Duration Dependence in Stock Prices: An Analysis of Bull and Bear Markets*

ASGER LUNDE

Department of Information Science,
The Aarhus School of Business, Fuglesangs Allé 4, DK-8210 Aarhus V, Denmark
Email: alunde@asb.dk

ALLAN TIMMERMANN

Department of Economics, University of California, San Diego,
9500 Gilman Drive La Jolla, CA 92093-0508, USA
Email: atimmerm@ucsd.edu

First version may 1999. This version July 9, 2002

ABSTRACT

This paper studies time series dependence in the direction of stock prices by modelling the (instantaneous) probability that a bull or bear market terminates as a function of its age and a set of underlying state variables such as interest rates. A random walk model is rejected both for bull and bear markets. Although it fits the data better, a GARCH model is also found to be inconsistent with the very long bull markets observed in the data. The strongest effect of increasing interest rates is found to be a lower bear market hazard rate and hence a higher chance of continued declines in stock prices.

*This paper benefitted from many insightsful comments from two anonymous referees. We thank Frank Diebold, Graham Elliott, Rob Engle, Essie Maasoumi, Adrian Pagan, Josh Rosenberg, Ruth Williams and seminar participants at New York University, Vanderbilt University, Arizona State University, Aalborg University, UC Davis, Federal Reserve Board, the CORE, the INQUIRE 2000 World meetings in San Diego, the CF/FFM2000 conference in London (we thank the Center for Analytical Finance (CAF) in Aarhus for supporting this presentation) and the Eigth Econometric Society World Congress in Seattle for many helpful suggestions. We are grateful to INQUIRE UK for financial support for this research.

1. INTRODUCTION

Since the seminal work by Samuelson (1965) and Leroy (1973), the random walk and martingale models of stock prices have formed the cornerstone of modern finance. Hence it is not surprising that an extensive empirical literature has considered deviations from these benchmark models. Several authors, including Lo & MacKinlay (1988), Fama & french (1988), Poterba & Summers (1988), Richardson & Stock (1989) and Boudokh & Richardson (1994) study long-run serial correlations in stock returns. Although this literature finds indications of a slowly mean reverting component in stock prices, deviations from normally distributed returns, time-varying volatility and small sample sizes have plagued existing tests and made it difficult to conclusively reject the random walk model.

This paper proposes a new approach to modelling time series dependence in stock prices that allow bull and bear hazard rates, i.e. the probability that a bull or bear market terminates next period, to depend on the age of the market. Inspection of these hazard rates yields new insights into long-run dependencies and deviations from parametric models of asset prices proposed in the literature, including the simple random walk model with a constant drift and models that allow for volatility persistence.

By explicitly focusing on duration dependence in stock prices, the proposed tests are very different from the tests based on autocorrelations previously adopted in the literature. The tests are more closely related to the duration dependence measures first proposed in the context of regime switching models used to analyze GDP growth (Durland & McCurdy (1994)) or stock market prices (Maheu & McCurdy (2000*a*)). Although the two approaches share some of the same objectives, namely to capture potential duration-dependence, they differ both in terms of the definition of the underlying states and in terms of econometric methodologies.

Our approach does not require that stock prices follow a low-order Markov process although this is a special case of our model when termination probabilities are memoryless. Faust (1992) demonstrated that existing tests for autocorrelation based on variance ratios have optimal power in testing the random walk hypothesis against certain classes of stationary ARMA models. However, these models only form a small subset of the alternatives that are interesting from an economic point of view, such as nonlinear speculative bubble processes or processes where the drift depends on past cumulated returns within a state. There is no result for the power of autocorrelation tests against nonlinear alternatives or processes with long memory. This is important since there is mounting evidence of nonlinearities such as a switching factor in the mean and volatility of stock returns, c.f. Maheu & McCurdy (2000*a*) and Perez-Quiros & Timmermann (2000).

We formalize bull and bear states in terms of a filter that tracks movements between local peaks and troughs. Earlier studies such as Fabozzi & Francis (1977), Kim & Zumwalt (1979) and Chen (1982) consider definitions of bull markets based simply on returns in a given month exceeding a certain threshold value. Fabozzi &

Francis (1977) also consider a definition of bull markets based on ’substantial’ up and down movements. In this definition, a substantial move in stock prices occurs whenever the absolute value of stock returns in a given month exceeds half of one standard deviation of the return distribution. Such definitions do not reflect long-run dependencies in stock prices and ignore information about the trend in stock price levels. According to our definition, the stock market switches from a bull to a bear state if stock prices have declined by a certain percentage since their previous (local) peak within that bull state. Likewise, a switch from a bear to a bull state occurs if stock prices experience a similar percentage increase since their local minimum within that state. This definition does not rule out sequences of negative (positive) price movements in stock prices during a bull (bear) market as long as their cumulative value does not exceed a certain threshold. By abstracting from the small unsystematic price movements that dominate time series as noisy as daily price changes this definition is designed to capture long-run dependencies in the underlying drift in stock prices.

We find evidence of distinctly different duration dependence in bull and bear states. Controlling for interest rates, the longer a bull market has lasted, the lower its hazard rate and hence the lower the probability that it terminates next period. In contrast, in bear markets the hazard rate tends to be highest at long durations. Interest rates are also found to have an important effect on hazard rates. Increases in the interest rate are associated with a small increase in the bull market hazard rate and a large decrease in the bear hazard rate. They are therefore associated with a higher probability of being in a bear market and a lower bull market probability. The finding of a hazard function that depends on the age of the state suggests that stock prices do not follow a low order Markov process. Instead the drift and the effect of interest rates on stock prices appear to be related to the market’s memory of the time spent in the current state. This means that the effect of an interest rate change on stock prices depends on the age and type of the state where the change occurs.

Most closely related to the current study are the papers by Maheu & McCurdy (2000*a*) and Pagan & Sossounov (2000). Pagan & Sossounov (2000) also consider a definition of bull and bear states based on cumulative price changes. Their study characterizes movements in stock prices through the average duration and amplitude of bull and bear markets. The Pagan-Sossounov dating method is based on a modification of the Bry-Boschan algorithm and seeks out local peaks and troughs within a predetermined window of data points subject to a set of censoring rules that restrict both the minimum length of the bull/bear market (4 months) as well as the minimum duration of a full cycle (16 months), c.f. appendix B of their paper. Our filter does not use restrictions on the minimal length of the bull/bear states but requires choosing the value of the initial state and setting threshold values for the size of cumulative movements in stock prices that trigger a switch between these states.

Maheu & McCurdy (2000*a*) treat the state as an unobserved variable and classify monthly stock returns into two latent states based on a Markov switching model extended to account for duration dependent state transitions. While our regimes are defined according to the cumulative movements in stock prices and thus track local peaks and troughs in stock prices, their Markov switching approach endogenously identifies a high mean, low variance bull state and a low mean, high variance bear state. This approach tends to identify turning points more frequently than ours. When the regime switching is restricted to only occur in the mean, in a 'decoupled' model, a state with a large negative mean return and a state with a small positive mean return are identified. Maheu & McCurdy (2000*b*) further study a model where regimes are present in the volatility but not in the mean. Whether an endogenous or an exogenous determination of the filter size is most appropriate depends, of course, on the purpose of the analysis.

The plan of the paper is as follows. Section 2 presents our definition of bull and bear market states. Section 3 characterizes the unconditional distribution of the durations and returns in bull and bear markets using more than a century of daily stock prices from the US. Section 4 considers formal test for duration dependence while Section 5 discusses estimation of bull and bear markets whose hazard rate may depend on the age of the market. Section 6 reports empirical results and undertakes a scenario analysis to investigate the effect of an interest rate change on bull and bear hazard rates. Section 7 briefly discusses further economic interpretations of our findings.

2. Definition of Bull and Bear Markets

Financial analysts and stock market commentators frequently classify the underlying trend in stock prices into bull and bear markets. Durations of bull and bear markets are key components of the risk and return characteristics of stock returns so it is clearly important to understand their determinants. Yet, despite their importance, little work has been done on formalizing these concepts and investigate whether bull and bear states provide a useful way of characterizing long-run dependencies in stock prices. It is not clear, for example, whether such states serve a purely descriptive purpose or whether the knowledge that stock prices have been in a particular state for a certain length of time affects the conditional distribution of future price movements. In the latter case, investment performance could conceivably be improved by conditioning on the type and age of the current state.

There is no generally accepted formal definition of bull and bear markets in the finance literature. One of the few sources that attempts a definition of bull and bear markets is Sperandeo (1990) who defines bull and bear markets as follows:

"**Bull market**: A long-term ... upward price movement characterized by a series of higher intermediate ... highs interrupted by a series of higher intermediate lows.

Bear market: A long-term downtrend characterized by lower intermediate lows interrupted by lower intermediate highs". (p. 102).

In a more recent contribution Chauvet & Potter (2000) offer a similar definition. To formalize the idea of a series of increasing highs interrupted by a series of higher intermediate lows, let I_t be a bull market indicator variable taking the value 1 if the stock market is in a bull state at time t, and zero otherwise. We assume that time is measured on a discrete scale and that the stock price at the end of period t is P_t. Let λ_1 be a scalar defining the threshold of the movements in stock prices that trigger a switch from a bear to a bull market, while λ_2 is the threshold for shifts from a bull to a bear market. Suppose that at t_0 the stock market is at a local maximum ($I_{t_0} = 1$) and set $P_{t_0}^{\max} = P_{t_0}$, where P_{t_0} is the value at time t_0 of the stochastic process tracking the stock price. Let $\tau_{\max}$ and $\tau_{\min}$ be stopping time variables defined by the following conditions:

$$\begin{aligned} \tau_{\max}(P_{t_0}^{\max}, t_0 | I_{t_0} &= 1) = \inf\{t_0 + \tau : P_{t_0+\tau} \geq P_{t_0}^{\max}\}, \\ \tau_{\min}(P_{t_0}^{\max}, t_0, \lambda_2 | I_{t_0} &= 1) = \inf\{t_0 + \tau : P_{t_0+\tau} < (1 - \lambda_2) P_{t_0}^{\max}\}, \end{aligned} \tag{1}$$

where $\tau \geq 1$. Then $\min(\tau_{\max}, \tau_{\min})$ is the first time the price process crosses one of the two barriers $\{P_{t_0}^{\max}, (1 - \lambda_2) P_{t_0}^{\max}\}$. If $\tau_{\max} < \tau_{\min}$, we update the local maximum in the current bull market state:

$$P_{t_0+\tau_{\max}}^{\max} = P_{t_0+\tau_{\max}}, \tag{2}$$

and the bull market continued between $t_0 + 1$ and $t_0 + \tau_{\max}$: $I_{t_{0+1}} = ... = I_{t_0+\tau_{\max}} = 1$.

Conversely, if $\tau_{\min} < \tau_{\max}$ so that the stock price at $t_0 + \tau_{\min}$ has declined by a fraction λ_2 since its local peak

$$P_{t_0+\tau_{\min}} < (1 - \lambda_2) P_{t_0}^{\max}, \tag{3}$$

then the bull market has switched to a bear market that prevailed from $t_0 + 1$ to $t_0 + \tau_{\min}$: I_{t_0+1} = ... = $I_{t_0+\tau_{\min}}$ = 0. In the latter case we set $P_{t_0+\tau_{\min}}^{\min} = P_{t_0+\tau_{\min}}$.

If the starting point at t_0 is a bear market state, the stopping times get defined as follows:

$$\begin{aligned} \tau_{\min}(P_{t_0}^{\min}, t_0 | I_{t_0} &= 0) = \inf\{t_0 + \tau : P_{t_0+\tau} \leq P_{t_0}^{\min}\}, \\ \tau_{\max}(P_{t_0}^{\min}, t_0, \lambda_1 | I_{t_0} &= 0) = \inf\{t_0 + \tau : P_{t_0+\tau} > (1 + \lambda_1) P_{t_0}^{\min}\} \end{aligned} \tag{4}$$

This definition of bull and bear states partitions the data on stock prices into mutually exclusive and exhaustive bull and bear market subsets based on the sequences of first passage times. The resulting indicator function, I_t, gives rise to a random variable, T, which measures the durations of bull or bear markets. These are simply given as the time between successive switches in I_t.

We consider a range of values for λ_1 and λ_2. The smaller the values these parameters are set at, the more bull and bear market spells we expect to see. This is likely to improve the power of our statistical tests as the sample size used in the duration analysis increases. However, there are also limits to how low λ_1 and λ_2 can be set since too small values will lead our analysis to capture short-term dynamics in stock price movements. A value of $\lambda_2 = 0.20$ is conventionally used in the financial press so we entertain this along with smaller values. Setting $\lambda_1 > \lambda_2$ provides a way to account for the upwards drift in stock prices which works against finding many bear markets. We consider the following four filters (λ_1, λ_2) expressed in percentage terms: (20,15), (20,10), (15,15) and (15,10). The (15,15) parameterization allows us to study the effect of using a symmetric filter.

The focus on local peaks and troughs allows us to concentrate on the systematic up and down movements in stock prices and to filter out short term noise. This is an important consideration for data as noisy as daily stock price changes. While some arguments can be made in favor of imposing an additional minimum duration constraint, this also adds an extra layer of complexity and means that the data has to be filtered through a recursive pattern recognition algorithm, as explained by Pagan & Sossounov (2000). Instead our approach models both short and long durations, but allows the hazard rate to differ across durations.

Naturally our filter is related to a long literature on technical trading rules that models local trends in stock prices, c.f. Brock, Lakonishok & Lebaron (1992), Brown, Goetzmann & Kumar (1998) and Sullivan, Timmermann & White (1999). However, the similarities between technical trading rules and duration measures are only superficial. Technical trading rules search for patterns in prices conditional on a time horizon that is typically quite short. For example, the value of a 100-day moving average of prices may be compared to the value of a 25-day moving average. In contrast, we do not condition on the time of a particular movement but instead explicitly treat this as a random variable whose distribution we are interested in modelling.

3. Durations of Bull and Bear Markets

3.1. Data

To investigate the properties of bull and bear market states along the definition proposed in Section 2, we construct a data set of daily stock prices in the US from 2/17/1885 to 12/31/1997. From 2/17/1885 to 2/7/1962 the nominal stock price index is based on the daily returns provided by Schwert (1990). These returns include dividends. From 3/7/1962 to 12/31/1997 the price index is constructed from daily returns on the Standard & Poors 500 price index, again including dividends and obtained from the CRSP tapes. Combining these series generates a time series of 31,412 daily nominal stock prices.

Inflation has varied considerably over the sample period and it can be argued that the drift in nominal prices does not have the same interpretation during low and high inflation periods. To deal with effects arising from this we consider both nominal and real stock prices. To construct real stock prices, we build a daily inflation index as follows. We use monthly data on the consumer price index taken from Shiller (2000) and convert it into daily inflation rates by solving for the daily inflation rate such that the daily price index grows smoothly - and at the same rate - between subsequent values of the monthly consumer price index. Finally we divide the nominal stock price index by the consumer price index to get a daily index for real stock prices. Since the volatility of daily inflation rates is likely to be only a fraction of that of daily stock returns, normalizing by the inflation rate has the effect of a time-varying drift adjustment. Lack of access to daily inflation data is unlikely to affect our results in any important way.

We will also consider the effect of time-varying interest rates on hazard rates. Since there is no continuous data series on daily interest rates from 1885 to 1997, we construct our data from four separate sources. From 1885 to 1889 the source is again Shiller (1989). From 1890 to 1925, we use the interest rate on 90-day stock exchange time loans as reported in Banking and Monetary Statistics, Board of Governors of the Federal Reserve System (1943). These rates are provided on a monthly basis and we convert them into a daily series by simply applying the interest rate reported for a given month to each day of that month. From 1926 to 1954, we use the one-month T-bill rates from the risk-free rates file on the CRSP tapes, again reported on a monthly basis and converted into a daily series. Finally, from July 1954 to 1997, we use the daily Federal Funds rate. These three sets of interest rates are concatenated to form one time series covering the full sample. We consider nominal interest rates in the analysis of nominal stock prices and real interest rates - computed as the difference between the daily interest and inflation rate - in the analysis of real stock prices.

Much of standard survival analysis in economics and finance assumes continuously measured data. However, since we use daily data and do not follow price movements continuously, our data is interval censored and the termination of our durations is only known to lie between consecutive follow ups. Effectively the measurement of T, the duration of bull and bear markets, is divided into A intervals

$$[a_0, a_1), [a_1, a_2), \dots, [a_{q-1}, a_q), [a_q, \infty) \text{ where } q = A - 1,$$

and only the discrete time duration $T \in \{1, \dots, A\}$ is observed, where $T = t$ denotes termination within the interval $[a_{t-1}, a_t)$. Although we draw on approaches from the literature on economic duration data (see, e.g., Kiefer (1988), Kalbfleisch & Prentice (1980) and Lancaster (1990)) this also means that we have to be careful in modifying the standard tools from continuous time analysis.

3.2. Random Walk Model

The natural benchmark that has underpinned much of the analysis in financial economics is the random walk model. This model, which is also considered by Pagan & Sossounov (2000), assumes constant mean and volatility of stock price changes. To capture the duration distribution generated by the random walk model, we simulate 2000 samples of 31,412 daily price observations from this model. Both the drift and volatility parameters were estimated to match the real or nominal stock price data.

Though simple, the random walk model has been surprisingly difficult to reject in many statistical tests and it is also quite flexible in fitting bull and bear durations, c.f. Pagan & Sossounov (2000). For example, it is able to generate longer bull market durations than bear durations simply by introducing a positive drift parameter.

3.3. Volatility Clustering

In many regards the pure random walk model for stock prices is too simple. Most notably, the volatility of daily stock price changes is strongly serially correlated and this will undoubtedly affect the duration distribution of our data. We are interested in documenting possible time-series dependencies in the drift of stock prices but also want to account for the effects of time-varying conditional volatility. To accomplish this and account for both short- and long-run persistence in volatility, we estimate the following components GARCH model proposed by Engle & Lee (1999) and extended to include an ARCH-in-mean effect:

$$\begin{aligned} r_t &= \mu + \beta r_{t-1} + \gamma \sigma_t + \varepsilon_t, \;\; \varepsilon_t \sim N(0, \sigma_t^2), \\ \sigma_t^2 &= q_t + \alpha(\varepsilon_{t-1}^2 - q_{t-1}) + \beta(\sigma_{t-1}^2 - q_{t-1}) \\ q_t &= \omega + \rho(q_{t-1} - \omega) + \phi(\varepsilon_{t-1}^2 - \sigma_{t-1}^2). \end{aligned} \tag{5}$$

We then compare the duration distribution of bull and bear market spells generated by this model to the duration distribution from the actual data. Again we simulate 2000 samples each with 31,412 observations. Comparing the duration distribution under the random walk model to that under the GARCH model provides a way to account for the effect of volatility clustering.

3.4. Bull and Bear Durations

Insight into how our definition partitions stock prices into bull and bear spells is gained from Figure 1 which uses the real and nominal price indices to show the sequence of consecutive bull and bear markets over the full sample period 1885-1997. To better illustrate the individual episodes, we plot the natural logarithm of the real and nominal stock price index in eight separate windows. The figure uses a 20/10 filter that splits the sample into 66 bull and bear markets. Many of the bull markets are very long, the longest lasting from 1990 to 1997. Most of the short durations occur from 1929 to 1934 around the time of the Great Depression. For this period the number of turning points identified by our method exceeds that found by Pagan & Sossounov (2000). This is a result of the increased stock market volatility during this period in conjunction with our turning point method that does not impose a minimum length on each state.

Table 1 presents descriptive statistics for the distribution of bull and bear market durations. Properties of bull and bear market states are reported in weeks although it should be recalled that our analysis was carried out using daily data. We report results for the four filter sizes (λ_1, λ_2) described in the previous section. As the filter size declines the means of the bull and bear durations decline as do their standard deviations. Since lower thresholds will be crossed more frequently, this is to be expected and this effect is also observed in the data generated by the random walk and GARCH models. Using nominal prices, the mean bull market duration is between two and three times as long as the mean bear duration, suggesting that between two thirds and three quarters of the time is spent in the bull state. This proportion declines somewhat once we use real stock prices which have a smaller positive drift. For real stock prices the mean bull market duration exceeds the mean bear market duration by a factor between 1.5 and 2.3, depending on the filter size.

The standard deviation of the bull market durations observed in the actual data is systematically greater than that generated by the random walk and GARCH models. Conversely, the standard deviation of the bear market durations in the actual data is lower than that produced by the benchmark models. This is reflected in the much larger (smaller) values of the maximum duration for the actual bull (bear) market sample than in the simulated data. Though by no means formally proving that the benchmark models are unable to match the duration distribution observed for the actual data, these observations nevertheless indicate that the longest bull markets last too long while the longest bear markets last too short to be compatible with our benchmark models.

The random walk model with a non-zero drift can capture some features of bull and bear markets such as the asymmetry between bull and bear durations when a symmetric 15/15 filter is used. However, consistent with the findings in Pagan & Sossounov (2000), this model also has clear shortcomings that are at odds with the data. Most notably, it cannot capture the behavior of the longest bull markets or the shortest bear markets such as those that occurred at the end of the 1920s.

An alternative look at bull and bear durations is provided in Figure 2 which, for the four filters, plots the estimated densities of bull and bear market durations using a Gaussian kernel smoother. The first two rows show results for nominal stock prices while the last two rows are based on real prices. We also plot the density for the simulated random walk and GARCH models. There are major differences in the duration densities of bull and bear markets relative to what one would expect under the benchmark models, even after adjusting for time-varying volatility as is done in the GARCH model. The bull market density is lower at short durations and higher at long durations compared to the simulated models. In contrast, at short durations the bear duration density is generally higher in the actual data than in the simulations. Notice how GARCH effects generally lead to a higher probability of very short bull or bear markets compared to the random walk model. This is consistent with the intuition that volatility effects are strongest at high frequencies.

To see how much returns vary across bull and bear states, Table 2 reports return statistics for these states.

Depending on the filter size, mean returns vary from 1.5 to 2.5 percent per week in bull markets. In bear markets the mean return varies from -3.4 to -1.5 percent per week. These figures are computed as the mean return per bull or bear market converted into a weekly number. The unconditional real stock return of six percent per annum can be computed as the duration-weighted average return per bull or bear market spell.

A larger asymmetry shows up in the median returns which are generally 50 to 100 percent larger in absolute terms for bear than for bull markets. Although bear markets are much shorter than bull markets, the downward drift in bear markets is thus stronger than the upward drift during bull markets. Interestingly, the extent of this asymmetry is not captured by the random walk or GARCH models which are far closer to symmetry in median returns.

4. Statistical Tests of Differences in Duration Distributions

So far we have uncovered a number of disparities between the duration distribution observed in the actual data and that generated by the simulated benchmark models, but we have not formally tested whether the two sets of duration distributions really differ. In this section we apply a battery of tests to compare the distributional properties of the data against those produced by the benchmark models.

Since there is no closed form distribution for any of the duration models, we apply nonparametric two-sample tests to compare the actual and simulated data. Detailed descriptions of these tests can be found in Hollander & Wolfe (1999). In all cases we have $N = m + n$ duration spells $x_1, \ldots, x_m$ and $y_1, \ldots, y_n$ from the two distributions. Under the null it is assumed that the x's and y's are mutually independent draws from continuous distribution functions F and G, respectively. To test for equality of the mean duration, we first apply the Wilcoxon, Mann and Whitney test:

$$H_0 : E(X) - E(Y) = 0.$$

The test is computed by ordering the combined sample of durations in ascending order. Denote the rank of y in the joint ordering by s_i for $i = 1, ..., n$ and let the rank sum of the Y-values be defined as $W = \sum_{j=1}^{n} s_j$. For test purposes we use the standardized version of W

$$W^* = \frac{W - E_0(W)}{\sqrt{\text{var}_0(W)}} \overset{a}{\sim} N(0, 1), \tag{6}$$

where

$$\begin{aligned} E_0(W) &= n(N+1)/2 \\ \text{var}_0(W) &= \frac{nm}{2}\left[N + 1 - \frac{\sum_{j=1}^{g}(t_j - 1)t_j(t_j + 1)}{N(N-1)}\right]. \end{aligned}$$

g is the number of tied groups and t_j is the size of the jth group.

To test for differences in either the dispersion or the location of the two duration distributions, we adopt the Lepage test. This test establishes whether there are differences in *either* the location parameters θ_1 and θ_2 *or* the scale parameters η_1 and η_2 of the two distributions. The null hypothesis is

$$H_0 : \theta_1 = \theta_2 \text{ and } \eta_1 = \eta_2.$$

To compute the test, suppose that a score of 1 is assigned to both the smallest and largest observations in the combined sample, a score of 2 is assigned to both the second smallest and second largest observation, and so forth. The resulting score sum of observations drawn from the Y-sample, denoted by R_j, is given by $C = \sum_{j=1}^{n} R_j$. This can be normalized to give a standardized test statistic

$$C^* = \frac{C - E_0(C)}{\sqrt{\text{var}_0(C)}}$$

where

$$E_0(C) = \begin{cases} n(N+2)/4 & \text{for } N \text{ even} \\ n(N+1)^2/(4N) & \text{for } N \text{ odd} \end{cases}$$

$$\text{var}_0(C) = \begin{cases} \frac{nm}{16N(N-1)} \left[16 \sum_{j=1}^{g} t_j r_j^2 - N(N+2)^2\right] & \text{for } N \text{ even} \\ \frac{nm}{16N^2(N-1)} \left[16N \sum_{j=1}^{g} t_j r_j^2 - (N+1)^4\right] & \text{for } N \text{ odd} \end{cases} .$$

The Lepage test statistic is simply the sum of the squares of W^* and C^*:

$$D = (W^*)^2 + (C^*)^2 \overset{a}{\sim} \chi_2^2. \tag{7}$$

Finally, to test for general differences in the two populations, we adopt the Kolmogorov-Smirnov test whose null hypothesis is that the two duration distributions are identical:

$$H_0 : F(t) = G(t)$$

for every t. The alternative hypothesis is that the two duration distributions differ, i.e. that $F(t) \neq G(t)$. The resulting test statistic is

$$J = \frac{mn}{d} \max_{(-\infty < t < \infty)} \{|F_m(t) - G_n(t)|\}, \tag{8}$$

where $F_m(t)$ and $G_n(t)$ are the empirical distribution functions of X and Y, and d is the greatest common divisor of m and n. Critical values of the sample distribution are given in, e.g., Hollander & Wolfe (1999).

4.1. Empirical Findings

The outcome of these tests is reported in Table 3. In bull markets there is very strong evidence against the random walk model which is rejected at the 10% level by all tests based on nominal stock prices. This model also gets rejected by the Lepage and Kolmogorov-Smirnov tests in the case of real stock prices. In the bear state the evidence is insufficiently strong to lead to a rejection of the random walk model for the larger filters (20/15 and 20/10) but this model is generally rejected when the smaller filters (15/15 and 15/10) are used. This difference is most likely related to the power of the duration tests since the smaller filters generate more duration spells than the larger ones, c.f. Table 1.

Although it its more difficult to reject the null of equality of the duration distributions generated by the GARCH model and the actual data, in bull markets the null is nevertheless rejected by the Lepage and Kolmogorov-Smirnov tests using the small filter sizes. In bear markets the null is not rejected at the 10% level. Failure to reject the null hypothesis for bear markets is again likely at least partially to reflect low power. Indeed, as the filter size declines and the tests become more powerful, the *p*-values decline from around 0.50 to around 0.20.

These results suggest that the Lepage and Kolmogorov-Smirnov tests produce significant evidence against the two parametric models particularly for the smaller filter sizes. The random walk model clearly cannot match the duration distribution of either bull or bear markets. While the GARCH model does a better job at fitting the data, it still does not fully capture the properties of bull market durations. Since the Lepage and Kolmogorov-Smirnov tests have power against differences in the dispersion of the duration distributions, our results suggest that these tests pick up differences related to the very long bull markets observed in the data.

5. Models of Duration Dependence in Bull and Bear Markets

Section 3 characterized the unconditional distribution of bull and bear market spells. However, if the age of a bull or bear market affects future price movements, investors will want to calculate expectations conditional on the path followed by stock prices up to a given point in time. For instance, during the long bull market of the nineties, concern was often expressed that this bull market was at greater risk of coming to an end because it had lasted 'too long' by historical standards. This indicates a belief that the bull market hazard rate depends positively on its duration. The opposite view is that bull markets gain momentum: the longer a bull market has lasted, the more robust it is, and hence the lower its hazard rate.

Testing these hypotheses requires that we go well beyond inspecting the unconditional probability of termination for the bull or bear markets. Instead the duration data needs to be characterized in terms of the conditional probability that the bull or bear state ends in a short time interval following some period t, given that the state lasted up to t. For the i'th duration, T_i, this is measured by the discrete hazard function

$$\lambda_i(t|\mathbf{X}_{it}) = \Pr(T_i = t|T_i \geq t, \mathbf{X}_{it}), \quad t = 1, \ldots, A, \tag{9}$$

which is the conditional probability of termination in the interval $[a_{t-1}, a_t)$ given that this interval was reached in the first place. $\mathbf{X}_{it} = \{\mathbf{x}_{i1}, ..., \mathbf{x}_{it}\}$ is a vector of additional conditioning information, which will depend on the particular duration. Hypotheses about the probability that a bull or bear market is terminated as a function of its age are naturally expressed in terms of the shape of this hazard function. For example, the natural null hypothesis is that the duration of the current state does not affect the hazard rate.

The probability that a bull or bear market lasts for a certain period of time can still be derived from these hazard rates. This is given as the discrete survivor function which measures the probability that a bull or bear market survives on the interval $[a_{t-1}, a_t)$:

$$S_i(t|\mathbf{X}_{it}) = \Pr(T_i > t|\mathbf{X}_{it}) = \prod_{s=1}^{t}(1 - \lambda_i(s|\mathbf{X}_{is})), \quad t = 1, \ldots, A. \tag{10}$$

Common choices of hazard models are the Probit, Logit and Double Exponential link. An extensive comparison of such link functions is found in Sueyoshi (1995). Throughout the paper we use a Logit link, i.e.,

$$\lambda_i(t|\mathbf{X}_{it}) = F(\mathbf{x}_{it}'\boldsymbol{\beta}) = \frac{\exp(\mathbf{x}_{it}'\boldsymbol{\beta})}{1 + \exp(\mathbf{x}_{it}'\boldsymbol{\beta})}. \tag{11}$$

We consider two separate models for these hazard rates. The first is a static model that assumes that the underlying parameters linking the covariates or state variables to the hazard rate do not vary over time and that the covariates are fixed from the point of entry into a state. This makes our results directly comparable to the large literature on univariate dynamics in stock prices surveyed in Chapter 2 in Campbell, Lo & MacKinlay (1997). Under this assumption the data takes the form of $\{t_i, \mathbf{x}_i; i = 1, \ldots, n\}$, where t_i is the survival time and $\mathbf{x}_i$ is a covariate (or state variable) observed at the beginning of the interval $[a_{t_i-1}, a_{t_i})$.

However, switches between bull and bear market states are likely to be caused by changes in the underlying economic environment. For example, the drift in stock prices may turn from positive to negative as a result of increased interest rates or worsening economic prospects. The effect of such covariates may well depend on the age of the current bull or bear market. To account for this possibility, our second model extends the setup and allows $\mathbf{x}_{it}$ to be a vector that incorporates time-varying covariates. Now the data for the i'th duration spell takes the form

$$\{\underbrace{t_i}_{\text{duration}}, \overbrace{\mathbf{x}_i(a_0), \mathbf{x}_i(a_1), \ldots \mathbf{x}_i(a_{t_i-1}), \mathbf{x}_i(a_{t_i})}^{\equiv \mathbf{X}_i\text{: time-varying covariates}}\}.$$

Since our data is discretely measured, the covariates follow a step function with jumps at the follow-up times, a_t. Within the interval $[a_{t-1}, a_t)$ the history of covariates

$$\mathbf{X}_{it} = (\mathbf{x}_{i1}, \mathbf{x}_{i2}, \ldots, \mathbf{x}_{it}),$$

is allowed to influence the hazard rate $\lambda_i(t|\mathbf{X}_{it})$.

To allow for the possibility that the effect on the hazard rate of these covariates could depend on the age of the current state, we consider an approach that allows the parameters to vary with duration:

$$\lambda_i(t|\mathbf{X}_{it}) = F(\mathbf{x}'_{it}\boldsymbol{\alpha}_t). \tag{12}$$

The vector $\boldsymbol{\alpha}_t = (\gamma_{0t}, \boldsymbol{\gamma}'_t)'$ comprises both the baseline and the covariance parameters. We use the first-order random walk as our choice of transition equation determining the evolution in $\boldsymbol{\alpha}_t$:

$$\boldsymbol{\alpha}_t = \Phi\boldsymbol{\alpha}_{t-1} + \boldsymbol{\xi}_t, \quad \Leftrightarrow \begin{pmatrix} \gamma_{0t} \\ \boldsymbol{\gamma}_t \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \gamma_{0t-1} \\ \boldsymbol{\gamma}_{t-1} \end{pmatrix} + \begin{pmatrix} \xi_{0t} \\ \boldsymbol{\xi}_{1t} \end{pmatrix}, \tag{13}$$

$$\text{where} \quad \begin{pmatrix} \xi_{0t} \\ \boldsymbol{\xi}_{1t} \end{pmatrix} \sim \mathcal{N}_{p+1}(\mathbf{0}, diag(\sigma_0^2, \sigma_1^2, \ldots, \sigma_p^2)), \quad p = \dim(\boldsymbol{\gamma}_t), \quad \boldsymbol{\alpha}_0 \sim \mathcal{N}_{p+1}(\mathbf{a}_0, \mathbf{Q}_0),$$

and $\mathcal{N}_{p+1}(\cdot, \cdot)$ is the $(p+1)$-dimensional standard normal distribution. This random walk specification has the advantage of not imposing mean reversion on the parameters which are allowed to differ across durations (although neighboring points cannot be too far from each other) if the data supports such variation.

5.1. Estimation

The log-likelihood function can be conveniently set up using notation from the literature on discrete choice models. Consider the following discrete indicator variable:

$$y_{is} = \begin{cases} 1, & i\text{'th bull or bear market terminates in } [a_{s-1}, a_s) \\ 0, & i\text{'th bull or bear market survives through } [a_{s-1}, a_s) \end{cases} \tag{14}$$

for $s = 1, \ldots, t_i$, and $i = 1, \ldots, n$. Each bull or bear spell, i, thus generates a string

$$\mathbf{y}_i = (y_{i1}, \ldots, y_{it_i}) = (0, \ldots, 0, 1), \qquad i = 1, \ldots, n.$$

Using this notation, the contribution to the likelihood function from the ith observation is

$$\mathcal{L}_i \propto \prod_{s=1}^{t_i} \lambda_i(s|\mathbf{x}_i)^{y_{is}}(1 - \lambda_i(s|\mathbf{x}_i))^{1-y_{is}}. \tag{15}$$

For every spell the bull or bear market lives through it therefore contributes to the likelihood with the survivor probability $1 - \lambda_i(s|\mathbf{x}_i)$. Summing across duration spells, the total log-likelihood function for the model $\lambda_i(t|\mathbf{x}_i) = F(\mathbf{x}_{it}'\boldsymbol{\beta})$ is given by

$$\ln \mathcal{L} \propto \sum_{i=1}^{n} \sum_{s=1}^{t_i} y_{is} \ln(\lambda_i(s|\mathbf{x}_i)) + (1 - y_{is}) \ln(1 - \lambda_i(s|\mathbf{x}_i)). \tag{16}$$

An approach of treating the covariate parameters as fixed effects is only appropriate if the number of intervals is very small. In applications such as ours without enough intervals to apply continuous time techniques, maximum likelihood estimates of a large number of parameters in the hazard functions of an unrestricted hazard model can be expected to have very poor sampling properties.

To get around this problem, we follow Fahrmeir (1994) and adopt state space techniques that treat the hazard function as the measurement equation. Since our measurement equation is non-normal it is complicated to solve for the posterior density of the hazard function conditional on the data, $p(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_q | \mathbf{y}_1, \ldots, \mathbf{y}_q, \mathbf{x}_1, \ldots, \mathbf{x}_q)$, which is required for writing down the likelihood function. We adopt the strategy, advocated by Fahrmeir (1992), of basing estimation on posterior modes subject to smoothing priors which penalize large changes in neighboring parameters, $\boldsymbol{\alpha}_t - \boldsymbol{\alpha}_{t-1}$. As detailed in appendix A, repeatedly applying Bayes' theorem to the posterior density, estimation of $\boldsymbol{\alpha}_t$ by posterior modes is equivalent to maximizing the following penalized log-likelihood function:

$$\begin{aligned} \ln \mathcal{L}(\boldsymbol{\alpha}_1, .., \boldsymbol{\alpha}_q) \quad = \quad & \sum_{i=1}^{n} \sum_{t=1}^{t_i} l_{it}(\boldsymbol{\alpha}_t) - \frac{1}{2}(\boldsymbol{\alpha}_0 - \mathbf{a}_0)'\mathbf{Q}_0^{-1}(\boldsymbol{\alpha}_0 - \mathbf{a}_0) \\ & - \frac{1}{2} \sum_{t=1}^{q} (\boldsymbol{\alpha}_t - \boldsymbol{\alpha}_{t-1})'\mathbf{Q}^{-1}(\boldsymbol{\alpha}_t - \boldsymbol{\alpha}_{t-1}), \end{aligned} \tag{17}$$

where,

$$l_{it}(\boldsymbol{\alpha}_t) = y_{it} \ln(F(\mathbf{x}_{it}'\boldsymbol{\alpha}_t)) + (1 - y_{it}) \ln(1 - F(\mathbf{x}_{it}'\boldsymbol{\alpha}_t)) \tag{18}$$

is the log-likelihood contribution of the i'th duration spell. The first term measures the goodness of fit of the model. The second and third terms - which are introduced by the smoothness priors, $\mathbf{Q}_0$, $\mathbf{Q}$, specified by the transition model - penalize large deviations between successive parameters and lead to smoothed estimates. Appendix B provides details on the numerical optimization of this penalized likelihood function through a generalized extended Kalman filter and smoother.

6. Empirical Results

Using the estimation techniques and hazard models from Section 5, we first estimate the hazard function for bull and bear markets in a model without time-varying covariates. The output from this exercise is the baseline hazard rates plotted in Figure 3 using filter sizes of 20/10 and 15/10. These baseline hazard rates measure the pure age dependence in bull and bear market termination probabilities. The left windows show the hazard rate estimated for the nominal price index, while the right windows plot estimates for real stock prices, in both cases surrounded by 90% confidence regions. To establish a benchmark we also plot the hazard rate under the GARCH model. Because it takes some time before the market moves the full distance of the filter, the hazard rate is initially very low under the GARCH model but it rapidly increases to a level of 2-4 percent. Once the initial threshold effect wears off, the GARCH hazard rates flatten out.

These baseline hazard rates are surrounded by considerable estimation uncertainty, particularly at the long end where the 90% confidence intervals are very wide. Even so, some patterns emerge from these plots. First, bear hazard rates increase systematically at long durations, while the bull market hazard rates follow no particular pattern. This is another way of detecting why long bear markets do not occur in the actual data, simply because of the rapid rise in bear hazard rates as their duration grows.

At short durations in the bull market the simulated GARCH hazard rates are distinctly below the hazard rates observed in the actual data. This pattern is reversed at longer durations where the GARCH hazard rates are above the actual hazard rates, at least for real stock prices. In bear markets the GARCH hazard rate is below its value in the actual data both at short and long durations irrespective of whether nominal or real stock prices are considered. This is consistent with our earlier observation that the longest bull market durations are longer than what we expect from the parametric models while the longest bear spells are shorter than what is consistent with these models.

To formally test whether the duration distribution for the bull and bear market states are identical we adopted the two-sample tests from Section 4 to the bull and bear durations observed in the actual data. Irrespective of which test we use, Table 4 shows that we generally reject that bull and bear durations are drawn from the same distribution. This might be expected given the upward drift in both real and nominal stock prices, but this is by no means the only explanation of our findings. For example, when the filter sizes are set so as to account for this drift, e.g. by using the 20/15 filter, the null of equality of the bull and bear duration distributions is strongly rejected by the actual data but not by the simulated data under the random walk model.

As a means of providing a single summary measure of the attrition rates in bull and bear markets, Figure 4 plots the survivor function (10) for bull and bear states against the survival function under the simulated GARCH model. Again the left column is based on nominal prices and the right column is based on real prices. While the individual hazard rates are estimated with considerable noise, the cumulated effects of differences in hazard rates at short and long durations becomes much clearer in this figure. In the bull market the survival probability at short durations is much lower in the actual data than in the simulated data. This changes at longer durations where the survival probability becomes higher in the actual than in the simulated data, particularly for real stock prices where the 95% confidence region lies entirely above the GARCH survivor function. A very different picture emerges in the bear state where the survival function based on the actual data is always at or below that from the simulated data. For the smaller 15/10 filter, the full 95% confidence region lies below the GARCH survival curve.

6.1. Interest Rate Effects

To shed light on how the hazard rates depend on the underlying state of the economy, we next include interest rates as a time-varying covariate. Interest rates have been widely documented to closely track the state of the business cycle and appear to be a key determinant of stock returns at the monthly horizon. See, e.g., Kandel & Stambaugh (1990), Fama & French (1989), Glosten, Jagannathan & Runkle (1993), Pesaran & Timmermann (1995) and Whitelaw (1994). Interest rate *levels*, i_t, may be affected by a low frequency component and therefore might not contain the same information over a sample as long as ours, while interest rate *changes*, Δi_t, are more likely to track business cycle variation across the full sample. For this reason we include both levels and changes in interest rates so the set of covariates is $\mathbf{x}_{it}^{\prime} = (1, i_t, \Delta i_t)$. In the analysis of nominal stock prices we use nominal interest rates while the analysis of real stock prices is based on real interest rates. Our hazard specification−which allows interest rate effects to vary with the age of the current state−is

$$\lambda_i(t|\mathbf{x}_i(t)) = F(\gamma_{0t} + \gamma_{1t} i_t + \gamma_{2t} \Delta i_t). \tag{19}$$

Figure 5 shows the sequence of baseline hazards and nominal interest rate effects using nominal stock prices and a 15/10 filter. Compared to Figure 3, it is clear that controlling for interest rates has a significant effect on the shape of the bull market baseline hazard. Controlling for interest rate effects, the market baseline hazard drops sharply from five to two percent per week as the bull market duration extends beyond six months. It remains flat at longer durations. Young bull markets thus appear substantially more at risk of termination than bull markets that have lasted for six months or longer.

Panel (b) shows that at very short durations, higher nominal interest rates are associated with a lower bull market hazard rate. However, this parameter is within one standard error of zero after 30 weeks. The initial negative sign should be interpreted with caution since interest rates tend to be high towards the beginning of a new expansion state and this is often the beginning of a bull market in stock prices. More importantly, perhaps, positive interest rate changes are associated with increases in the bull hazard rate that are much larger in magnitude than the effects from the interest rate level, c.f. panel (c).

The baseline hazard in the bear market also changes as a result of controlling for interest rate effects. The baseline hazard rate now increases from five to about eight percent per week at durations longer than six months. Both the interest rate level and interest rate changes are associated with negative effects on bear hazard rates that are more than one standard error away from zero in most cases. Notice in particular the very large negative effect of interest rate increases on hazard rates in young bear markets. An environment of high and increasing interest rates thus leads to lower bear market hazard rates which means a higher chance of remaining in the bear state.

Results for the real stock price index are provided in Figure 6. The baseline hazard rate declines in the bull market while it increases after an initial slight decline in the bear state. The real interest rate level has a small positive correlation with both bull and bear hazard rates for durations up to 20 weeks and becomes negative thereafter. While real interest rate changes are insignificantly correlated with hazard rates in the bull market, they are strongly negatively correlated with hazard rates in the bear market, thus confirming our earlier finding that increasing interest rates lead to a higher probability of remaining in the bear state.

6.2. Interest Rate Changes and Bull and Bear Survival Rates

A large literature has found strong negative effects of nominal interest rates on stock returns. Mostly such effects have been documented in the context of single-period regressions of stock returns on interest rates. However, our results suggest that hazard rates vary even at long horizons so that a low-order Markov representation fails to properly capture the dynamics of stock prices. They suggest that interest rate effects will depend both on the state (bull or bear) and on its age when the shock occurs. We therefore need to account for these factors when computing the effect of an interest rate change on the two state probabilities.

To do so, consider a scenario where the current interest rate level is permanently raised from 5% to 7% after 52 weeks in a bull market. In the baseline or non-raise scenario the covariate matrix is therefore

$$\mathbf{X}_i^{\text{non-raise}} = \begin{pmatrix} \underbrace{\begin{matrix} 1 & \cdots & 1 \\ 5 & \cdots & 5 \\ 0 & \cdots & 0 \end{matrix}}_{\text{obs 1 to 51}} & \underbrace{\begin{matrix} 1 \\ 5 \\ 0 \end{matrix}}_{52} & \underbrace{\begin{matrix} 1 & \cdots & 1 \\ 5 & \cdots & 5 \\ 0 & \cdots & 0 \end{matrix}}_{\text{obs 52 to 90}} \end{pmatrix},$$

while in the raise scenario the covariate matrix is

$$\mathbf{X}_i^{\text{raise}} = \begin{pmatrix} \underbrace{\begin{matrix} 1 & \cdots & 1 \\ 5 & \cdots & 5 \\ 0 & \cdots & 0 \end{matrix}}_{\text{obs 1 to 51}} & \underbrace{\begin{matrix} 1 \\ 7 \\ 2 \end{matrix}}_{52} & \underbrace{\begin{matrix} 1 & \cdots & 1 \\ 7 & \cdots & 7 \\ 0 & \cdots & 0 \end{matrix}}_{\text{obs 53 onwards}} \end{pmatrix}.$$

In both cases the hazard rates are computed as

$$\lambda_i(t|\mathbf{X}_{it}) = F\left(\gamma_{0t} + \gamma_{1t} i_t + \gamma_{2t}\Delta i_t\right) = \frac{\exp(\gamma_{0t} + \gamma_{1t} i_t + \gamma_{2t}\Delta i_t)}{1 + \exp(\gamma_{0t} + \gamma_{1t} i_t + \gamma_{2t}\Delta i_t)}, \tag{20}$$

while the impact of a raise on stock prices depends on

$$\begin{aligned} \frac{\partial \lambda_i(t|\mathbf{X}_{it})}{\partial i_t} &= \gamma_{1t}\lambda_i(t|\mathbf{X}_{it})(1-\lambda_i(t|\mathbf{X}_{it})) \\ \frac{\partial \lambda_i(t|\mathbf{X}_{it})}{\partial \Delta i_t} &= \gamma_{2t}\lambda_i(t|\mathbf{X}_{it})(1-\lambda_i(t|\mathbf{X}_{it})). \end{aligned} \tag{21}$$

The duration dependence of these effects is indicated through their t subscripts. Figure 7 shows the effect on the bull market hazard rate of the 2% increase in the interest rate after 26 and 52 weeks, respectively. We concentrate on the nominal price index since the outcome of the analysis for the real price index is very similar. The spike in the hazard rate arises because of the one-off change in Δi_t. Following this impact, the hazard rate varies randomly around its value in the baseline scenario. Overall the bull market survival probability is marginally lower in the interest rate raise scenario, with the strongest effect showing up immediately after the interest rate raise. Figure 7 also shows that the interest rate raise leads to a strong immediate decline in the bear market hazard rate. The long-run effects on the bear market hazard and survival rates appear to be stronger than in the bull market.

Turning to the real stock price index, Figure 8 shows that a change in the real interest rate does not have much of an impact on the survival probability of a bull market or a bear market when the raise occurs after 26 weeks. The main effect is in the form of a higher survival probability of the bear market when the raise occurs after 52 weeks.

7. Conclusion

This paper has proposed a new approach to identify dependence in the direction of stock prices based on the probability of exiting from bull or bear states. Since the length of time spent in these states is a key determinant of the mean and risk of stock returns, it is important to characterize bull and bear durations. We find evidence contradicting standard models of stock prices even after accounting for time-varying volatility and state variables such as interest rates. At short durations bull market hazard rates are well above their values under the random walk or GARCH models. However, at long durations bull market hazard rates fall below their values under these benchmark models. In bear markets, hazard rates observed in the actual data are above the hazard rates generated by the benchmark models both at short and long durations. This means that long bull market spells are more likely and long bear market spells less likely than one would expect from the benchmark models.

Such evidence of deviations from the random walk model does not imply a rejection of the efficient market hypothesis. Long-run dependencies in stock prices do, however, have important implications for both risk management and for interpretation of the sources of movements in stock prices. It is beyond this paper to propose an economic model that can explain duration dependence in stock prices. Instead we briefly consider alternative economic explanations of duration dependence based on speculative bubbles or fundamentals.

McQueen & Thorley (1994) study speculative bubbles that take the form of sequences of small positive abnormal returns interrupted by rare but large negative abnormal returns in a crash state. Their bubble implies that the probability that a run of positive abnormal returns comes to an end declines with the length of the sequence. In empirical tests on monthly stock returns over the period 1927-1991, they find evidence of negative duration dependence for positive runs while there appears to be no duration dependence in negative runs. Data limitations mean that they consider runs of at most six months' duration. We use daily data over a much longer period (1885-1997) which allows us to consider hazard rates at both much shorter and longer durations. Our finding of a declining bull market hazard rate is consistent with McQueen and Thorley's result. It is more difficult to appeal to a bubble-related explanation for the U-shaped pattern in the bear market hazard rate.

Duration dependence in stock prices may alternatively be driven by information effects or by fundamentals such as dividend payoffs and time-varying risk premiums. Wang (1993) models asymmetry of information between noise traders and rational investors which leads uninformed traders to rationally behave like price chasers. This introduces serial correlation in stock returns. If such effects are linked to the underlying state of the economy it is possible that they could affect the duration distribution of stock returns. Campbell & Cochrane (1999) propose an asset pricing model in which consumption growth follows a lognormal process with habit formation effects. Pagan & Sossounov (2000) find that this model has some promise for matching the average duration of bull and bear states, although matching the hazard function may be a more difficult test to pass. Cecchetti, Lam & Mark (2000) introduce belief distortions that vary over expansions and contractions and lead to systematic predictability in returns while Gordon & St-Amour (2000) present a model where preferences change according to an exogenous regime switching process. These models all seem to have some promise for explaining bull and bear durations which we intend to explore in future work.

Bull and bear markets could also be related to recession and expansion states. In the most systematic work to date, Diebold & Rudebusch (1990) and Diebold, Rudebusch & Sichel (1993) investigate duration dependence in the US business cycle. Although duration analyses of aggregate data must be tempered by the infrequency of such data, these papers nevertheless find evidence of positive duration dependence in pre-war expansions and post-war contractions. Their finding of a very strong rise in the hazard rate for post-war recessions is likely to be closely related to the rise in the bear market hazard rate that we found for stock prices.

Appendix A: Assumptions of the Estimations

The assumptions underlying our estimation approach are most appropriately stated by reordering the observations and using risk set notation. Suppose that all bull or bear durations have been lined up so that they start at the same point in time and let risk indicators r_{it} $(i, t \geq 1)$ be defined by

$$r_{it} = \begin{cases} 1, & \text{if the } i\text{'th bull or bear market is at risk in } [a_{t-1}, a_t) \\ 0, & \text{otherwise.} \end{cases} \tag{A.1}$$

Furthermore, define the risk vector $\mathbf{r}_t = (r_{it}, i \geq 1)$, and the risk set $\mathcal{R}_t = \{i : t \leq t_i\}$ at time t, i.e., the set of duration spells that are at risk in the interval $[a_{t-1}, a_t)$. Using this notation the log-likelihood function (16) can be written as

$$\ln \mathcal{L} \propto \sum_{t=1}^{q} \sum_{i \in \mathcal{R}_t} y_{it} \ln(\lambda(t|\mathbf{X}_{it})) + (1 - y_{it}) \ln(1 - \lambda(t|\mathbf{X}_{it})). \tag{A.2}$$

As t increases, fewer bull or bear markets continue to survive and thus the dimension of $\mathbf{r}_t$ declines. Covariates and failure indicators are collected in the vectors

$$\begin{aligned} \mathbf{x}_t &= (x_{it}, i \in \mathcal{R}_t) \\ \mathbf{y}_t &= (y_{it}, i \in \mathcal{R}_t). \end{aligned} \tag{A.3}$$

Finally the histories of covariates, failure and risk indicators up to period $t-1$ are given by

$$\begin{aligned} \mathbf{x}^*_{t-1} &= (\mathbf{x}_1, \ldots, \mathbf{x}_{t-1}) \\ \mathbf{y}^*_{t-1} &= (\mathbf{y}_1, \ldots, \mathbf{y}_{t-1}) \\ \mathbf{r}^*_{t-1} &= (\mathbf{r}_1, \ldots, \mathbf{r}_{t-1}). \end{aligned} \tag{A.4}$$

The following set of assumptions are required for the maximum likelihood estimation, c.f. Fahrmeir (1992):

(A1) Given $\boldsymbol{\alpha}_t$, $\mathbf{y}^*_{t-1}$, $\mathbf{r}^*_t$ and $\mathbf{x}^*_t$, current $\mathbf{y}_t$ is independent of $\boldsymbol{\alpha}^*_{t-1} = (\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_{t-1})$:

$$p(\mathbf{y}_t|\boldsymbol{\alpha}^*_t, \mathbf{y}^*_{t-1}, \mathbf{r}^*_t, \mathbf{x}^*_t) = p(\mathbf{y}_t|\boldsymbol{\alpha}_t, \mathbf{y}^*_{t-1}, \mathbf{r}^*_t, \mathbf{x}^*_t), \;\; t = 1, 2, \ldots$$

This assumption is standard in state space modelling. It simply states that the conditional information in $\boldsymbol{\alpha}^*_t$ about $\mathbf{y}_t$ is exclusively contained in the current parameter $\boldsymbol{\alpha}_t$.

(A2) Conditional on $\mathbf{y}^*_{t-1}$, $\mathbf{r}^*_{t-1}$ and $\mathbf{x}^*_{t-1}$, the covariate $\mathbf{x}_t$ and risk vector $\mathbf{r}_t$ are independent of $\boldsymbol{\alpha}^*_{t-1}$:

$$p(\mathbf{x}_t, \mathbf{r}_t|\boldsymbol{\alpha}^*_{t-1}, \mathbf{y}^*_{t-1}, \mathbf{r}^*_{t-1}, \mathbf{x}^*_{t-1}) = p(\mathbf{x}_t, \mathbf{r}_t|\mathbf{y}^*_{t-1}, \mathbf{r}^*_{t-1}, \mathbf{x}^*_{t-1}), \;\; t = 1, 2, \ldots$$

Thus we assume that the covariate processes contain no information on the parameter process.

(A3) The parameter transitions follow a Markov process:

$$p(\boldsymbol{\alpha}_t|\boldsymbol{\alpha}^*_{t-1}, \mathbf{y}^*_{t-1}, \mathbf{x}^*_t) = p(\boldsymbol{\alpha}_t|\boldsymbol{\alpha}_{t-1}), \quad t = 1, 2, \ldots$$

This assumption is implied by the transition model and the assumption on the error sequence.

(A4) Given $\boldsymbol{\alpha}_t$, $\mathbf{y}^*_{t-1}$, $\mathbf{r}^*_t$ and $\mathbf{x}^*_t$, individual responses y_{it} within $\mathbf{y}_t$ are conditionally independent:

$$p(\mathbf{y}_t|\boldsymbol{\alpha}_t, \mathbf{y}^*_{t-1}, \mathbf{x}^*_t, \mathbf{r}^*_t) = \prod_{i\in\mathcal{R}_t} p(y_{it}|\boldsymbol{\alpha}_t, \mathbf{y}^*_{t-1}, \mathbf{x}^*_t, \mathbf{r}^*_t), \quad t = 1, 2, \ldots$$

This assumption is much weaker than an assumption of unconditional independence.

To estimate the parameters, $\boldsymbol{\alpha}^*_q$, we repeatedly apply Bayes' theorem to get the posterior density:

$$\begin{aligned} p(\boldsymbol{\alpha}^*_q|\mathbf{y}^*_q, \mathbf{x}^*_q, \mathbf{r}^*_q) &= \prod_{t=1}^{q} p(\mathbf{y}_t|\mathbf{y}^*_{t-1}, \mathbf{x}^*_t, \mathbf{r}^*_t; \boldsymbol{\alpha}^*_t) \prod_{t=1}^{q} p(\boldsymbol{\alpha}_t|\boldsymbol{\alpha}^*_{t-1}, \mathbf{y}^*_{t-1}, \mathbf{x}^*_t, \mathbf{r}^*_t) \\ &\cdot \prod_{t=1}^{q} p(\mathbf{x}_t, \mathbf{r}_t|\boldsymbol{\alpha}^*_{t-1}, \mathbf{y}^*_{t-1}, \mathbf{x}^*_{t-1}, \mathbf{r}^*_{t-1}) \frac{p(\boldsymbol{\alpha}_0)}{p(\mathbf{y}^*_q, \mathbf{x}^*_q, \mathbf{r}^*_q)}. \end{aligned} \tag{A.5}$$

Under assumptions (A1)-(A4), this now simplifies to

$$p(\boldsymbol{\alpha}^*_q|\mathbf{y}^*_q, \mathbf{x}^*_q, \mathbf{r}^*_q) \propto \prod_{t=1}^{q} \prod_{i\in\mathcal{R}_t} p(y_{it}|\mathbf{y}^*_{t-1}, \mathbf{x}^*_t, \mathbf{r}^*_t; \boldsymbol{\alpha}_t) \prod_{t=1}^{q} p(\boldsymbol{\alpha}_t|\boldsymbol{\alpha}_{t-1}) \cdot p(\boldsymbol{\alpha}_0), \tag{A.6}$$

which is the expression used in the calculation of the log-likelihood function (18).

Appendix B: Maximization of the Generalized Kalman Filter and Smoother

This appendix briefly explains some of the details of the numerical optimizations. To perform numerical optimization of the penalized log-likelihood function, we use the generalized extended Kalman filter and smoother suggested by Fahrmeir (1992). Let $d_{it}(\boldsymbol{\alpha}_t)$ denote the first derivative $\partial F(\eta)/\partial\eta$ of the response function $F(\eta)$ evaluated at $\eta = \mathbf{x}'_{it}\boldsymbol{\alpha}_t$. The contribution to the score of the failure indicator y_{it} is given by

$$\mathbf{u}_{it}(\boldsymbol{\alpha}_t) = \frac{\partial l_{it}(\boldsymbol{\alpha}_t)}{\partial \boldsymbol{\alpha}_t} = \mathbf{x}_{it} \frac{d_{it}(\boldsymbol{\alpha}_t)}{F(\mathbf{x}'_{it}\boldsymbol{\alpha}_t)\{1 - F(\mathbf{x}'_{it}\boldsymbol{\alpha}_t)\}} \{y_{it} - F(\mathbf{x}'_{it}\boldsymbol{\alpha}_t)\}, \tag{B.1}$$

and the contribution of the expected information matrix is

$$\mathbf{U}_{it}(\boldsymbol{\alpha}_t) = -E\left[\frac{\partial^2 l_{it}(\boldsymbol{\alpha}_t)}{\partial\boldsymbol{\alpha}'_t\partial\boldsymbol{\alpha}_t}|\boldsymbol{\alpha}_t, y^*_{t-1}, \mathbf{x}_{t-1}\right] = \mathbf{x}_{it}\mathbf{x}'_{it} \frac{(d_{it}(\boldsymbol{\alpha}_t))^2}{F(\mathbf{x}'_{it}\boldsymbol{\alpha}_t)\{1 - F(\mathbf{x}'_{it}\boldsymbol{\alpha}_t)\}}. \tag{B.2}$$

The contributions of the risk set to the score vector and the expected information matrix in the interval $[a_{t-1}, a_t)$ can be obtained by summing over the durations in the risk set at time t. This means computing $\mathbf{u}_t(\boldsymbol{\alpha}_t) = \sum_{i\in\mathcal{R}_t} \mathbf{u}_{it}(\boldsymbol{\alpha}_t)$ and $\mathbf{U}_t(\boldsymbol{\alpha}_t) = \sum_{i\in\mathcal{R}_t} \mathbf{U}_{it}(\boldsymbol{\alpha}_t)$.

Let $\mathbf{a}_{t|q}$ $(t = 0, \ldots, q)$ denote the smoothed estimates of $\boldsymbol{\alpha}_t$. These estimates can be obtained as numerical approximations to posterior modes given all the data $(\mathbf{y}^*, \mathbf{x}^*, \mathbf{r}^*)$ up to q. Approximate error covariance matrices $\mathbf{V}_{t|q}$ are obtained as the corresponding numerical approximations to curvatures, i.e. inverses of expected negative second derivatives of $\ln \mathcal{L}(\boldsymbol{\alpha}^*)$, evaluated at the mode. Finally $\mathbf{a}_{t|t-1}$ and $\mathbf{a}_{t|t}$ are the prediction and filter estimates of $\boldsymbol{\alpha}_t$ given the data up to $t-1$ and t, with corresponding error matrices $\mathbf{V}_{t|t-1}$ and $\mathbf{V}_{t|t}$.

Filtering and smoothing of our sample data proceed in the following steps:

1. <u>INITIALIZATION:</u>

$$\begin{aligned} \mathbf{a}_{0|0} &= \mathbf{a}_0, \\ \mathbf{V}_{0|0} &= \mathbf{Q}_0. \end{aligned} \tag{B.3}$$

2. <u>FILTER PREDICTION STEPS:</u>

For $t = 1, \ldots, q$:

$$\begin{aligned} \mathbf{a}_{t|t-1} &= \Phi \mathbf{a}_{t-1|t-1}, \\ \mathbf{V}_{t|t-1} &= \Phi \mathbf{V}_{t-1|t-1} \Phi + \mathbf{Q}. \end{aligned} \tag{B.4}$$

3. <u>FILTER CORRECTION STEPS:</u>

For $t = 1, \ldots, q$ use the global scoring steps:

$$\begin{aligned} \mathbf{a}_{t|t} &= \mathbf{a}_{t|t-1} + \mathbf{V}_{t|t} \mathbf{u}_t, \\ \mathbf{V}_{t|t} &= (\mathbf{V}_{t|t-1}^{-1} + \mathbf{U}_t)^{-1}. \end{aligned} \tag{B.5}$$

4. <u>BACKWARD SMOOTHING STEPS:</u>

For $t = 1, \ldots, q$:

$$\begin{aligned} \mathbf{a}_{t-1|q} &= \mathbf{a}_{t-1|t-1} + \mathbf{B}_t (\mathbf{a}_{t|q} - \mathbf{a}_{t|t-1}), \\ \mathbf{V}_{t-1|q} &= \mathbf{V}_{t-1|t-1} + \mathbf{B}_t (\mathbf{V}_{t|q} - \mathbf{V}_{t|t-1}) \mathbf{B}_t', \end{aligned} \tag{B.6}$$

where

$$\mathbf{B}_t = \mathbf{V}_{t-1|t-1} \Phi' \mathbf{V}_{t|t-1}^{-1}. \tag{B.7}$$

The algorithm relies on having initial values $\mathbf{a}_0$, $\mathbf{Q}_0$ and error covariances $\mathbf{Q}$ of the transition equation. In practice hyper-parameters $\boldsymbol{\alpha}_0, \ldots \boldsymbol{\alpha}_q$, $\mathbf{a}_0$, $\mathbf{Q}_0$ and $\mathbf{Q}$ can be jointly estimated by the following EM-type algorithm, applied until some convergence point is reached:

1. Select initial values $\mathbf{a}_0^{(0)}$, $\mathbf{Q}_0^{(0)}$ and $\mathbf{Q}^{(0)}$.

   Iterate on steps 2 and 3 for $p = 1, 2, \ldots$

2. SMOOTHING:

   Compute $\mathbf{a}_{t|q}^{(p)}$ and $\mathbf{V}_{t|q}^{(p)}$ $(t = 1, \ldots, q)$ by the generalized Kalman filter and smoother replacing the unknown parameters by the current estimates. These are given by $\mathbf{a}_0^{(p)}$, $\mathbf{Q}_0^{(p)}$ and $\mathbf{Q}^{(p)}$.

3. EM STEP:

   Compute $\mathbf{a}_0^{(p+1)}$, $\mathbf{Q}_0^{(p+1)}$ and $\mathbf{Q}^{(p+1)}$ as follows

$$\begin{aligned} \mathbf{a}_0^{(p+1)} &= \mathbf{a}_{0|q}^{(p)} \\ \mathbf{Q}_0^{(p+1)} &= \mathbf{V}_{0|q}^{(p)} \\ \mathbf{Q}^{(p+1)} &= \frac{1}{q}\sum_{t=1}^{q}\{(\mathbf{a}_{t|q}^{(p)} - \Phi\mathbf{a}_{t-1|q}^{(p)})(\mathbf{a}_{t|q}^{(p)} - \Phi\mathbf{a}_{t-1|q}^{(p)})' + \mathbf{V}_{t|q}^{(p)} \\ &\quad -\Phi\mathbf{B}_t^{(p)}\mathbf{V}_{t|q}^{(p)} - \mathbf{V}_{t|q}^{(p)\prime}\mathbf{B}_t^{(p)\prime}\Phi + \Phi\mathbf{V}_{t-1|q}^{(p)}\Phi'\} \end{aligned} \tag{B.8}$$

   where $\mathbf{B}_t^{(p)}$ is defined in (B.7).

REFERENCES

Boudokh, J. & Richardson, M. (1994), 'The statistics of long-horizon regressions revisited', *Mathematical Finance* **4**, 103–119.

Brock, W., Lakonishok, J. & Lebaron, B. (1992), 'Simple technical trading rules and the stochastic properties of stock returns.', *Journal of Finance* **47**, 1731–1764.

Brown, S., Goetzmann, W. & Kumar, A. (1998), 'The dow theory: William peter hamilton's track record reconsidered', *Journal of Finance* **53**, 1311–1334.

Campbell, J. & Cochrane, J. (1999), 'By force of habit: A consumption-based explanation of aggregate stock market behavior', *Journal of Political Economy* **107**, 205–251.

Campbell, J. Y., Lo, A. W. & MacKinlay, A. C. (1997), *The Econometrics of Financial Markets*, Princeton University Press.

Cecchetti, S., Lam, P. & Mark, N. (2000), 'Asset pricing with distorted beliefs: Are equity returns too good to be true?', *Mimeo, Ohio State University* .

Chauvet, M. & Potter, S. (2000), ‘Coincident and leading indicators of the stock market’, *Journal of Empirical Finance* **7**, 87–111.

Chen, S. N. (1982), ‘An examination of risk return relationship in bull and bear markets using time varying betas’, *Journal of Financial and Quantitative Analysis; 17(2)* pp. 265–286.

Diebold, F. X. & Rudebusch, G. D. (1990), ‘A nonparametric investigation of duration dependence in the American business cycle’, *Journal of Political Economy* **98**(3), 596–616.

Diebold, F. X., Rudebusch, G. D. & Sichel, D. E. (1993), Further evidence on business cycle duration dependence, *in* J. H. Stock & M. W. Watson, eds, ‘Business Cycles, Indicators and Forecasting’, University of Chicago Press for the NBER, pp. 255–280.

Durland, J. & McCurdy, T. (1994), ‘Duration-dependent transitions in a markov model of u.s. gnp growth’, *Journal of Business and Economic Statistics* (12), 279–288.

Engle, R. F. & Lee, G. J. (1999), A permanent and transitory component model of stock return volatility, *in* R. F. Engle & H. White, eds, ‘Cointegration, Causality, and Forecasting: A Festschrift in Honor of Clive W.J. Granger’, Oxford University Press, pp. 475–497.

Fabozzi, F. J. & Francis, J. C. (1977), ‘Stability tests for alphas and betas over bull and bear market conditions’, *Journal of Finance; 32(4)* pp. 1093–1099.

Fahrmeir, L. (1992), ‘Posterior mode estimation by extended Kalman filtering for multivariate dynamic generalized linear models’, *Journal of the American Statistical Association* **87**(418), 501–509.

Fahrmeir, L. (1994), ‘Dynamic modelling and penalized likelihood estimation for discrete time survival data’, *Biometrika* **81**(2), 317–330.

Fama, E. F. & french, K. R. (1988), ‘Permanent and temporal components of stock prices’, *Journal of Political Economy* (2), 246–273.

Fama, E. F. & French, K. R. (1989), ‘Business conditions and expected returns on stocks and bonds’, *Journal of Financial Economics* **25**, 23–49.

Faust, J. (1992), ‘When are variance ratio tests for serial dependence optimal?’, *Econometrica* **60**, 1215–1226.

Glosten, L. R., Jagannathan, R. & Runkle, D. (1993), ‘On the relation between the expected value and the volatility of the nominal excess return on stocks’, *Journal of Finance* **48**, 1779–1801.

Gordon, S. & St-Amour, P. (2000), ‘A preference regime model of bull and bear market’, *American Economic Review* **90**, 1019–1033.

Hollander, M. & Wolfe, D. A. (1999), *Nonparametric Statistical Methods*, John Wiley & Sons, Inc.

Kalbfleisch, J. & Prentice, R. (1980), *The Statistical Analysis of Failure Time Data*, John Wiley & Sons, New York.

Kandel, S. & Stambaugh, R. (1990), ‘Expectations and volatility of consumption and asset returns’, *Review of Financial Studies* **3**, 207–232.

Kiefer, N. M. (1988), ‘Economic duration data and hazard functions’, *Journal of Economic Literature* **26**, 646–679.

Kim, M. K. & Zumwalt, J. K. (1979), ‘An analysis of risk in bull and bear markets’, *Journal of Financial and Quantitative Analysis; 14(5)* pp. 1015–1025.

Lancaster, T. (1990), *The Econometric Analysis of Transition Data*, Cambridge university press.

Leroy, S. (1973), ‘Risk aversion and the martingale property of stock returns’, *International Economic Review* **14**, 436–446.

Lo, A. & MacKinlay, A. C. (1988), ‘Stock market prices do not follow random walks: evidence from a simple specification test’, *Review of Financial Studies* **1**(1), 41–66.

Maheu, J. & McCurdy, T. (2000*a*), ‘Identifying bull and bear markets in stock returns’, *Journal of Business and Economic Statistics* **18**, 100–112.

Maheu, J. & McCurdy, T. (2000*b*), ‘Volatility dynamics under duration-dependent mixing’, *Journal of Empirical Finance* **7**, 345–372.

McQueen, G. & Thorley, S. (1994), ‘Bubbles, stock returns and duration dependence’, *Journal of Financial and Quantitative Analysis* **29**, 379–401.

Pagan, A. & Sossounov, K. (2000), ‘A simple framework for analyzing bull and bear markets’, *Journal of Applied Econometrics (forthcoming)* .

Perez-Quiros, G. & Timmermann, A. (2000), ‘Firm size and cyclical variations in stock returns’, *Journal of Finance* **55**, 1229–1262.

Pesaran, M. H. & Timmermann, A. (1995), ‘Predictability of stock returns: robustness and economic significance’, *Journal of Finance* **50**, 1201–1228.

Poterba, J. & Summers, L. (1988), ‘Mean reversion in stock returns: evidence and implications’, *Journal of Financial Economics* **22**, 27–60.

Richardson, M. & Stock, J. (1989), ‘Drawing inferences from statistics based on multi-year asset returns’, *Journal of Financial Economics* **25**, 323–348.

Samuelson, P. (1965), ‘Proof that properly anticipated prices fluctuate randomly’, *Industrial Management Review* **6**, 41–49.

Schwert, G. W. (1990), ‘Indexes of united states stock prices from 1802 to 1987’, *Journal of Business* **63**, 399–426.

Shiller, R. J. (1989), *Market Volatility*, MIT Press.

Shiller, R. J. (2000), *Irrational Exuberance*, Princeton University Press.

Sperandeo, V. (1990), *Principles of professional speculation*, John Wiley.

Sueyoshi, G. T. (1995), ‘A class of binary response models for grouped duration data’, *Journal of Applied Econometrics* **10**, 411–431.

Sullivan, R., Timmermann, A. & White, H. (1999), ‘Data-snooping, technical trading rule performance, and the bootstrap’, *Journal of Finance* **54**, 1647–1691.

Wang, J. (1993), ‘A model of intertemporal asset prices under asymmetric information’, *Review of Economic Studies* **60**, 249–282.

Whitelaw, R. F. (1994), ‘Time variations and covariations in the expectation and volatility of stock market returns’, *Journal of Finance* **49**, 515–541.

Table 1: Summary Statistics for Bull and Bear Market Durations.

Panel A: Nominal Stock Price Index

			Bull					Bear				
Filter	Series	#dur.	mean	med.	std.	min	max	mean	med.	std.	min	max
	Actual	38.0	119.8	103.0	117.2	1.00	492.0	42.63	36.00	29.75	3.00	111.0
20/15	RW	47.7	79.26	61.22	63.48	9.11	299.9	54.51	41.39	43.93	8.28	208.6
	GARCH	39.6	112.6	80.86	104.0	7.29	461.4	48.84	35.38	44.10	4.53	199.5
	Actual	66.0	61.82	51.00	60.50	1.00	354.0	33.71	25.50	30.65	1.00	115.0
20/10	RW	70.3	46.52	39.56	29.68	6.49	152.9	43.60	28.93	42.71	3.84	214.1
	GARCH	95.9	58.82	46.13	46.92	4.33	231.3	37.75	23.26	41.08	2.08	208.1
	Actual	48.0	98.53	67.00	116.2	1.00	522.0	30.60	19.00	26.39	1.00	111.0
15/15	RW	56.2	70.05	50.87	62.36	6.10	299.8	42.68	34.50	30.26	7.60	150.5
	GARCH	44.5	102.7	69.49	102.3	4.94	461.3	40.95	31.26	34.13	4.09	158.1
	Actual	86.0	48.97	27.00	56.59	1.00	354.0	24.42	17.00	23.17	1.00	111.0
15/10	RW	91.7	37.29	30.03	27.13	4.44	146.1	31.85	22.72	28.29	3.43	150.9
	GARCH	81.1	48.76	35.89	43.20	2.84	223.4	29.82	19.89	29.96	1.84	159.5

Panel B: Real Stock Price Index

			Bull					Bear				
Filter	Series	#dur.	mean	med.	std.	min	max	mean	med.	std.	min	max
	Actual	41.0	101.8	84.50	105.4	1.00	454.0	49.37	37.00	42.17	3.00	190.0
20/15	RW	47.4	79.57	61.46	63.80	9.14	301.3	54.25	41.21	43.70	8.27	207.6
	GARCH	41.5	95.84	70.28	85.54	6.76	386.6	57.79	40.77	54.03	4.68	246.2
	Actual	66.0	55.83	35.50	60.31	1.00	354.0	39.77	28.00	39.15	1.00	190.0
20/10	RW	70.5	46.59	39.60	29.74	6.49	153.1	43.37	28.76	42.50	3.83	213.2
	GARCH	64.0	53.99	43.01	41.85	4.20	205.0	45.41	27.01	50.85	2.16	254.2
	Actual	51.0	85.13	56.50	100.9	1.00	454.0	36.84	25.00	35.18	1.00	136.0
15/15	RW	56.2	70.38	51.09	62.68	6.12	301.2	42.54	34.39	30.13	7.59	149.7
	GARCH	48.1	86.17	59.42	83.67	4.54	386.6	46.74	35.00	40.12	4.17	187.9
	Actual	83.0	47.94	27.00	56.76	1.00	354.0	28.18	18.00	27.79	1.00	119.0
15/10	RW	91.7	37.36	30.08	27.20	4.45	146.5	31.73	22.63	28.17	3.43	150.3
	GARCH	81.2	44.11	33.22	37.84	2.76	196.1	34.38	22.27	35.68	1.88	188.7

This table reports summary statistics for bull and bear market durations. In Panel A the durations are derived from the nominal stock price index while in Panel B the durations are derived from the real price index. The simulated durations are generated from 2000 random walk (RW) or GARCH series, each with 31412 observations.

Table 2: Summary Statistics for Bull and Bear Market Returns.

Panel A: Nominal Stock Price Index

		Bull					Bear				
Filter	Series	mean	med.	std.	min	max	mean	med.	std.	min	max
	Actual	1.55	0.47	3.30	0.21	20.43	-1.57	-0.90	1.57	-6.11	-0.20
20/15	RW	0.70	0.59	0.40	0.23	2.22	-0.78	-0.65	0.48	-2.47	-0.21
	GARCH	0.84	0.55	0.90	0.19	4.89	-1.11	-0.70	1.29	-6.80	-0.17
	Actual	1.58	0.79	2.68	0.20	20.43	-2.00	-1.06	2.96	-20.36	0.00
20/10	RW	0.86	0.75	0.44	0.26	2.64	-0.91	-0.72	0.67	-3.58	-0.16
	GARCH	1.05	0.71	1.11	0.20	6.93	-1.29	-0.79	1.58	-9.47	-0.13
	Actual	2.35	0.62	4.79	0.21	24.65	-3.32	-1.23	7.22	-45.99	-0.20
15/15	RW	0.76	0.63	0.48	0.22	2.66	-0.86	-0.73	0.49	-2.65	-0.26
	GARCH	0.91	0.57	1.05	0.18	5.86	-1.20	-0.77	1.37	-7.48	-0.20
	Actual	2.47	0.85	4.86	0.20	29.63	-2.91	-1.26	5.89	-45.99	0.00
15/10	RW	0.96	0.83	0.52	0.28	3.23	-1.02	-0.83	0.68	-3.93	-0.23
	GARCH	1.18	0.78	1.29	0.21	8.55	-1.41	-0.88	1.67	-10.61	-0.17

Panel A: Real Stock Price Index

		Bull					Bear				
Filter	Series	mean	med.	std.	min	max	mean	med.	std.	min	max
	Actual	1.49	0.56	3.19	0.19	20.46	-1.57	-0.93	1.57	-5.96	-0.14
20/15	RW	0.70	0.59	0.40	0.23	2.21	-0.78	-0.65	0.48	-2.47	-0.21
	GARCH	0.85	0.57	0.91	0.18	4.98	-1.04	-0.64	1.25	-6.71	-0.16
	Actual	1.58	0.79	2.67	0.32	20.46	-1.87	-0.94	2.84	-20.96	0.00
20/10	RW	0.86	0.75	0.44	0.26	2.64	-0.92	-0.72	0.67	-3.58	-0.16
	GARCH	1.07	0.73	1.13	0.20	6.95	-1.23	-0.73	1.56	-9.32	-0.12
	Actual	2.10	0.69	4.40	0.19	24.65	-3.41	-1.12	7.32	-46.15	-0.23
15/15	RW	0.76	0.63	0.48	0.22	2.66	-0.86	-0.73	0.49	-2.65	-0.26
	GARCH	0.92	0.59	1.05	0.17	5.98	-1.12	-0.71	1.32	-7.42	-0.19
	Actual	2.49	0.85	4.82	0.23	29.63	-3.01	-1.22	6.18	-46.15	0.00
15/10	RW	0.96	0.83	0.52	0.28	3.23	-1.03	-0.83	0.68	-3.93	-0.23
	GARCH	1.20	0.80	1.30	0.21	8.59	-1.35	-0.82	1.64	-10.48	-0.16

This table reports summary statistics for bull and bear market returns. In Panel A the returns are derived from the nominal stock price index while in Panel B the returns are derived from the real price index. The simulated durations are generated from 2000 random walk (RW) or GARCH series, each with 31412 observations.

Table 3: Two-Sample Tests for Equality of Duration Distributions.

Panel A: Nominal Stock Price Index

		Wilcoxon		Lepage		Kol.-Sm.	
Filter	Series	Bull	Bear	Bull	Bear	Bull	Bear
20/15	Actual-RW	**0.08**	0.32	<**.01**	0.40	**0.02**	0.31
	Actual-GARCH	0.38	0.57	0.29	0.53	0.39	0.60
20/10	Actual-RW	0.06	0.22	<**.01**	0.26	<**.01**	0.14
	Actual-GARCH	0.25	0.55	**0.09**	0.31	0.14	0.41
15/15	Actual-RW	**0.07**	**0.03**	<**.01**	0.17	<**.01**	**0.02**
	Actual-GARCH	0.17	0.25	**0.05**	0.53	**0.05**	0.25
15/10	Actual-RW	**0.06**	**0.07**	<**.01**	0.15	<**.01**	**0.03**
	Actual-GARCH	0.20	0.38	**0.04**	0.37	**0.05**	0.29

Panel B: Real Stock Price Index

		Wilcoxon		Lepage		Kol.-Sm.	
Filter	Series	Bull	Bear	Bull	Bear	Bull	Bear
20/15	Actual-RW	0.26	0.30	**0.03**	0.16	**0.07**	0.20
	Actual-GARCH	0.43	0.53	0.22	0.49	0.30	0.52
20/10	Actual-RW	0.13	0.39	<**.01**	0.17	<**.01**	0.20
	Actual-GARCH	0.29	0.61	**0.08**	0.44	0.11	0.52
15/15	Actual-RW	0.18	**0.05**	<**.01**	**0.02**	**0.02**	**0.01**
	Actual-GARCH	0.26	0.19	**0.08**	0.27	0.10	0.16
15/10	Actual-RW	**0.06**	**0.05**	<**.01**	**0.01**	<**.01**	<**.01**
	Actual-GARCH	0.20	0.23	**0.03**	0.19	**0.04**	0.17

Two-Sample Tests for Equality of Duration Distributions.

Table 4: Two-Sample Tests for Equality of Bull and Bear Durations.

Panel A: Nominal Stock Price Index

Filter	Wilcoxon	Lepage	Kolmogorov-Smirnov
20/15	**<.01**	**<.01**	**<.01**
20/10	**0.01**	0.26	**<.01**
15/15	**<.01**	**<.01**	**<.01**
15/10	**<.01**	**0.07**	**<.01**

Panel B: Real Stock Price Index

Filter	Wilcoxon	Lepage	Kolmogorov-Smirnov
20/15	**<.01**	**0.03**	**<.01**
20/10	0.32	0.89	0.24
15/15	**0.02**	**0.02**	**<.01**
15/10	**0.09**	0.38	**0.06**

This table reports p-values of two-sample tests comparing bull and bear market durations. p-values below 0.1 are highlighted in boldface.

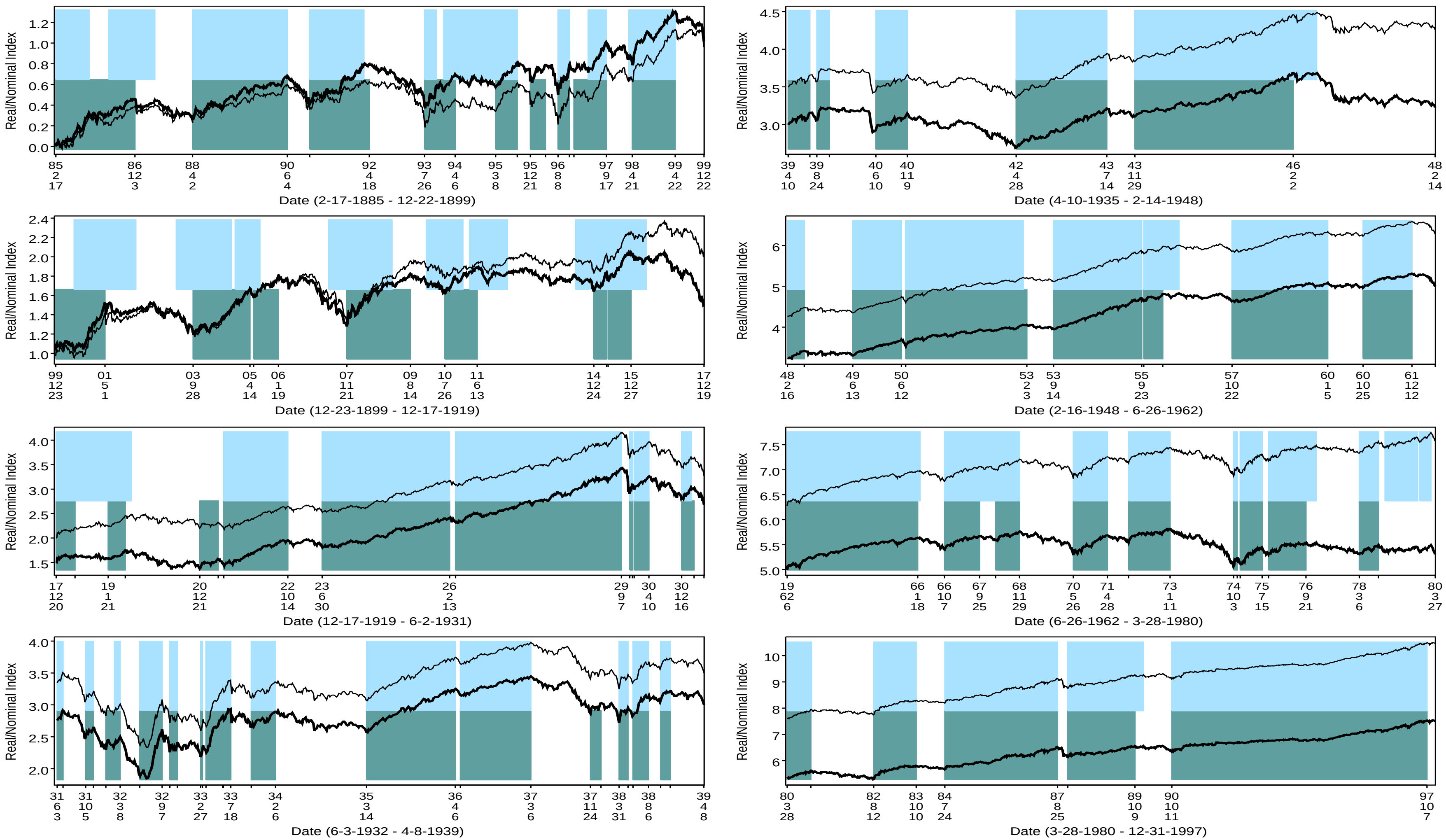

Figure 1: Bull and Bear markets based on the nominal/real S&P500 stock price index (on log-scale) using a 20/10 filter. The fat line shows the real price index and the bottom shadings track the Bull markets derived from this index. Likewise, the thin line shows the nominal price index and the top shadings track the Bull markets derived from this index.

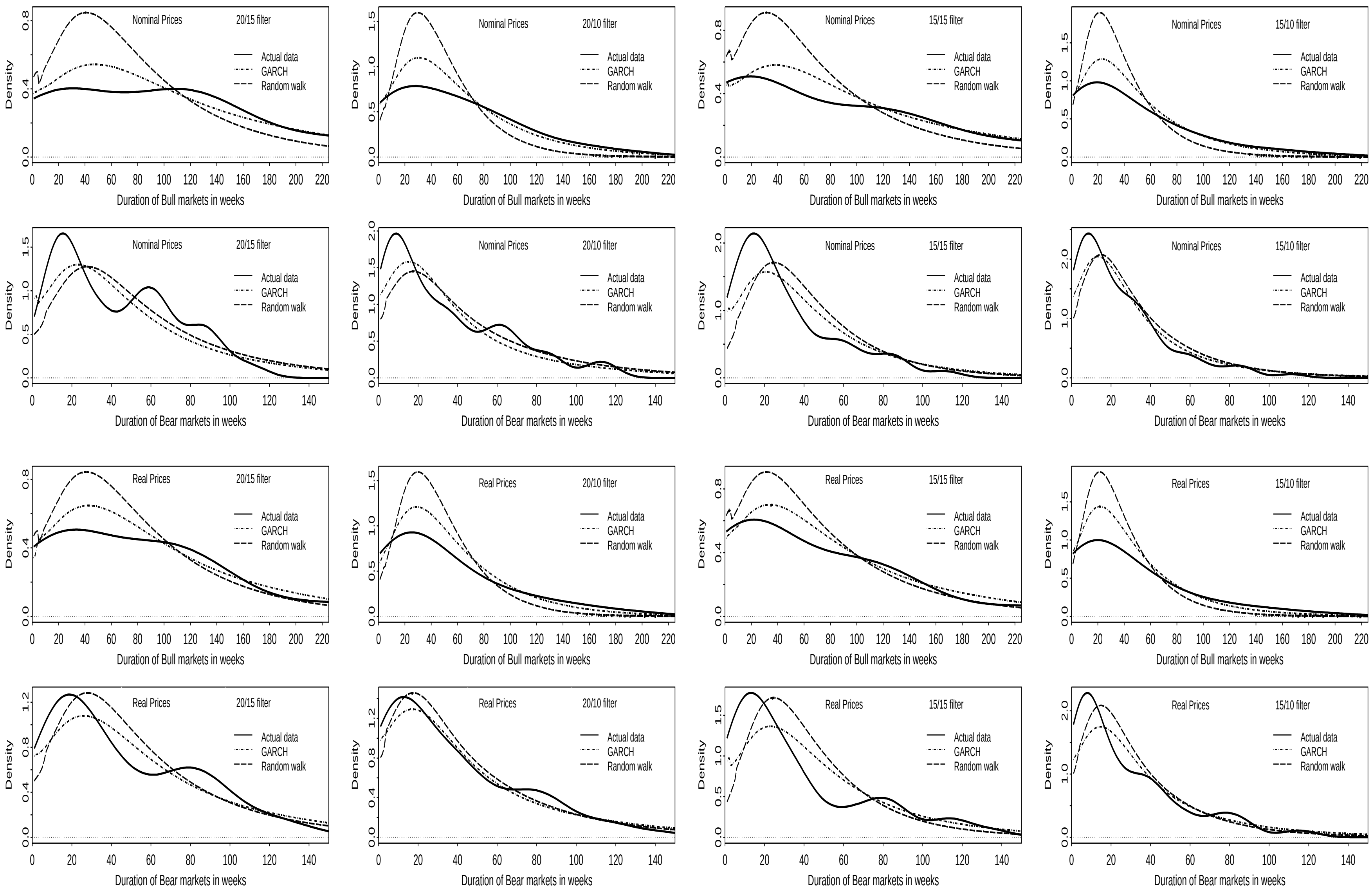

Figure 2: This figure plots kernel smoothed densities of Bull and Bear market durations derived from different filters. The simulated durations are generated from 2000 random walk or GARCH series each with 31412 observations.

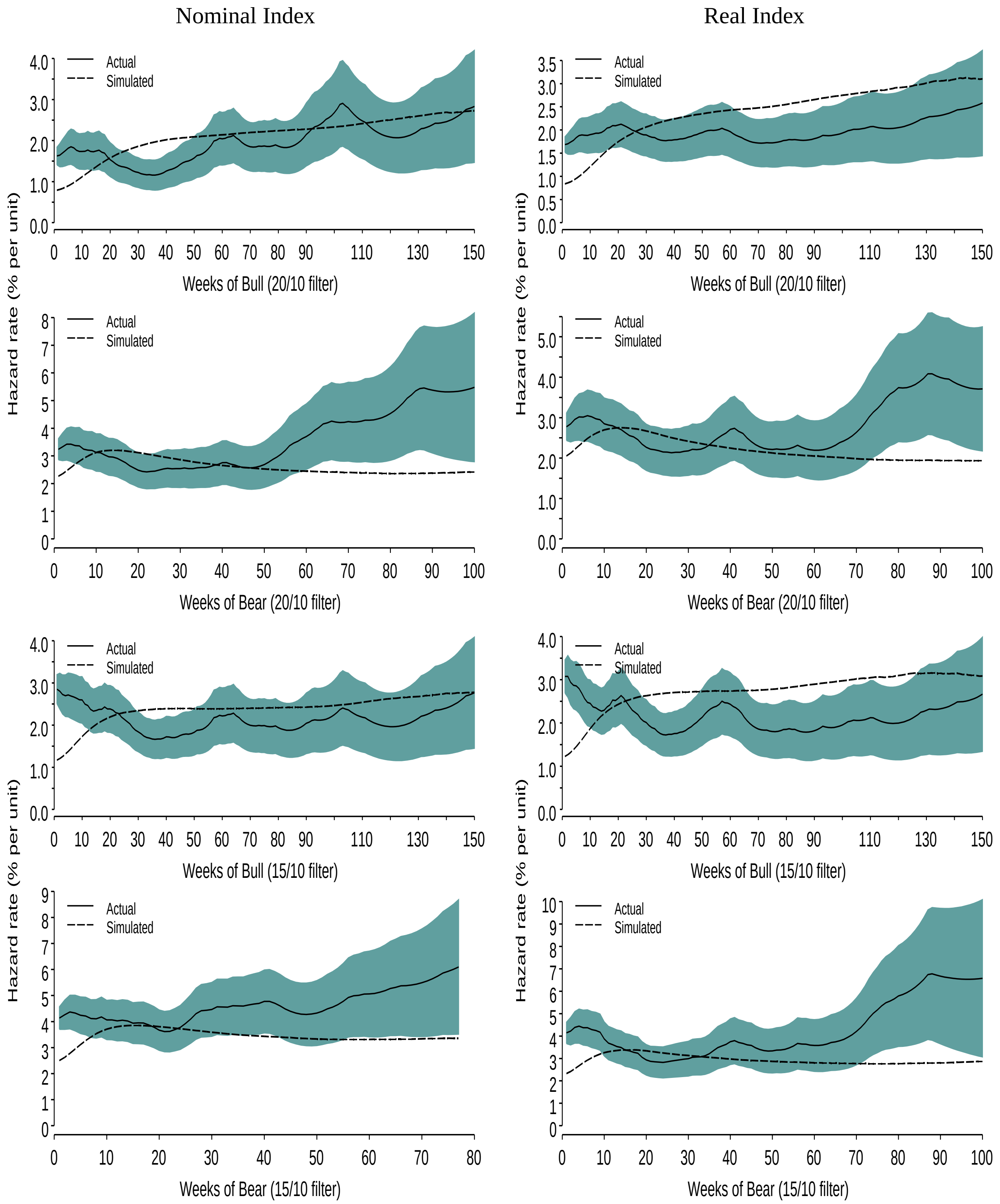

Figure 3: Unconditional hazard rates for Bull and Bear market durations derived from the nominal and real stock price index and from 2000 GARCH series, each with 31412 observations. Plots on the left are based on the nominal price index while the right plots are derived from the real price index. In all cases the hazard rate model assumes a logit link function: $\lambda(t|\boldsymbol{X}_i(t)) = F(x'_{it}\alpha_t)$, where $x'_{it} = 1$ and $\alpha_t = \gamma_{0t}$, $\gamma_{0t} = \gamma_{0t-1} + \xi_{0t}$, $\xi_{0t} \sim N(0,\sigma_1^2)$, and $\gamma_{00} \sim N(g_0, \sigma_0^2)$. The shaded area is 90% confidence region.

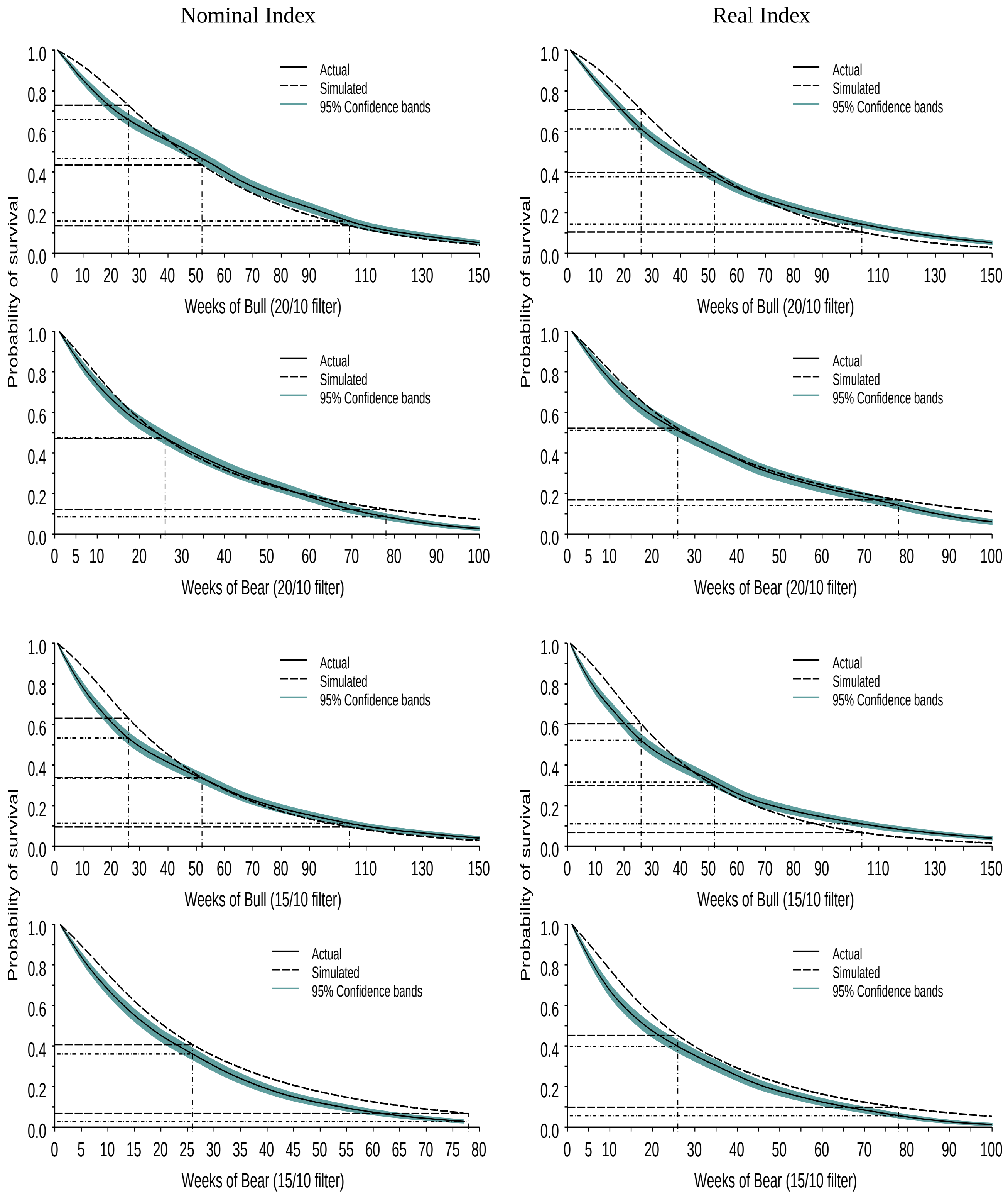

Figure 4: Survivor functions for Bull and Bear markets estimated from the unconditional hazard rates shown in Figure 3. Plots on the left are based on the nominal stock price index while plots on the right are derived from the real stock price index.

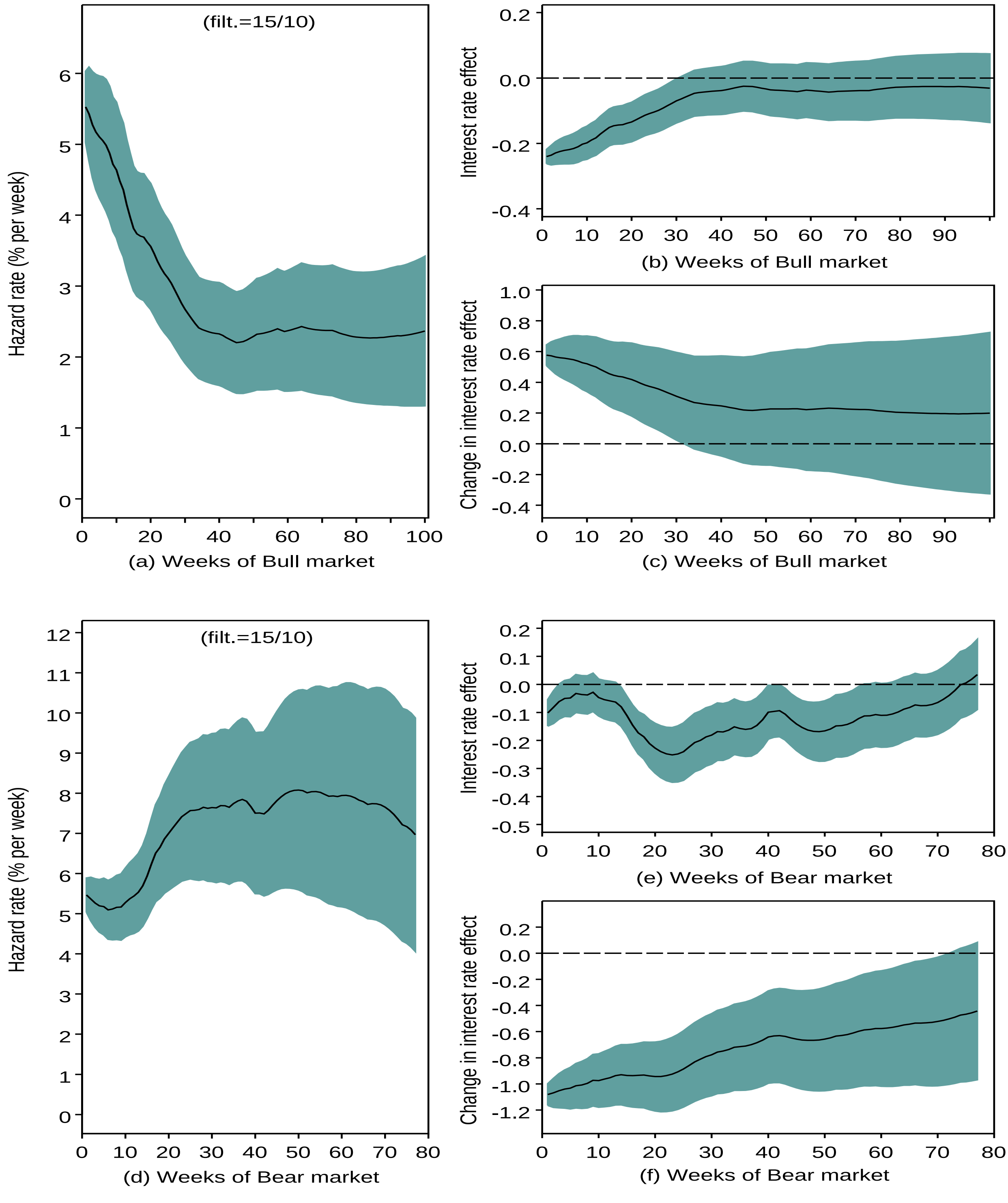

Figure 5: Interest rate effects on hazard rates: 15/10 filter, nominal stock prices and nominal interest rates. Panel (a/d) presents the baseline hazard rates for Bull/Bear markets, controlling for interest rate and interest rate change effects. Panel (b/e) shows the interest rate effect on the Bull/Bear hazard rate, and panel (c/f) plots the interest rate change effect on the Bull/Bear hazard rate. The confidence bands are $\pm$ 1 standard error. The hazard rate model is the logit link function: $\lambda(t|\boldsymbol{X}_i(t)) = F(\boldsymbol{x}'_{it}\boldsymbol{\alpha}_t)$, where $\boldsymbol{x}'_{it} = (1, i_{it}, \Delta i_{it})$ and $\boldsymbol{\alpha}'_t = (\gamma_{0t}, \boldsymbol{\beta}_t)$, $\boldsymbol{\alpha}_t = \boldsymbol{\alpha}_{t-1} + \boldsymbol{\xi}_t$, $\boldsymbol{\xi}_t \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q})$, and $\boldsymbol{\alpha}_0 \sim \mathcal{N}(\boldsymbol{g}_0, \boldsymbol{Q}_0)$. i_{it} is the nominal interest rate at the beginning of the week in question, and Δi_{it} is the change in the nominal interest rate.

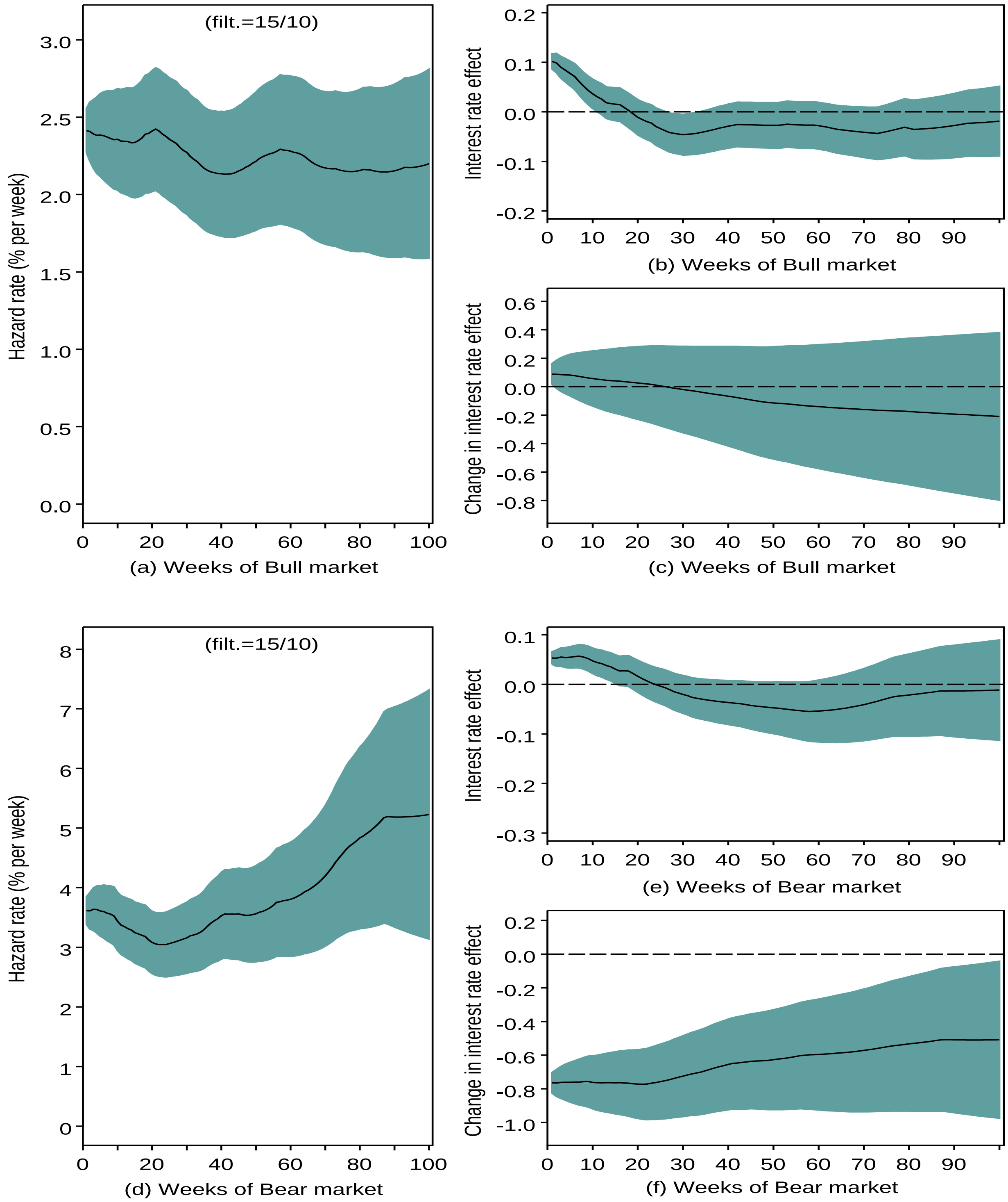

Figure 6: Interest rate effects on hazard rates: 15/10 filter, real stock prices and real interest rates. Panel (a/d) presents the baseline hazard rates for Bull/Bear markets, controlling for interest rate and interest rate change effects. Panel (b/e) shows the interest rate effect on the Bull/Bear hazard rate, and panel (c/f) plots the interest rate change effect on the Bull/Bear hazard rate. The confidence bands are $\pm$ 1 standard error. The hazard rate model is the logit link function: $\lambda(t|\boldsymbol{X}_i(t)) = F(\boldsymbol{x}'_{it}\boldsymbol{\alpha}_t)$, where $\boldsymbol{x}'_{it} = (1, i_{it}, \Delta i_{it})$ and $\boldsymbol{\alpha}'_t = (\gamma_{0t}, \boldsymbol{\beta}_t)$, $\boldsymbol{\alpha}_t = \boldsymbol{\alpha}_{t-1} + \boldsymbol{\xi}_t$, $\boldsymbol{\xi}_t \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q})$, and $\boldsymbol{\alpha}_0 \sim \mathcal{N}(\boldsymbol{g}_0, \boldsymbol{Q}_0)$. i_{it} is the real interest rate at the beginning of the week in question, and Δi_{it} is the change in the real interest rate.

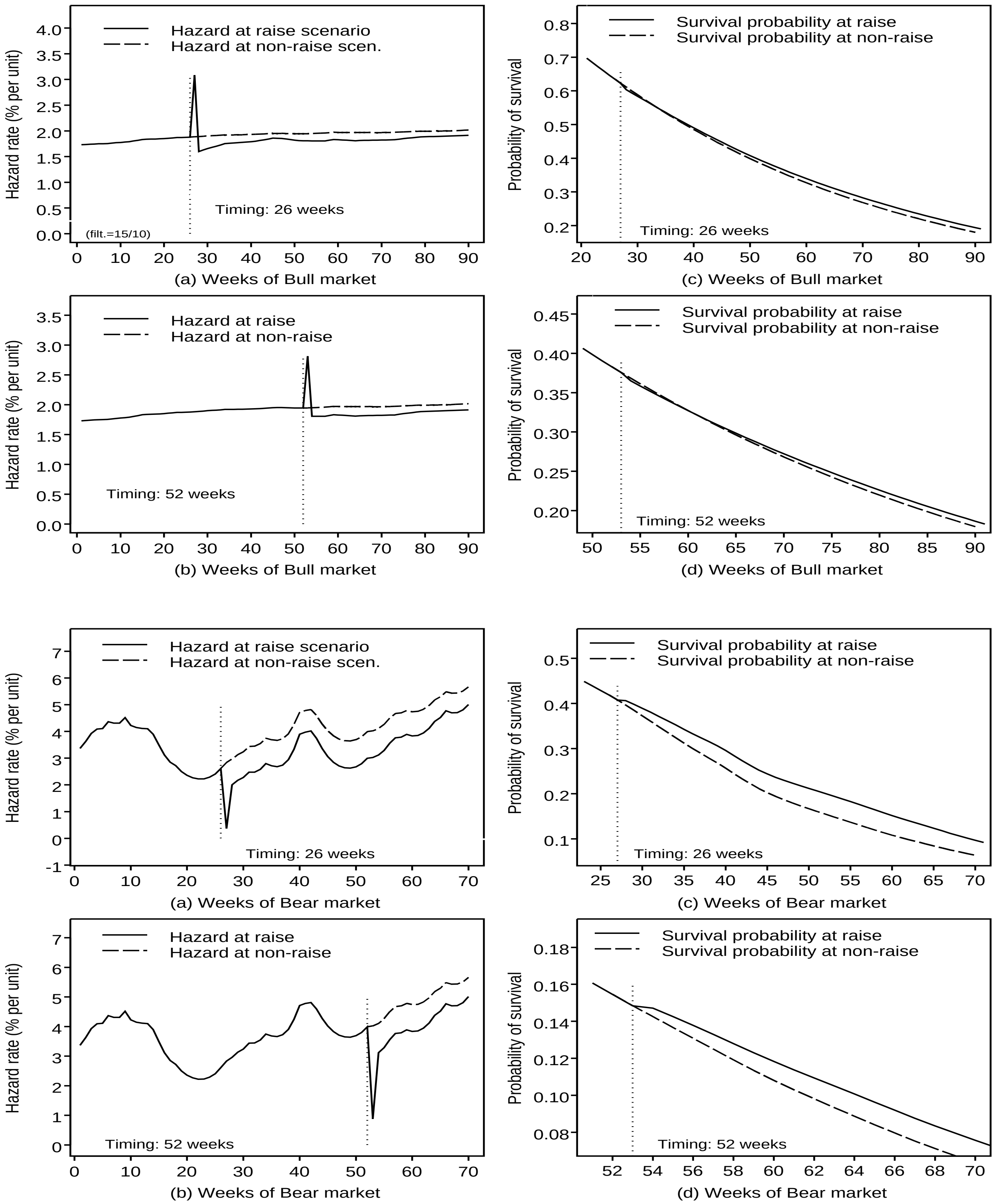

Figure 7: Scenario analysis for the effects of an interest rate increase on bull and bear hazard rates: 15/10 filter, nominal stock prices and nominal interest rates. The figures present scenarios where the nominal interest rate is raised from 5% to 7% in a bull/bear market. Panel (a) shows the effect on the hazard rate for a raise occurring after 26 weeks and panel (b) has the raise occurring after 52 weeks. Panels (c) and (d) present the corresponding survivor functions.

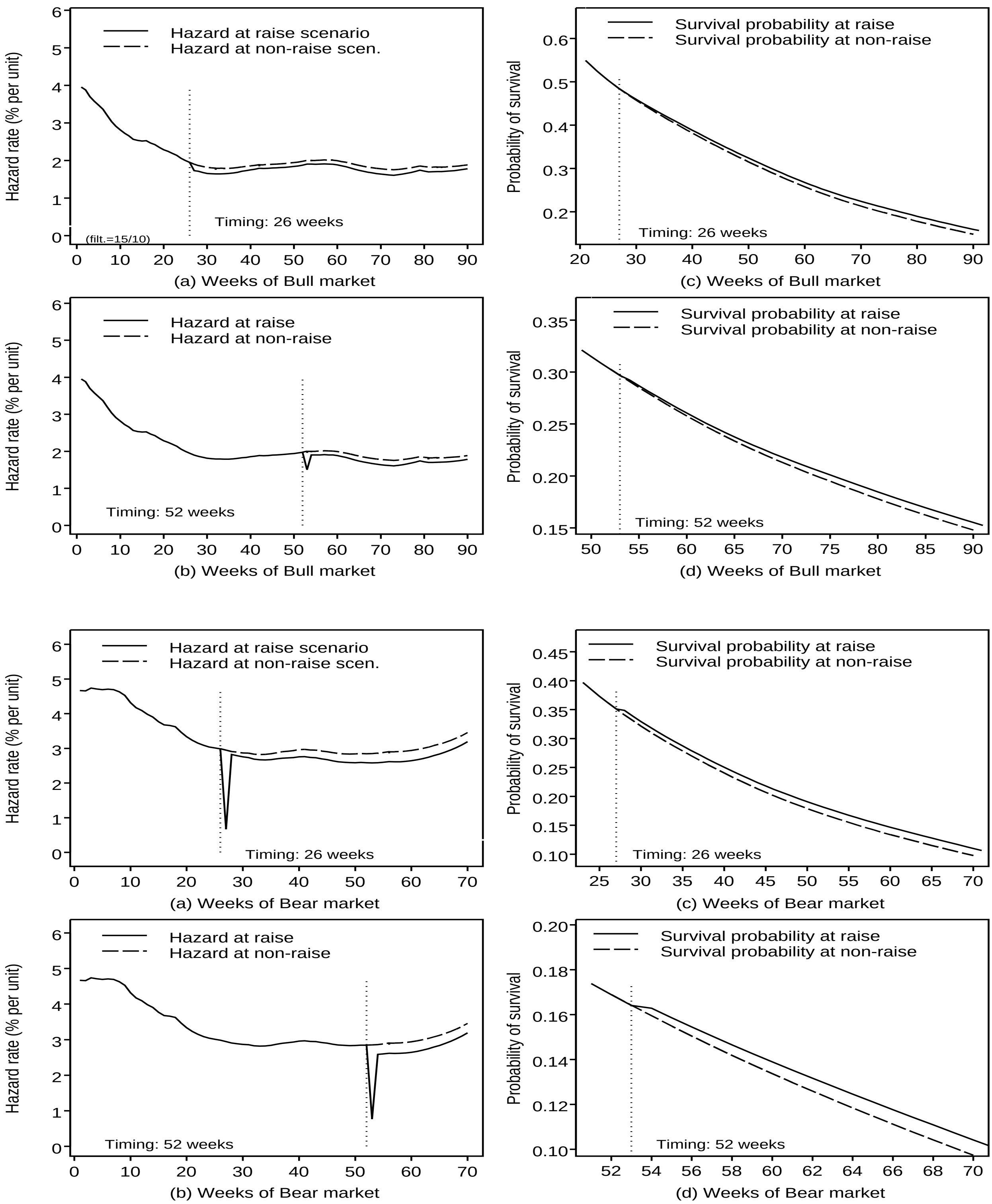

Figure 8: Scenario analysis for the effects of an interest rate increase on bull and bear hazard rates: 15/10 filter, real stock prices and real interest rates. The figures present scenarios where the real interest rate is raised from 5% to 7% in a bull/bear market. Panel (a) shows the effect on the hazard rate for a raise occurring after 26 weeks and panel (b) has the raise occurring after 52 weeks. Panels (c) and (d) present the corresponding survivor functions.