Predicting and Capitalizing on Two Types of Stock Bear Markets in the U.S.

Wan-Jung Hsu*

Department of Economics, University of Washington, Savery Hall 305, Box 353330, Seattle, WA, 98195-3330, USA.

This version: August 1st, 2016

Abstract

Forecasting the states of the stock market is of interest to policy makers and investors. While previous literature classifies the stock market into binary states (bull and bear markets), I further classify U.S. stock bear markets into good bear and bad bear markets. The latter are the bear markets associated with contraction phases of future cash flows, while the former are not. Most bad bear markets are accompanied by NBER declared recessions, whereas good bear markets are not accompanied by serious depressions in the real economy. Commonly used macroeconomic predictors also signal differently in forecasting these two types of bear markets. The value premium has distinct magnitude across the two types of bear markets. By applying a multinomial logit model with three alternatives (bull, good bear, and bad bear markets) to predict stock market states, I provide richer information about stock market states which is beneficial for policy makers and investors.

JEL Classification: C25,C53, E30, G11

Keywords: Bear markets, Multinomial logit model, Value premium, Asset Allocation

* Tel.: +1 206 619 1581
E-mail address: wjh5@uw.edu

1. Introduction

The predictability of stock market states is of interest to policy makers and investors. For policy makers, large stock price fluctuations could be an early warning signal of countrywide recessions. Barro and Ursua (2009) study the relationship between stock market crashes and economic depressions for a sample of 30 countries from 1869 to 2006. They find stock market crashes provide useful information about the prospects of a depression. Claessens, Kose and Terrones (2012) examine the relationship between business and financial cycles for a large number of countries over the past fifty years. They find recessions associated with asset price busts tend to be longer and deeper than other recessions. For investors, the stock price shifts between regimes or states of economy. The ability to identify these regimes is crucial for investment. Sigel (1991) suggests investment portfolios can be improved by switching between short-term fixed-income securities and equities before turning points in the economic cycle. Guidolin and Timmermann (2007) find the optimal asset allocations vary significantly across the business cycles as weights on various asset classes strongly depend on the perception of the state of the economy. Chen (2009) proposes that by predicting bull and bear stock market states, investors can implement a simple switching trading strategy to gain higher returns than a passive buy-and-hold trading strategy.

However, stock market crashes are more frequent than economic depressions. Barro and Ursua (2009) find that conditional on a stock market crash (return of -25% or worse) in a non-war environment, the probability of a minor depression (macroeconomic decline of 10% or more) is 22%. In reverse, conditional on a minor depression, the probability of a stock market crash is 67%. Hence, major economic depressions are particularly likely to be accompanied with stock market crashes, whereas a stock market crash could be a false alarm to the economy. In U.S. history, there are observations where the decline of stock markets does not precede or coincide with economic contractions. For example, the sharp decline in the stock market in 1962 did little to unsettle economic recovery. Also, the stock market crash of 1987 did not significantly affect economic activities. Fama (1981) and Harvey (1989) find stock returns generally don't have substantial in-sample predictive content for future output. Stock and Watson (1989, 1999a, and 2003) find equity prices are usually poor predictors of output growth. The Conference Board Leading Economic Index (LEI) looks at ten indicators, where the stock market (S&P 500 index) only constitutes one of the ten indexes with a small weighting. Samuelson's (1966) famous epigram: "The stock market has forecast nine of the last five recessions." can be a summary of above findings that not all the busts of stock markets are followed by recessions or significant economic downturns.

A present value discount model explains why the stock price moves. The fundamental source of an asset value derives from the expected cash flows that can be obtained by owning that asset. For the value of a company's equity, these cash flows come from dividends or from cash distributions resulting from earnings,

$$P_t = \sum_{\tau=1}^{\infty} E(D_{t+\tau})/(1 + r_{t+\tau})^{\tau}, \tag{1}$$

where P_t is the stock price at period t and $E(D_{t+\tau})$ denotes the expected dividends paid during period $t + \tau$, and r is the discount rate or the internal rate of return. From equation (1), it is clear that the movement in stock price level is caused by the movement in expected future dividends (expected future cash flows), or caused by the movement in discount rate. Hence, the decline of stock price is caused by either lower expected future cash flows or by higher discount rates.

In this paper, I separate stock bear markets into two types. More specifically, I use the concept of stock present value model to classify the states of stock bear markets. From equation (1), the downturns in the stock market (stock bear markets) should be associated with contraction phases of future cash flows or with higher discount rates. Hence, if a stock bear market is accompanied with a contraction phase of future cash flows, I classify it as a bad bear market; otherwise, I consider that the bear market is mainly driven by a higher discount rate and classify it as a good bear market[1]. More importantly, I show how these two types of bear markets interact with the real economy disparately, which is essential for general business and policy makers.

Previous studies have discussed how stock bear markets driven by different forces can have diverse implications for investors. Campbell and Vuolteenaho (2004) and Campbell, Giglio and Polk (2013) explain that stock market fluctuations mainly driven by movements in future cash flows or by movements in future discount rates can have very different impacts on long run investors' wealth. Stock market downturns mainly driven by cash flow news are particularly hard, whereas the downturns mainly driven by discount rate news are temporary. For example, they identify that the stock bear market of 2007-2009 is mainly driven by bad news of future corporate profits and particularly hard for investors, whereas the stock market crash of 1987 is identified as a "pure sentiment" episode which is exclusively driven by higher discount rate.

Further, Campbell and Vuolteenaho (2004) discuss the sensitivity of growth stocks and value stocks to cash flow news and discount rate news. They find value stocks are more sensitive to bad cash flow news, while growth stocks are more sensitive to bad discount rate news. Based on their findings, I suspect the behaviors of growth stocks and value stocks should be different across the two types of bear markets. Specifically, spreads between growth stocks and value stocks (value premiums) could be more notable at good bear markets, which will be profitable for investors if they can time the types of bear markets and exploit value premiums. Moreover, it is well known that the value effect is stronger among small stocks (Fama and French, 1993, 2012). Israel and Moskowitz (2013) find the value premium is largely concentrated among small stocks and is insignificant among the largest two quintiles of stocks (largest 40% of NYSE stocks). Also, Novy-Marx (2013, 2014) finds that by controlling for profitability, measured by profits-to-assets, investors can improve their trading performances substantially relative to traditional value strategies. Given these findings that size and profitability can impact the value

[1] The role of stock present value model is providing a guide to separate stock bear market into two types. However, this study is not aimed to estimate expected cash flows or discount rate precisely.

premium, combining size, value, and profitability into portfolio constructions could help investors gain higher returns.

This study is different from the previous literature in several respects. First, most literature forecasting stock market states only considers binary-state (bull and bear markets) models or only discuss certain periods of stock bear markets without clear classification standards[2]. Chen (2009) uses both the two-state Markov-switching model and the static binary probit model to predict stock market states. Nyberg (2013) finds adding dynamic structures into the binary probit model can improve predictability. Candelon et al. (2014) find binary choice models (probit or logit) with or without dynamics generally perform better than the two-state Markov-switching model. The present study classifies the stock market index into three states with clear classification methods. Second, previous studies consider univariate or multivariate forecasting models with only few macroeconomic or financial variables as predictors. The present study further includes technical indicators as candidate predictors and uses common factors estimated by principal components analysis to predict stock market states. Finally, with only two states in the stock market, previous studies have limited discussions about the implication of stock market states predictability for trading strategies. The present study, with three states of classification, can provide more sophisticated trading strategies to exploit the spreads between growth and value stocks which act differently across the two types of bear markets.

With the classification of three states (bull, good bear and bad bear markets) in the stock market, I use a multinomial logit model to forecast stock market states. The economic implications of this model for investors can be analyzed into three layers. First, in the previous literature where the forecasting models are binary-state models, the proposed trading strategy is that the investor holds the market portfolio if the model forecast is bull market state, and switches to the short term bond market if the forecast is bear market state. Hence, following the same strategy, whether the multinomial logit model can improve the trading performance by higher accuracy of predicting bear markets than a conventionally used binary logit model[3] is of interest. Second, in addition to the market portfolio and the short term bond market, the value premium is an alternative profitable investment opportunity. Therefore, unlike the previous strategy, where the investor switches to the short term bond market when the model forecast is the bear market state, whether the investor can gain higher returns by implementing a value strategy instead is of interest. In the first two trading strategies, information about bear market types is not used in the trading decisions. Third, through using information about bear market types, can the investor implement a more sophisticated trading strategy and gain higher returns? Specifically, the trading strategy I propose is that the investor holds market portfolio if the

[2] Campbell and Vuolteenaho (2004) only analyze stock bear markets for certain periods. Campbell, Giglio and Polk (2013) analyze stock markets as event studies. Neither study aim to classify stock markets into discrete states and so does not distinguish stock markets with specific standards.

[3] Previous literatures only investigate the predictability of stock market states in the binary class frameworks (i.e., bull and bear markets), such as binary probit or binary logit model. For the consistency of comparison, I choose binary logit model as the alternative to the multinomial logit model proposed in this study.

forecast is bull market state, switches to the short term bond market if the forecast is bad bear market state, but performs various value strategies, combining size, value, and profitability if the forecast is good bear market state.

Overall, this paper answers three questions. First, do different types of stock bear markets interact with the real economy differently? Second, compared to the commonly used binary logit model, does the multinomial logit model improve the ability of forecasting bear markets? Third, what is the economic implication for investors?

The empirical analysis shows that, first, real economic activities behave distinctly between the two types of bear markets. While all the real economic activities severely deteriorate at bad bear markets, most of them still mildly expand during good bear markets. This implies real activity indicators (and macroeconomic variables that predict real economy states) actually signal differently across the two types of bear markets. Without distinguishing bear markets into two types, information contained in these macroeconomic variables cannot be revealed and used efficiently in a binary class model. Second, since the dataset forming the forecasting models includes real economic activity indicators, the forecasting models suffer real-time data availability and revision issues. While the out-of-sample (with the ex-posted revised data) evaluation shows the dynamic multinomial logit model has better classification ability than either the static or the dynamic binary logit model, it becomes statistical indifferent under the real-time out-of-sample examination. However, the dynamic multinomial logit model still reserves the better ability in predicting bear markets. Third, by taking information about bear market types into trading decisions, the multinomial logit model can improve trading performances dramatically through exploiting value premiums (monthly Sharpe ratio increases from benchmark Buy-and-Hold strategy 0.16 to maximum 0.32) and decreasing the maximum dropdown from 51% to 23%. This information superiority is prominent either under ex-post revised data or real-time data out-of-sample evaluations.

The rest of the paper is organized as follows. Section 2 presents the method of determining stock market states. Section 3 discusses the interaction of stock markets with real economic indicators across three states. Here, I also investigate the magnitudes of different value premiums that combine size, value and profitability characteristics across three stock market states. Section 4 presents the multinomial logit model and compares its predicting performance with the binary logit model under in-sample and out-of-sample (with the ex-post revised data) tests. I consider both static and dynamic specifications for each model. Section 5 shows economic implications of the multinomial logit model for investors. Section 6 performs the real-time out-of-sample tests as a robustness check. Section 7 concludes.

2. Classifying states of the stock market

Different from the commonly used two-state binary model (bull and bear markets), I decompose stock bear markets into good bear and bad bear markets using the stock present value model Eq. (1). The procedure is: (1) classify the stock market and its cash flows into contraction and expansion phases; (2) based on whether a stock bear market (the contraction phase in the stock market) is associated with a contraction phase of cash flows classifies bear markets into bad bear and good bear markets.

2.1. Identifying contraction and expansion phases in the stock market and its cash flows

Following Pagan and Sossounov (2003) and Chen (2009), I use the U.S. monthly S&P 500 index as the price level of the stock market. To measure the stock market's future cash flows and being aware of smoothing policy in dividends[4] and the reporting frequency of other cash flows alike measurements[5], I choose 12-month moving average of earning on S&P 500 index as the proxy of cash flows.[6] Miller and Modigliani (1961), Dechow et al. (1998), Kim and Kross (2005) and Chen, Da, and Priestley (2012) also choose earnings as a more informative measurement to reflect future cash flows. The data sample period is from January 1967 to December 2013[7].

Chen (2009) uses two approaches to identify bull and bear stock markets. One is based on a two-state Markov-Switching model and another is based on the nonparametric Bry-Boschan (1971) dating rule. The Bry-Boschan dating rule has been extensively used in the business cycle literature. Harding and Pagan (2003) compare the Markov-Switch model implied business cycles with that generated by Bry-Boschan dating rule. They conclude that Bry-Boschan dating rule is preferable in terms of transparency, simplicity, and replicability. Nyberg (2013) compares the Markov-Switching model with the Bry-Boschan dating rule and concludes that the Bry-Boschan dating rule can estimate the states of stock market more accurately and fit the data better in terms of higher log-likelihood value and lower Akaike (AIC) and Schwarz (BIC) information criteria values. Candelon et al.(2014) find the binary class model based on Bry-Boschan dating rule outperforms the Markov-switching model in terms of in-sample and out-of-sample fit.

Following Chauvet and Potter (2000), Pagan and Sossounov (2003), Candelon et al. (2008, 2014), and Nyberg (2013), I apply the Bry-Boschan dating rule to the S&P 500 index and its earnings to generate their chronology, respectively. Based on the assumptions made by Candelon et al. (2008) and the modification made by Claessens et al. (2012) for dating financial markets, the Bry-Boschan dating algorithm searches for maxima and minima in the series over a two-side window of 6 months length. Then, it selects pairs of adjacent, locally absolute maxima and minima meeting certain censoring rules. In particular, it requires a complete cycle and each

[4] Chen, Da, and Priestley (2012) find the divided smoothing policy, which is most appearing in postwar periods, can bury its' signal about future cash flows.
[5] Earnings are reported shortly after the quarter-end, whereas cash flow statements are not quarterly reported and are not required-reported until 1988.
[6] Since earning has strong seasonality, without notification, all the earnings data used in this study are 12-month moving averaged.
[7] The availability of predictors used in the forecasting model determines the beginning of the sample period.

phase to last at least 15 months and 6 months[8], respectively. Assuming y_t is the time series data examined, the turning point would be a peak at time t if, $y_{t-6}, \cdots, y_{t-1} < y_t > y_{t+1}, \cdots y_{t+6}$, and a trough if, $y_{t-6}, \cdots, y_{t-1} > y_t < y_{t+1}, \cdots y_{t+6}$. Periods from trough to peak are classified as the expansion states ($S_t = 0$), while periods from peak to trough are the contraction states ($S_t = 1$), where S_t is a binary index to indicate the expansion and contraction phases in the series.

2.2. Classifying stock bear markets into good bear and bad bear markets

The resulting chronology of the stock market in this study is consistent with Chauvet and Potter (2000), Pagan and Sossounov (2003) and Nyberg (2013). Next I compare the chronologies of the stock market and its earnings. To be consistent with Bry-Boschan dating algorithm, I choose a window of 6 months ahead to decide whether a bear market is associated with a contraction phase in earnings. That is, if a stock bear market is accompanied with a contraction phase of earnings within 6 months[9], it is classified as a bad bear market; otherwise it is classified as a good bear market. Figure 1 plots the time series of S&P 500 index and its earnings. The shaded bars indicate stock bear markets identified by the Bry-Boschan dating algorithm, where pink indicates good bear markets and grey indicates bad ones. Based on this classification rule, one original stock bear market phase could comprise two types of stock bear markets. For example, during the dot.com bubble bear market (2000/09 to 2002/09), the first 19 months of this period are associated with contraction phases in earnings and classified as bad bear markets, whereas the last 6 months of this period are associated with strong expansion phases in earnings and identified as good bear markets.

[Insert Figure 1]

3. Interaction of stock market states with the real economy

3.1. Real economic indicators across three stock market states

According to Campbell and Vuolteenaho (2004) and Campbell, Giglio and Polk (2013), stock returns mainly driven by cash flow news or discount rate news are accompanied with different economic states. Here, I consider four real economic indicators: employees on nonfarm payrolls (EMP), real manufacturing and trade sales (MTS), personal income less transfer payments (PIX) and industrial production (IP). These four indicators are variables that constitute the Conference Board's Index of Coincident Indicators and also as of the primary series that NBER Business Cycle Committee uses to establish its business cycle chronology (Hall 2002). To characterize the changes in these indicators along stock market states, following Claessens et al. (2012), I calculate amplitude and cumulative loss to address the dynamics of these variables. Following

[8] Since financial variables are much more volatile than economic business series, the duration constraint for a contraction phase reduces to at least three months if the series declines more than 20% in three months, a threshold used in Pangan and Sossounov (2003) and Claessens et al. (2011b).

[9] Empirically, in my data sample, current state of the stock market has the highest correlation with future 6 to 10 months earning states. Using a window of 10 months doesn't affect the implications of this study.

formulas from Claessens et al. (2012), the amplitude of a bear market, A_c, measures the change in y_t from a peak (y_0) to the next trough (y_k) ($A_c = y_k - y_0$). The amplitude of a bull market, A_u, measures the change in y_t from a trough (y_k) to the level reached one year after the trough[10] ($A_u = y_{k+12} - y_k$). For bear market only, another widely used measurement, cumulative loss, combines information on duration and amplitude to proxy for an overall cost of a bear market. The cumulative loss, F_c, of a bear market, with duration k, is calculated as[11]: $F_c = \sum_{j=1}^{k}(y_j - y_0) - \frac{A_c}{2}$. For comparison, I standardize all the indicators with each one's sample mean and standard deviation before calculating any measurements. Moreover, since the bust of stock market is likely to lead a recession, I also calculate how the frequency of a certain type of bear market that is followed by a NBER recession in 6 months.

Table 1 reports the results. Comparing all three states in the stock market, bull markets constitute 76% of stock market states; bad bear markets constitute 15%, while good bear markets constitute only 9% of stock market states. Correspondingly, the duration of bull market is longest, 43.6 months on average, while the duration of good bear market is shortest, 8.7 months on average. Moreover, 88% of bad bear markets are followed by NBER recessions, while only 19% of good bear markets lead NBER recessions. This implies knowing the type of a bear market is critical to decide whether it is an early warning of economic downturns. For the fluctuations of stock prices and real economic indicators across stock market states, all the variables deteriorate much more in bad bear markets relative to good bear markets. Particularly, the four economic indicators still moderately expand during good bear markets. Comparing the amplitude measurement between bull and good bear market states, some real economic indicators (IP, PIX and EMP) increase even less in bull markets than they do in good bear markets. This is caused by the lower recovery rate of speed for real economic indicators after bad bear markets. Figure 2 depicts the time series of four standardized real economic indicators. Apparently, none of them declines at good bear market states, while most drop at bad bear market states with certain lags. For example, EMP (employees on nonfarm payrolls) usually declines later than other indicators at bad bear states.

[Insert Table 1]

[Insert Figure 2]

3.2. Value premiums across stock markets states

Campbell and Vuolteenaho (2004) shows value stocks are more sensitive to cash flow news (the main driving force of bad bear market), while growth stocks are more sensitive to discount rate news (the main driving force of good bear market). Based on their findings, it is interesting to investigate the returns of a value strategy investment rule (buy value stocks and short growth

[10] Since the recovery after a contraction phase is the matter of interest, only a certain period after the stock market trough is considered in analysis.
[11] This formula is based on a triangular approximation of lost output during a contraction phase.

stocks) across the two types of bear markets. Previous studies also find a firm's size and profitability can impact the returns of value strategy[12]. Specifically, value strategies conditional on small size and high profitability firms have higher average returns. Therefore, in this section, I investigate the returns of several value strategies that combine size, value, and profitability characteristics constructed from 32 stock portfolios sorted by (2x4x4) size, value, and profitability. Value strategies considered include: HML (value strategy conditional on size and profitability), HML_S (value strategy conditional on profitability within small firms), HML_RMW (value and profitability combined strategy conditional on size), HML_RMW_S (value and profitability combined strategy within small firms). Appendix A gives the details in construction of these value strategies.

Table 2 shows monthly sample mean, volatility and Sharpe ratio of value-weighted market excess return, 3-month Treasury bill rate, and returns of different value strategies across stock market states. Interestingly, for excess market return, differences in sample mean and volatility between two types of bear markets are quite small. This shows the good and bad bear market classifications have distinct meaning in their interaction with real economy but not in the aggregate stock market itself. For value strategies, the returns are consistent with Campbell and Vuolteenaho (2004)'s findings, Sharpe ratios of all value strategies are much higher at good bear market states than they are at the other two states. Among value strategies, combining profitability with value strategies does increase the average returns substantially. Monthly average return of the HML_RMW strategy under full sample period is 0.95%, nearly twice as high as that of HML strategy (0.5%). However, the higher return of HML_RMW strategy comes with higher exposure to volatility. Monthly volatility is 41.88% in HML_RMW strategy versus 12.83% in HML strategy, which results the Sharpe ratios between HML_RMW and HML strategies are very close (the difference is 0.01). On the other hand, restricting value or value and profitability combined strategies in small firms seems to be more profitable without increasing volatility risk much. Under full sample period, monthly Sharpe ratio of HML_S strategy is 0.04 higher than that of HML strategy, whereas the Sharpe ratio of HML_RMW_S is 0.06 higher than that of HML_RMW strategy.

Overall, the statistics from Table 2 imply that value strategies provide better investment opportunities for investors at good bear market states which will be feasible if investors can predict bear market types. For example, if an investor currently invests in the market portfolio and knows that the next month will be a good bear market state, then instead of holding the market portfolio or switching to the 3-month Treasury bill market, he could implement HML strategy (buy value stocks and short growth stocks conditional on profitability and size) at the end of current month. In this case, at the end of next month, on average, the investor will earn 1.95% rate of return (assuming no transaction cost) which is much higher than holding the

[12] Fama and French (1993, 2012), Israel and Moskowitz (2013) find the value premium is largely concentrated among small stocks. Novy-Marx (2013, 2014) finds controlling for profitability, investors can improve trading performances substantially.

market portfolios (-2.08%) or switching to the 3-months Treasury bill market (0.53%). Combining value strategies with profitability and size characteristics can be more profitable but with the threat of higher volatility risk. More comprehensive evaluations of implementing various value strategies will be discussed in section 5.

[Insert Table 2]

4. Model specifications and evaluations

Previous literature predicts stock market states with two-state models. Chen (2009) predicts bull and bear markets with a static binary probit model. He finds macroeconomic variables are useful in predicting stock market states both in-sample and out-of-sample. Nyberg (2013) finds the dynamic autoregressive probit model can improve predictability substantially. Candelon et al. (2014) find binary choice models (probit or logit) with or without dynamics generally perform better than the two-state Markov-Switching model, while the dynamic binary choice models (probit or logit) perform best. Unlike the previous literature, I emphasize the importance of distinguishing bear markets into good and bad. With three states in the stock markets, I use a multinomial logit model to predict stock market states and compare its forecasting performance to the conventional binary choice model (binary logit model for consistent comparison). To comprehend the previous literature, I also investigate the predictabilities of dynamic multinomial logit model and the dynamic binary logit model.

4.1. Binary logit model

The binary logit model assumes the stock market can be modeled as a binary state variable S_t, that the stock market is either in a bull state ($S_t = 0$) or in a bear state ($S_t = 1$). Denoting a vector of explanatory variables (predictors) as X_t, the information set at time *t* is given by $\Omega_t = \sigma[(S_s, X_s), s \leq t]$. Denoting the conditional expectation given information set Ω_{t-1} as $E_{t-1}(.)$, the conditional probability of a bear market state at time *t* can be written as:

$$p_t = E_{t-1}(S_t) = P_{t-1}(S_t = 1) = \Lambda(\pi_t). \quad (2)$$

In this express, π_t is a linear function of the variables included in Ω_{t-1} and $\Lambda(.)$ is the cumulative distribution function of a logistic distribution. The linear function π_t should be determined to complete the model for future states of the stock market. In the static model, π_t is specified as:

$$\pi_t = X_{t-h}'\beta, \quad (3)$$

where vector X_{t-h} contains predictors, and h denotes the forecasting horizon. Under this specification, with forecasting horizon one month ahead, the conditional probability of a bear market state at time *t* is as:

$$P_{t-1}(S_t = 1) = \Lambda(\pi_t) = \frac{e^{(X_{t-1}{}'\beta)}}{1+e^{(X_{t-1}{}'\beta)}}. \tag{4}$$

Parameters in eq. (4) can be estimated by the maximum likelihood (ML) method. In addition, the odds ratio in this model is defined as:

$$\frac{P_{t-1}(S_t=1)}{P_{t-1}(S_t=0)} = e^{X_{t-1}{}'\beta}. \tag{5}$$

Therefore, the effects of predictors on the probability of a bear market state relative to the probability of a bull market state are measured by β[13]. In this setting, no matter what the underlying bear market type is, the model assumes and estimates the same β for both good and bear markets.

To add dynamic structures in the conditional probability p_t, or equivalently , π_t, the *k*-period lagged state, S_{t-k} , can be simply added as:

$$\pi_t = X_{t-1}'\beta + \delta_k S_{t-k} \ , \tag{6}$$

In Nyberg (2013) and Candelon et al. (2014), their dynamic binary probit/logit models are specified with one period lagged stock market state. However, due to the rules of Bry-Boschan dating algorithm, current state of the stock market can only be certain 6 months later. I therefore use S_{t-6} in my dynamic binary logit model specification. Unlike Nyberg (2013) who assumes a 6-month information lag in the value of stock market states in his out-of-sample test, using the 6-month lagged state as the dynamic specification does not need any assumption of the lag in investors' information about stock market states.

4.2. Multinomial logit model

On the other hand, being aware that predictors could signal differently across two types of bear markets, I use a multinomial logit model with 3 states, the bull market state (S_t= 0), the good bear market state (S_t = 1), and the bad bear market state (S_t = 2), to predict stock market states. Conditional on information set Ω_{t-1}, S_t has a distribution function with probabilities $p_{jt} = P_{t-1}(S_t = j) = \Lambda(\pi_{jt})$; $j = 0, 1, 2,$ where $\sum_{j=0}^{2} p_{jt} = 1$. Under these specifications, using the bull market state ($S_t = 0$) as the base state, the conditional probabilities of one month ahead stock market states at time *t* are as:

$$P_{t-1}(S_t = j) = \frac{\exp(\pi_{jt})}{1+\sum_{i=1}^{2}\exp(\pi_{it})} \ , j = 1, 2 \tag{7}$$

$$P_{t-1}(S_t = 0) = \frac{1}{1+\sum_{i=1}^{2}\exp(\pi_{it})} \tag{8}$$

---

[13] The marginal effect of a change in predictors on the probability of outcome S_t is not constant but depends on the precise values of predictors in X_{t-1}.

where $\pi_{jt} = X'_{t-1}\beta_j$.

The odds ratios between states are:

$$\frac{P_{t-1}(S_t=1)}{P_{t-1}(S_t=0)} = e^{\pi_{1t}} = e^{(X_{t-1}{}'\beta_1)} \tag{9}$$

$$\frac{P_{t-1}(S_t=2)}{P_{t-1}(S_t=0)} = e^{\pi_{2t}} = e^{(X_{t-1}{}'\beta_2)} \; . \tag{10}$$

In this specification, $\beta_1$measures the effect of a change in predictors X_{t-1} on the probability of S_t being in the good bear market state relative to the probability of being in the bull market state. Accordingly, β_2 measures the effect of a change in predictors X_{t-1} on the probability of S_t being in the bad bear market state relative to the probability of being in the bull market state. The key advantage of the multinomial logit model is that it allows the model to explicitly distinguish between three states, and enables the predictors X_{t-1} to have different impacts β_1 and β_2 across states.

To add dynamic structures in the multinomial logit model, it is useful to define indicator function I_{jt} , such that I_{jt}=1, if $S_t = j$, or $I_{jt} = 0$, otherwise; $j = 0, 1, 2$. The conditional probability p_{jt}, or equivalently , π_{jt}, can be specified as:

$$\pi_{jt} = X'_{t-1}\beta_j + \sum_{i=1}^{2} \delta_{ji}^{k}\, I_{it-k} \; , j = 1{,}2 \; . \tag{6}$$

Because of my classification method of bear market types, current state of the stock market can only be certain 12 months later. I therefore use $I_{jt-12}, j = 0, 1, 2$ in my dynamic multinomial logit model specification.

About the predictors used in this study, previous studies use macroeconomics or financial variables. Chen (2009) finds term spread and inflation are the most useful predictors. Nyberg (2013) further finds the past stock return and the dividend-price ratio also have significant ability in predicting stock market states. Candelon et al. (2014) show that term spreads, inflation, and industrial production yield better predicting results. Among these studies, because of relatively rare bear market periods[14], univariate models or multivariate models with few individual variables are used in their forecasting tests.

Asset pricing theory posits that stock return predictability could result from its exposure to time-varying aggregate risk, which depends on the states of economy or business-cycle fluctuations. Variables that measure and/or predict the states of economy should help predict stock returns (Fama and French, 1989; Campbell and Cochrane, 1999; Cochrane, 2007, 2011). Besides common financial and macroeconomic variables related to stock and bond markets, Berge (2015) and Fossati (2015) find variables that describe real economic activities provide

[14] The bear markets contribute 23% of the whole stock market in my sample period.

clearer signals about the states of economy. On the other hand, the technical analysis has long been applied in industry practice. Many brokerage firms publish technical commentary on the market and many advices are based on the technical analysis. Schwager (1993, 1995) finds many top traders and fund managers use it. Covel (2005) advocates the use of technical analysis by citing examples of large and successful hedge funds. Faber (2007) proposes a simple technical asset allocation rule among multiple asset classes which can improve trading performance substantially. Interestingly, Neely et al. (2014) find technical indicators can provide complementary information about the business cycle beyond macroeconomic variables. Hence, in this study, I consider 14 monthly macroeconomic variables, 14 monthly technical indicators, and 4 monthly real economic activity indicators from January 1967 to December 2013[15] as candidate predictors.

Table 3 provides descriptions and data sources of all variables in details. Table 4 reports prescriptive statistics summaries. With a total of 32 variables, I first examine each individual variable's predictive power, and then use factors estimated by principal components analysis from all candidate predictors in the forecasting model. With 32 highly correlated variables, using factors extracted from principal components analysis can efficiently reduce the dimensionality of a dataset. This method has been proven successful in many forecasting studies (Stock and Waston, 1991, 2002a, 2006; Ludvigson and Ng, 2007, 2009). Similar approaches have been used in Chen et al. (2011), Bellego and Ferrara (2012), Fossati (2014), and Christiansen et al. (2014) in dating or forecasting recessions.

[Insert Table 3 and Table 4]

4.3. Evaluation measures

To assess the performance of models, several forecast metrics are used for in-sample and out-of-sample evaluations. For in-sample evaluation, in addition to addressing the significance of predictors' coefficients, pseudo-R^2 and Schwarz Information Criterion (BIC) are considered to measure model fits and especially to decide the number of factors from principal components analysis to be used in the out-of-sample test. For out-of-sample evaluation, two conventional measures are used: quadratic probability score (QPS) and log probability score (LPS), proposed by Diebold and Rudebusch (1989). The QPS statistic is simply a mean square error measure comparing the predicted bear market probability with the true stock market state:

$$QPS = \frac{2}{T}\sum_{t=1}^{T}(\widehat{P}_t - S_t)^2, \quad (7)$$

where $\widehat{P}_t$ represents the predicted probability of bear market at time t and S_t is the state variable which is equal to 1 if the realized state is bear market state, or 0 , otherwise. LPS statistic corresponds to a loss function that penalizes large errors more heavily:

[15] The availability of the real economic activity variables determines the beginning of the sample period.

$$LPS = \frac{-1}{T}\sum_{t=1}^{T}[\,(1-S_t)\ln\big(1-\widehat{P}_t\big) + S_t\ln(\widehat{P}_t)]. \tag{8}$$

QPS and LPS range from 0 to 2 and from 0 to ∞, respectively, where score 0 for both QPS and LPS represents perfect predicting accuracy. However, these evaluation measures focus on the model's fit but not specifically classification ability. Recently, the Receiver Operating Characteristic (ROC)[16] curve has been applied to evaluate financial crisis early-warning systems (EWS) (Candelon, Dumitrescu, and Hurlin, 2012) and evaluating the classification of states of the economy (Berge and Jorda`, 2011). In particular, by using the area under the ROC curve (AUC), one can measure the categorization ability of a model over the entire spectrum of different cut-offs determining bear markets, instead of any one arbitrary threshold. It is also a model-free method that can assess the forecasts issued from different model specifications. Hence, in addition to QPS and LPS, I provide AUC measures to give a more appropriate and comprehensive evaluation[17]. A perfect classification has an AUC of 1, whereas a coin-toss classification has an AUC of 0.5.

Several test statistics are provided for model comparisons emphasized in the out-of-sample examination. For comparisons based on forecasting errors of two competing models, $\{e_{1,t}\}_{t=1}^{T}$and $\{e_{2,t}\}_{t=1}^{T}$, with $e_{j,t} = S_t - \hat{p}_{j,t}$ for $j = 1, 2$, the null hypothesis of equal predictive accuracy is conditional on a loss function, $g(e_{j,t}) = (S_t - \hat{p}_{j,t})^2$. For non-nested models, Diebold and Mariano (1995) propose a test statistic *DM*:

$$DM = \frac{\sqrt{T}\bar{d}}{\sigma_{\bar{d},0}} \xrightarrow{d} N(0,1),$$

where $d_t = (S_t - \hat{p}_{1,t})^2 - (S_t - \hat{p}_{2,t})^2$, $\bar{d} = (\frac{1}{T})\sum_{t=1}^{T} d_t$, and $\sigma_{\bar{d},0}^2$ is the asymptotic long-run variance of the loss differential. For nested models, assuming model 1 is a restricted model whereas model 2 is the more unrestricted model, Clark and West (2007) suggest a test statistic *CW*:

$$CW = \frac{\sqrt{T}\bar{f}}{\sigma_{\bar{f},0}} \xrightarrow{d} N(0,1),$$

where $f_t = (S_t - \hat{p}_{1,t})^2 - [(S_t - \hat{p}_{2,t})^2 - \big(\hat{p}_{1,t} - \hat{p}_{2,t}\big)^2]$, $\bar{f} = (\frac{1}{T})\sum_{t=1}^{T} f_t$, and $\sigma_{\bar{f},0}^2$ is the sample variance of $f_t - \bar{f}$. Last, to formally compare AUC of two models, DeLong, DeLong, and Clarke-Pearson (1998) propose a test statistic W_{AUC} :

[16] The ROC curve is a graphical tool which reveals the predictive abilities of an EWS (Early-Warning System). More exactly, it represents the trade-off between *sensitivity (* S_e) and *1-specifcity (1-* S_p) for every possible cut-off. The ROC curve is thus obtained by representing all the couples $\{Se\ (c); 1 - Sp(c)\}$ corresponding to each value of the cut-off c ranging from 0 to 1. See Candelon, Dumitrescu, and Hurlin (2012) for a complete introduction.

[17] Candelon, Dumitrescu, and Hurlin (2012) provides a Matlab toolbox to calculate various measures to evaluate an Early Warning System.

$$W_{AUC} = \frac{(AUC_1 - AUC_2)^2}{\mathbb{V}(AUC_1 - AUC_2)} \xrightarrow{d} \chi^2(1),$$

where $\mathbb{V}$ is the variance-covariance matrix of the vector $(AUC_1\ AUC_2)'$. The null hypothesis of W_{AUC} corresponds to the equality of areas under the ROC curves, which is H_0: AUC_1=AUC_2. As mentioned previously, AUC is more appropriate for evaluating a EWS model. I uses W_{AUC} as the main comparison test statistic and provide *DM* and *CW* test statistics as supplements.

4.4. In-sample performance

The full sample period is from January 1967 to December 2013. First, each variable is examined individually as a predictor in the static univariate forecasting model. Then, factors estimated by principal components analysis from all variables are selected as predictors in the static multivariate forecasting model[18]. Given that dynamic structures can improve predictability in previous literature, I investigate the predictability both in static and dynamic multivariate forecasting models.

4.4.1. Univariate models

Table 5 summarizes the results of each individual predictor, where the left panel presents the estimates from the static multinomial logit model and the right panel presents the estimates from the static binary logit model. Overall, with the advantage of multinomial logit model, essential information contained in predictors is more efficiently revealed and used. Among macroeconomic variables, all the variables exhibit significant predictive power. Consistent with previous literature, dividend-price ratio (DP), term spread (TMS), inflation (INFL), and industrial production (IP) are relatively stronger predictors. However, some variables have very different impacts on the probability of good bear markets and of bad bear markets. For example, the increase of term spread (TMS) significantly decreases the probability of bad bear markets but has no significant effects on the probability of good bear markets. This result is intuitively consistent with how the good bear and bad bear markets interact with the real economy. According to Estrella and Mishkin (1998), term spread is a leading indicator of future recessions that it turns negative in advance of recessions. Recall in Section 3.1, I show that most bad bear markets are accompanied or followed by NBER recessions, while only 20% of good bear markets lead NBER recessions. This can explain why term spread has significant and larger effects on the probability of bad bear markets. Dividend-price ratio (DP), dividend yield (DY), and earning-price ratio (EP) are significant in predicting good bear markets but not in predicting bad bear markets, which is consistent with the findings of Campbell and Vuolteenaho (2004) and Campbell, Gigli and Polk (2013) that innovation in earning-price ratio is highly correlated to discount rate news (the main driving force of good bear markets) but only weakly correlated with cash flow news (the main driving force of bad bear markets). Cochrane (2011) also finds that the

[18] To maintain a parsimonious model, I consider at most up to first 5 principal components. The first 5 factors contribute 70% of the whole data set variation, and individual factors after 5th contribute less than 5% each.

variation of dividend-price ratio exclusively corresponds to variation in discount rates. On the other hand, default yield spread (DFY) is only significant in predicting bad bear markets which is consistent with Campbell, Gigli and Polk (2013) that default spread plays a significant role in the determination of cash flow news. Ng (2014) also finds default yield spread is the most robust predictor of recessions which usually happen right after the bad bear markets.

Turning to real economic activity variables, all are significant in forecasting bad bear markets but are not or have opposite signals in forecasting good bear markets. For example, payroll employment (EMP) positively predicts the probability of good bear markets but negatively predict the probability of bad bear markets. Again, this is intuitive with how the two types of bear markets interact with the economy. As Berge (2015) and Owyang et al (2013) find that employment growth rates substantially improve very short-horizon forecasts of business cycle phases, the way bad bear markets precede NBER recessions can justify this result.

Finally, all technical indicators have similar magnitudes, signs and significances in predicting the two types of bear markets. Strikingly, the predictive power of a technical indicator is much stronger than a macroeconomic variable or a real economic indicator. The monthly pseudo-R^2 for individual technical indicator is 17.7% on average, whereas individual macroeconomic or real economic variable is 3.3% on average. This result is even stronger than the results from Neely et al. (2014) who focus on predicting stock return value[19]. With the identification methodology of bear markets used in this paper (the stock price is on a downward trend for at lease 6 months.), this in-sample test implies technical indicators, based on the past price or trading volume trends, are very significant in predicting future price trend of the stock market.

Comparing the left panel to the right panel, all the estimates in binary logit model are closer to the estimates for bad bear markets in the multinomial logit model, both in magnitudes and signs. Some variables that are only but strongly significant in predicting good bear markets in the multinomial logit model, such as earning-price ratio (EP) and equity risk premium volatility (RVOL), do not reveal significant ability in forecasting bear markets in binary logit model. This implies with only binary class model to forecast stock bear markets, investors could miss relevant information crucial in forecasting stock bear markets.

[Insert Table 5]

4.4.2. Multivariate models

As all the candidate predictors are relevant in predicting either good or bad bear markets and are highly correlated to each other, using principal components analysis (PCA) with all variables to extract common factors can efficiently incorporate information from multivariate variables and reduce data dimensionality. Moreover, due to the construction of the common components,

[19] In Neely et al. (2014), the monthly R^2 for individual technical indicator is 0.62% on average, whereas individual macroeconomic variable is 0.34% on average.

factors estimated by PCA automatically capture the lag dynamics of the underlying factors in forecasting (Stock and Watson, 2002a; 2002b). The common factors estimated from large datasets are also less affected by the structure changes or data revisions of the original variables, which might erode model performance over time (Chen et al., 2011; Fossati, 2015).

Using Schwarz Information Criterion (BIC) as factor selection criteria as suggested by Bai and Ng (2002), both the static multinomial logit model and the static binary logit model choose the first 3 principal components as predictors. Table 6, Model (1) and Model (3) represent the static multinomial logit model and the static binary logit model, respectively. Under static multinomial logit model, Model (1), the first principal component (PC1) is significant in forecasting both good and bad bear markets with similar magnitude and same sign. The second (PC2) and the third (PC3) principal components only contain significant information in forecasting good bear markets. For static binary logit model, Model (3), only the first (PC1) and the third (PC3) principal components are significant in predicting bear markets with similar magnitudes and signs of those for bad bear markets in multinomial logit model. To understand the economic contents of these extracted principal components, Figure 3 presents the dynamic of first 3 principal components, where gray shaded bars indicate bad bear market states and pink shaded bars indicate good ones. Figure 4 gives the corresponding loadings on each individual variable of the principal component. The first principal component has uniformly large loading on all technical indicators which means the first principal component mostly represents the simple average of all technical indicators. Indeed, in Figure 3, the first principal component moves closely with good bear and bad bear markets, which is consistent with previous in-sample forecasting results that all technical indicators strongly predict both two types of bear markets with similar magnitudes and same sign. The second principal component has large loading on stock value ratios (DP, EY, EP, and BM), which means this factor mostly represents fluctuations in stock value ratios. The third principal component has large loadings on real economic activity variables (IP, EMP, MTS and PIX) and on the stock market volatility (RVOL), which means this factor would increase dramatically during financial crisis.

As regard to the dynamic models, Table 6, Model (2) and Model (4) represent the dynamic multinomial logit model and the dynamic binary logit model, respectively. For the dynamic multinomial logit model, Model (2), the significances of the first three principal components stay the same. The 12-month lagged stock market state indicators, indicated by $I_{1\ t-12}$ and $I_{2\ t-12}$, are significant in predicting future stock market states. For the dynamic binary logit model, Model (4), the 6-month lagged stock market state, indicated by S_{t-6}, also has significant predicting power, in line with previous literature. In the next section, I use the first 3 principal components as predictors in the out-of-sample evaluation. That is:

$$X_{t-1} = (PC1_{t-1}, PC2_{t-1}, PC3_{t-1}).$$

I investigate both the static and the dynamic specifications for the multinomial logit model and the binary logit model.

[Insert Table 6]

4.5. Out-of-sample test

In this section, I evaluate the out-of-sample (with the ex-post revised data)[20] forecasting performance of the multinomial logit model. To perform out-of-sample evaluation, the forecasting model uses up to current data to extract principal components to predict the next period's stock market state. The model is estimated recursively with an expanding window at every period. However, because of the rules of dating algorithm and the specific classification method used in this study, the final[21] identification of stock market states cannot be realized in real-time. To be clarified, in this out-of-sample exercise, I only use up to current period information to decide the classification of stock market states and estimate model parameters, while the classification might be modified latter with the expansion of forecasting window.

To compare the multinomial logit model (with three states) with binary logit model, the probability of a bear market (either good bear or bad bear market) from the multinomial logit model is:

$$\rho_t = P_{t-1}(S_t = 1) + P_{t-1}(S_t = 2)$$

The predicted probability, $\widehat{\rho_t}$, is evaluated and compared with the probability generated from the binary logit model. I first examine each model's performance in predicting bear markets (without considering bear markets types) through the value of evaluation statistics and then investigate hitting rates[22] across three stock market states to better understand each model's predictability.

4.5.1. Performances of predicting bear markets

Using sample from Janurary1967 to December 1976 as the initial estimation period, the out-of-sample forecasting evaluation period spans from January 1977 to December 2013. Table 7, Panel (a), reports the statistical results. Under the multinomial logit models, comparing the static and the dynamic model specifications, Model (1) and Model (2), all the evaluation criteria, QPS, LPS and AUC values, are better in the dynamic model specification. However, under the binary logit models, the static model, Model (3), performs better than the dynamic model, Model (4), in all evaluation criteria. This is different from previous literature which finds adding dynamic structures can improve out-of-sample predictability (Nyberg 2013, and Candelon et al. 2014). One reason might be that in my dynamic binary logit model specification, the lagged stock market state is 6-month lagged which is not very close to current state. More importantly, unlike Nyberg (2013) who assumes a 6-month information lag in the value of stock market states in his

[20] The out-of-sample tests using real-time data are investigated in section 6.
[21] The final classification means using full sample information to decide good bear and bad bear market states.
[22] Hitting rate is the fraction of true states that is forecasted correctly.

out-of-sample test, in my out-of-sample evaluation, it doesn't have any assumption of the information lag in the stock market states, which might impact the result as well[23].

Table 7, Panel (b), presents the p-value of various model comparison tests. For the comparison in forecasting accuracy, the *CW* (*DM*) test for nested (non-nested) models is presented. For the comparison in classification ability, the W_{AUC} test is presented. For multinomial logit models, the dynamic specification is superior to the static specification both in the forecasting accuracy and in the classification ability. For binary logit models, the static specification is significantly better than the dynamic specification in the forecasting accuracy but not in the classification ability. To compare the multinomial and binary logit models, better performed model specifications are chosen for each, which are the dynamic multinomial logit model, Model (2), and the static binary logit model, Model (3), as the two competing models[24]. For the forecasting accuracy, there is no significant difference between Model (2) and Model (3). For the classification ability, Model (2) is significantly better than Model (3). Figure 5 presents the predicted probability of bear markets under Model (2) and Model (3). In the next section, I further look into the resulting hitting rates of the two competing models across three stock market states.

[Insert Table 7]

[Insert Figure 5]

4.5.2. Optimal cut-off and predicting performance across three stock market states

To decide the stock market state at time *t*, the probability forecasted from the model needs to be transformed into a discrete categorical variable. Formally, the forecasted state at time *t*, $\widehat{S_t}$, is computed as follows:

$$\widehat{\rho_t} = P_{t-1}\widehat{(S_t} = 1) + P_{t-1}\widehat{(S_t} = 2)$$

$$\widehat{S_t} = \begin{cases} 0 \; if \; \widehat{\rho_t} < \; C \\ 1 \; if \; \widehat{\rho_t} \geq \; C \; and \; P_{t-1}\widehat{(S_t} = 1) \geq \; P_{t-1}\widehat{(S_t} = 2) \\ 2 \; if \; \widehat{\rho_t} \geq \; C \; and \; P_{t-1}\widehat{(S_t} = 1) < \; P_{t-1}\widehat{(S_t} = 2) \end{cases},$$

where $C\epsilon[0,1]$ represents the cut-off. The choice of cut-off value is important since it determines type I and type II errors, errors associated with a misidentified bear market and a false alarm. For example, if the cut-off value is very low, bear markets will be more accurately detected (lower

[23] In an unreported evaluation, under the same out-of-sample forecasting period but with the in-sample stock market state information, the dynamic model has better predictability than the static model both in the multinomial logit model and the binary logit model.

[24] The rest of the evaluations in this paper, including the hitting rate (section 4.5.2) and trading performances evaluations (section 5), are based on the dynamic multinomial logit model and the static binary logit model.

type I error), but at the same time, the number of false alarm will increase (higher type II errors). Surprisingly, many previous literatures set an ad-hoc cut-off value[25] which makes their model evaluations potentially questionable. Being aware of this and not arbitrary deciding the weighting between type I and type II errors, I use *Youden Index*[26] which optimally considers both type I and type II errors equivalently to decide cut-off values (Candelon, Dumitrescu, and Hurlin, 2012) for the dynamic multinomial logit model and the static binary logit model, respectively.

Table 8 reports the accuracy of the dynamic multinomial logit model, Model (2), and the static binary logit model, Model (3), in terms of hitting rate across three stock market states. The hitting rates of bull markets and of bear markets without considering the types in the dynamic multinomial logit model (the static binary logit model) are 83.42 % (89.14 %) and 72.34% (67.02%) respectively. This means comparing to the static binary logit model, the dynamic multinomial model has better ability in predicting bear markets but is more likely to have a false alarm. However, in economic theory, people are risk averse. The cost of failing to identify a bear market is likely higher than that of failing to identify a bull market. In this sense, the multinomial logit model can be more useful for investors. Moreover, given that the true state is a good bear (bad bear) market, the probability to identify it as a bear market without considering types, is 66.66% (75.41%) for the dynamic multinomial logit model, and is 60.6% (70.49%) for the static binary logit model. Hence, the superiority of the dynamic multinomial logit model comes from its ability in detecting both types of bear markets.

[Insert Table 8]

5. Economic values of the multinomial logit model in predicting bear markets

In the empirical stock return predictability literature[27], another way to evaluate the economic significance of a model's predictability is to examine the profitability an investor can obtain with asset allocation decisions based on the forecasting model. That is, given a trading strategy specified for stock market states, whether the investor can time the market correctly and switch investment portfolios to gain higher returns is of particular interest. Most previous stock market states predictability literature use binary state models to forecast stock market states. Chen (2009) proposes a trading strategy based on a two-state Markov-Switching model with a cut-off value of 30% that can generate higher monthly return than the benchmark Buy-and-Hold

[25] Chen (2009) sets the cut-off value as 30%. Nyberg (2013) assumes a 50% rule and also a sample average of bear market months which turns out to be approximately 30%. Candelon et al. (2014) assume three cut-off values: 40%, 50%, and 70%.

[26] According to the accuracy measure, the optimal cut-off satisfies: $C^* = \arg \max_{C\in[0,1]} J(c), \quad J = Se(c) + Sp(c) - 1$, where $J(c)$ is the Youden Index, and sensitivity $Se(c)$, also known as hit rate, is the proposition of bear market states correctly identified by the forecasting model, whereas $Sp(c)$is the proposition of bull market states correctly identified by the model.

[27] Barberis (2000), Stambaugh (1999), and a vast of stock return predictability studies discuss the economic implication of return predictability for an investor.

strategy. Nyberg (2013) finds the trading strategy based on a dynamic binary probit model with a cut-off value of sample average of bear market months (close to 30%) can generate higher Sharpe ratio than the Buy-and-Hold strategy. Candelon et al. (2014) also find a trading strategy based on a dynamic probit model with a cut-off value of 40% can generate superior monthly returns. In these studies, the proposed trading strategies entail switching portfolios between the market portfolio and the short term bond market across the forecasted stock market states (hold the market portfolio if the model forecast is a bull market state; otherwise switch to short term bond market). Following this convention, I compare the trading performances of strategies based on the dynamic multinomial logit model and the static binary logit model to the conventional buy-and-hold benchmark strategy.

As mentioned earlier, the economic value of this dynamic multinomial logit model can be analyzed in three layers: first, as the trading strategy proposed in the previous literature, by forecasting next period stock market state, an investor will hold the market portfolio if the bear market predictability generated from the forecasting model is lower than the threshold C or otherwise switch his portfolio to the 3-month Treasury bill market. In this strategy, called type 1 trading strategy, the performances between the dynamic multinomial logit model and the static binary logit model are of interest. In this type of trading strategy, the investor only needs to know whether the next period will be a bull or a bear market. Information about bear market types is not used in this trading strategy.

Second, the value strategy (value premium) is an alternative investment opportunity. It would be interesting to know, instead of switching portfolios to 3-month Treasury bill market, whether implementing the value strategy (HML) when the model prediction is the bear market state (either good bear or bad bear) can improve trading performances. In this strategy, called type 2 trading strategy, the investor does not use information about bear market types in his trading decisions either.

Finally, as shown in section 3.2 that various value premiums are much higher at good bear market states than at any other states, information about bear market types should be crucial to exploit value premiums. Therefore, to use information about bear market types, called type 3 trading strategy, the investor holds the market portfolio if the forecast is a bull market state, switches to the 3-month Treasury bill market if the forecast is a bad bear market state, but performs HML (value strategy conditional on size and profitability), HML_S (value strategy conditional on profitability within small firms), HML_RMW (value and profitability combined strategy conditional on size), or HML_RMW_S (value and profitability combined strategy within small firms), respectively, if the forecast is a good bear market state. Hence, as with the type 3 trading strategies, there are 4 potential trading strategies that can gain higher return. I investigate each of them. The optimal cut-off value for each forecasting model is decided by the *Youden Index*. I also assume transaction costs to make the evaluation closer to reality. The transaction costs are proportional to the wealth, which are 25 basis points per dollar of portfolio value traded in each transaction when going long in stock markets but only 10 basis points for the Treasury bill market. (Pearan and Timmermann, 1995; Balduzzi and Lynch, 1999; Han et al., 2011). To short stocks, the cost is 100 basis points per dollar, which is often costly.

Table 9 (A) reports the trading performance of an investor who invests $1 at the beginning of the out-of-sample period. Column (1) reports the performance under conventional benchmark Buy-and-Hold strategy; Columns (2) and (3) report the performances of the type 1 trading strategies based on the dynamic multinomial logit model and the static binary logit model, respectively. Columns (4) and (5) show the performances of type 2 trading strategies based on the dynamic multinomial logit model and the static binary logit model, respectively. While in type 1 trading strategy, the investor switches portfolios to the 3-month Treasury bill market when the model predicts bear market states, in type 2 trading strategies, the investor implements HML strategy instead. To illustrate the importance of timing to implement value strategy, I include the performance of another benchmark strategy (A_HML) in Column (6), which the investor always implements HML strategy at each month without timing. Columns (7) to (10) show the performances of type 3 trading strategies. In this type of strategy, the investor uses the dynamic multinomial logit model to forecast stock market states and invests in the market portfolio when the model predicts bull market states; switches to 3-month Treasury bill markets when the model predicts bad bear market states; but performs value strategy of HML, HML_S, HML_RMW or HML_RMW_S, respectively, when the market predicts good bear market states.

Panel (A) reports the performances over the full out-of-sample period; Panel (B) reports the performances across three stock market states. Rows (1) to (4) report monthly average excess returns, monthly standard deviations, monthly Sharpe ratios, and final wealth at the end of the trading period. Row (5) gives monthly average turnover, which is the percentage of wealth traded each month. Row (6) gives the maximum monthly drawdown, which is the maximum of percentage drops in the cumulative returns from the peak during the trading period. To evaluate strategies in terms of the investor's utility that incorporates risk aversion, with mean-variance preferences, Rows (7) and (8) report the certainty equivalent returns (CER). The CER represents the risk-free rate of return that an investor is willing to accept instead of adopting a given trading strategy, calculated as:

$$CER = E[r_p] - A\frac{1}{2}Var[r_p],$$

where A is the risk-aversion parameter, r_p is the monthly portfolio return under the associated trading strategy. I choose risk-aversion value as 2 and 5 to represent different degrees of risk aversion. Rows (9) to (12) report the Jensen's alpha under Fama-French 3 factor (FF3) model and Fama-French 5 factor (FF5)[28] model with the corresponding t statistics values to evaluate risk-adjusted returns (abnormal returns).

[28]Fama and French (2015) shows the five-factor model performs better in capturing the size, value, profitability, and investment patterns in average stock returns than the three-factor model (FF, 1993).

Overall, all the trading strategies based on either the dynamic multinomial logit model or the static binary logit model perform better than the benchmark Buy-and-Hold strategy. Under type 1 trading strategies, the monthly Sharpe ratio increases from 0.16 (Buy-and-Hold) to 0.18 (dynamic multinomial logit model) and 0.2 (static binary logit model). In this type of trading strategy, under the full out-of-sample period, based on the Sharpe ratios, the multinomial logit model slightly underperforms the binary logit model. However, across three stock market states, Panel (B), trading based on the dynamic multinomial logit model performs poorer in bull market states but better in good bear and bad bear market states. This corresponds to the multinomial logit model's strength and weakness in predicting bear markets and bull markets relative to the binary logit model. The upper panel of Figure 6 shows the growths of one dollar invested at the beginning of January 1977 that follows the benchmark Buy-and-Hold (B&H) strategy, and the two type 1 trading strategies based on the dynamic multinomial logit model (M) and the static binary logit model (B), respectively. Shaded bar indicates identified stock bear markets, where pink represents good bear states and gray represents bad bear states. The lower panel of Figure 6 shows the drawdowns of these trading strategies. Obviously, type 1 trading strategies effectively reduce the drawdown of the Buy-and-Hold strategy, especially during bear market states, though the improvement in final wealth is mild.

[Insert Figure 6]

Among type 2 trading strategies, columns (4) and (5), in which the investor performs HML strategy instead of switching to the 3-month Treasury bill market when the model forecasts bear market states (either good bear or bad bear). Trading strategies based on both the dynamic multinomial logit model and the static binary logit model perform better than the Buy-and-Hold strategy. Sharpe ratios increase to 0.22 in both forecasting models which are also higher than the two type 1 trading strategies. However, the two type 2 trading strategies all suffer from larger drawdowns than the Buy-and-Hold strategy and type 1 trading strategies. Compared to where investors implement the HML strategy each month without timing, Column (6) A_HML, the two type 2 trading strategies underperform A_HML strategy (Sharpe ratio 0.22 vs. 0.24), though the A_HML strategy has larger drawdowns. The upper panel of Figure 7 shows the growths of one dollar invested at the beginning of January 1977 that follows the benchmark Buy-and-Hold (B&H), A_HML, and the two type 2 trading strategies based on the dynamic multinomial logit model (M_HML) and the static binary logit model (B_HML), respectively. The lower panel shows the drawdowns of these trading strategies. Obviously, all the trading strategies suffer from large drawdowns especially during the 2009 stock bear market. Benchmark strategy A_HML has large drawdowns more frequently than other strategies.

[Insert Figure 7]

Finally, using information about bear market types, type 3 trading strategy, columns (7) to (10), the trading performances are improved substantially. The monthly Sharpe ratio increases from 0.16 (Buy-and-Hold) to around 0.3, also higher than another benchmark, A_HML (0.24).

More strikingly, the maximum drawdown decreases dramatically from 51% (Buy-and-Hold) or 73% (A_HML) to the lowest 23% (HML_S). The upper panel of Figure 8 shows the performances of two benchmarks Buy-and-Hold (B&H), A_HML, and the four type 3 trading strategies involving HML, HML_S, HML_RMW, and HML_RMW_S, respectively. Overall, improvements in type 3 trading strategies are remarkable. Strategies that combined value and profitability characteristics, column (9) and (10), (HML_RMW and HML_RMW_S) do increase trading profits further than the strategy involving value strategy alone[29] (HML, column (7)) which is consistent with Novy-Marx (2013, 2014) that cheap and profitable firms tend to outperform firms that are just cheap. However, trading strategies involving HML_RMW and HML_RMW_S are also exposed to higher volatility risks, which make Sharpe ratios of the two involved strategies close to that of the strategy involving HML, column (7). Also, the higher return of strategy involving HML_S, column (8), over strategy involving HML, column (7), is consistent with previous literature[30] in that the value premium is concentrated in small firms. The lower panel of Figure 8 shows that type 3 trading strategies not only increase trading profits but also decrease drawdown dramatically, which is significantly different from the investor who always implements the value strategy at each month without timing (A_HML). This result further proves knowing bear market types is valuable for a passive Buy-and-Hold investor but also crucial for an active investor who want to exploit value premiums.

[Insert Figure8]

With respect to CER, an alternative measurement addressing the effects of risk aversion in utility, type 3 trading strategies still show their superiorities over Buy-and-Hold strategy, type 1, and type 2 trading strategies, both in less (CER_2) or more risk-averse (CER_5) cases. Examining the abnormal returns (Jensen's alpha), under the conventional FF3 model, both type 1 and type 3 trading strategies exhibit significant abnormal returns. However, under the FF5 model, only type 3 trading strategies show significantly positive abnormal monthly returns from 0.56% to 0.71%, which implies trading strategies based on the dynamic multinomial logit model can generate excess returns beyond the Fama-French 5 risk factors. Assessing performances across stock market states, Panel (B), though the type 3 trading strategies slightly underperform Buy-and Hold strategy and type 2 strategies during bull market states, they perform much better than the Buy-and-Hold strategy, type 1, and type 2 trading strategies during bear market states.

Clearly, type 3 trading strategies, exploiting various value premiums in the prediction of good bear markets, dramatically improve the trading performances in both increasing returns and decreasing unfavorable drawdowns. This out-of-sample trading performance evaluation therefore shows that the multinomial logit model which provides the valuable information about bear market types, is helpful for investors and much more profitable than the binary logit model.

---

[29] The HML strategy here is not a pure value strategy but conditional on firms' profitability and size. However, this strategy doesn't exploit the spread between high profitable firms and low profitable firms.

[30] Fama and French (1993, 2012) and Israel and Moskowitz (2013) show the value premium is largely concentrated among small stocks.

[Insert Table 9]

6. Real-time out-of-sample robustness test

In the forecasting models discussed above, three principal component factors are estimated from 32 variables including 14 monthly macroeconomic variables, 14 monthly technical indicators, and 4 monthly real economic activity indicators (ex-posted revised data). Since the ex-posted revised real economic activity indicators are not available in real-time, the out-of-sample tests could be biased. To examine the robustness of the results obtained above, I use real-time vintage data for real activity indicators (i.e., data that was available at the time the prediction was made) with the other real-time available 28 variables to perform a real-time out-of-sample test. Among 4 real economic activity indicators, the real-time vintage data for real manufacturing and trade sales (MTS) is not available in my out-of-sample testing period and hence is discarded. For the reminding 3 variables[31], to deal with the issue of one month publishing lag (i.e., month *t* data is released at month *t+1*), I use two univariate benchmark methods[32] commonly applied in the real-time nowcasting literature to predict the value for month *t* in order to construct a balanced dataset with the other 28 variables. The first model is an autoregressive (AR) model with lag *p* selected recursively using Bayesian information criteria with maximum $p=6$. The second one is a constant growth model (RW) of no predictability (random walk with drift in levels), which simply uses the average of past growths as the predicted value. The real-time out-of-sample period is as the same as previous out-sample tests, from January 1977 to December 2013. The results from the AR model and from the RW model are qualitatively the same[33]. First, unlike the out-of-sample tests using ex-post revised data, there is no statistically significant difference between the static binary logit model and the dynamic multinomial logit model either in classification ability or in forecasting accuracy[34]. However, comparing to the static binary logit model, the dynamic multinomial logit model has better ability in predicting bear markets but is more likely to have a false alarm, which is consistent with previous tests using ex-post revised data. Second, evaluating the profitability from investments based on the real-time forecasting models further confirms that the multinomial logit provides crucial information for investors to exploit the value premium. Overall, while the results based on real-time data don’t support the superiority of multinomial logit model in stock market classification ability, they provide robust evidence of better accuracy in predicting bear markets and substantial benefits in forming higher profits trading strategies.

---

[31] The real-time vintage data for non-farm payroll employment (EMP) and industrial production (IP) are from the Federal Reserve Bank of Philadelphia real-time database; real personal income (PIX) is from the Federal Reserve Bank of St Louis ALFRED database.
[32] Using more sophisticated nowcasting methods to handle the lag releasing issue in real-time data could possibly improve the model’s performance but is beyond of the scope of this study.
[33] For brevity, the results are not tabulated in the paper but are available upon request.
[34] The static binary logit model is better in classification ability than the dynamic binary logit model and the static multinomial logit model.

7. Conclusion

The present value model shows that the stock price moves either because of movement in expected future cash flows or because of movement in discount rate. Using the U.S. monthly S&P 500 index as the price level of the stock market and its 12-month moving average of earnings as the proxy of cash flows, stock bear markets associated with contraction phases of earnings are classified as bad bear markets or are otherwise classified as good bear markets. During good bear market states, most real economic activities do not significantly deteriorate or even still positively expand, and only about 20% of good bear markets are followed by NBER declared recessions within 6 months. In contrast, during bad bear market states, most real economic indicators show economic downturns, and NBER recessions usually happen right away.

Based on the out-of-sample evaluation (either using ex-post revised data or real-time data), for a multinomial logit model with three states (bull, good bear and bad bear stock markets), adding the dynamic structure into the model improves its predictability. For a binary logit model, the dynamic structure deteriorates the model's predictability instead. Comparing the multinomial logit model and the binary logit model, the dynamic multinomial logit model has better ability in predicting bear markets than the conventional binary logit model (either in static or in dynamic).

For investors, without incorporating information about bear market types into their trading decisions, a trading strategy based on the dynamic multinomial logit model does not perform better than that based on the static binary logit model under the full sample period, though it is preferred during bear market states. By using information about bear market types provided by the multinomial logit model (either in static or in dynamic), investors can gain much higher returns and decrease unfavorable large drawdowns by properly exploiting value premiums.

Reference

Barberis, N. C. (2000). Investing for the Long Run when Returns are Predictable. *Journal of Finance,* 55, 225-264.

Barro,R. and Ursua, J. F. (2009). Stock-market crashes and depressions. NBER *working paper* No. 14760.

Bellego, C. and Ferrara, L. (2012). Macro-financial linkages and business cycles: a factor-augmented Probit approach. Economic Modelling, 29, 1793-1797.

Berge, T. J. and Jorda, O. (2011). Evaluating the Classification of Economic Activity into Recessions and Expansions. *American Economic Journal: Macroeconomics*, 3 (2), 246-277.

Berge, T. J. (2015), Predicting recessions with leading indicators: model averaging and selection over business cycle. *Journal of Forecasting*, 34 (6), 455-471.

Bry, G. and Boschan, C. (1971). *Cyclical Analysis of Time Series: selected procedures and Computer Programs*, Columbia University Press, National Bureau of Economic Research, New York.

Campbell, J. Y., and Shiller, R. J. (1988). Stock Prices, Earnings and Expected Dividends. *Journal of Finance*, 43, 661–676.

Campbell, J. Y., (1991). A Variance Decomposition for Stock Returns. *Economic Journal*, 101, 157–179.

Campbell, J. Y. and Vuolteenaho, T. (2004). Bad Beta , Good Beta. *The American Economic Review*, 94 (5), 1249–1275.

Campbell, J. Y., Giglio, S., and Polk, C. (2013). Hard Times. *Review of Asset Pricing Studies*, 3 (1), 95–132.

Candelon, B., Piplack, J., and Straetmans, S. (2008). On measuring synchronization of bulls and bears: The case of East Asia. *Journal of Banking & Finance*, 32 (6), 1022–1035.

Candelon, B., Ahmed, J., and Straetmans, S. (2014). Predicting and Capitalizing on Stock Market Bears in the U.S. *Working paper*, IPAG Business School.

Candelon, B., Dumitrescu, E.-I., and Hurlin, C. (2012). How to Evaluate an Early-Warning System: Toward a Unified Statistical Framework for Assessing Financial Crises Forecasting Methods. *IMF Economic Review*, 60 (1), 75-113.

Chauvet, M. and Piger, J. (2008). A cpmparison of the real-time performance of business cycle dating methods. *Journal of Business and Economic Statistics*, 26 (1), 42-29.

Chauvet, M. and Potter, S. (2000). Coincident and leading indicators of the stock market. *Journal of Empirical Finance*, 7 (1), 87-111.

Chen, S.-S. (2009). Predicting the bear stock market: Macroeconomic variables as leading indicators. *Journal of Banking and Finance*, 33 (2), 211–223.

Chen, L., Da, Z., and Priestley, R. (2012). Dividend Smoothing and Predictability. *Management Science*, 58 (10), 1834 – 1853.

Chen, Z., Iqbal, A., and Lai, H. (2011). Forecasting the probability of recessions: a probit and dynamic factor modelling approach. *Canadian Journal of Economics,* 44 (2), 651-672.

Christiansen, C., Eriksen, J. N., and Moller, S. V. (2014). Forecasting U.S. recessions: the role of sentiment. *Journal of Banking and Finance*, 49, 459-468.

Claessens, S., Kose, M. A., and Terrones, M. E. (2012). How do business and financial cycles interact? *Journal of International Economics*, 87 (1), 178–190.

Cochrane, J. H. (2008). The Dog that Did Not Bark: A Defense of Return Predictability. *Review of Financial Studies*, 21 (4), 1533-1575.

Cochrane, J. H. (2011). Discount Rates. *Journal of Finance*, 66, 1047-1108.

Cowles, A. (2005). Trend following : How great traders make millions in up or down markets. Prentice-Hall, New York.

Dechow, P. M., Kothari, S. P., and Watts, R. L. (1998). The relation between earnings and cash flows. *Journal of Accounting and Economics*, 25, 133-168.

Diebold F. X. and Rudebusch, G. D. (1989). Scoring the leading indicators. *The Journal of Business*, 62 (3), 369-391.

Estrella, A. and Mishkin, F. S. (1998). Predicting U.S Recessions: Financial Variables as Leading Indicators. *The Review of Economics and Statistics*, 80 (1), 45-61.

Faber, M. (2007). A quantitative approach to technical asset allocation. *Journal of Wealth Management*, 9, 69-79.

Fama, E. F. (1981). Stock returns, Real activity, Inflation and Money. *American Economic Review*, 71, 545-65.

Fama, E. F. and French, K. R. (1988). Dividend Yields and Expected Stock Returns. *Journal of Financial Economics*, 22, 3-25.

Fama, E. F. and French, K. R. (1989). Business Conditions and Expected Returns on Stocks and Bonds. *Journal of Financial Economics*, 25, 23-49.

Fama, E. F. and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33, 3-56.

Fama, E. F. and French, K. R. (2012). Size, value, and momentum in international stock returns. *Journal of Financial Economics*, 105, 457-472.

Fama, E. F. and French, K. R. (2015). The five-factor asset pricing model. *Journal of Financial Economics*, 116, 1-22.

Fossati, S. (2014). Dating U.S. Business Cycles with Macro Factors. *Studies in Nonlinear Dynamics & Econometrics, forthcoming.*

Fossati, S. (2015). Forecasting U.S. Recessions with Macro Factors. *Applied Economics, 47 (53), 5726-5738.*

Goyal, A. and Welch, I. (2003). Predicting the Equity Premium with Dividend Ratios. *Management Science*, 49 (5), 639-654.

Goyal, A. and Welch, I. (2008). A Comprehensive Look at the Empirical Performance of Equity Premium Prediction. *Review of Financial Studies*, 21 (4), 1455-1508.

Guidolin, M. and Timmermann, A. (2007). Asset allocation under multivariate regime switching. *Journal of Economic Dynamics and Control*, 31 (11), 3503-3544.

Hall, R. E. (2002). Dating Business Cycles: A Perspective. Manuscript, Stanford University.

Harding, D. and Pagan, A. R. (2003). A Comparison of Two Business Cycle Dating Methods. *Journal of Economic Dynamics and Control*, 27 (9), 681-1690.

Harvey, C. R. (1989). Forecasts of Economic Growth from the Bond and Stock Markets. *Financial Analysts Journal*, 45 (5), 38-45.

Israel, R. and Moskowitz, T. (2013). The role of shorting, firm size, and time on market anomalies. *Journal of Financial Economics*, 108, 275-301.

Kauppi, H., Saikkonen, P. (2008). Predicting U.S. recessions with dynamic binary response models. *The Review of Economics and Statistics*, 90 (4), 777-791.

Kim, M. and Kross, W. (2005). The ability of earnings to predict future operation cash flows has been increasing – not decreasing. *Journal of Accounting Research*, 43, 753-780.

Lettau, M. and Nieuwerburgh, S. V. (2008). Reconciling the Return Predictability Evidence. *The Review of Financial Studies*, 21 (4), 1607-1652.

Lewellen, J. (2004). Predicting returns with financial ratios. *Journal of Financial Economics*, 74, 209-235.

Ludvigson, S. C., and Ng, S. (2009). A factor analysis of bond risk premia. *forthcoming* in Handbook of Applied Econometrics.

Miller, M. H. and Modigliani, F. (1961). Dividend Policy, Growth, and the Valuation of Shares. *The Journal of Business*, 34 (4), 411-433.

Neely, C. J., Rapach, D. E., Tu, J., and Zhou, G. (2014). Forecasting the equity risk premium: The of technical indicators. *Management Science*, 60 (7), 1772-1791.

Ng, S. (2014). Boosting Recessions. *Canadian Journal of Economics*, 47 (1), 1-34.

Novy-Marx, R. (2013). The other side of value: the gross profitability premium. *Journal of financial economics*, 108, 1-28.

Novy-Marx, R. (2014). Quality Investing. *Working paper*.

Nyberg, H. (2013). Predicting bear and bull stock markets with dynamic binary time series models. *Journal of Banking and Finance*, 37 (9), 3351–3363.

Owyang M. T., Piger, J., and Wall, H. J. (2015). Forecasting national recessions using state level data. *Journal of Money, Credit and Banking*, 47 (5), 847–866.

Pagan, A. R. and Sossounov, K. A. (2003). A simple framework for analyzing bull and bear markets. *Journal of Applied Econometrics*, 18 (1), 23–46.

Perez-Quiros, G. and Timmermann, A. (2000). Firm Size and Cyclical Variations in Stock Returns. *The Journal of Finance*, 55 (3), 1229–1262.

Samuelson, P. (1996). Science and Stocks. *Newsweek*, September 19, p.92.

Schwager, J. (1993). Market Wizards: interviews with top traders. New York: Collins.

Schwager, J. (1995). The New Market Wizards: Conversations with America's Top Traders., New York: Wiley.

Siegel, J. J. (1991). Does it pay stock investors to forecast the business cycle? *Journal of portfolio management*, 18 (1), 27-34.

Stambaugh, R. F. (1999). Predictive regressions. *Journal of Financial Economics*, 54, 375-421.

Stock, J. H. and Watson, M. W. (1991). A probability model of the coincident economic indicators. in K. Lahiri and G, Moore (eds.), *Leading Economic Indicators: New approaches and forecasting records*, Cambridge, MA: Cambridge University Press.

Stock, J. H. and Watson, M. W. (2002a). Forecasting Using Principal Components From a Large Number of Predicators. *Journal of the American Statistic Association*, 97, 1167-1179.

Stock, J. H. and Watson, M. W. (2003). Forecasting Output and Inflation: The Role of Asset Prices. *Journal of Economic Literature*, 41 (3), 788-829.

Stock, J. H. and Watson, M. W. (2006). Forecasting with many predicators. In G. Elliott, C. Granger, and Timmermann, A. 1, (eds*), Handbook of Economic Forecasting* (pp.515-554), Elsevier.

Appendix A. Value strategies construction

The 32 value-weighted portfolios jointly sorted by size, book-to-market ratio, and operating profitability (2 x 4 x 4) are indicated as:

"s_bm1_op1, s_bm1_op2, s_bm1_op3, s_bm1_op4, s_bm2_op1, s_bm2_op2, s_bm2_op3, s_bm2_op4, s_bm3_op1, s_bm3_op2, s_bm3_op3, s_bm3_op4, s_bm4_op1, s_bm4_op2, s_bm4_op3, s_bm4_op4, b_bm1_op1, b_bm1_op2, b_bm1_op3, b_bm1_op4, b_bm2_op1, b_bm2_op2, b_bm2_op3, b_bm2_op4, b_bm3_op1, b_bm3_op2, b_bm3_op3, b_bm3_op4, b_bm4_op1, b_bm4_op2, b_bm4_op3, b_bm4_op4",

where s_bm1_op1 represents the 50th size, 25th book-to-market ratio, and 25th operating profitability jointly sorted portfolio. The rest of portfolios are named in the same way. The data is provided by Kenneth R. French data library. Please refer: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html for details.

Following are the procedures to construct the returns of four value strategies (value premiums): HML, HML_S, HML_RMW, and HML_RMW_S.

HML = *(s_bm4_op1+ s_bm4_op2+ s_bm4_op3+ s_bm4_op4+ b_bm4_op1+ b_bm4_op2+ b_bm4_op3+ b_bm4_op4)*(1/8) - (s_bm1_op1+ s_bm1_op2+ s_bm1_op3+ s_bm1_op4+ b_bm1_op1+ b_bm1_op2+ b_bm1_op3+ b_bm1_op4)*(1/8).*

HML_S = *s_bm4_op1+ s_bm4_op2+ s_bm4_op3+ s_bm4_op4)*(1/4) – (s_bm1_op1+ s_bm1_op2+ s_bm1_op2+ s_bm1_op3+ s_bm1_op4)*(1/4).*

HML_RMW = *(s_bm4_op4+ b_bm4_op4)*(1/2) - (s_bm1_op1+ b_bm1_op1)*(1/2).*

HML_RMW_S = *s_bm4_op4 - s_bm1_op1.*

Table 1: Economic indicators across three stock market states, Jan. 1967 to Dec. 2013

1967m1-2013m12																
states of stock markets					S&P 500		E12		IP		PIX		MTS		EMP	
	Obs	(%)	duration (months)	followed by a recession	A	F	A	F	A	F	A	F	A	F	A	F
Bull	430	76.2	43.6		12		16.8		0.9		1.2		2.1		0.6	
Good bear	52	9.2	8.7	19%	-7.6	-22.1	5.1	22.9	2.6	20.5	1.5	9.8	2.0	13.9	2.0	13.4
Bad bear	82	14.5	13.2	88%	-27.4	-221.7	-34.2	-274.4	-4.0	-25.7	-0.8	3.3	-4.0	-27.3	-0.8	-0.9
full sample	564															

E12 is 12-month moving average of aggregate earnings of firms on S&P 500 list; IP is seasonal adjusted industrial production; PIX is real personal income less transfers; MTS is real manufacturing and trade sales; and EMP is nonfarm payroll employment. In the third row, A indicates a variable's changes measured in amplitude, while F indicates the changes measured in cumulative loss. For bull stock markets, only the change of amplitude is provided by definition.

Table 2: Monthly market excess return and value premiums across stock market states

1967 m1- 2013 m12

Bull		EX_MKT	HML	HML_S	HML_RMW	HML_RMW_S	3m_tbl
	Mean (%)	1.55	0.17	0.29	0.52	0.93	0.39
	Volatility (%)	16.03	11.33	12.17	37.12	42.72	0.21
	Sharpe ratio		0.05	0.08	0.09	0.14	
good bear							
	Mean (%)	-2.61	1.95	2.06	2.54	2.47	0.53
	Volatility (%)	21.54	12.07	12.85	22.67	24.43	0.32
	Sharpe ratio		0.56	0.57	0.53	0.5	
bad bear							
	Mean (%)	-2.97	1.32	1.67	2.21	3.29	0.57
	Volatility (%)	24.27	21.19	20.32	78.97	76.18	0.47
	Sharpe ratio		0.29	0.37	0.25	0.38	
full sample							
	Mean (%)	0.51	0.5	0.65	0.95	1.41	0.43
	Volatility (%)	21.15	12.83	13.42	41.88	45.9	0.26
	Sharpe ratio	0.11	0.14	0.18	0.15	0.21	

This table shows monthly sample means, variances, and Sharpe ratios of value-weight market excess return and value premiums constructed from 32 value weighted portfolios sorted by size, book-to-market ratio, and operating profitability (2 x 4 x 4). Please see Appendix A for details.

Table 3: predictor description and transformation

Short Name	full Name	Description
Macroeconomic variables		
DP	dividend-price ratio (log)	log of a 12-month moving average of dividends paid on the S&P500 index minus the log of stock prices (S&P500 index).
DY	dividend yield (log)	log of 12-month moving average of dividends minus log of lagged stock price.
EP	earning-price ratio	log of a 12-month moving average of earnings on the S&P500 index minus the log of stock prices (S&P500 index).
DE	dividend - payout ratio (log)	log of 12-month moving average of dividends minus log of a 12 month moving average of earnings.
RVOL	equity risk premium volatility	based on a 12- month moving standard deviation estimator.
BM	book-to-market ratio	book to market value ratio for the Dow Jones Industrial Average.
NTIS	net equity expansion	ratio of a 12-month moving average of net equity issues by NYSE-listed stocks to the total end-of-year market capitalization of New York Exchange (NYSE) stocks.
TBL	treasure bill rate	interest rate on the three month treasury bill (secondary market).
LTY	long-term yield	long-term government bond yield.
LTR	long-term return	return on long-term government bonds
TMS	term spread	long term yield minus the treasury bill rate
DFY	default yield spread	difference between Moody's BAA and AAA rated corporate bond yields.
DFR	default return spread	long-term corporate bond return minus the long-term government bond return.
INFL	inflation	caculated from the CPI for all urban consumers; use lag one period data to account for the delay in CPI releases.
real economic variables		
IP	industrial production	first difference in log of industrial production index-total index
PIX	personal income	first difference in log of personal income less transfer payments
MTS	sales	first difference in log of manufacturing and trade sales
EMP	employment	first difference in log of employees on nonfarm payrolls: total privates

To be continued.

Table 3: predictor description and transformation

Technical indicator	description
MA(1,9)	moving average trading rule
MA(1,12)	-
MA(2,9)	-
MA(2,12)	-
MA(3,9)	-
MA(3,12)	-
MOM(9)	momentum trading rule
MOM(12)	-
VOL(1,9)	volume-based trading rule
VOL(1,12)	-
VOL(2,9)	-
VOL(2,12)	-
VOL(3,9)	-
VOL(3,12)	-

The data source of macroeconomic variables and technical indicators is from Neely et al. (2014). Real economic indicators are from Fed. of St. Louis and Fed. of Philadelphia website. Please refer Neely et al. (2014) for details in construction of technical indicators.

Table 4: Summary statistics

variable	mean	std	Min	max	auto-corr		(buy signal)%
Macroeconomic variable						technical indicator	
DP	-3.59	0.42	-4.52	-2.75	0.99	MA(1,9)	0.67
DY	-3.58	0.42	-4.53	-2.75	0.99	MA(1,12)	0.71
EP	-2.82	0.46	-4.84	-1.9	0.99	MA(2,9)	0.68
DE	-0.78	0.33	-1.24	1.38	0.99	MA(2,12)	0.7
RVOL	0.15	0.05	0.06	0.32	0.96	MA(3,9)	0.69
BM	0.51	0.28	0.12	1.21	0.99	MA(3,12)	0.71
NTIS	0.01	0.02	-0.06	0.05	0.98	MOM(9)	0.71
TBL	5.18	3.2	0.01	16.3	0.99	MOM(12)	0.73
LTY	7.11	2.55	2.06	14.82	0.99	VOL(1,9)	0.67
LTR	0.65	3.1	-11.24	15.23	0.04	VOL(1,12)	0.69
TMS	1.93	1.53	-3.65	4.55	0.95	VOL(2,9)	0.67
DFY	1.09	0.45	0.55	3.38	0.96	VOL(2,12)	0.69
DFR	0.01	1.51	-9.75	7.37	-0.07	VOL(3,9)	0.68
INFL	0.35	0.36	-1.92	1.79	0.61	VOL(3,12)	0.69
real economic variable							
IP	0.19	0.75	-4.36	2.38	0.35		
PIX	0.22	0.63	-6.8	4	-0.06		
MTS	0.22	0.66	-2.67	2.94	0.24		
EMP	0.13	0.22	-0.78	1.23	0.61		

This table shows descriptive statistics of all predictors under full sample period Jan. 1967 to Dec. 2013, 564 monthly data points. Please see Table 3 for details in data descriptions, resource, and transformations.

Table 5: In-sample univariate test, Jan. 1967 to Dec. 2013

Panel A	multinomial logit								Binary logit				
	β1			β2			R^2(%)	BIC	β			R^2(%)	BIC
macroeconomis variables													
DP	1.41	[3.59]	***	0.3	[1.05]		1.8	807	0.7	[2.88]	***	1.4	621.47
DY	1.21	[3.15]	***	0.12	[0.41]		1.4	811	0.52	[2.15]	**	0.8	625.34
EP	0.99	[2.79]	***	-0.1	[-0.38]		1.1	813	0.29	[1.30]		0.3	628.31
DE	0.37	[0.88]		0.58	[1.85]	*	0.4	818	0.51	[1.83]	*	0.5	626.79
RVOL	-15.02	[-4.04]	***	3.17	[1.41]		2.9	798	-2.52	[-1.28]		0.3	628.37
BM	2.44	[4.79]	***	0.95	[2.22]	**	3.2	796	1.53	[4.40]	***	3.1	610.63
NTIS	71.02	[6.03]	***	-10.94	[-1.87]	*	7.2	764	11.09	[2.13]	**	0.8	625.3
TBL	0.16	[3.39]	***	0.22	[5.61]	***	4.9	782	0.19	[5.88]	***	6.1	592.14
LTY	0.19	[3.42]	***	0.16	[3.37]	***	2.5	802	0.17	[4.36]	***	3.1	610.75
LTR	-0.08	[-1.67]	*	-0.03	[-0.80]		0.4	819	-0.05	[-1.55]		0.4	627.59
TMS	-0.12	[-1.17]		-0.51	[-6.13]	***	5	780	-0.36	[-5.32]	***	4.8	600.14
DFY	-0.28	[-0.73]		0.95	[4.15]	***	2.3	803	0.59	[2.86]	***	1.3	622.08
DFR	0.03	[0.28]		-0.13	[-1.70]	*	0.4	819	-0.07	[-1.11]		0.2	628.81
INFL	1.27	[3.08]	***	0.64	[1.85]	*	1.4	810	0.89	[3.10]	***	1.6	620.14
real economic variable													
IP	0.7	[2.99]	***	-0.85	[-5.17]	***	5.5	778	-0.35	[-2.72]	***	1.2	622.6
EMP	4.99	[5.34]	***	-3.65	[-6.23]	***	10.9	735	-0.82	[-1.8]	*	0.5	626.82
MTS	0.26	[1.04]		-1.35	[-6.84]	***	7.4	763	-0.77	[-4.95]	***	4	603.92
PIX	0.13	[0.51]		-0.52	[-2.61]	***	1.1	813	-0.3	[-1.87]	*	0.6	626.32

to be continued.

Table 5: In-sample univariate test, Jan. 1967 to Dec. 2013

	multinomial logit								Binary logit				
technical signal	β1			β2			R^2(%)	BIC	β			R^2(%)	BIC
MA1_9	-2.25	[-6.93]	*******	-2.94	[-9.44]	*******	19.1	670	-2.64	[-11.06]	***	24.1	481.01
MA1_12	-2.5	[-7.59]	***	-2.88	[-9.82]	*******	19.8	664	-2.72	[-11.47]	***	25.4	**473.22**
MA2_9	-2.04	[-6.49]	***	-2.9	[-9.51]	*******	17.9	679	-2.53	[-10.85]	***	22.4	491.55
MA2_12	-2.18	[-6.91]	***	-2.9	[-9.76]	*******	18.6	674	-2.6	[-11.13]	***	23.4	485.44
MA3_9	-1.83	[-5.93]	***	-2.71	[-9.22]	*******	15.7	696	-2.33	[-10.29]	***	20	509.59
MA3_12	-1.97	[-6.35]	***	-2.65	[-9.33]	*******	15.8	696	-2.37	[-10.46]	***	20	507.29
MOM_9	-2.35	[-7.28]	***	-2.7	[-9.48]	*******	17.8	680	-2.56	[-11.03]	***	23	489.06
MOM_12	-1.87	[-6.06]	***	-2.66	[-9.50]	*******	15.4	699	-2.34	[-10.35]	***	19	512.18
OBV1_9	-2.33	[-7.06]	***	-2.77	[-9.26]	*******	18	677	-2.58	[-10.92]	***	23	486.23
OBV1_12	-2.32	[-7.19]	***	-3.12	[-9.82]	*******	20	662	-2.72	[-11.39]	***	25	473.55
OBV2_9	-2.31	[-7.02]	***	-2.75	[-9.22]	*******	18	678	-2.6	[-10.87]	***	23	487.65
OBV2_12	-2.36	[-7.23]	***	-2.68	[-9.34]	*******	17.7	680	-2.55	[-10.96]	***	22.7	489.35
OBV3_9	-2.18	[-6.80]	***	-2.72	[-9.27]	*******	17.3	684	-2.49	[-10.75]	***	22	494.39
OBV3_12	-2.12	[-6.69]	***	-2.68	[-9.26]	*******	16.7	688	-2.45	[-10.65]	***	21.2	498.67

This table shows in-sample forecasting results by using individual predictor. Left panel is the results from static multinomial logit model, while right panel is the results from static binary logit model. The full sample period is from Jan. 1967 to Dec. 2013. The forecasting horizon is one month ahead. β1 indicates the predictor's impacts for the probability of good bear markets relative to the probability of bull markets, whereas β2 indicates the predictor's impacts for the probability of bad bear markets relative to the probability of bull markets. t-statistics are provided in square bracket. R^2 and BIC represent pseudo-R2 and Schwarz Information Criterion (BIC) of the regression. "*","**",and "***" denote significance at 10%, 5%, and 1% level respectively.

Table 6: In-sample multivariate test, Jan. 1967 to Dec. 2013.

Panel B	Multinomial logit								Binary logit			
	Model 1				Model 2				Model 3		Model 4	
	static				dynamic				static		dynamic	
	beta1		beta2		beta1		beta2		beta		beta	
Const	-2.95	***	-2.28	***	-3.21	***	-2.13	***	-1.65	***	-1.89	***
	[-11.36]		[-11.54]		[-9.56]		[-10.05]		[-11.40]		[-11.17]	
PC1	0.31	***	0.49	***	0.31	***	0.55	***	0.45	***	0.35	***
	[5.57]		[10.46]		[4.75]		[10.07]		[11.72]		[6.89]	
PC2	-0.23	***	-0.01		-0.34	***	-0.06		-0.06		-0.04	
	[-2.25]		[-0.17]		[-4.02]		[-1.01]		[-1.12]		[-0.86]	
PC3	-0.89	***	-0.08		-1.14	***	-0.07		-0.23	***	-0.29	***
	[-5.89]		[-1.16]		[-6.01]		[-0.81]		[-3.48]		[-4.02]	
$I_{1\ t-12}$					-1.59	**	-1.54	***				
					[-2.45]		[-2.98]					
$I_{2\ t-12}$					1.54	**	-0.37					
					[2.25]		[-0.81]					
S_{t-6}											1.06	**
											[2.93]	
log_L			-270.34				-260.1				-201.05	
R2(%)			31.59				34.18		33.2		34.58	
QPS			0.227				0.224		0.23		0.22	
BIC			591.17				595.92		442.18		433.71	

This table shows in-sample forecasting results by using the first three principal components extracted from all candidate variables as predictors. Left panel is the results from the multinomial logit model, while right panel is the results from binary logit model. The full sample period is from Jan. 1967 to Dec. 2013. The forecasting horizon is one month ahead. β1 indicates predictors' impacts for the probability of good bear markets relative to the probability of bull markets, whereas β2 indicates predictors' impacts for the probability of bad bear markets relative to the probability of bull markets. t-statistics are provided in square bracket. R^2 and BIC represent pseudo-R2 and Schwarz Information Criterion (BIC) of the regression. "*","**",and "***" denote significance at 10%, 5%, and 1% level respectively.

Table 7: Out-of-sample performances, Jan. 1977 to Dec. 2013

	QPS	LPS	AUC	cut-off
Panel (a)				
multinomial logit				
Model (1)	0.262	0.472	0.768	0.261
Model (2)	0.259	0.469	0.788	0.31
binary logit				
model (3)	0.253	0.457	0.769	0.364
model (4)	0.274	0.485	0.766	0.265
Panel (b)				
model comparisons				
		CW	W_{AUC}	
Model (1) vs. Model (2)		**0.027**	**0.012**	
Model (3) vs. Model (4)		**0.019**	0.576	
		DM	W_{AUC}	
Model (2) vs. Model (3)		0.493	**0.041**	

This table reports out-of-sample results of using the first three principal components as predictors in the multinomial logit model and in the binary logit model. Model (1) and Model (3) are the static models, whereas Model (2) and Model(4) are the dynamic models. Panel (a) presents the evaluation measures. Panel (b) reports the p-vale of various model comparison tests. For forecasting accuracy, the p-values of CW (nested models) and DM (non-nested models) statistics are presented; for classification ability, the p-value of W_{AUC} statistics is presented. Bold type denotes reject null hypothesis of equal forecasting accuracies or equal classification abilities.

Table 8: Hitting rate (%) across three stock market states, Jan. 1977 to Dec. 2013

	Stock market states		1977m1-2013m12		
	bull	good bear	bad bear	bear	total
obs (%)	78.82	7.43	13.74	21.17	100

model(2) multinomial model		true states			model(3) binary model		true states		
	bull	good bear	bad bear	bear		bull	good bear	bad bear	bear
bull	83.42	33.33	24.59	27.66	bull	89.14	39.39	29.51	32.97
good bear	9.14	48.48	29.51		bear	10.86	60.60	70.49	67.02
bad bear	7.43	18.18	45.90						
bear	16.57	66.66	75.41	72.34					

This table shows the hitting rate across three stock market states. The left panel is the result based on the dynamic multinomial logit model, whereas the right panel is the result based on the static binary logit model. The forecasting period is from Jan. 1977 to Dec. 2013. All the values are presented in percentage terms.

Table 9: Out-of-sample trading performance, Jan. 1977 to Dec. 2013

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
		B&H	M	B	M_HML	B_HML	A_HML	HML	HML_S	HML_RMW	HML_RMW_S
Panel (A)	Full sample										
(1)	mean (%)	0.7	0.62	0.681	1.18	1.01	1.57	1.23	1.34	1.48	1.57
(2)	sd(%)	4.48	3.35	3.491	5.44	5.32	6.6	4.09	4.21	5.16	5.45
(3)	Sharpe ratio	0.16	0.18	0.195	0.22	0.19	0.24	0.3	0.32	0.29	0.29
(4)	final_W	88.65	76.07	98.935	579.16	281.16	2391.94	1014.57	1573.12	2467.62	3425.97
(5)	turnover		11.36	8.81	21.05	16.56	6.62	27.85	28.98	29.24	29.29
(6)	Max drawdown	50.66	23.24	23	69.87	69.94	73.27	23.31	22.79	26.65	27.87
(7)	CER_2 (%)	0.92	0.93	0.98	1.3	1.14	1.56	1.49	1.58	1.64	1.69
(8)	CER_5 (%)	0.62	0.76	0.8	0.85	0.72	0.9	1.24	1.32	1.24	1.25
(9)	alpha_ff3 (%)	0.02	0.3**	0.32**	0.25	0.11	0.2**	0.74***	0.84***	0.97***	1.03***
(10)		[0.58]	[2.12]	[2.43]	[1.31]	[0.6]	[2.19]	[3.86]	[4.12]	[4.15]	[4.46]
(11)	alpha_ff5 (%)	-0.11***	0.13	0.17	0.19	0.05	0.12	0.56***	0.69***	0.71***	0.68***
(12)		[-3.95]	[1.03]	[1.4]	[0.98]	[0.29]	[1.24]	[3.69]	[4.14]	[4.06]	[3.6]
Panel (B)											
	bull										
(13)	mean (%)	1.53	0.99	1.12	1.69	1.77	1.98	1.45	1.5	1.68	1.71
(14)	sd	3.98	3.39	3.52	4.39	4.55	6.04	3.9	4.03	4.47	4.73
(15)	Sharpe ratio	0.39	0.29	0.32	0.39	0.39	0.33	0.37	0.37	0.37	0.36
	good bear										
(16)	mean (%)	-2.88	-1.15	-1.36	-0.66	-0.75	0.71	-0.18	0.4	0.08	0.66
(17)	sd	5.01	4	4.03	6.73	6.82	6.84	5.56	5.6	6.67	6.42
(18)	Sharpe ratio	-0.58	-0.29	-0.34	-0.1	-0.11	0.1	-0.03	0.07	0.01	0.1
	bad bear										
(19)	mean (%)	-2.17	-0.58	-0.74	-0.8	-0.9	-0.29	0.76	0.92	1.15	1.26
(20)	sd	4.83	1.79	1.89	8.67	8.66	8.92	4.13	4.32	7.43	8.17
(21)	Sharpe ratio	-0.45	-0.32	-0.39	-0.09	-0.1	-0.03	0.19	0.21	0.15	0.15

This table reports out-of-sample trading performances based on the dynamic multinomial logit model and the static binary logit model. Column (1) reports the performance under Buy-and-Hold strategy; Columns (2) to (3) report the performances of type 1 trading strategies based on the dynamic multinomial logit model and the dynamic binary logit model, respectively. Columns (4) to (5) show the performances of type 2 trading strategies based on the dynamic multinomial logit model and the dynamic binary logit model, respectively. Column (6) shows the performance of always performing the value strategy (HML) at each month without timing. Columns (7) to (10) present the performances of type 3 trading strategies based on the dynamic multinomial logit model, where the investor performs value strategy of HML, HML_S, HML_RMW or HML_RMW_S respectively when the model prediction is the good bear market state. Panel (A) reports the performances over full out-of-sample period, whereas Panel (B) reports the performances across three stock market states. Rows (1) to (4) report monthly average excess returns, monthly standard deviations, monthly Sharpe ratios, and final wealth at the end of trading period. Row (5) gives monthly average turnover. Row (6) gives the maximum monthly drawdown. Rows (7) to (8) report the monthly certainty equivalent returns (CER), where risk aversion is set to 2 and 5 to represent different degrees of risk-averse. Rows (9) to (12) report the Jensen's alpha under Fama-French 3 factor (FF3) model and Fama-French 5 factor (FF5) model respectively, with the corresponding t statistics values provided in square bracket. "*","**",and "***" denote significance at 10%, 5%, and 1% level respectively.

Figure 1: S&P 500 index and aggregate earnings, Jan. 1967 to Dec. 2013.

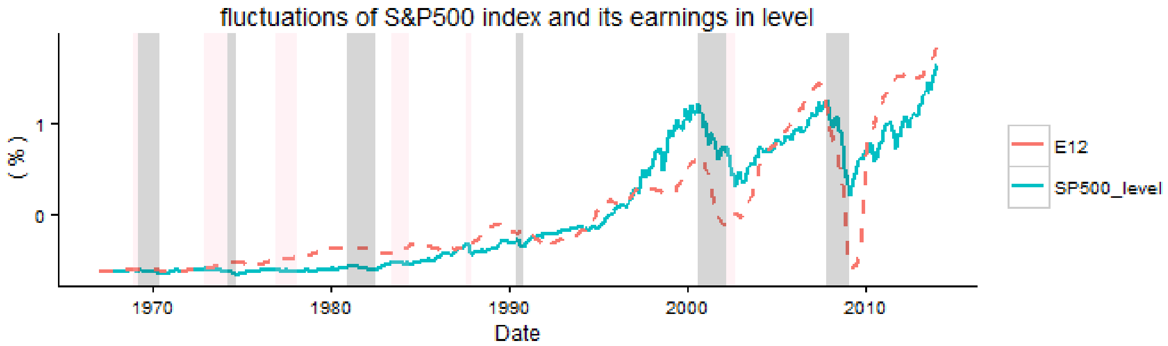


Figure 1 depicts time series of S&P 500 index and its 12-month moving average of earnings. For comparison, both series are standardized. The shaded bars are the stock bear markets generated through Bry-Boschan dating rule algorithm, where pink indicates good bear markets and gray indicates bad bear markets.

Figure 2: Four real economic activity indicators, Jan. 1967 to Dec. 2013.

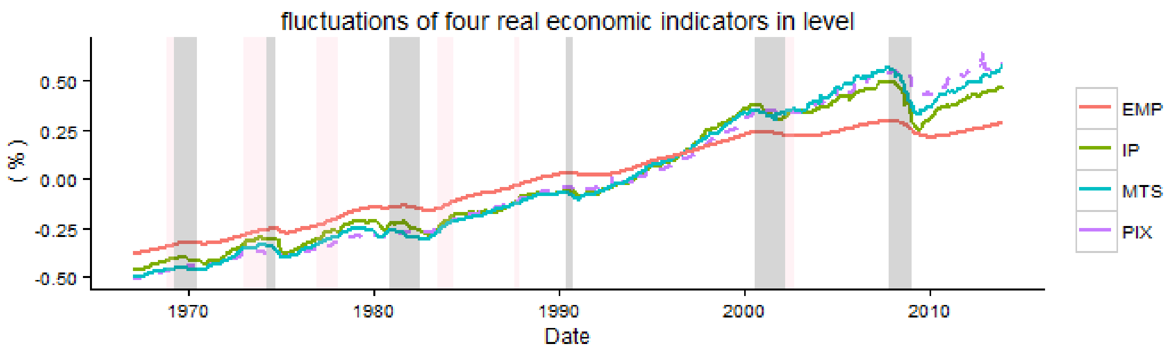
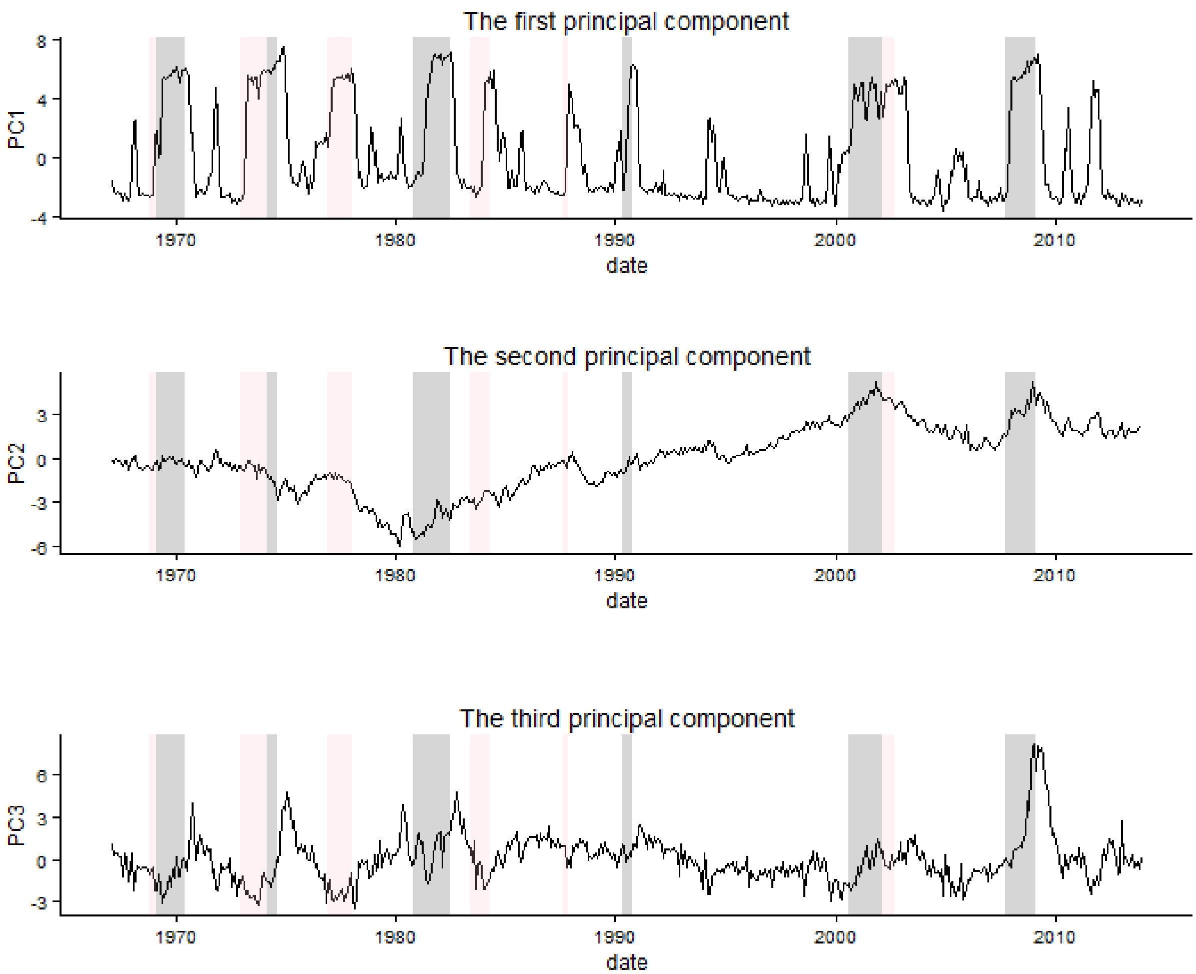


Figure 2 depicts time series of four real economic indicators. For comparison, all series are standardized. The shaded bars are the stock bear markets generated through Bry-Boschan dating rule algorithm, where pink indicates good bear markets and gray indicates bad bear markets.

Figure 3: The first 3 principal components extracted from macroeconomic, real activity variables and technical indicators (32 variables), Jan. 1967 to Dec. 2013.



The first 3 principal components contribute 60% of total variation of whole dataset. The shaded bars are the stock bear markets generated through Bry-Boschan dating rule algorithm, where pink indicates good bear markets and grey indicates bad bear markets.

Figure 4: Loadings on principal components extracted from 32 variables, Jan 1967 to Dec. 2013.

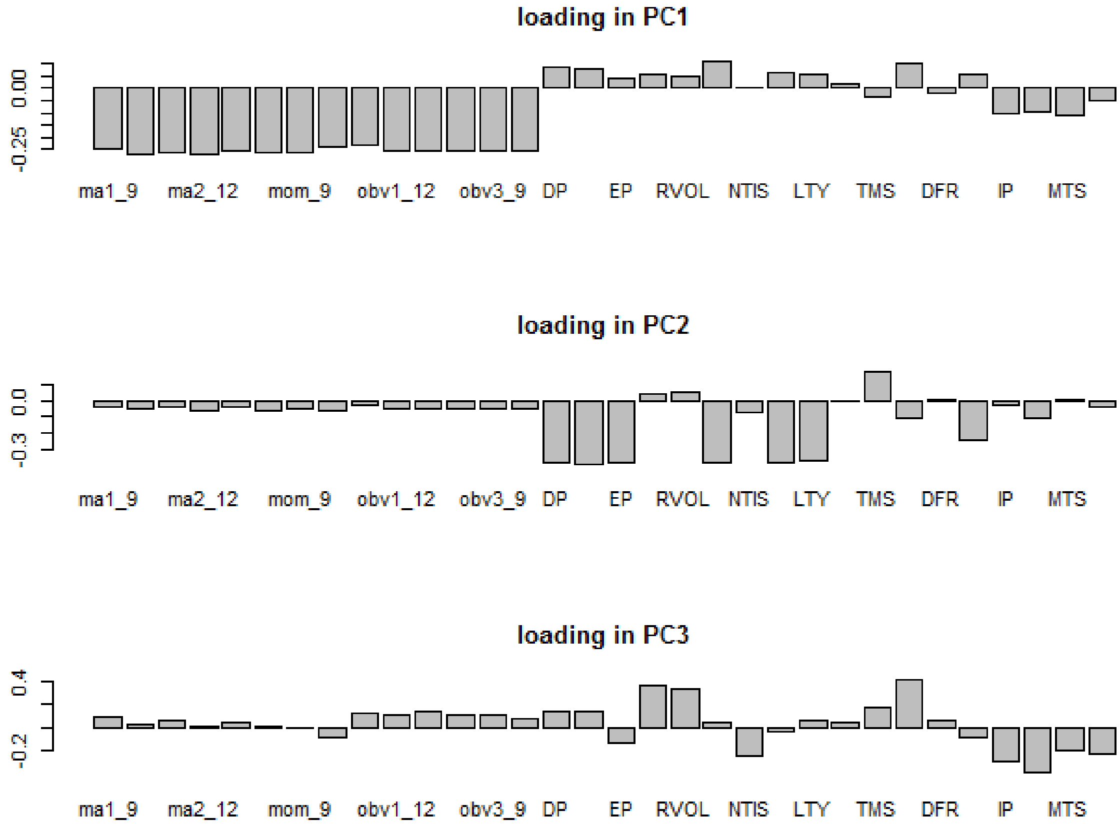
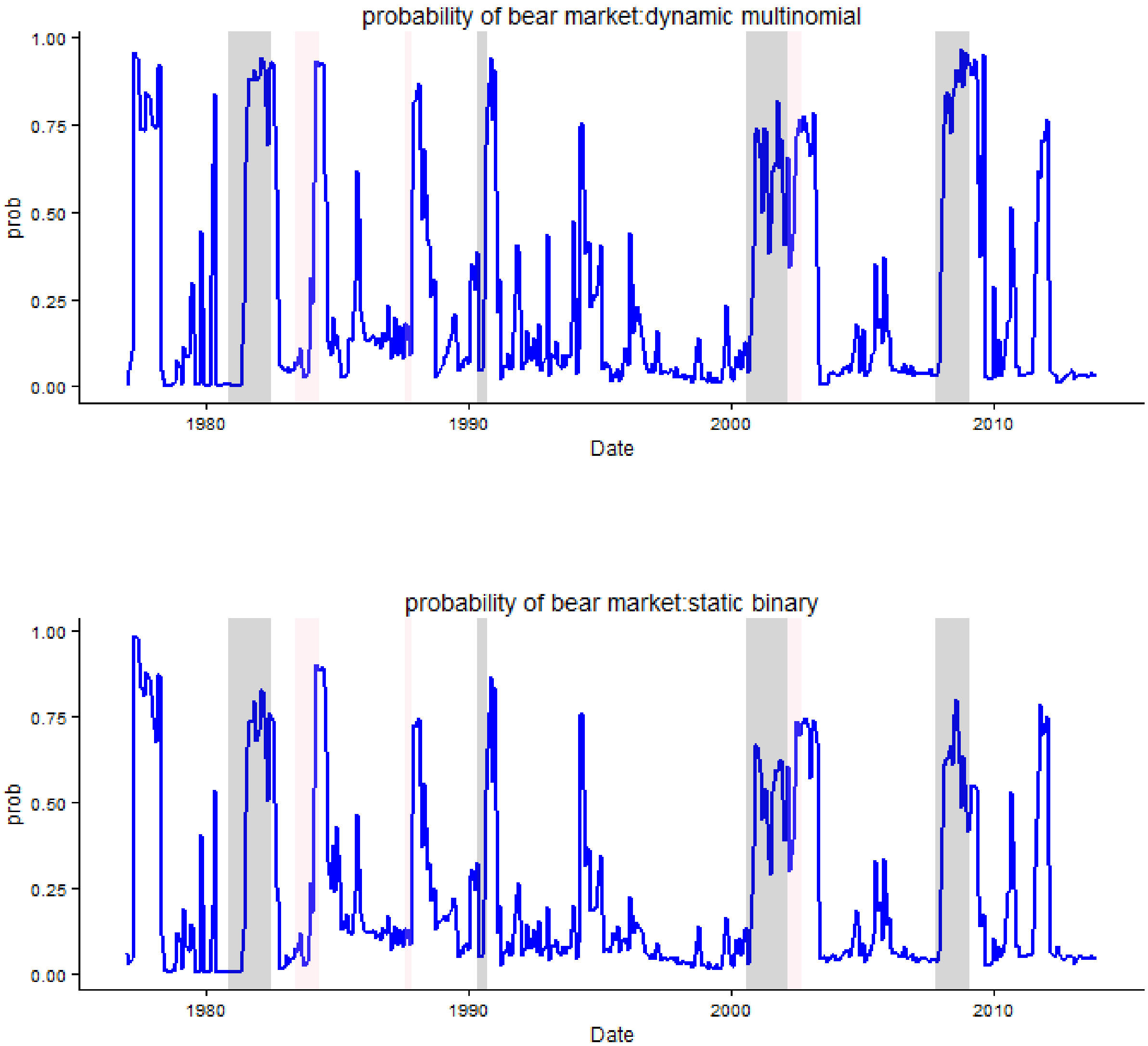
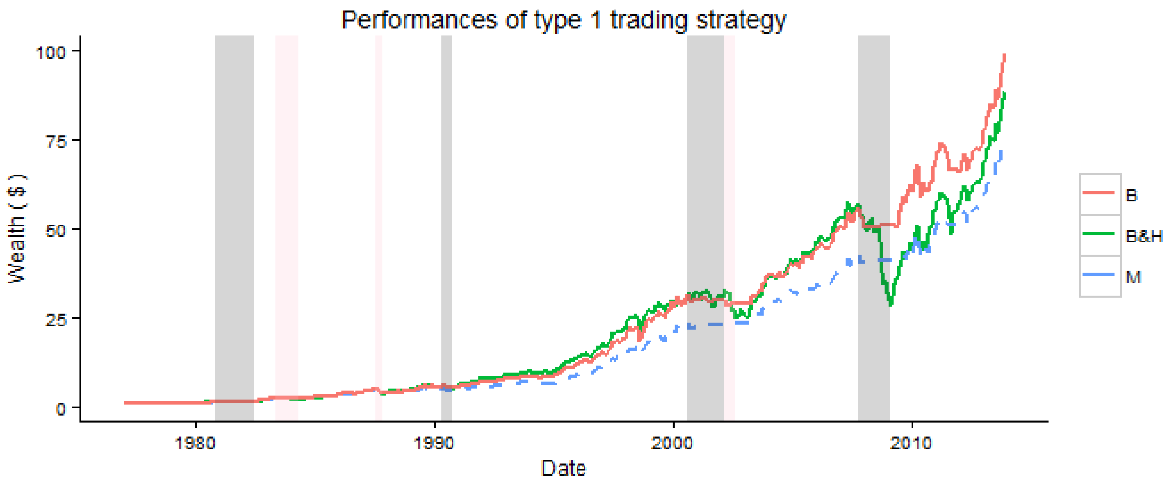


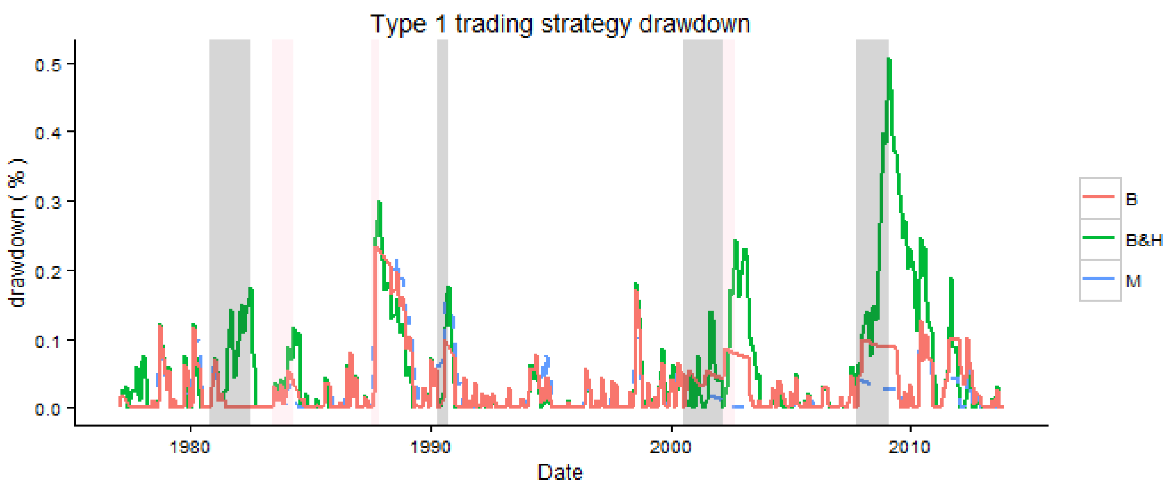
Figure 5: Out-of-sample predicted probabiities of bear markets under the dynamic multinomial logit model and the static binary logit model, Jan. 1977 to Dec. 2013.



The upper panel presents the out-of-sample predicted probability from the dynamic multinomial logit model, whereas the lower panel presents the out-of-sample predicted probability from the static binary logit model. Predictors are the first three principal components extracted from 32 variables. The shaded bars are the stock bear markets generated through Bry-Boschan dating rule algorithm, where pink indicates good bear markets and gray indicates bad bear markets.

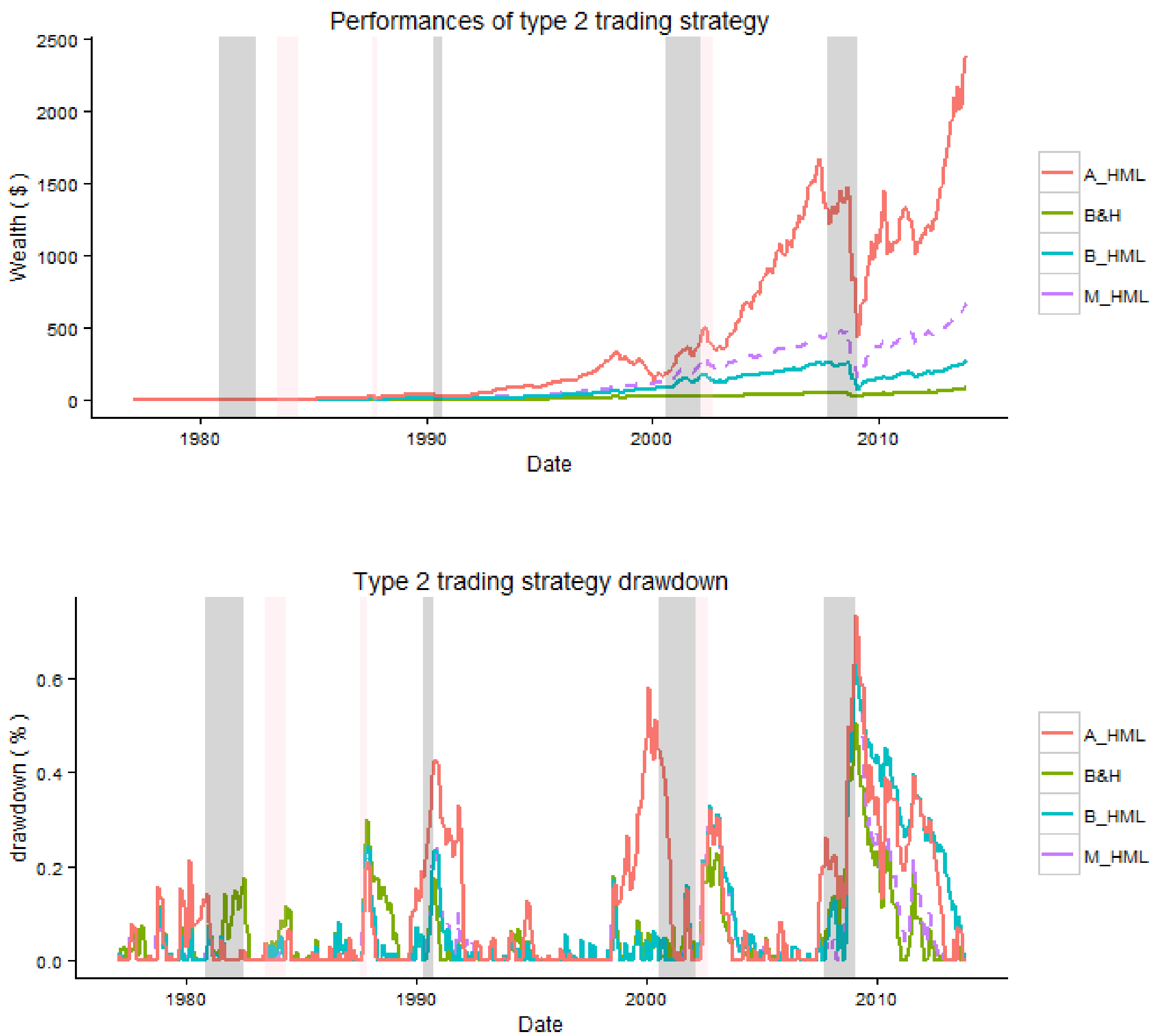
Figure 6: Cumulative wealth for type 1 trading strategies based on the dynamic multinomial logit and the static binary logit model, Jan. 1977 to Dec. 2013.





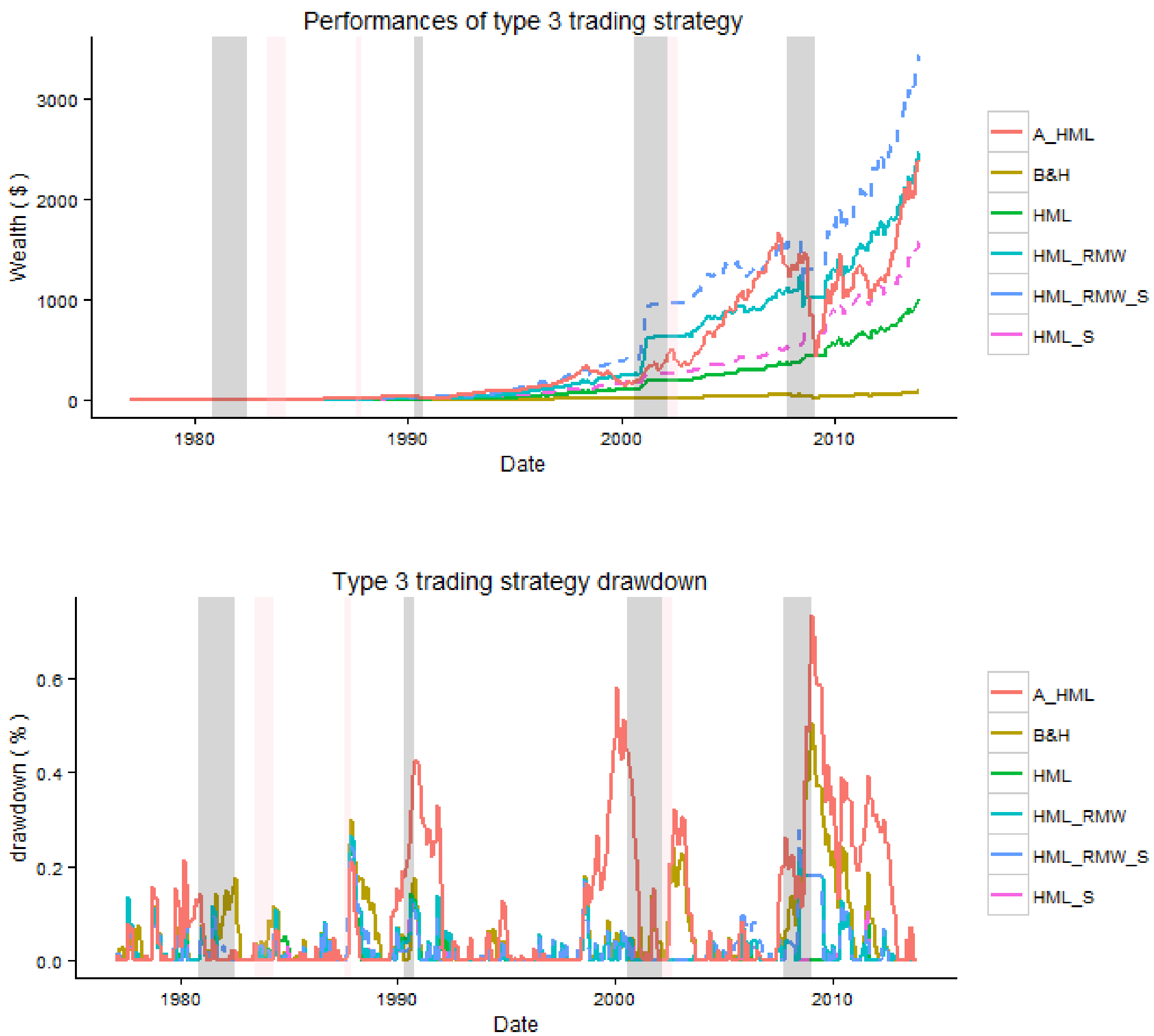
This figure presents out-of-sample performances of two type 1 trading strategies based on dynamic multinomial logit model (M) and static binary logit model (B). The performance of Buy-and-Hold (B&H) benchmark strategy is provided for comparison. The upper panel depicts the cumulative wealth of investing $1 at the beginning. The lower panel depicts the percentage drops in the cumulative returns from the peak along the trading period. The shaded bars are the stock bear markets generated through Bry-Boschan dating rule algorithm, where pink indicates good bear markets and gray indicates bad bear markets.

Figure 7: Cumulative wealth for type 2 trading strategies based on the dynamic multinomial logit model and the static binary logit model, Jan. 1977 to Dec. 2013.



This figure presents out-of-sample performances of two type 2 trading strategies based on multinomial logit model (M) and binary logit model (B). The performances of benchmark Buy-and-Hold (B&H) strategy and always implementing HML strategy (A_HML) are provided for comparison. The upper panel depicts the cumulative wealth of investing $1 at the beginning. The lower panel depicts the percentage drops in the cumulative returns from the peak along the trading period. The shaded bars are the stock bear markets generated through Bry-Boschan dating rule algorithm, where pink indicates good bear markets and gray indicates bad bear markets.

Figure 8: Cumulative wealth for type 3 trading strategies based on the dynamic multinomial logit model, Jan. 1977 to Dec. 2013.



This figure presents out-of-sample performances of four type 3 trading strategies (HML, HML_S, HML_RMW, and HML_RMW_S) based on the dynamic multinomial logit model. The performances of benchmark Buy-and-Hold (B&H) strategy and always implementing HML strategy (A_HML) are provided for comparison. The upper panel depicts the cumulative wealth of investing $1 at the beginning. The lower panel depicts the percentage drops in the cumulative returns from the peak along the trading period. The shaded bars are the stock bear markets generated through Bry-Boschan dating rule algorithm, where pink indicates good bear markets and gray indicates bad bear markets.