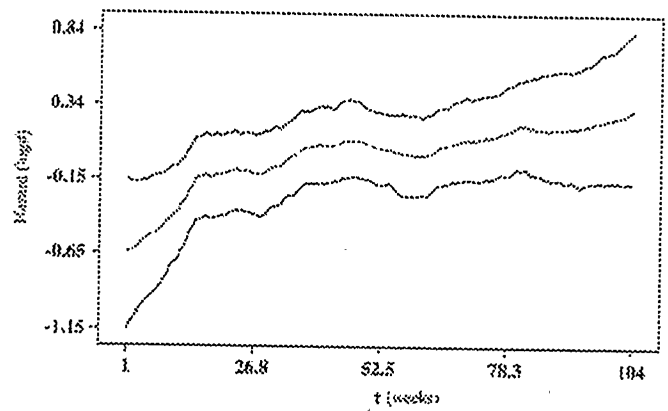


(panel b: Autocorrelation Functions)

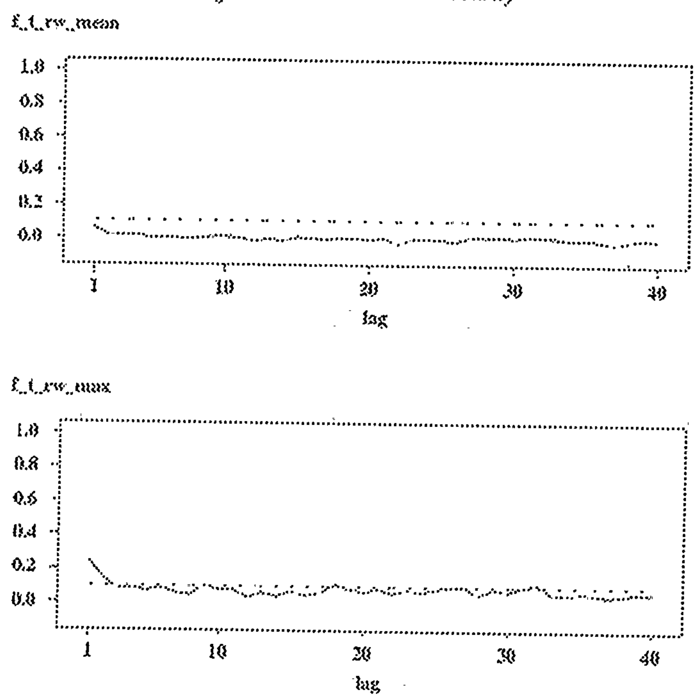
Note: Panel b contains two autocorrelation functions for the time parameters. The upper one is the mean autocorrelation function while the lower one is the max autocorrelation function.

Figure 4.4

Post WWII Bull Market Hazard Function with a=*0.001* and b=*0.001*.



(panel a: Hazard Function)

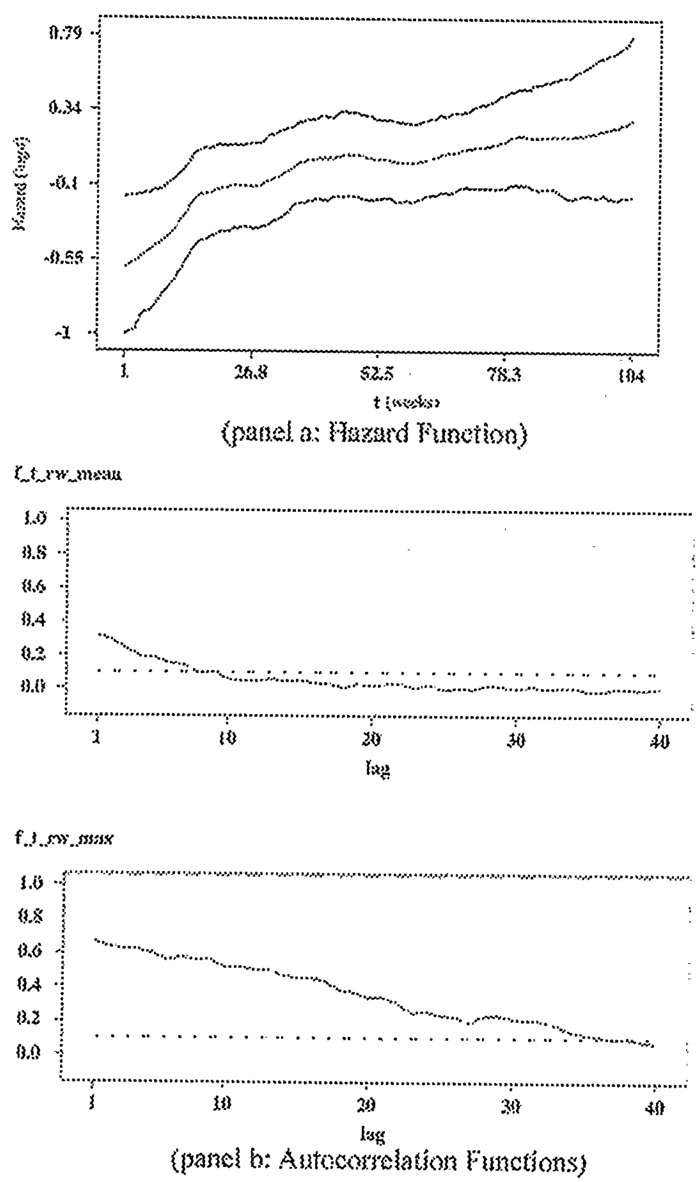


(panel b: Autocorrelation Functions)

Note: Panel b contains two autocorrelation functions for the time parameters. The upper one is the mean autocorrelation function while the lower one is the max autocorrelation function.

Figure 4.5

Post WWII Bull Market Hazard Function with $a=1$ and $b=0.005$.



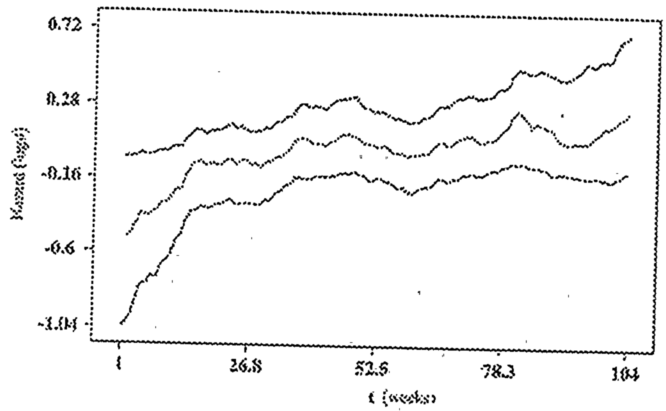
(panel a: Hazard Function)

(panel b: Autocorrelation Functions)

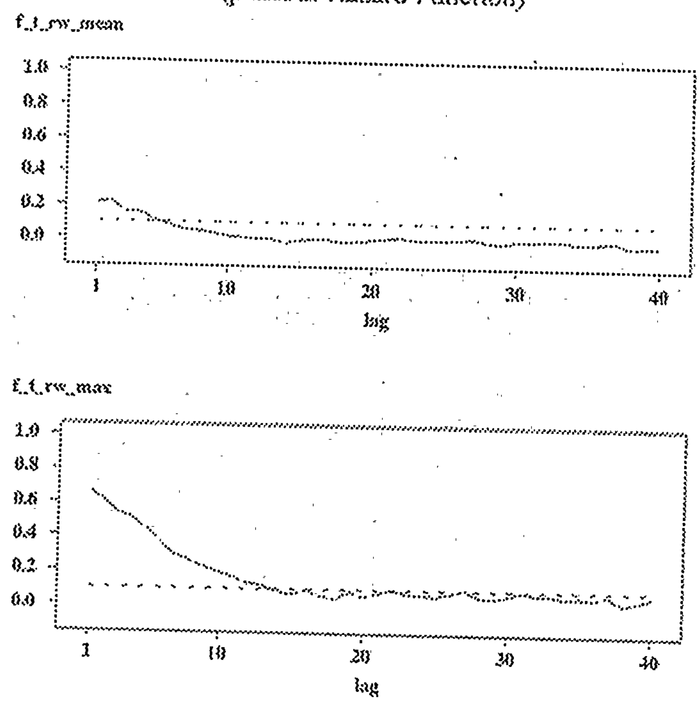
Note: Panel b contains two autocorrelation functions for the time parameters. The upper one is the mean autocorrelation function while the lower one is the max autocorrelation function.

Figure 4.6

Pre WWII Bull Market Hazard Function with a=*0.0001* and b=*0.0001*.



(panel a: Hazard Function)

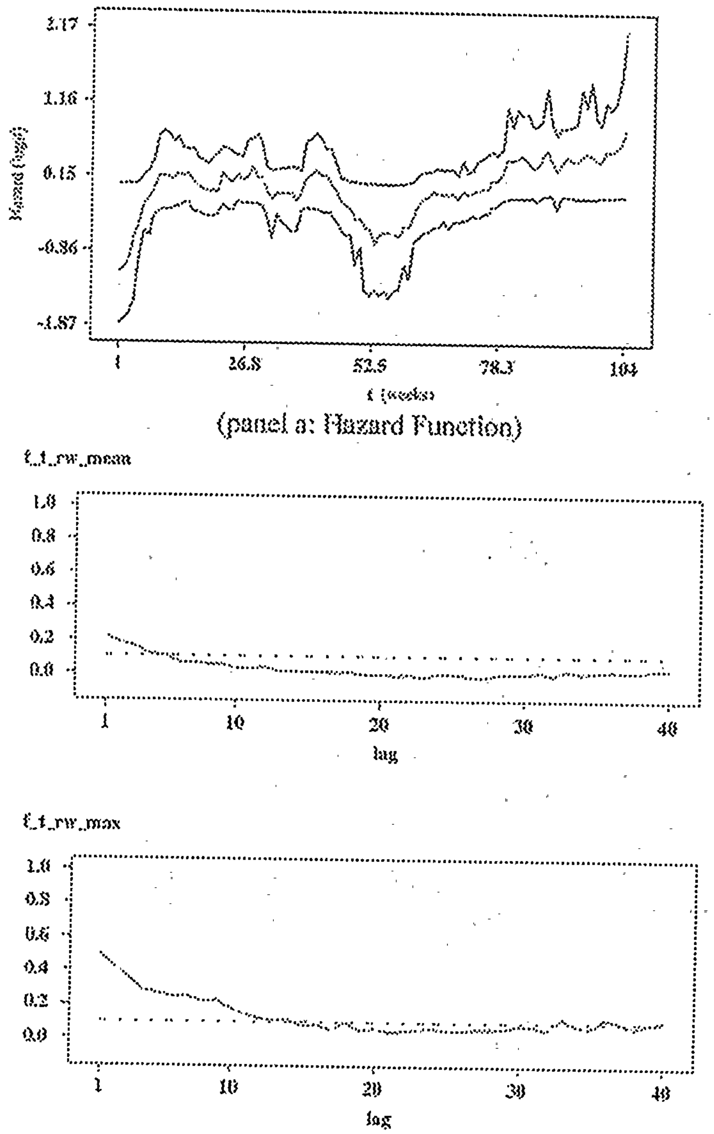


(panel b: Autocorrelation Functions)

Note: Panel b contains two autocorrelation functions for the time parameters. The upper one is the mean autocorrelation function while the lower one is the max autocorrelation function.

Figure 4.7

Full Sample Bull Market Hazard Function with a=0.001 and b=0.001.

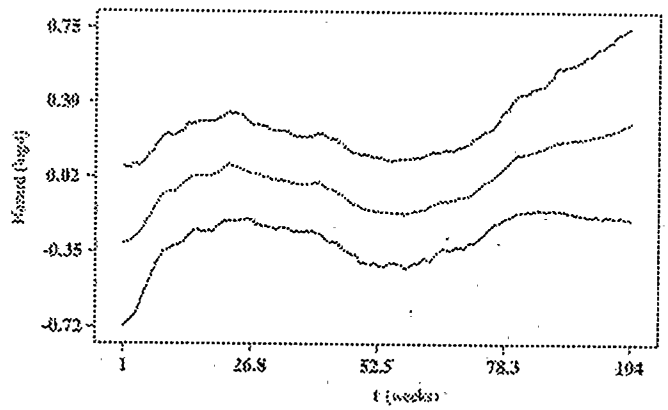


(panel a: Hazard Function)

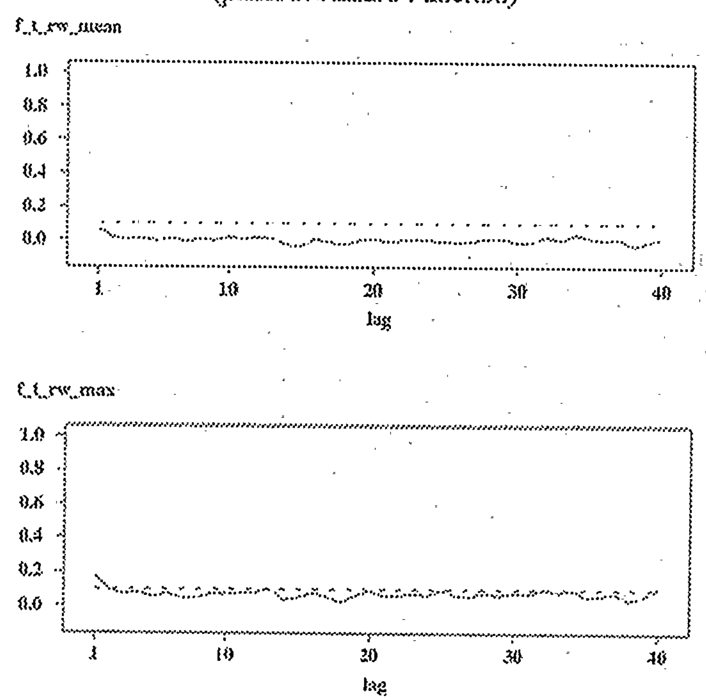
(panel b: Autocorrelation Functions)

Note: Panel b contains two autocorrelation functions for the time parameters. The upper one is the mean autocorrelation function while the lower one is the max autocorrelation function.

Figure 4.8
Full Sample Bull Market Hazard Function with $a=1$ and $b=0.005$.



(panel a: Hazard Function)

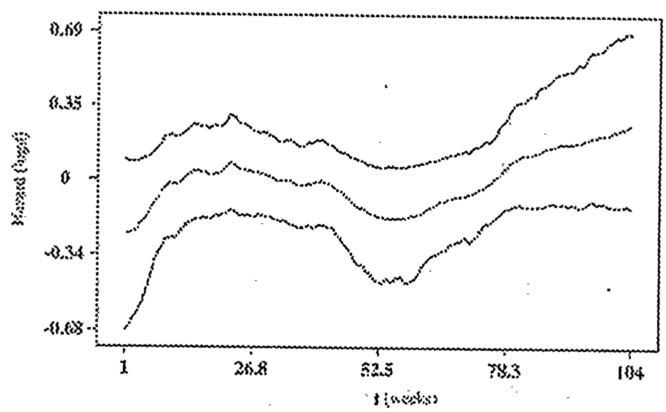


(panel b: Autocorrelation Functions)

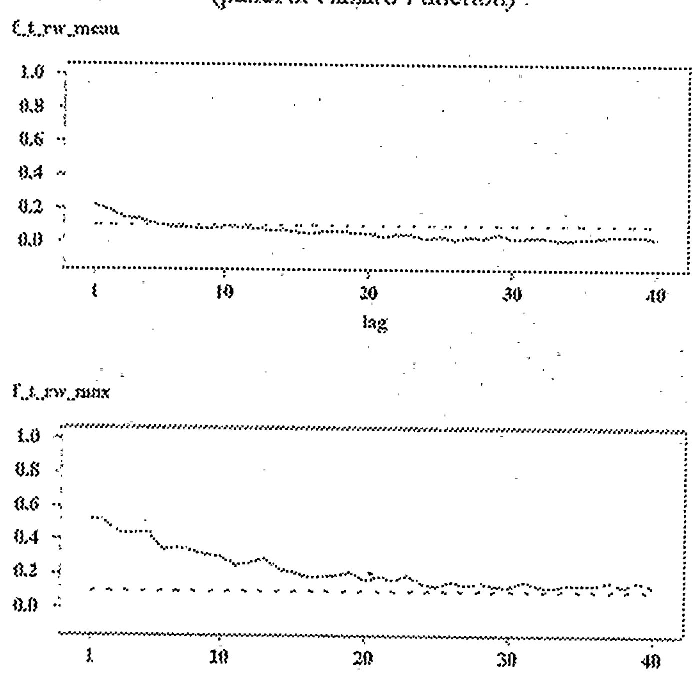
Note: Panel b contains two autocorrelation functions for the time parameters. The upper one is the mean autocorrelation function while the lower one is the max autocorrelation function.

Figure 4.9

Full Sample Bull Market Hazard Function with $a=0.0001$ and $b=0.0001$.



(panel a: Hazard Function)

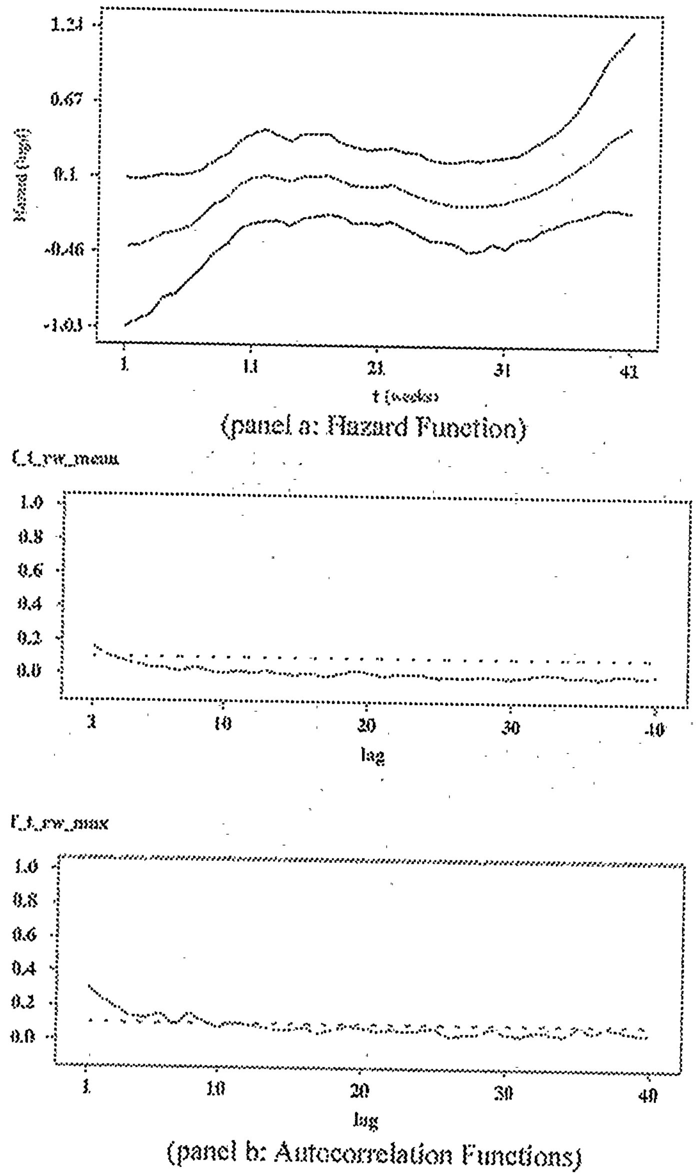


(panel b: Autocorrelation Functions)

Note: Panel b contains two autocorrelation functions for the time parameters. The upper one is the mean autocorrelation function while the lower one is the max autocorrelation function.

Figure 4.10

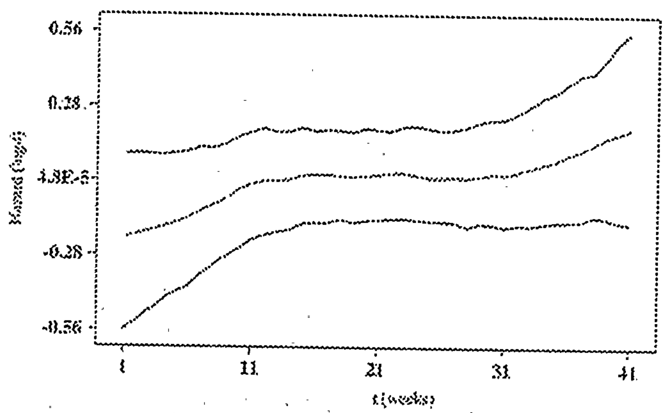
Pre WWII Bear Market Hazard Function with $a=0.001$ and $b=0.001$.



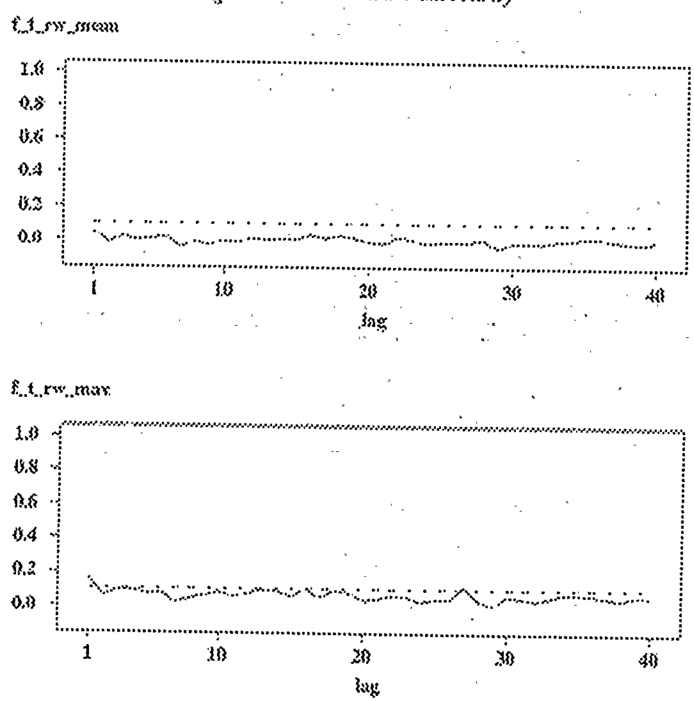
Note: Panel b contains two autocorrelation functions for the time parameters. The upper one is the mean autocorrelation function while the lower one is the max autocorrelation function.

Figure 4.11

Pre WWII Bear Market Hazard Function with $a=1$ and $b=0.005$.



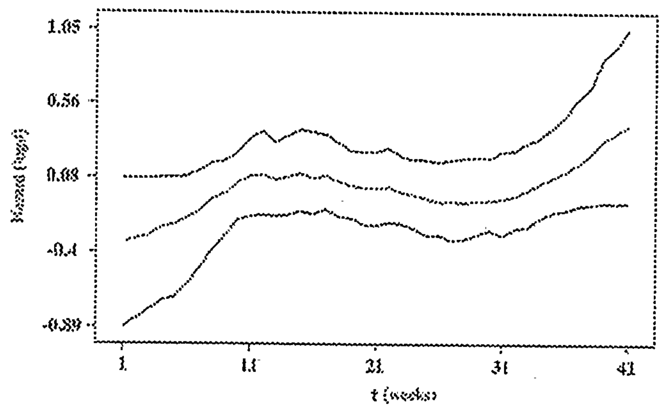
(panel a: Hazard Function)



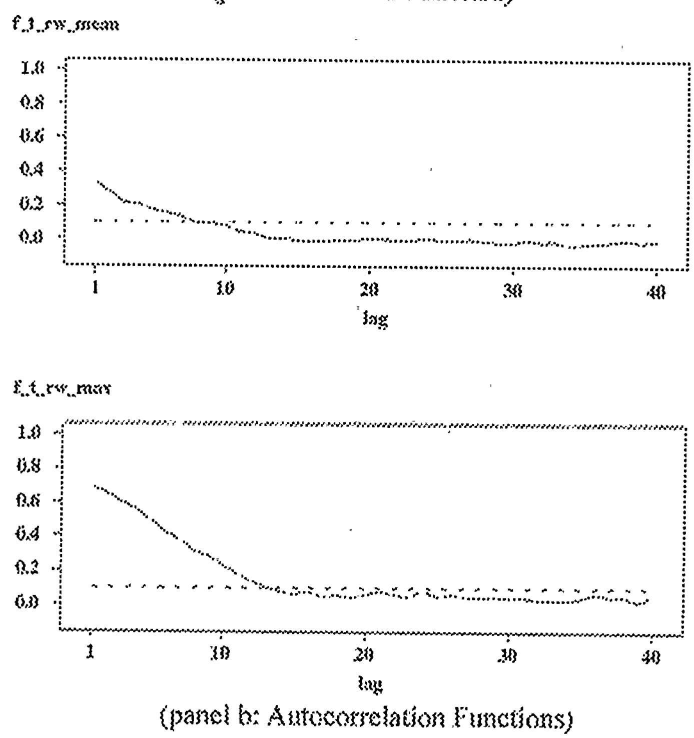
(panel b: Autocorrelation Functions)

Note: Panel b contains two autocorrelation functions for the time parameters. The upper one is the mean autocorrelation function while the lower one is the max autocorrelation function.

Figure 4.12
Pre WWII Bear Market Hazard Function with a=0.0001 and b=0.0001.



(panel a: Hazard Function)

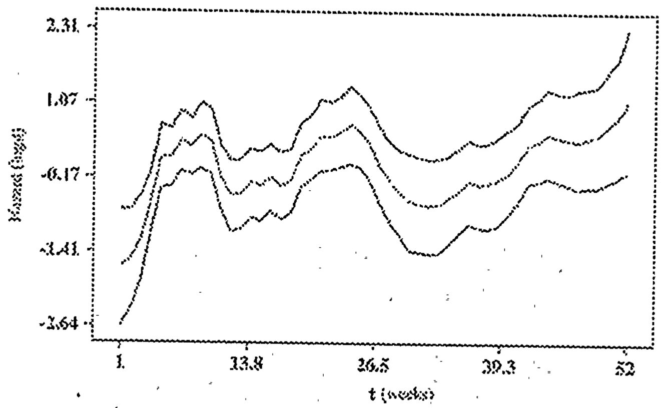


(panel b: Autocorrelation Functions)

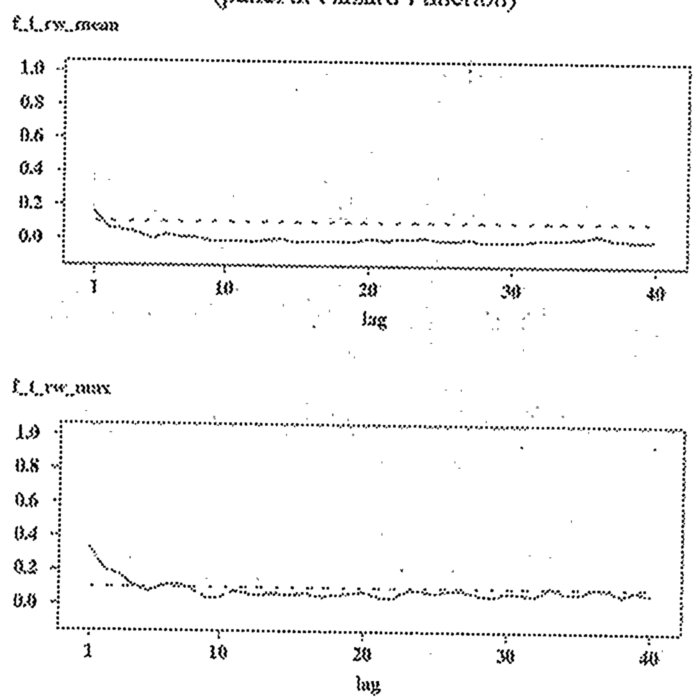
Note: Panel b contains two autocorrelation functions for the time parameters. The upper one is the mean autocorrelation function while the lower one is the max autocorrelation function.

Figure 4.13

Post WWII Bear Market Hazard Function with a=*0.001* and b=*0.001*.



(panel a: Hazard Function)

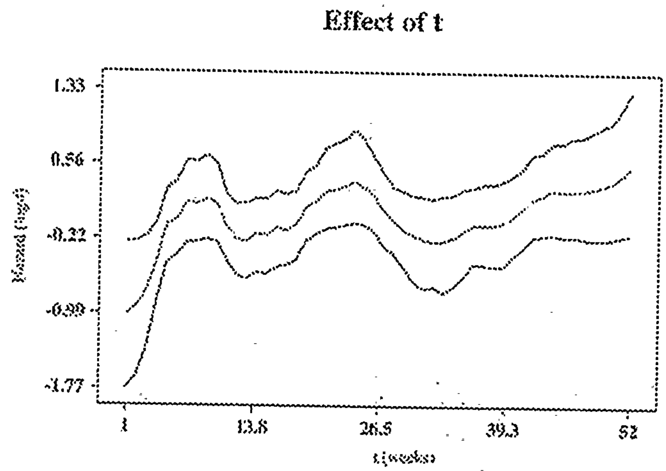


(panel b: Autocorrelation Functions)

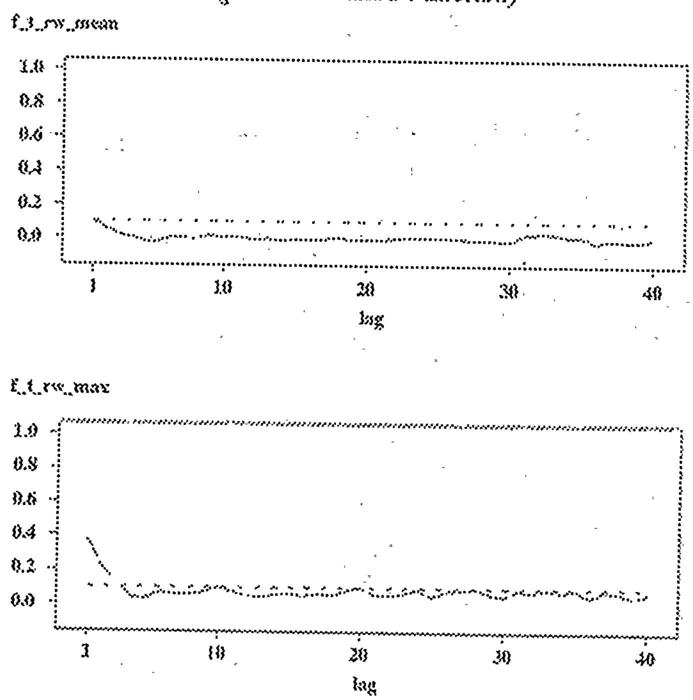
Note: Panel b contains two autocorrelation functions for the time parameters. The upper one is the mean autocorrelation function while the lower one is the max autocorrelation function.

Figure 4.14

Post WWII Bear Market Hazard Function with $a=1$ and $b=0.005$.



(panel a: Hazard Function)

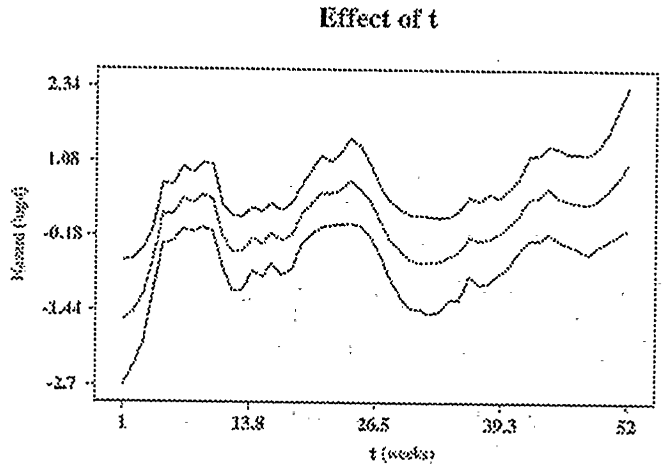


(panel b: Autocorrelation Functions)

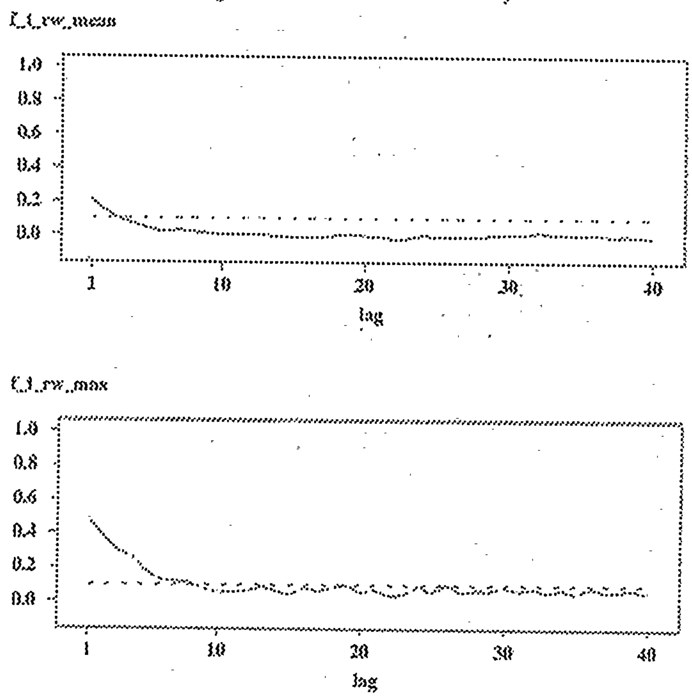
Note: Panel b contains two autocorrelation functions for the time parameters. The upper one is the mean autocorrelation function while the lower one is the max autocorrelation function.

Figure 4.15

Post WWII Bear Market Hazard Function with $a=0.0001$ and $b=0.0001$.



(panel a: Hazard Function)

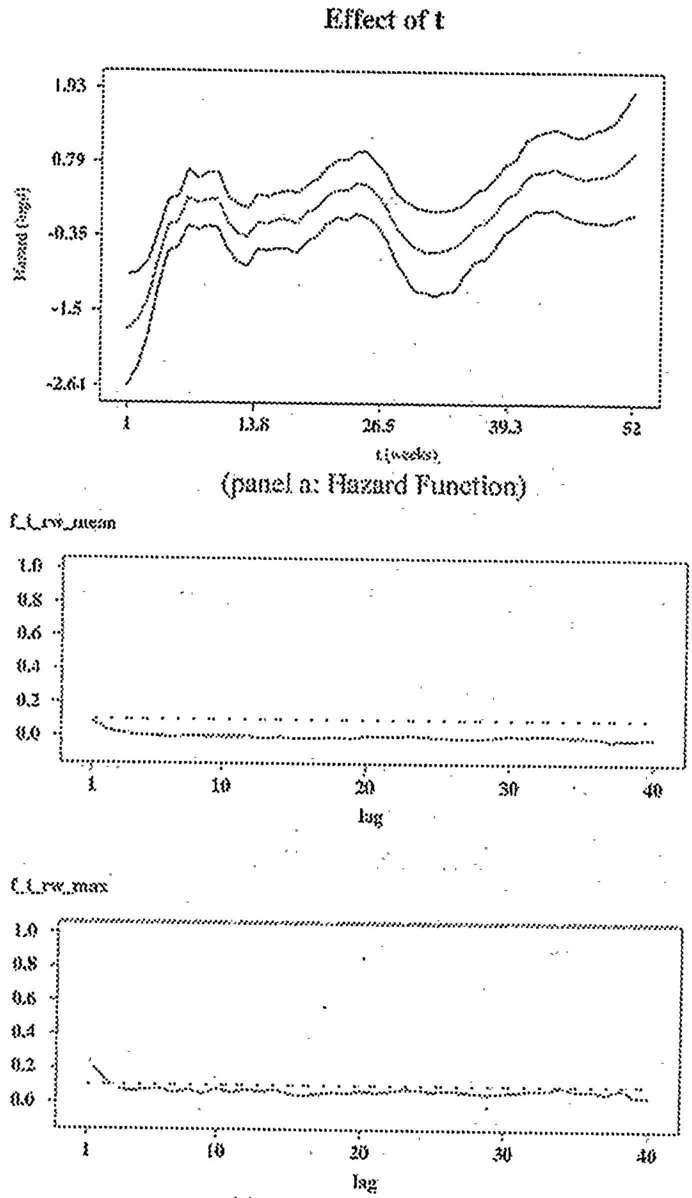


(panel b: Autocorrelation Functions)

Note: Panel b contains two autocorrelation functions for the time parameters. The upper one is the mean autocorrelation function while the lower one is the max autocorrelation function.

Figure 4.16

Full Sample Bear Market Hazard Function with a=0.001 and b=0.001.

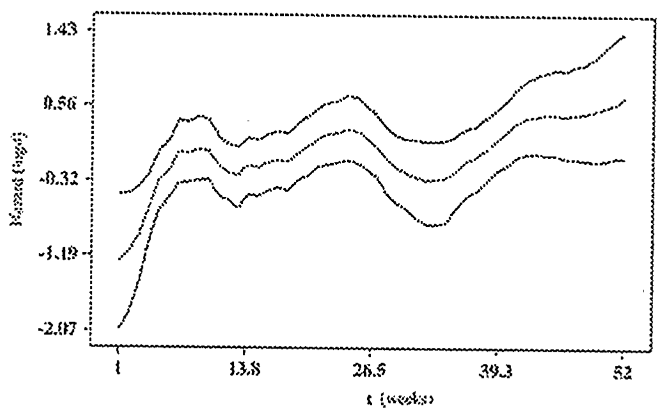


(panel a: Hazard Function)

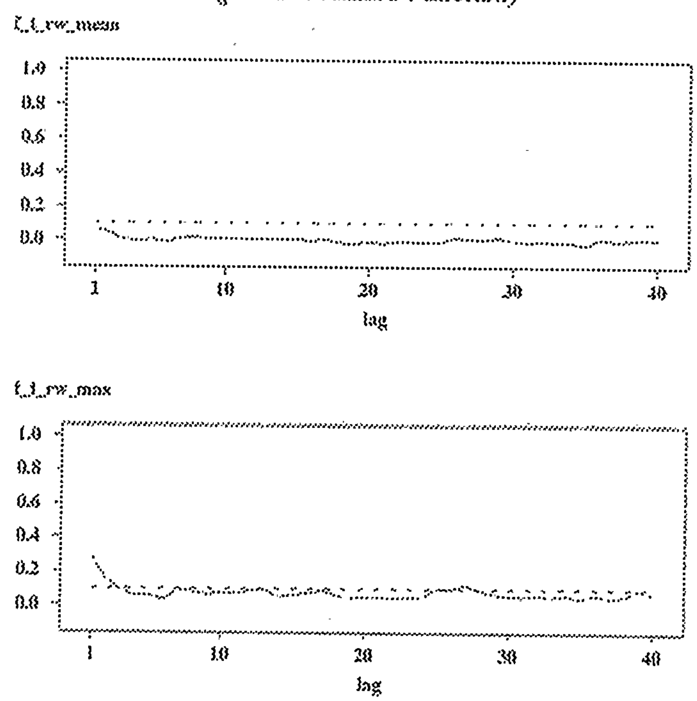
(panel b: Autocorrelation Functions)

Note: Panel b contains two autocorrelation functions for the time parameters. The upper one is the mean autocorrelation function while the lower one is the max autocorrelation function.

Figure 4.17
Full Sample Bear Market Hazard Function with $a=1$ and $b=0.005$.



(panel a: Hazard Function)

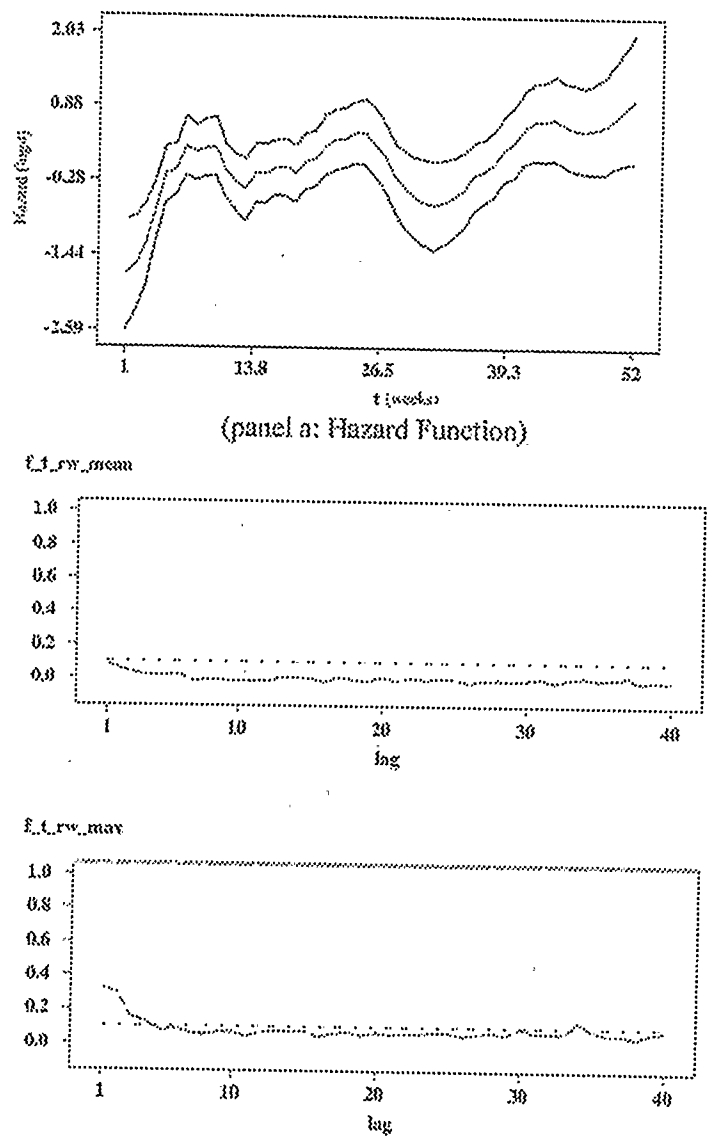


(panel b: Autocorrelation Functions)

Note: Panel b contains two autocorrelation functions for the time parameters. The upper one is the mean autocorrelation function while the lower one is the max autocorrelation function.

Figure 4.18

Full Sample Bear Market Hazard Function with a=0.0001 and b = 0.0001.



(panel a: Hazard Function)

(panel b: Autocorrelation Functions)

Note: Panel b contains two autocorrelation functions for the time parameters. The upper one is the mean autocorrelation function while the lower one is the max autocorrelation function.

Figure 4.19
A Comparison of Bull Market Hazard Functions Across the Full, Post WWI and Pre WII Samples.

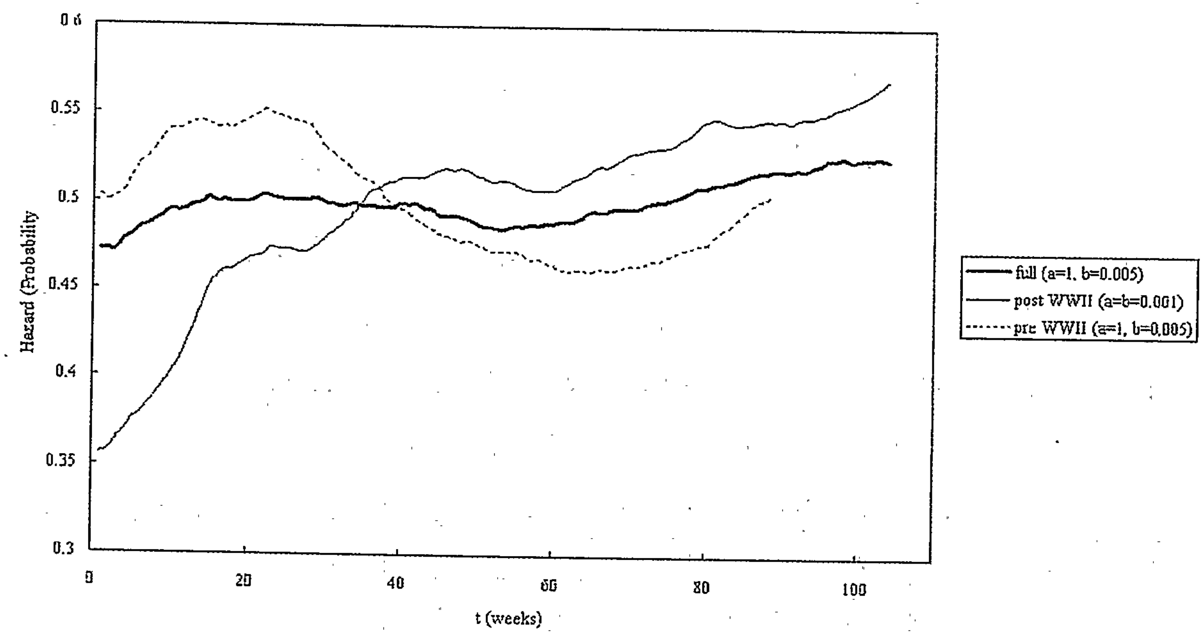


Figure 4.20

A Comparison of Bear Market Hazard Functions Across the Full, Post WWI and Pre WII Samples.

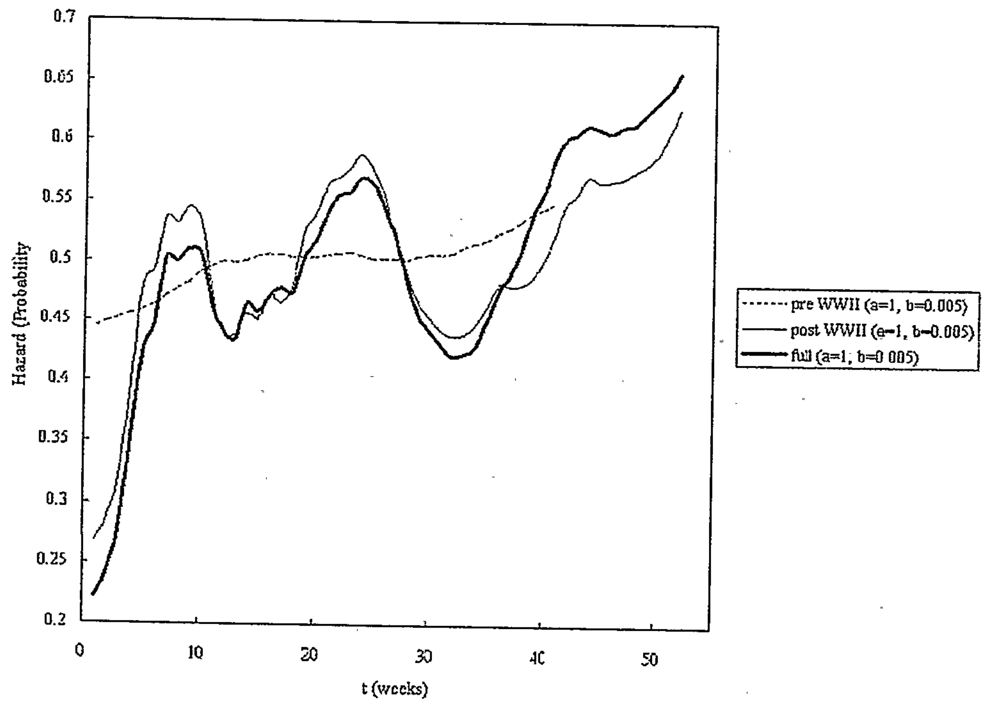
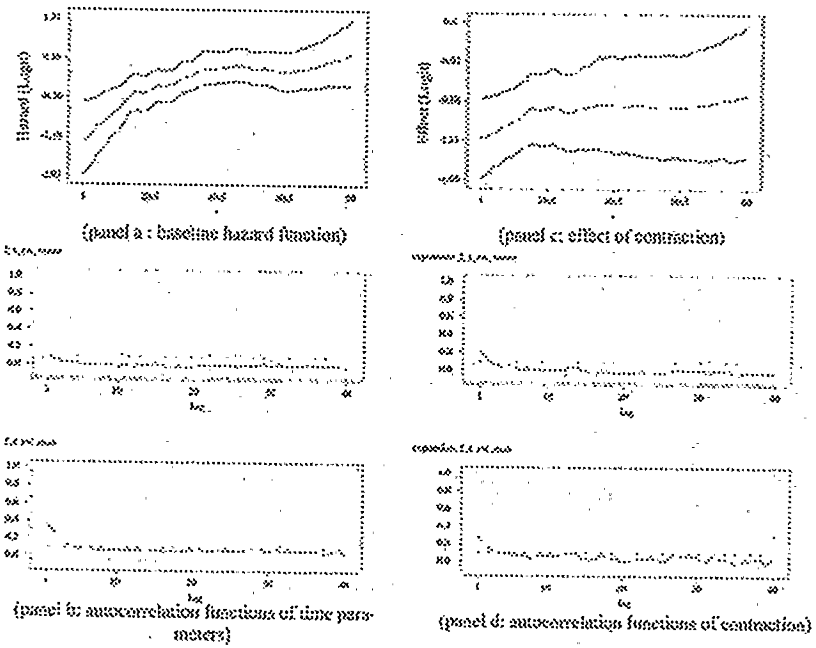


Figure 4.21

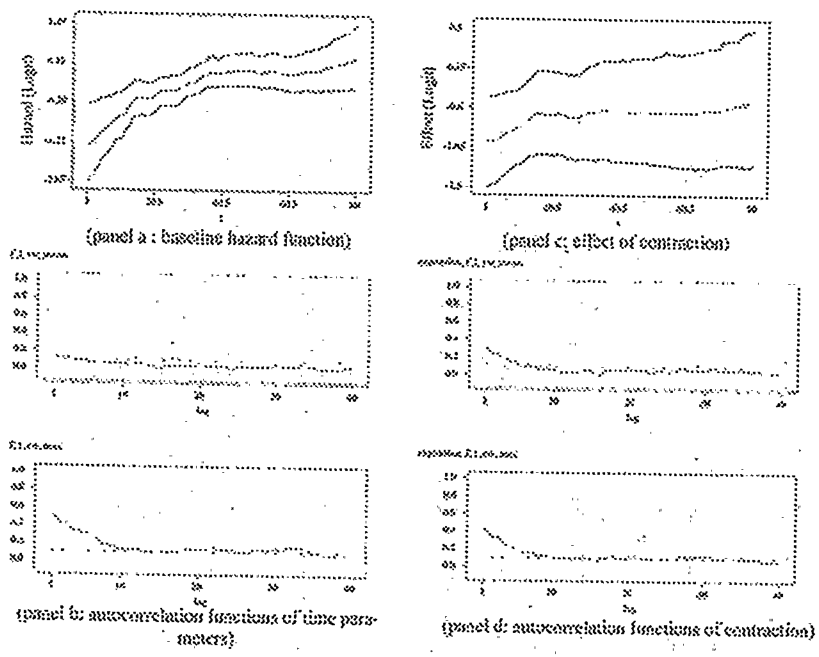
The Impact of Economic Expansion on Bull Market Duration for $a=0.001$ and $b=0.001$.



Note: Panels b and d each contain two graphs; the upper and lower graphs are the mean and maximum autocorrelation functions respectively for the corresponding parameters.

Figure 4.22

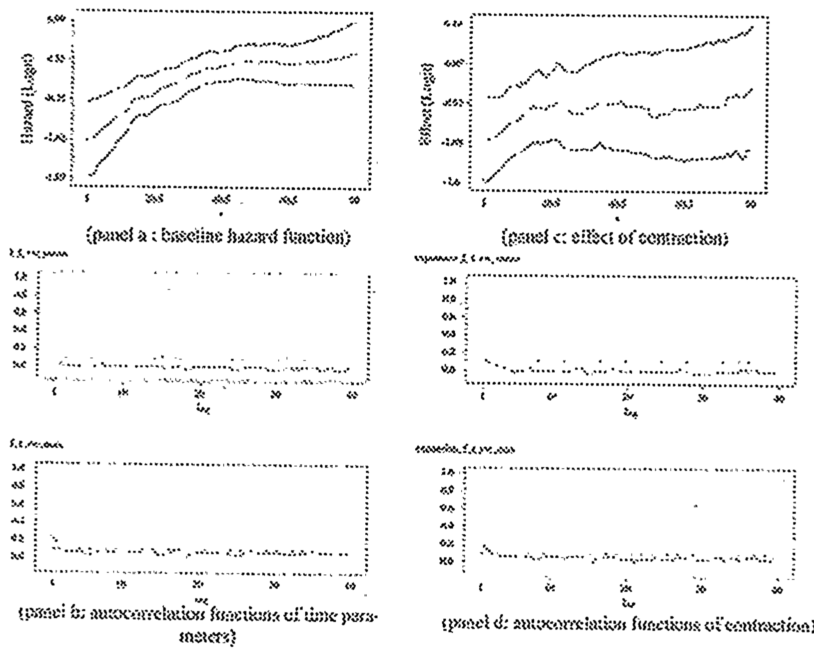
The Impact of Economic Expansion on Bull Market Duration for $a=1$ and $b=0.005$.



Note: Panels b and d each contain two graphs; the upper and lower graphs are the mean and maximum autocorrelation functions respectively for the corresponding parameters. Also, t denotes time and is measured in weeks.

Figure 4.23

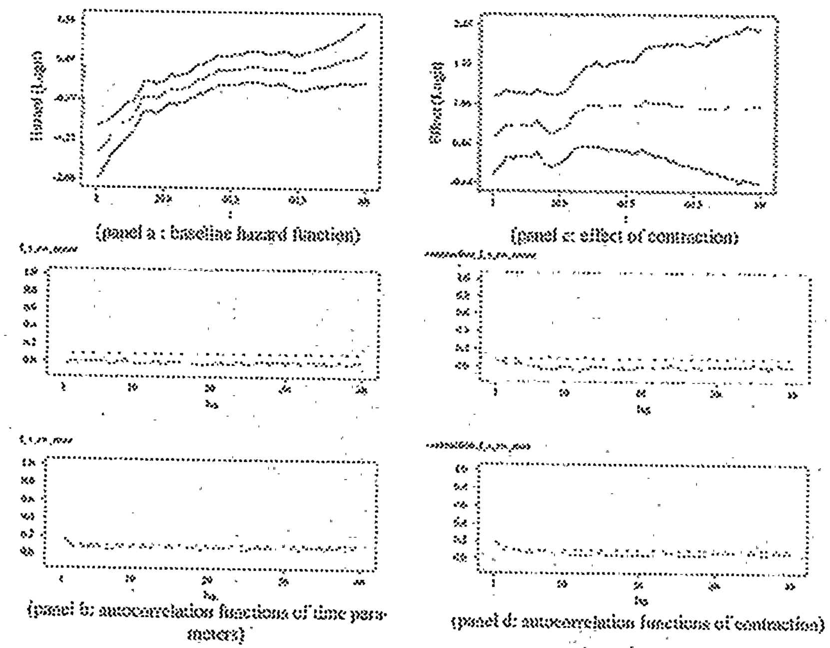
The Impact of Economic Expansion on Bull Market Duration for $a=0.0001$ and $b=0.0001$.



Note: Panels b and d each contain two graphs; the upper and lower graphs are the mean and maximum autocorrelation functions respectively for the corresponding parameters. Also, t denotes time and is measured in weeks.

Figure 4.24

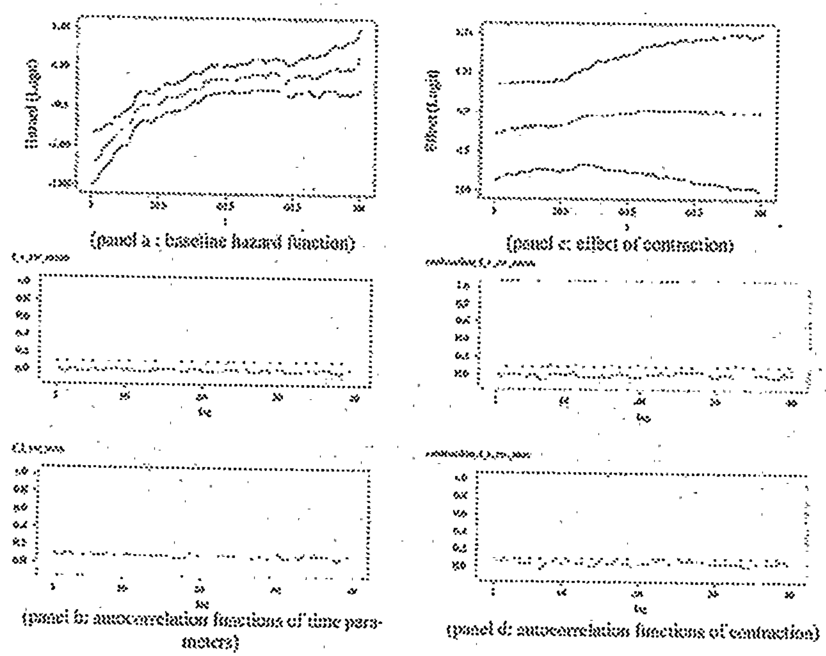
The Impact of Economic Contraction on Bull Market Duration for $a=0.001$ and $b=0.001$.



Note: Panels b and d each contain two graphs; the upper and lower graphs are the mean and maximum autocorrelation functions respectively for the corresponding parameters. Also, t denotes time and is measured in weeks.

Figure 4.25

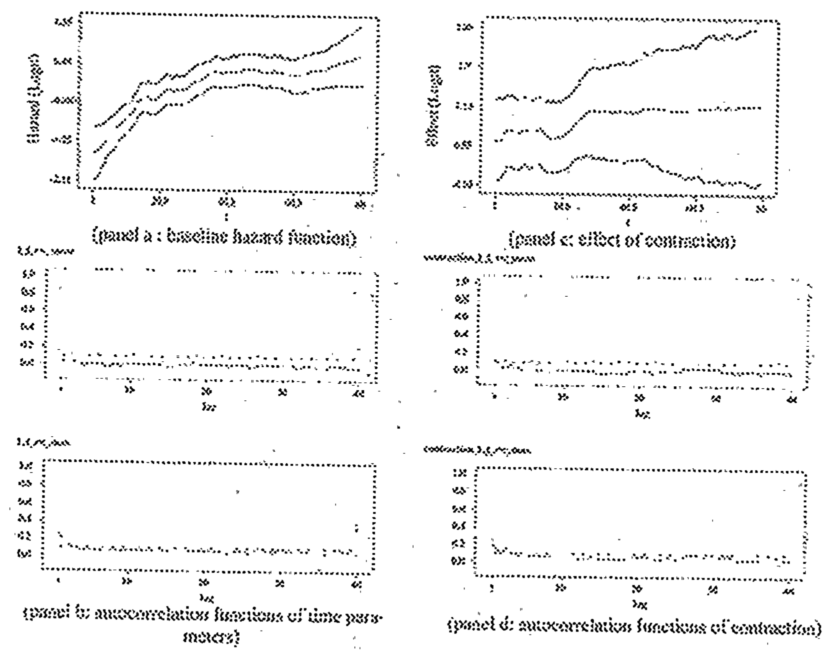
The Impact of Economic Contraction on Bull Market Duration for $a=1$ and $b=0.005$.



Note: Panels b and d each contain two graphs; the upper and lower graphs are the mean and maximum autocorrelation functions respectively for the corresponding parameters. Also, t denotes time and is measured in weeks.

Figure 4.26

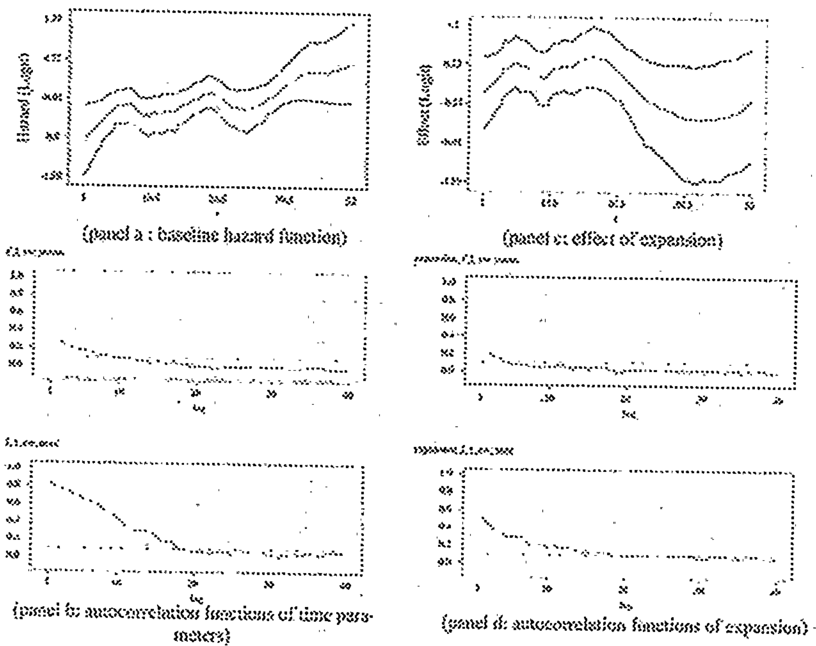
The Impact of Economic Contraction on Bull Market Duration for *a=0.0001* and *b=0.0001*.



Note: Panels b and d each contain two graphs; the upper and lower graphs are the mean and maximum autocorrelation functions respectively for the corresponding parameters. Also, *t* denotes time and is measured in weeks.

Figure 4.27

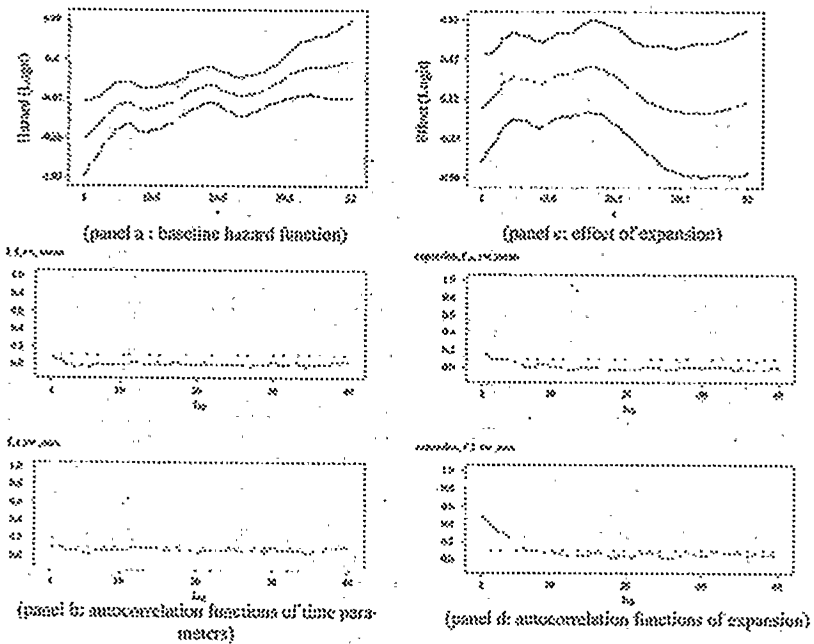
The Impact of Economic Expansion on Bear Market Duration for $a=0.001$ and $b=0.001$.



Note: Panels b and d each contain two graphs; the upper and lower graphs are the mean and maximum autocorrelation functions respectively for the corresponding parameters. Also, t denotes time and is measured in weeks.

Figure 4.28

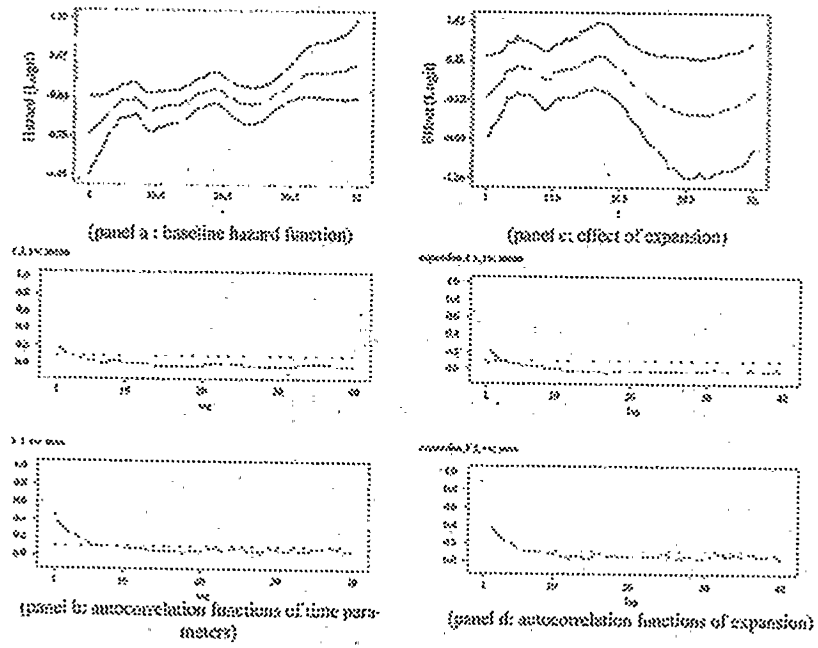
The Impact of Economic Expansion on Bear Market Duration for $a=1$ and $b=0.005$.



Note: Panels b and d each contain two graphs; the upper and lower graphs are the mean and maximum autocorrelation functions respectively for the corresponding parameters. Also, t denotes time and is measured in weeks.

Figure 4.29

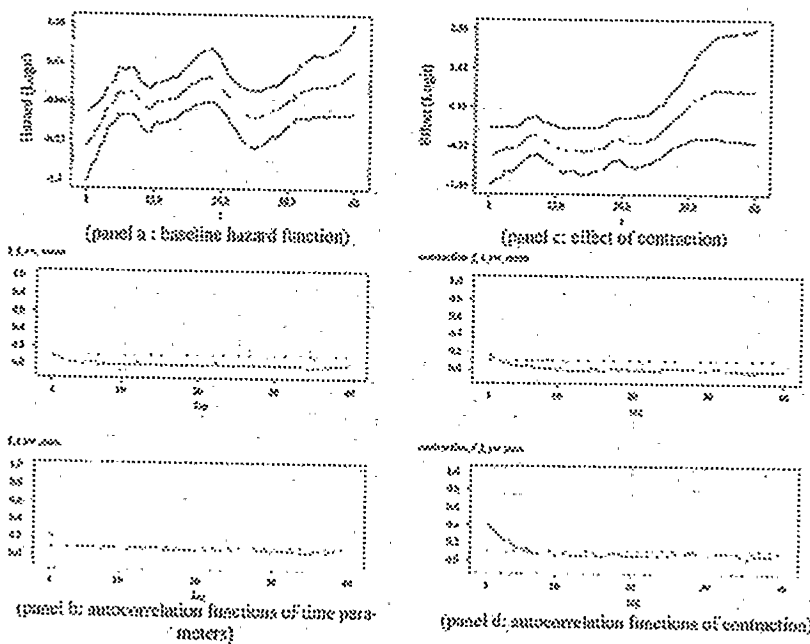
The Impact of Economic Expansion on Bear Market Duration for $a=0.0001$ and $b=0.0001$.



Note: Panels b and d each contain two graphs; the upper and lower graphs are the mean and maximum autocorrelation functions respectively for the corresponding parameters. Also, t denotes time and is measured in weeks.

Figure 4.30

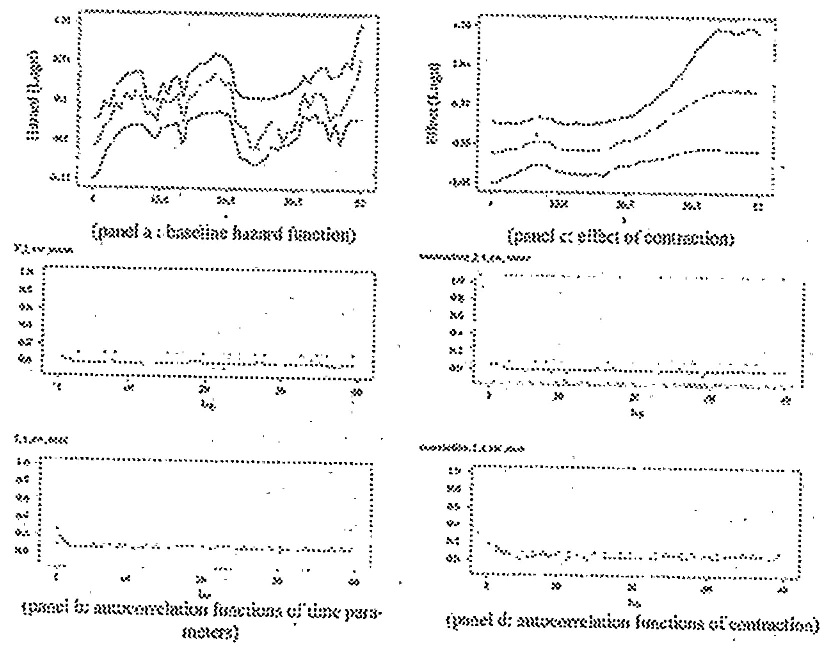
The Impact of Economic Contraction on Bear Market Duration for $a=0.001$ and $b=0.001$.



Note: Panels b and d each contain two graphs; the upper and lower graphs are the mean and maximum autocorrelation functions respectively for the corresponding parameters. Also, t denotes time and is measured in weeks.

Figure 4.31

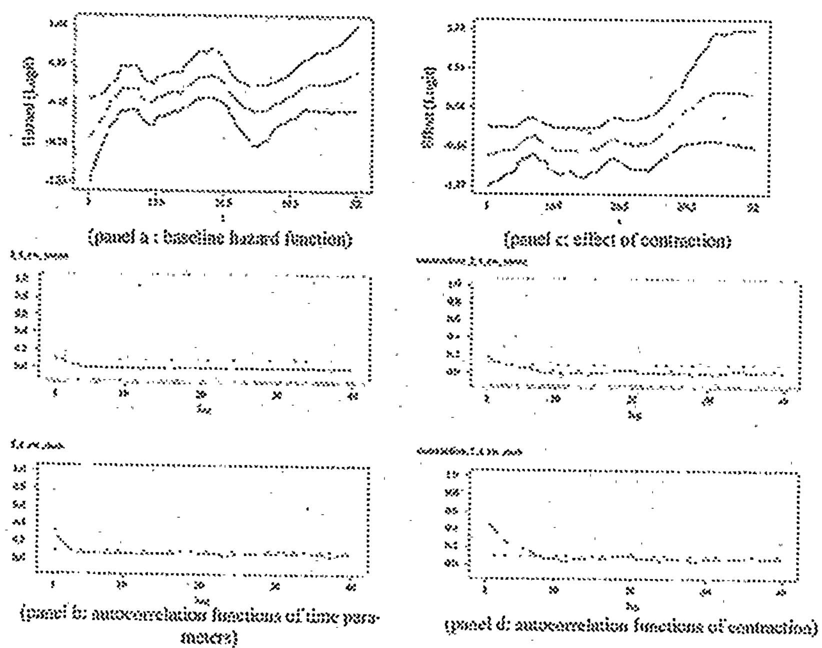
The Impact of Economic Contraction on Bear Market Duration for $a=1$ and $b=0.005$.



Note: Panels b and d each contain two graphs; the upper and lower graphs are the mean and maximum autocorrelation functions respectively for the corresponding parameters. Also, t denotes time and is measured in weeks.

Figure 4.32

The Impact of Economic Contraction on Bear Market Duration for $a=0.0001$ and $b=0.0001$.



Note: Panels b and d each contain two graphs; the upper and lower graphs are the mean and maximum autocorrelation functions respectively for the corresponding parameters. Also, t denotes time and is measured in weeks.

Chapter 5

Conclusion

The objective of this thesis is to investigate the behavior of stock prices by examining the duration aspects of bull and bear markets. Chapter 2 uses a selection algorithm that dissects a time series of the Dow Jones Industrial Average into bull and bear components. The summary statistics of the bull and bear markets suggest that bull markets are generally longer and stronger than bear markets.

In Chapter 3, we investigate whether stock prices exhibit mean reverting behavior by checking for the presence of positive duration dependence in bull and bear markets using continuous time and discrete time tests. From the continuous time realm, we use the Weibull test. From the discrete time realm, we use the zero plim test, method of moments test, generalized method of moments test, SB test and the Chi-Square goodness of fit test. In general, we find some evidence of positive duration dependence in bear and bull markets for the full sample as well as the post WWII sample. However, we fail to find any evidence for positive duration dependence for the pre WWII sample; we attribute this to the small sample size of pre WWII bull and bear markets.

In order to further check for positive duration dependence in bull and bear markets, Chapter 4 turns to the dynamic duration model to model the markets' hazard functions. In general, we found that the hazard functions of bull and bear markets are upward sloping functions suggesting positive duration dependence. Given that the estimation technique for the hazard functions involves the MCMC, it was necessary to check for proper mixing and the results' sensitivity to the choice of hyperparameters. We found that, in most cases, the shapes of the estimated hazard functions were not sensitive to the choice of hyperparameters whereas the mixing properties differ quite significantly across different sets of hyperparameters. In summary, together with the results from chapter 3, we identify evidence for positive duration dependence in bull and bear markets. This in turn suggests that stock prices do not follow the random walk but are mean reverting.

The other objective of this thesis was to examine the switching behavior of stock prices around business cycle turning points. Using the NBER business cycle dates we generated dummy variables that define economic expansions and contractions. Using these variables, we then used dynamic duration models and the MCMC technique to investigate whether the baseline hazard functions of bull and bear markets are responsive to changes in the underlying economic state.

We found that bull markets are sensitive to changes in the economy, even after accounting for statistical significance. In particular, a change from economic expansion to contraction tend to increase the hazard of bull markets, which would make stock returns more likely to be negative. On the other hand, a change from economic contraction to expansion tends decrease the hazard of bull markets, which would make stock returns more likely to be positive. These findings are conceptually similar to previous well documented studies using traditional time series techniques.

As for bear markets, we find quite the opposite. Ignoring statistical significance, we find that a change from economic expansion to contraction generally decreases the hazard of bear markets. Similarly, we find that a change from economic expansion to contraction tend to increase the hazard of bear markets. However, once statistical significance is accounted for, we find that the effect of changes in the economy is insignificant on the hazard of bear markets.

Bibliography

[1] Allison, P. D., 1982. Discrete-time methods for the analysis of event histories. Sociological Methodology 13, 61-89.

[2] Argas, E., Haara, P., 1987. A Logistic Regression Model for Hazard: Asymptotic Results. Scandinavian Journal of Statistics 14, 1-18.

[3] Boudohk, J., Richardson, M., 1994. The Statistics of Long-Horizon Regressions Revisited. Mathematical Finance 4, 103-119.

[4] Brown, C. C., 1975. On the use of indicator variables for studying the time dependence of parameters in a response-time model. Biometrics 31, 863-872.

[5] Bry, G., Boschan, C., 1971. Cyclical Analysis of Time Series: Selected Procedures and Computer Programs. NBER, New York.

[6] Casella, G., George, E., 1992. Explaining the Gibbs Sampler. American Statistician 46, 167-174.

[7] Chauvet, M., Petter, S., 2000. A Preference Regime Model of Bull and Bear Market. Journal of Empirical Finance 7, 87-111.

[8] Chen, N. F., Roll, R., Ross, S. A., 1986. Economic Forces and the Stock Market. Journal of Business 59.

[9] Chib, S., Greenberg, E., 1995. Understanding the Metropolis-Hastings Algorithm. The Statistical Association 90, 1313-1321.

[10] Cochran, S. J., DeFina, R. H., 1995a. Mean Reversion in Stock Prices: Tests Using Duration Models. Managerial Finance 21, 2-24.

[11] Cochran, S. J., DeFina, R. H., 1995b. Duration Dependence in the US Stock Market Cycle: A Parametric Approach. Applied Financial Economics 5, 309-318.

[12] Cox, D. R., 1972. Regression Models and Life Tables. Journal of the Royal Statistical Society B 34, 187-220.

[13] Davidson, R., Mackinnon, J., 2004. Econometric Theory and Methods. Oxford University Press, New York.

[14] DeStefano, M., 2004. Stock Returns and the Business Cycle. The Financial Review 39, 527-547.

[15] Diebold, F. X., Rudebusch, G. D., 1990. A Nonparametric Investigation of Duration Dependence in the American Business Cycle. Journal of Political Economy 98, 596-616.

[16] Durland, J. M., McCurdy, T. H., 1994. Duration Dependent Transitions in a Markov Model of U.S. GNP Growth. Journal of Business and Economic Statistics 12, 279-88.

[17] Efron, B., 1988. Logistic Regression, Survival Analysis, and the Kaplan Meier Curve. Journal of the American Statistical Association 83, 414-425.

[18] Engle, R., Russell, J., 1998. Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data. Econometrica 66, 1127-1162.

[19] Fahrmeir, L., 1994. Dynamic Modeling and Penalized Likelihood Estimation for Discrete Time Survival Data. Biometrika 81, 317-330.

[20] Fahrmeir, L., Wagenpfeil, S., 1996. Smoothing Hazard Functions and Time-Varying Effects in Discrete Duration and Competing Risks Models. Journal of Statistical Association 91, 1584-1594.

[21] Fahrmeir, L., 1992. Posterior Mode Estimation by Extended Kalman Filtering for Multivariate Dynamic Generalized Linear Models. Journal of American Statistical Association 87, 501-509.

[22] Fahrmeir, L., Kaufmann, H., 1991. On Kalman Filtering, Posterior Mode Estimation and Fisher Scoring in Dynamic Exponential Family Regression. Metrika 38, 37-60.

[23] Fahrmeir, L., Goss, M., 1992. "On Filtering and Smoothing in Dynamic Models for Categorical Longitudinal Data." Statistical Modeling. Editted by P. Heijden, W. Jansen, B. Francis and G. Seeber, North Holland, Amsterdam.

[24] Fahrmeir, L., Lang, S., 2001. Bayesian Semiparametric Regression Analysis of Multicategorical Time-Space Data. Annals of the Institute of Statistical Mathematics 53, 10-30.

[25] Fahrmeir, L., Knorr-Held, L., 1997. Dynamic Discrete-Time Duration Models: Estimation Via Markov Chain Monte Carlo. Sociological Methodology 27, 417-452.

[26] Fahrmeir, L., Tutz, G., 2001. Multivariate Statistical Modeling Based on Generalized Linear Models. Springer-Verlag, New York.

[27] Fama, E., French, K., 1988. Permanent and Temporal Components of Stock Prices. Journal of Political Economy 96, 246-273.

[28] Geman, D., Geman, S., 1984. Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. IEEE transations on Pattern Analysis and Machine Intelligence 6, 721-741.

[29] Gilks, W. R., Richardson, S., Spiegelhalter, D. J., 1996. Markov Chain Monte Carlo in Practice. Chapman and Hall, London.

[30] Gonzalas, L., Powell, J., Shi, J., Wilson, A., 2005. Two Centuries of Bull and Bear Market Cycles. International Review of Economics and Finance 14, 469-486.

[31] Hamilton, W. P., 1919. Stock Market Analysis. Wall Street Journal 9 August: 1919. Reprinted in Rhea R. 1932 The Dow Theory. Barron's, New York.

[32] Harding, D., Pagan, A., 2002. Dissecting the Cycle: a Methodological Investigation. Journal of Monetary Economics 49, 365-381.

[33] Kalfleisch, J., Prentice, A., 1980. The Statistical Analysis of Failure Time Data. Wiley, New York.

[34] Kennan, J., 1985. The Duration of Contract Strikes in US Manufacturing. Journal of Econometrics 28, 5-28.

[35] Kiefer, N. M., 1988. Economic Duration Data and Hazard Functions. Journal of Economic Literature 26, 646-679.

[36] King, R., Plosser, C., 1994. Real Business Cycles and the Test of the Adelmans. Journal of Monetary Economics 33, 405-438.

[37] King, G., Alt, J. E., Laver, M., Burns, N., 1990. A Unified Model of Cabinet Dissolution in Parliamentary Democracies. American Journal of Political Science 34, 847-871.

[38] Lancaster, T., 1979. Econometric Methods for Duration of Unemployment. Econometrica 47, 939-955.

[39] Lancaster, T., 1990. The Econometric Analysis of Transition Data. Cambridge University Press, New York.

[40] Lawless, J. F., 1982. Statistical Models and Methods for Lifetime Data. John Wiley and Sons, New York.

[41] Leroy, S., 1973. Risk Aversion and the Martingale Property of Stock Returns. International Economic Review 14, 436-446.

[42] Lo, A., MacKinlay, A., 1988. Stock Market Prices Do Not Follow Random Walks: Evidence From a Simple Specification Test. Review of Financial Studies 1, 41-66.

[43] McQueen, G., Thorley, S., 1994. Bubbles, Stock Returns, and Duration Dependence. Journal of Financial and Quantitative Analysis 29, 79-401.

[44] McQueen, G., Roley, V. V., 1993. Stock Prices, News, and Business Conditions. Review of Financial Studies 6, 683-707.

[45] Mantel, N., Hankey, B. F., 1978. A Logistic Regression Analysis of Response-Time Data Where the Hazard Function is Time Dependent. Communications in Statistics - Theory and Methods A7, 333-347.

[46] Mudambi, R., Taylor, L., 1991. A Nonparametric Investigation of Duration Dependence in the American Business Cycle: A Note. Journal of Political Economy 99, 654-656.

[47] Mudambi, R., Taylor, L., 1995. Some Nonparametric Test for Duration Dependence: An Application To UK Business Cycle Data. Journal of Applied Statistics 22, 163-177.

[48] Nickel, S., 1979. Estimating the Probability of Leaving Unemployment. Econometrica 47, 1249-1266.

[49] Pakes, A., Schankerman, M., 1984. The Rate of Obsolescence of Knowledge, Research Lags, and the Private Rate of Return to Research Resources. in Z. Griliches, Ed., Patents, R&D and Productivity. University of Chicago Press, Chicago.

[50] Ohn, J., Taylor, L., Pagan, A., 2004. Testing for Duration Dependence in Economic Cycles. Econometrics Journal 7, 528-549.

[51] Pagan, A., Sossounov, K., 2003. A Simple Framework For Analyzing Bull and Bear Markets. Journal of Applied Econometrics 18, 23-46.

[52] Partington, G., Stevenson, M., 2001. The Probability and Timing of Price Reversals in the Property Market. Managerial and Decision Economics 22, 389-398.

[53] Perex-Quiros, G., Timmermann, A., 1998. "Variations in the Mean and Volatility of Stock Returns Around Turning Points of the Business Cycle." in J. Knight and S. Satchell, eds., Forecasting Volatility in the Financial Markets. Butterworth-Heinemann, Woburn, Massachusetts.

[54] Poterba, J., Summers, L., 1988. Mean Reversions in Stock Prices: Evidence and Implications. Journal of Financial Economics 22, 27-59.

[55] Samuelson, P., 1965. Proof that Properly Anticipated Prices Fluctuate Randomly. Industrial Management Review 6, 41-49.

[56] Schaller, H., Van Norden, S., 1977. Regime Switching in Stock Market Returns. Applied Financial Economics 7, 177-191.

[57] Schwart, G. W., 1989. Business Cycles, Financial Crises and Stock Volatility. Carnegie-Rochester Conference Series on Public Policy 31, 83-126.

[58] Shiller, R. K., 1989. Market Volatility. MIT Press,Cambridge, MA.

[59] Singer, J. D., Willet, J. B., 2003. Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence. Oxford University Press, New York.

[60] Tauchen, G., 1985. Diagnostic Testing and Evaluation of Maximum Likelihood Models. Journal of Econometrics 30, 415-443.

[61] Thompson, W. A., 1977. On the treatment of grouped observations in life studies. Biometrics 33, 463-470.

[62] Timmermann, A., Lunde, A., 2004. Duration Dependence in Stock Prices: An Analysis of Bull and Bear Markets. Journal of Business and Economic Statistics 22(3), 253-273.

[63] Tsay, R. S., 2002. Analysis of Financial Time Series. John Wiley and Sons, New York.

[64] Turner, C. M., Startz, R., Nelson, C. R., 1989. A Markov Model of Heteroscedasticity, Risk and Learning in the Stock Market. Journal of Financial Economics 25, 3-22.

[65] Vilcassim, N. J., Jain, D. C., 1991. Modeling Purchase Timing and Brand Switching Behavior Incorporating Explanatory Variables and Unobserved Heterogeneity. Journal of Marketing Research 28, 29-41.

[66] Watson, M., 1994. Business Cycle Durations and Postwar Stabilization of the U.S. Economy. American Economic Review 84, 24-46.

[67] Yamaguchi, K., 1993. Modeling Time-Varying Effects of Covariates in Event History Analysis Using Statistics from the Saturated Hazard Rate Model. Sociological Methodology 23, 279-317.

[68] Zuehlke, T. W., 2003. Business Cycle Duration Dependence Reconsidered. Journal of Business and Economic Statistics 21, 564-569.