

# Essays in Financial Economics

Wan-Jung Hsu

A dissertation  
submitted in partial fulfillment of the  
requirements for the degree of

Doctor of Philosophy

University of Washington

2017

Reading Committee:

Eric Zivot, Co-Chair

Levis Kochin, Co-Chair

Thomas Gilbert

Chang-Jin Kim

Program Authorized to Offer Degree:

Economics

©Copyright 2017

Wan-Jung Hsu

University of Washington

**Abstract**

Essays in Financial Economics

Wan-Jung Hsu

Chairs of the Supervisory Committee:

Robert R. Richards Professor Eric Zivot

Associate Professor Levis Kochin

Department of Economics

This dissertation aims to investigate the interaction between financial markets and the real economy both in the short-run and the long-run. The first two chapters study the distinct interactions between different states of stock markets and the real economy at monthly frequencies. The third chapter studies the causality between financial development and the economic growth at business cycle frequencies (i.e., three or five-year spells).

The first chapter focuses on forecasting the states of the stock market. While previous literature classifies the stock market into binary states (i.e., 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 with NBER declared recessions, whereas good bear markets are not accompanied with 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.

In the second chapter, I examine the reliability and timeliness of using stock bad bear markets as early warning signals of economic recessions. I find that bad bear markets are much more reliable to predict recessions than conventional stock bear markets or the forecasting model that targets recessions directly. The forecasting model that predicts bad bear markets also provide timely information about the starts and the ends of economic recessions over NBER announcements.

In the third chapter, I revisit the debate of “too much finance” on economic growth. I use different econometric methods in the dynamic panel data framework to address potential biases induced by the dynamic nature of economic growth and financial development but control the heterogeneity across countries. Particularly, I conduct a battery of robustness tests to examine weak instrument problems in the system GMM estimator, developed by Blundell and Bond (1998), and use Half-Panel Jackknife Fixed-Effect estimator, developed by Chudik, Pesaran, and Yang (2016), as an alternative method. I also take care of the outlier issue, which is particularly sensitive when there is nearly multicollinearity among explanatory variables. My empirical results find no sufficient evidence to support a positive causal effect, nor do I find a quadratic effect of financial development on economic growth.

## ACKNOWLEDGMENTS

I wish to express my sincere gratitude to many people who helped me complete this dissertation. Firstly, I would like to thank my committee chairs, Eric Zivot and Levis Kochin, for their guidance and supports. I also thank Prof. Thomas Gilbert and Prof. Chang-Jin Kim for all their helpful suggestions on my work. My thanks also go to Prof. Philip Bond for giving many insightful comments. Last but not least, I thank my parents for full supports throughout my doctoral study. Without them I'd not have gotten this far in life.

DEDICATION

to my parents.

## TABLE OF CONTENTS

	Page
List of Figures .....	ix
List of Tables .....	x
Chapter 1: Predicting and Capitalizing on Two Types of Stock Bear Markets in the U.S. ....	1
1.1 Introduction .....	1
1.2 Classifying states of the stock market .....	7
1.3 Interaction of stock market states with the real economy .....	10
1.4 Model specifications and evaluations .....	14
1.5 Economic values of the multinomial logit model in predicting bear markets .....	30
1.6 Real-time out-of-sample robustness test .....	37
1.7 Conclusion .....	38
1.8 References .....	40
1.9 Appendix .....	44
Chapter 2: Predicting recessions with bad bear markets .....	64
2.1 Introduction .....	64
2.2 Stock markets and business cycles .....	66
2.3 Out-of-sample test of signaling NBER recessions .....	67

2.4 Performances in providing timely information of business cycles .....	70
2.5 Conclusion .....	74
2.6 References .....	76
Chapter 3: Revisit the “Too Much Finance” in dynamic panel data analysis .....	87
3.1 Introduction .....	87
3.2 Empirical model, data, and methodology .....	92
3.3 Alternative estimation .....	95
3.4 Empirical results .....	100
3.5 Robustness tests .....	104
3.6 Conclusion .....	106
3.7 References .....	109

## LIST OF FIGURES

Figure Number	Page
1.1 S&P 500 index and aggregate earnings .....	57
1.2 Four real economic activity indicators .....	57
1.3 The first 3 principal components extracted from macroeconomic, real activity variables and technical indicators (32 variables) .....	58
1.4 Loadings on principal components extracted from 32 variables .....	59
1.5 Out-of-sample predicted probabilities of bear markets .....	60
1.6 Cumulative wealth for type 1 trading strategies.....	61
1.7 Cumulative wealth for type 2 trading strategies .....	62
1.8 Cumulative wealth for type 3 trading strategies .....	63
2.1 Out-of-sample bear market probabilities and the identified bad bear market phases under the multinomial logit model .....	84
2.2 Out-of-sample bear market probabilities and identified bear market phases under the binary logit model.....	84
2.3 Out-of-sample recession probabilities and identified recession phases under 4 scenarios ...	85
3.1 Estimates based on the simulations with $k = 0.1$ and $T = 10$ .....	121
3.2 Estimates based on the simulations with $k = 0.3$ and $T = 10$ .....	123
3.3 Estimates based on the simulations with $k = 0.1$ and $T = 15$ .....	125
3.4 Estimates based on the simulations with $k = 0.3$ and $T = 15$ .....	127
3.5 Leverage plot based on the Pool OLS estimates of column (1), Table 3.4 .....	129
3.6 Partial correlations between credits to the private sector (PC) and growth .....	129
3.7 Partial correlations between the square of credits to the private sector (PC2) and growth ..	130

## LIST OF TABLES

Table Number	Page
1.1 Economic indicators across three stock market states .....	45
1.2 Monthly market excess return and value premiums across stock market states .....	46
1.3 Predictor description and transformation .....	48
1.4 Summary statistics .....	49
1.5 In-sample univariate test .....	50
1.6 In-sample multivariate test .....	52
1.7 Out-of-sample performances .....	53
1.8 Hitting rate (%) across three stock market states .....	54
1.9 Out-of-sample trading performance .....	55
2.1 In-sample stock bear market classification and business cycles .....	77
2.2 Out-of-sample identified bear markets phases .....	78
2.3 Out-of-sample identified recession phases under binary logit model .....	79
2.4 Out-of-sample performances in timeliness (peak) over the last 5 NBER recessions .....	80
2.5 Out-of-sample performances in timeliness (peak) of binary logit model that predicts recessions directly over the last 5 NBER recessions .....	81
2.6 Out-of-sample performances in timeliness (trough) over the last 5 NBER recessions .....	82
2.7 Out-of-sample performances in timeliness (trough) of binary logit model that predicts recessions directly over the last 5 NBER recessions .....	83
3.1 Summary statistics: 5-year spell panel dataset .....	111
3.2 Monte Carlo Simulation sets: $k = 0.1$ .....	112
3.3 Monte Carlo Simulation sets: $k = 0.3$ .....	113

3.4 Dynamic Panel data estimations with 5-year spell dataset: Baseline model .....	114
3.5 Dynamic Panel data estimation with 5-year spell dataset: Model including the square term .....	115
3.6 Dynamic Panel data estimations with 5-year spell dataset excluding the outliers .....	117
3.7 Dynamic Panel data estimations with 3-year spell dataset excluding the outliers .....	119

## Chapter 1

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

#### **1.1 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,

$$P_t = \sum_{\tau=1}^{\infty} E(D_{t+\tau}) / (1 + r_{t+\tau})^{\tau}, \quad (1.1)$$

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 Eq. (1.1), 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<sup>1</sup>. 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

---

<sup>1</sup> 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.

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 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<sup>2</sup>. 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

---

<sup>2</sup> 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 aims to classify stock markets into discrete states and so does not distinguish stock markets with specific standards.

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<sup>3</sup> 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

---

<sup>3</sup> 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.

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 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 1.2 presents the method of determining stock market states. Section 1.3 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 1.4 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 1.5 shows economic implications of the multinomial logit model for investors. Section 1.6 performs the real-time out-of-sample tests as a robustness check. Section 1.7 concludes.

## **1.2 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.

### *1.2.1. 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<sup>4</sup> and the reporting frequency of other cash flows alike measurements<sup>5</sup>, I choose 12-month moving average of earning on S&P 500 index as the proxy of cash flows.<sup>6</sup> 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<sup>7</sup>.

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.

---

<sup>4</sup> 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.

<sup>5</sup> Earnings are reported shortly after the quarter-end, whereas cash flow statements are not quarterly reported and are not required-reported until 1988.

<sup>6</sup> Since earning has strong seasonality, without notification, all the earnings data used in this study are 12-month moving averaged.

<sup>7</sup> The availability of predictors used in the forecasting model determines the beginning of the sample period.

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 local 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 phase to last at least 15 months and 6 months<sup>8</sup>, respectively. Assuming  $y_t$  is the time series data examined, the turning point would be a peak at time  $t$  if,  $y_{t-6}, \dots, y_{t-1} < y_t > y_{t+1}, \dots, y_{t+6}$ , and a trough if,  $y_{t-6}, \dots, y_{t-1} > y_t < y_{t+1}, \dots, y_{t+6}$ . 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.

### *1.2.2. 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<sup>9</sup>, it is classified as a bad bear market; otherwise it is classified as a good bear market. Figure 1.1 plots the time series of S&P 500 index and its earnings. The

---

<sup>8</sup> 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 Pagan and Sossounov (2003) and Claessens et al. (2011b).

<sup>9</sup> 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.

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.

### **1.3 Interaction of stock market states with the real economy**

#### *1.3.1. 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 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<sup>10</sup> ( $A_u = y_{k+12} - y_k$ ). 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<sup>11</sup>:  $F_c = \sum_{j=1}^k (y_j - y_0) - \frac{A_c}{2}$ . 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 1.1 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 1.2 depicts the time series of four standardized real economic indicators.

---

<sup>10</sup> 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.

<sup>11</sup> This formula is based on a triangular approximation of lost output during a contraction phase.

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.

### *1.3.2. 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 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<sup>12</sup>. 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 1.2 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

---

<sup>12</sup> 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.

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 1.2 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

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 1.5.

## 1.4 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.

### 1.4.1. 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  $\Omega_{t-1}$  as  $E_{t-1}(\cdot)$ , the conditional probability of a bear market state at time  $t$  can be written as:

$$p_t = E_{t-1}(S_t) = P_{t-1}(S_t = 1) = \Lambda(\pi_t). \quad (1.2)$$

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

$$\pi_t = X_{t-h}'\beta, \quad (1.3)$$

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:

$$P_{t-1}(S_t = 1) = \Lambda(\pi_t) = \frac{e^{(X_{t-1}'\beta)}}{1+e^{(X_{t-1}'\beta)}}. \quad (1.4)$$

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

$$\frac{P_{t-1}(S_t=1)}{P_{t-1}(S_t=0)} = e^{X_{t-1}'\beta}. \quad (1.5)$$

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  $\beta$ <sup>13</sup>. In this setting, no matter what the underlying bear market type is, the model assumes and estimates the same  $\beta$  for both good and bear markets.

---

<sup>13</sup> 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  $X_{t-1}$ .

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

$$\pi_t = X'_{t-1}\beta + \delta_k S_{t-k}, \quad (1.6)$$

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.

#### 1.4.2. 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  $\Omega_{t-1}$ ,  $S_t$  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:

$$P_{t-1}(S_t = j) = \frac{\exp(\pi_{jt})}{1 + \sum_{i=1}^2 \exp(\pi_{it})}, j = 1, 2 \quad (1.7)$$

$$P_{t-1}(S_t = 0) = \frac{1}{1 + \sum_{i=1}^2 \exp(\pi_{it})} \quad (1.8)$$

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

The odds ratios between states are:

$$\frac{P_{t-1}(S_t=1)}{P_{t-1}(S_t=0)} = e^{\pi_{1t}} = e^{(X_{t-1}'\beta_1)} \quad (1.9)$$

$$\frac{P_{t-1}(S_t=2)}{P_{t-1}(S_t=0)} = e^{\pi_{2t}} = e^{(X_{t-1}'\beta_2)} . \quad (1.10)$$

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,  $\beta_2$  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  $\beta_1$  and  $\beta_2$  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,  $\pi_{jt}$ , can be specified as:

$$\pi_{jt} = X'_{t-1}\beta_j + \sum_{i=1}^2 \delta_{ji}^k I_{it-k}, j = 1, 2 . \quad (1.11)$$

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  $I_{j,t-12}, j = 0, 1, 2$  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<sup>14</sup>, 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 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. Coval (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,

---

<sup>14</sup> The bear markets contribute 23% of the whole stock market in my sample period.

and 4 monthly real economic activity indicators from January 1967 to December 2013<sup>15</sup> as candidate predictors.

Table 1.3 provides descriptions and data sources of all variables in details. Table 1.4 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 Watson, 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.

#### *1.4.3. 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:

$$QPS = \frac{2}{T} \sum_{t=1}^T (\hat{P}_t - S_t)^2, \quad (1.12)$$

---

<sup>15</sup> The availability of the real economic activity variables determines the beginning of the sample period.

where  $\hat{P}_t$  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:

$$LPS = \frac{-1}{T} \sum_{t=1}^T [(1 - S_t) \ln(1 - \hat{P}_t) + S_t \ln(\hat{P}_t)]. \quad (1.13)$$

QPS and LPS range from 0 to 2 and from 0 to  $\infty$ , 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)<sup>16</sup> 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 Jordà, 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<sup>17</sup>. 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

---

<sup>16</sup> 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-specificity* ( $1 - S_p$ ) 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.

<sup>17</sup> Candelon, Dumitrescu, and Hurlin (2012) provides a Matlab toolbox to calculate various measures to evaluate an Early Warning System.

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  $DM$ :

$$DM = \frac{\sqrt{T}\bar{d}}{\sigma_{\bar{d},0}} \xrightarrow{d} N(0,1), \quad (1.14)$$

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  $CW$ :

$$CW = \frac{\sqrt{T}\bar{f}}{\sigma_{\bar{f},0}} \xrightarrow{d} N(0,1), \quad (1.15)$$

where  $f_t = (S_t - \hat{p}_{1,t})^2 - [(S_t - \hat{p}_{2,t})^2 - (\hat{p}_{1,t} - \hat{p}_{2,t})^2]$ ,  $\bar{f} = (\frac{1}{T}) \sum_{t=1}^T f_t$ , 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  $W_{AUC}$ :

$$W_{AUC} = \frac{(AUC_1 - AUC_2)^2}{\mathbb{V}(AUC_1 - AUC_2)} \xrightarrow{d} \chi^2(1), \quad (1.16)$$

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.

#### 1.4.4. 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<sup>18</sup>. Given that dynamic structures can improve predictability in previous literature, I investigate the predictability both in static and dynamic multivariate forecasting models.

#### *1.4.4.1. Univariate models*

Table 1.5 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

---

<sup>18</sup> 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 5<sup>th</sup> contribute less than 5% each.

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 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<sup>19</sup>. With the identification methodology of bear markets used in this paper (the stock price is on a downward trend for at least 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.

#### *1.4.4.2. 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, 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).

---

<sup>19</sup> 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.

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 1.6, 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 1.3 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 1.4 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 1.3, 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 1.6, 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  $I_{1t-12}$  and  $I_{2t-12}$ , 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:

$$X_{t-1} = (PC1_{t-1}, PC2_{t-1}, PC3_{t-1}). \quad (1.17)$$

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

#### 1.4.5. Out-of-sample test

In this section, I evaluate the out-of-sample (with the ex-post revised data)<sup>20</sup> 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<sup>21</sup> 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.

---

<sup>20</sup> The out-of-sample tests using real-time data are investigated in section 6.

<sup>21</sup> The final classification means using full sample information to decide good bear and bad bear market states.

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:

$$\rho_t = P_{t-1}(S_t = 1) + P_{t-1}(S_t = 2). \quad (1.18)$$

The predicted probability,  $\hat{\rho}_t$ , 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<sup>22</sup> across three stock market states to better understand each model's predictability.

#### *1.4.5.1. Performances of predicting bear markets*

Using sample from January 1967 to December 1976 as the initial estimation period, the out-of-sample forecasting evaluation period spans from January 1977 to December 2013. Table 1.7, 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

---

<sup>22</sup> 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<sup>23</sup>.

Table 1.7, 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<sup>24</sup>. 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 1.5 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.

#### *1.4.5.2. 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  $t$ ,  $\hat{S}_t$ , is computed as follows:

---

<sup>23</sup> 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.

<sup>24</sup> 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.

$$\hat{\rho}_t = P_{t-1}(\widehat{S}_t = 1) + P_{t-1}(\widehat{S}_t = 2) \quad (1.19)$$

$$\widehat{S}_t = \begin{cases} 0 & \text{if } \hat{\rho}_t < C \\ 1 & \text{if } \hat{\rho}_t \geq C \text{ and } P_{t-1}(\widehat{S}_t = 1) \geq P_{t-1}(\widehat{S}_t = 2), \\ 2 & \text{if } \hat{\rho}_t \geq C \text{ and } P_{t-1}(\widehat{S}_t = 1) < P_{t-1}(\widehat{S}_t = 2) \end{cases}$$

where  $C \in [0,1]$  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 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<sup>25</sup> 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*<sup>26</sup> 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 1.8 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

<sup>25</sup> 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%.

<sup>26</sup> According to the accuracy measure, the optimal cut-off satisfies:  $C^* = \arg \max_{C \in [0,1]} J(c)$ ,  $J = Se(c) + Sp(c) - 1$ , 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.

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.

### **1.5 Economic values of the multinomial logit model in predicting bear markets**

In the empirical stock return predictability literature<sup>27</sup>, 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 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

---

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

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 1.9 (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:

$$CER = E[r_p] - A \frac{1}{2} Var[r_p], \quad (1.20)$$

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)<sup>28</sup> model with the corresponding t statistics values to evaluate risk-adjusted returns (abnormal returns).

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 1.6 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 1.6 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.

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

---

<sup>28</sup> 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).

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 1.7 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.

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 1.8 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<sup>29</sup> (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<sup>30</sup> in that the value premium is concentrated in small firms. The lower panel of Figure 1.8 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.

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

---

<sup>29</sup> 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.

<sup>30</sup> Fama and French (1993, 2012) and Israel and Moskowitz (2013) show the value premium is largely concentrated among small stocks.

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.

### **1.6 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<sup>31</sup>, 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

---

<sup>31</sup> 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.

methods<sup>32</sup> 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<sup>33</sup>. 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<sup>34</sup>. 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.

## 1.7 Conclusion

---

<sup>32</sup> 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.

<sup>33</sup> For brevity, the results are not tabulated in the paper but are available upon request.

<sup>34</sup> The static binary logit model is better in classification ability than the dynamic binary logit model and the static multinomial logit model.

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.

## 1.8 References

- 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 comparison 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 predictors. In G. Elliott, C. Granger, and Timmermann, A. 1, (eds), *Handbook of Economic Forecasting* (pp.515-554), Elsevier.

## 1.9 Appendix: 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 50<sup>th</sup> size, 25<sup>th</sup> book-to-market ratio, and 25<sup>th</sup> 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](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.

$$\begin{aligned} \text{HML} = & (s\_bm4\_op1 + s\_bm4\_op2 + s\_bm4\_op3 + s\_bm4\_op4 + b\_bm4\_op1 + b\_bm4\_op2 + \\ & b\_bm4\_op3 + b\_bm4\_op4) * (1/8) - (s\_bm1\_op1 + s\_bm1\_op2 + s\_bm1\_op3 + s\_bm1\_op4 + \\ & b\_bm1\_op1 + b\_bm1\_op2 + b\_bm1\_op3 + b\_bm1\_op4) * (1/8). \end{aligned}$$

$$\begin{aligned} \text{HML\_S} = & s\_bm4\_op1 + s\_bm4\_op2 + s\_bm4\_op3 + s\_bm4\_op4) * (1/4) - (s\_bm1\_op1 + \\ & s\_bm1\_op2 + s\_bm1\_op3 + s\_bm1\_op4) * (1/4). \end{aligned}$$

$$\text{HML\_RMW} = (s\_bm4\_op4 + b\_bm4\_op4) * (1/2) - (s\_bm1\_op1 + b\_bm1\_op1) * (1/2).$$

$$\text{HML\_RMW\_S} = s\_bm4\_op4 - s\_bm1\_op1.$$

Table 1.1 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 1.2 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	
<hr/>							
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	
<hr/>							
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	
<hr/>							
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 for details.

Table 1.3 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	calculated 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 1.3 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 1.4 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 1.3 for details in data descriptions, resource, and transformations.

Table 1.5 In-sample univariate test

Panel A	multinomial logit							Binary logit					
	$\beta_1$			$\beta_2$		$R^2(\%)$	BIC	$\beta$		$R^2(\%)$	BIC		
macroeconomic 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 1.5 In-sample univariate test

technical signal	multinomial logit								Binary logit				
	$\beta 1$			$\beta 2$			$R^2(\%)$	BIC	$\beta$	$R^2(\%)$	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	<b>473.22</b>
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.  $\beta 1$  indicates the predictor's impacts for the probability of good bear markets relative to the probability of bull markets, whereas  $\beta 2$  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- $R^2$  and Schwarz Information Criterion (BIC) of the regression. “\*”, “\*\*”, and “\*\*\*” denote significance at 10%, 5%, and 1% level respectively.

Table 1.6 In-sample multivariate test

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.  $\beta_1$  indicates predictors' impacts for the probability of good bear markets relative to the probability of bull markets, whereas  $\beta_2$  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 1.7 Out-of-sample performances

Panel (a)	QPS	LPS	AUC	cut-off
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)		<b>0.027</b>	<b>0.012</b>	
Model (3) vs. Model (4)		<b>0.019</b>	0.576	
		DM	$W_{AUC}$	
Model (2) vs. Model (3)		0.493	<b>0.041</b>	

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 1.8 Hitting rate (%) across three stock market states

obs (%)	Stock market states				total					
	bull	good bear	bad bear	bear		1977m1-2013m12				
	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	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						

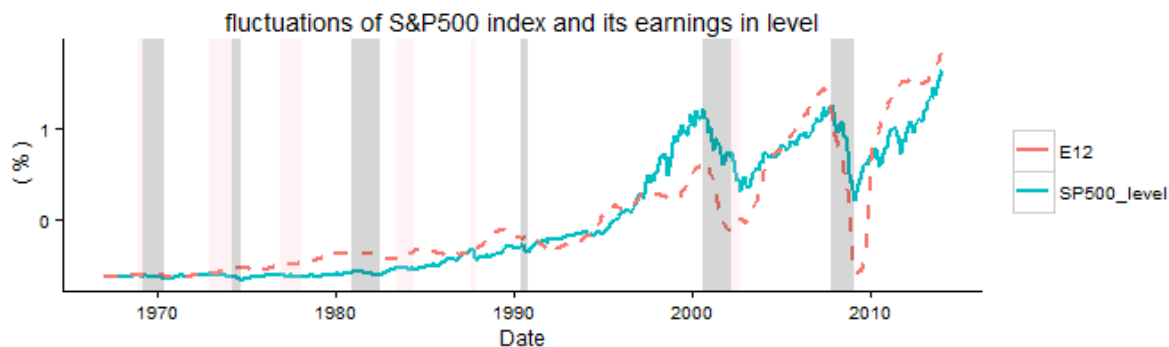
This table shows the hitting rate across three stock market states. The left panel is the result based on the dynamic multinomial logit model, whereas the right panel is the result based on the static binary logit model. The forecasting period is from Jan. 1977 to Dec. 2013. All the values are presented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

Table 1.9 Out-of-sample trading performance

		(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

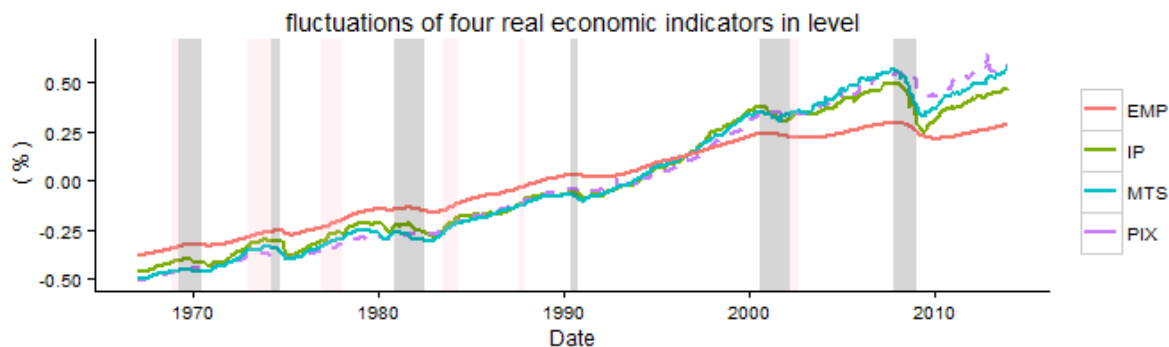
This table reports out-of-sample trading performances based on the dynamic multinomial logit model and the static binary logit model. Column (1) reports the performance under Buy-and-Hold strategy; Columns (2) to (3) report the performances of type 1 trading strategies based on the dynamic multinomial 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 1.1 S&P 500 index and aggregate earnings, Jan. 1967 to Dec. 2013.



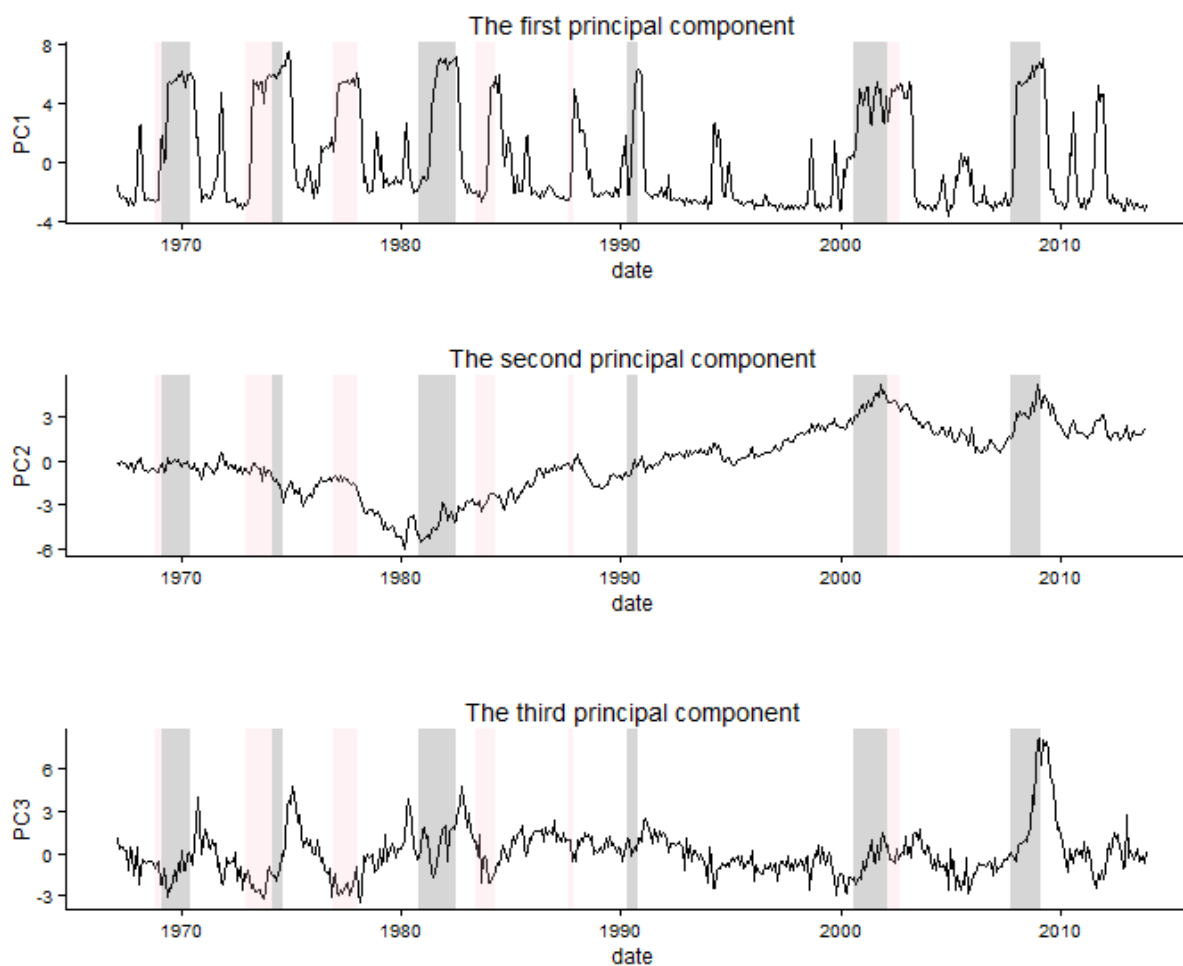
This figure depicts time series of S&P 500 index and its 12-month moving average of earnings. For comparison, both series are standardized. The shaded bars are the stock bear markets generated through Bry-Boschan dating rule algorithm, where pink indicates good bear markets and gray indicates bad bear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

Figure 1.2 Four real economic activity indicators, Jan. 1967 to Dec. 2013.



This figure 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 1.3 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 1.4 Loadings on principal components extracted from 32 variables, Jan 1967 to Dec. 2013.

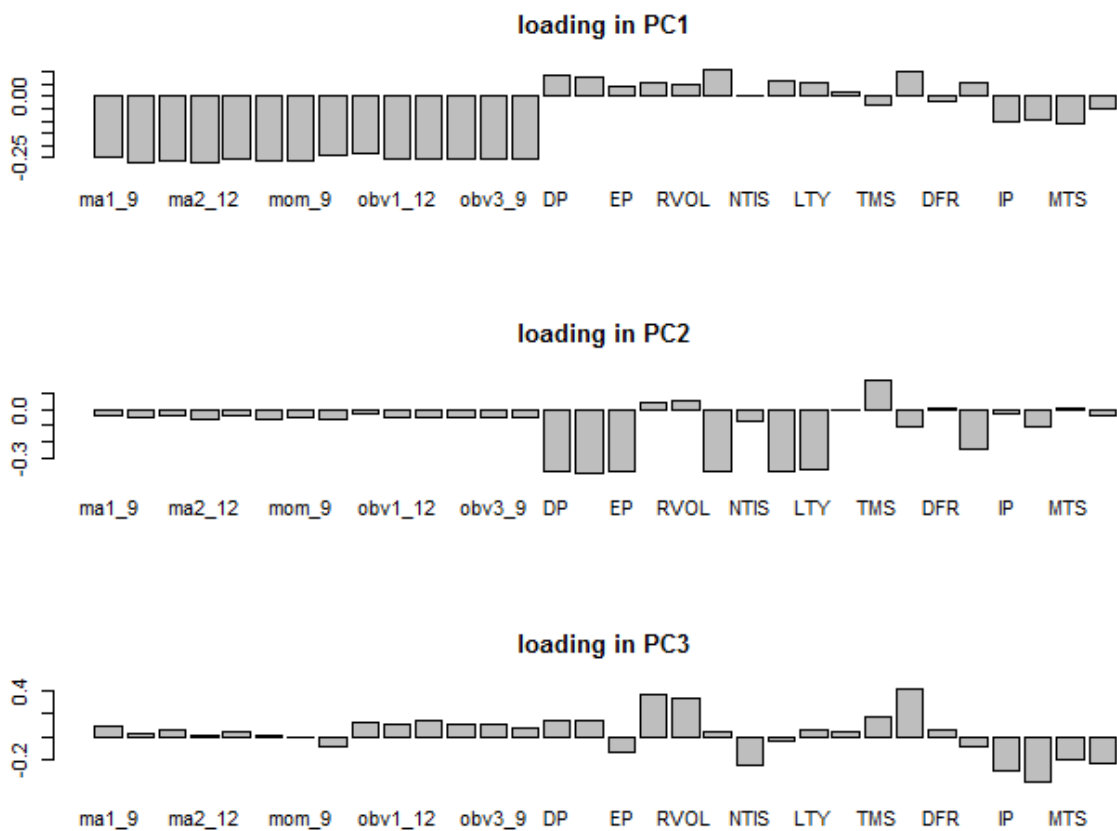
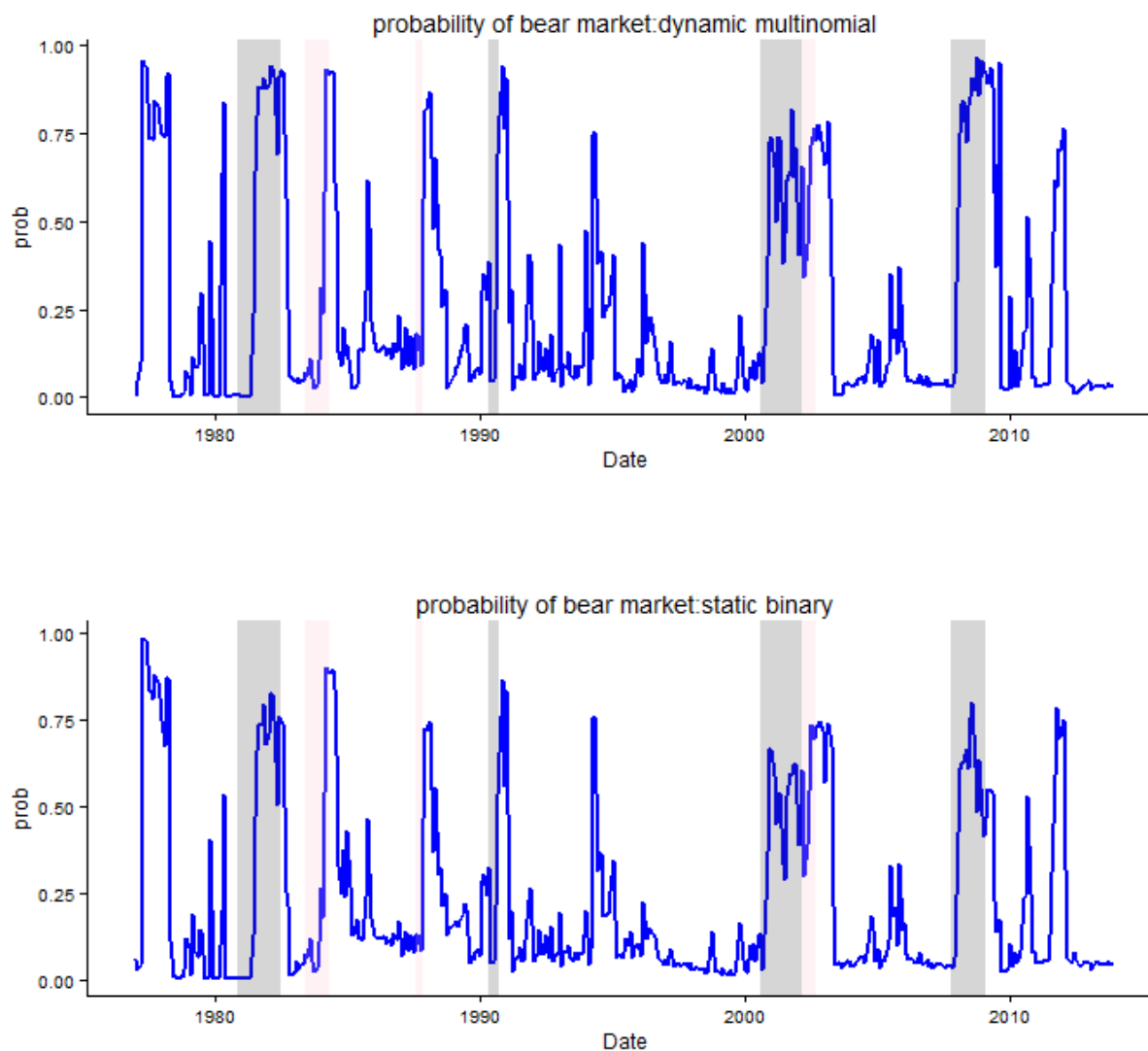
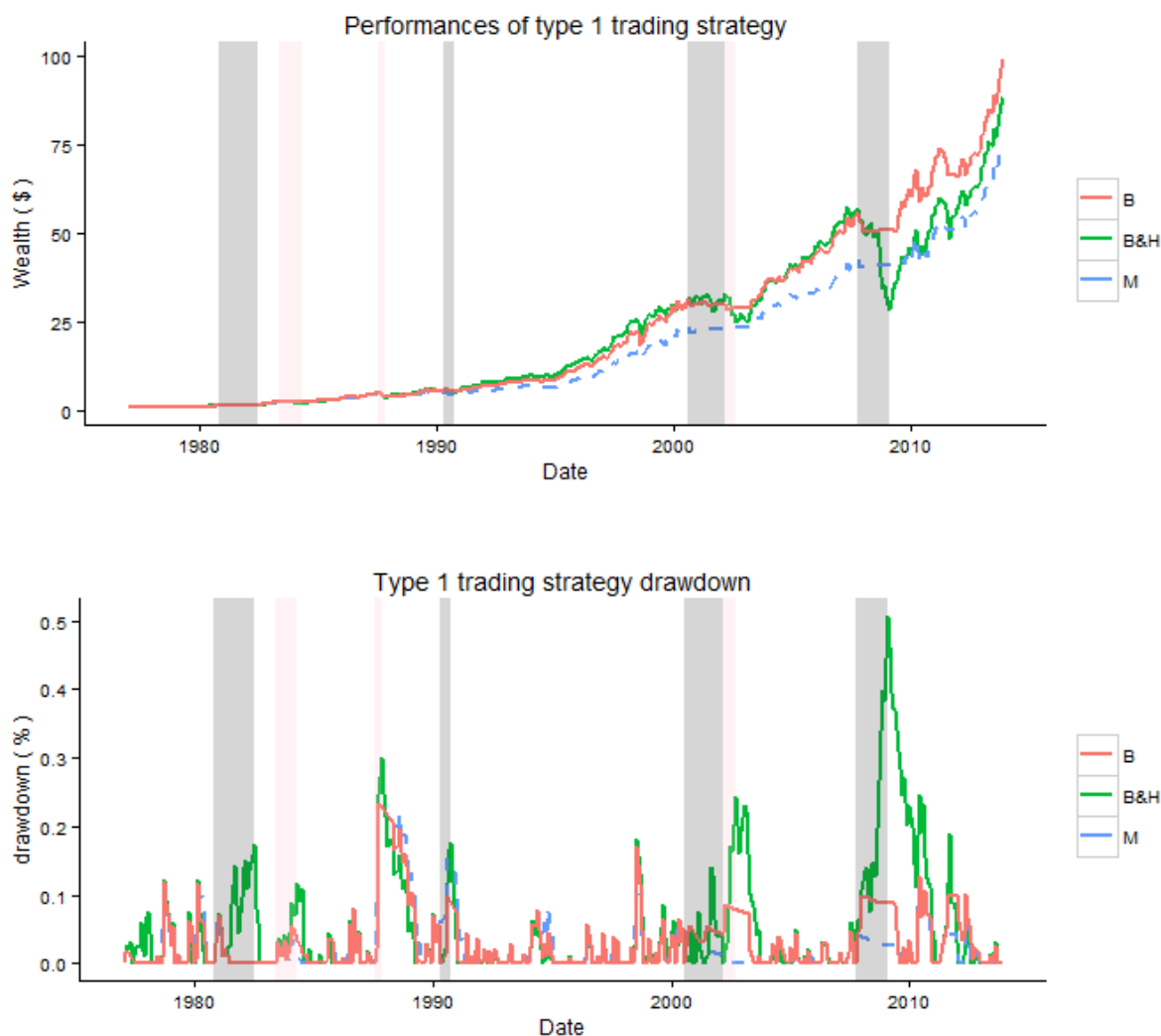


Figure 1.5 Out-of-sample predicted probabilities of bear markets, Jan. 1977 to Dec. 2013.



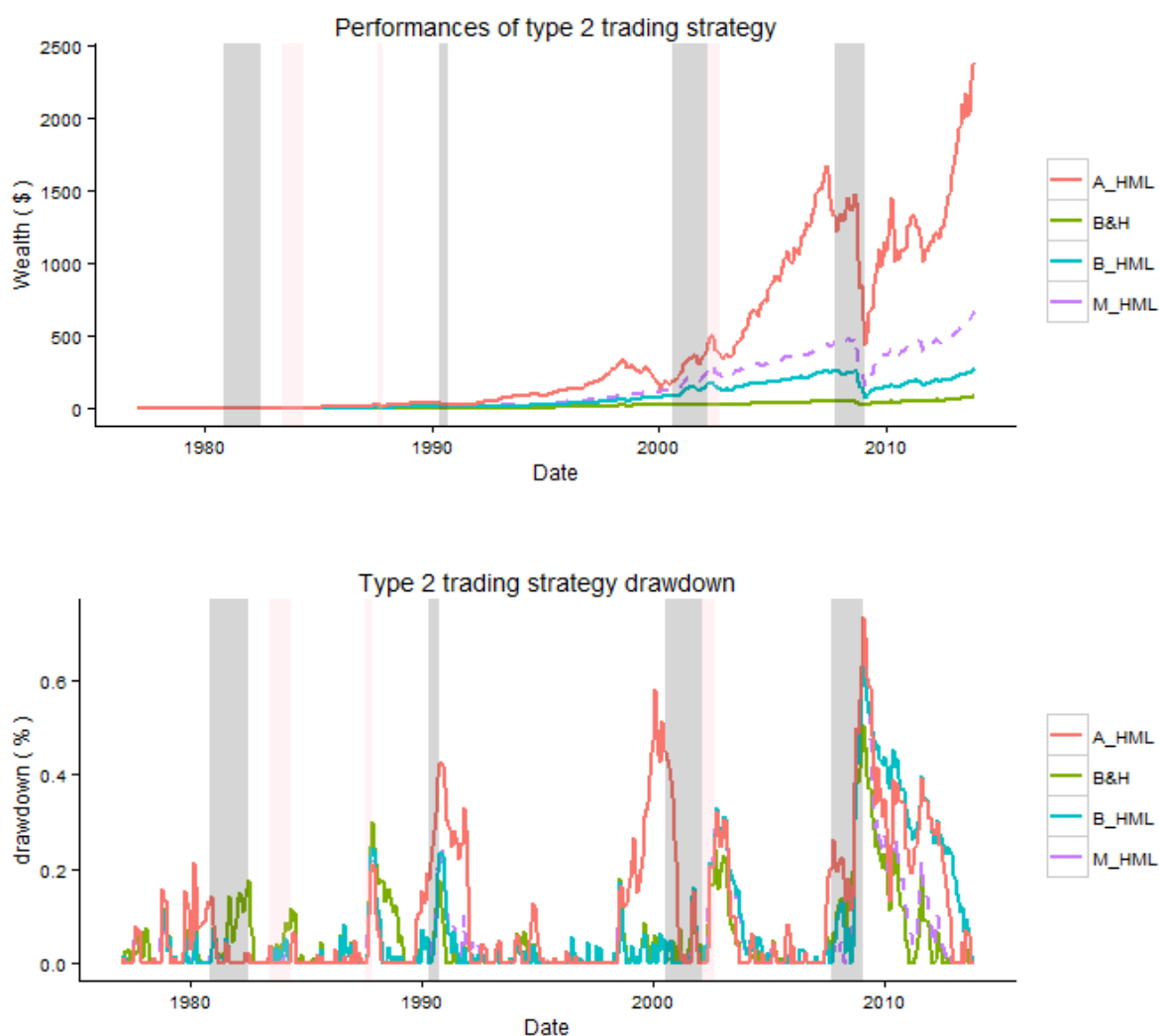
The upper panel presents the out-of-sample predicted probability from the dynamic multinomial logit model, whereas the lower panel presents the out-of-sample predicted probability from the static binary logit model. Predictors are the first three principal components extracted from 32 variables. The 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 1.6 Cumulative wealth for type 1 trading strategies based on the dynamic multinomial logit and the static binary logit model, Jan. 1977 to Dec. 2013.



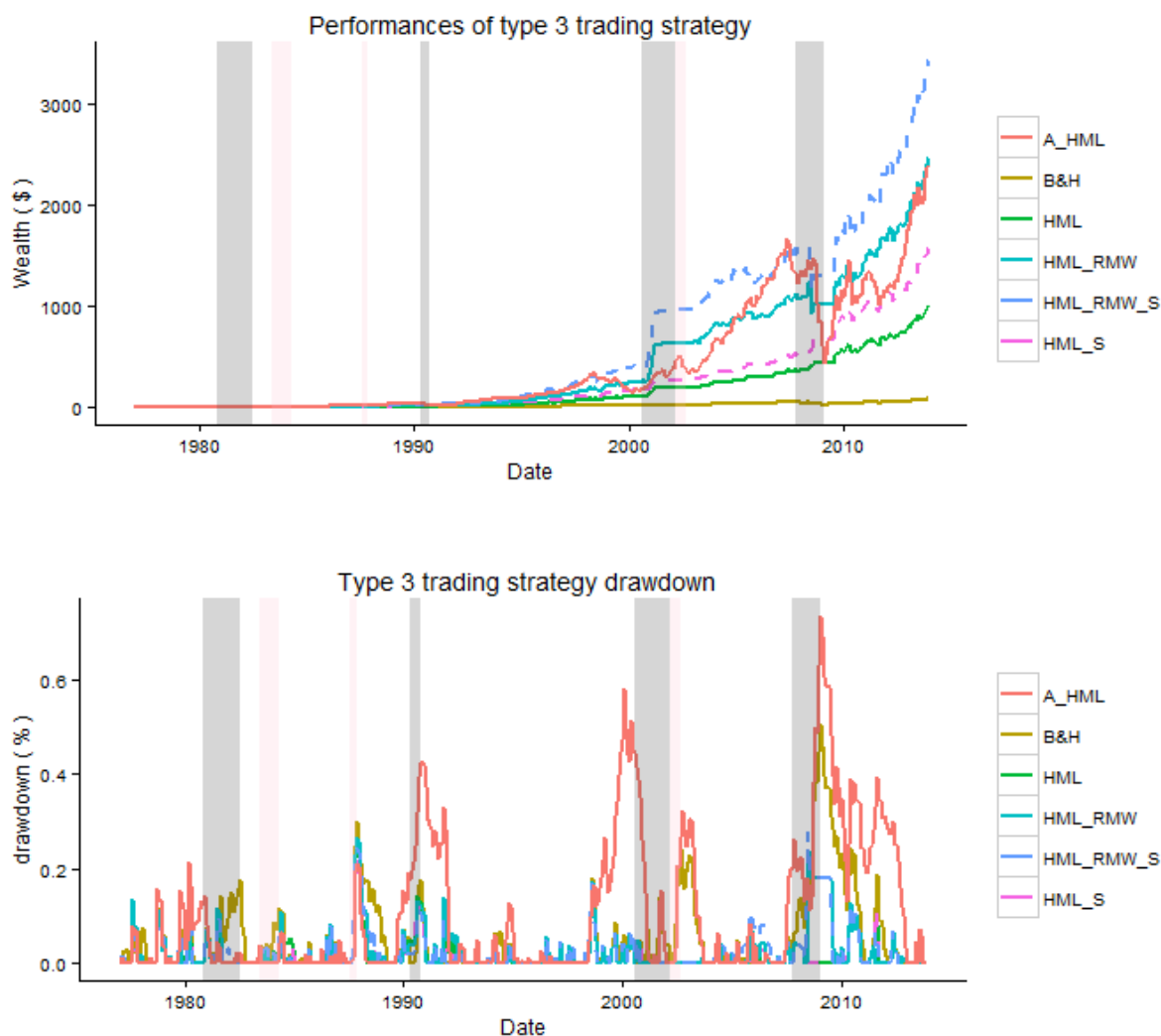
This figure presents out-of-sample performances of two type 1 trading strategies based on dynamic multinomial logit model (M) and static binary logit model (B). The performance of Buy-and-Hold (B&H) benchmark strategy is provided for comparison. The upper panel depicts the cumulative wealth of investing \$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 1.7 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 1.8 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 market.

## Chapter 2

### Predicting recessions with bad bear markets

#### 2.1 Introduction

Aggregate business conditions are of central importance for business and policy communities national-wide. In the U.S., the National Bureau of Economic Research (NBER) establishes a chronology of “turning point” dates at which the shifts between expansion and recession phases occur. However, NBER business cycle peak and trough dates are often determined with a substantial lag. For example, the March 1991 trough was not called until 21 months later, December 1992, and the November bottom of the 2001 recession was not called until December 2008. On the other hand, stock market prices have been thought to be one of the most important leading indicators of economic downturn. 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 prices busts tend to be longer and deeper than other recessions.

However, in U.S. history, there are observations where the declines of the stock market do not precede or coincided with economic contractions. For example, the sharp decline in the stock market in 1962 did little to unsettle the economic recovery. Also, stock market crash of 1987 didn't significantly affect economic activities. Samuelson (1966)'s 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.

Interestingly, instead of classifying stock markets into binary states (i.e., bull and bear markets), in the first chapter, I further distinguishes bear markets into good bear and bad bear markets. As I find, all the real economic indicators deteriorate at bad bear market states, while most of them still expand at good bear market states. Hence, the type of bear market should be critical to determine whether a bear market is an early warning signal of an economic recession. In this study, I investigate the performance of using bad bear markets as signals of recessions declared by NBER. More specifically, I use the classification methods and forecasting models proposed in the first chapter to test whether the bad bear markets can provide reliable and timely information about NBER recessions under in-sample and out-of-sample examinations. The questions I try to answer are as followed:

- 1) Is the bad bear market a more reliable signal of NBER recessions than the bear market under the conventional binary classification (under in-sample test)?
- 2) By predicting peaks and troughs turning points of bad bear markets, can the multinomial logit model with three states (bull, good bear, and bad bear) provide timely information over NBER announcements?
- 3) How does the bad bear market perform comparing to the conventional bear market under binary logit model and the binary logit model that predicts NBER recessions directly (under out-of-sample test)?

The empirical results indicate that the bad bear market is a more reliable signal of a NBER recession than the conventional bear market or even the forecasting model that targets recessions directly. Both peak and trough turning points of bad bear markets can also be identified before NBER announcements, which provides timely information about business cycles. Even though

the turning points of conventional bear markets are closest to the business cycle turning points declared by NBER, conventional bear markets are prone to be false alarms.

The rest of paper is organized as follows. Section 2.2 presents the good bear and bad bear markets identified in chapter 1 and the in-sample relationship between two types of bear markets with NBER recessions. I also present the relationship between conventionally binary classified stock bear markets and NBER recessions. Section 2.3 investigates the effectiveness of using bad bear markets to signal NBER recessions under out-of-sample test. For comparison, I also present the results under the binary classified bear market forecasting model and the results from the model forecasting NBER recessions directly. Section 2.4 further discusses the timeliness of information providing by forecasting models discussed in section 2.3. Specially, the peak and trough turning points of bad bear markets and the time when these turning points are available are compared with business cycle turning points, NBER announcements, and also other forecasting models. Section 2.5 concludes.

## **2.2 Stock markets and business cycles**

Given the controversial role of stock bear markets in predicting NBER recessions in U.S. history, I further classify stock bear markets into good bear and bad bear stocks. The latter are the bear markets associated with contraction phases of future cash flows, while the former are not. In chapter 1, I find the two types of bear markets have distinct interaction with real economy. Here, I focus on whether the stock bad bear market is a more robust signal of NBER recessions. Following the same method<sup>35</sup>, I classify stock market states into bull, good bear, and

---

<sup>35</sup> Please refer chapter 1 for the details in classifying stock markets into three states.

bad bear market states. The sample period is from Jan. 1967 to Dec. 2013. Table 2.1, Panel (a) presents the resulting good bear and bad bear markets and their relationship with NBER recessions; Panel (b) presents the relationship in the case of conventional bear markets (i.e., under binary classification). Column (3) and (4) list all the identified peak and trough dates of bear markets. Column (5) indicates the bear market types. A bear market is recognized as being accompanied with a NBER recession, if it is overlapping or followed by a NBER recession in 6 months. Column (1) and (2) list the peak and trough dates of corresponding NBER recessions if the bear market is accompanied with a NBER recession. Regarding to the timeliness between bear markets and NBER recessions, column (6) to (9) show the leading time between peaks and troughs of bear markets and NBER recessions. Following classifications in chapter 1, there are 6 good bear markets and 6 bad bear markets. Only 1 out of 6 good bear markets is accompanied with NBER recession, while all the bad bear markets are overlapping with NBER recessions. Bad bear markets lead NBER recessions most of the time both in peaks and troughs with similar leading period around 3.4 months. On the other hand, under conventional binary classification, Panel (b), there are 6 out of 9 bear markets are accompanied with NBER recessions. The leading time of peak turning points is 7.1 months on average, while leading time between troughs is 2.3 months on average. This in-sample comparison implies that, by classifying bear markets into good bear and bad bear markets, the bad bear market is a more reliable early warning signal of an economic recession, while the bear market under conventional classification is prone to be a false alarm. Next section, I investigate the bad bear market's effectiveness of signaling NBER recessions in out-of-sample test.

### **2.3 Out-of-sample test of signaling NBER recessions**

With three states in stock markets (bull, good bear, and bad bear), I apply the forecasting model used in chapter 1 to investigate the bad bear markets' robustness in signaling NBER recessions under out-of-sample test, where the sample period is from Jan. 1977 to Dec. 2013. The forecasting model is multinomial logit model with three alternatives. The predictors are the first 3 principal components extracted from 14 macroeconomics variables, 3 real economic indicators, and 14 stock technical indicators.<sup>36</sup> . Because of the rules of dating algorithm and the specific classification method used in chapter 1, the final<sup>37</sup> identification of stock market states can't 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 to estimate model parameters, while the classification might be modified latter with the expansion of forecasting window. To identify stock bear market phases, I follow the approach similar to Chauvet and Piger (2008) with some revisions to suit the multinomial logit model<sup>38</sup>. That is, to identify a good (bad) bear market phase, firstly I require the probability of bear market (sum of probabilities of good bear and bad bear market) move from below to above the optimal threshold value and remain above threshold value for at least three consecutive months. More, these three consecutive months have to be identified as good (bad) bear markets. The peak turning point is the month prior to this good (bad) bear market phase. The trough of this good (bad) bear market phase is then the month prior to three consecutive months being identified as other states. *Youden Index*<sup>39</sup>, which optimally considers both type I and type II errors equivalently, is used to decide a threshold value for each model. For comparison, I also investigate the bear market phases

---

<sup>36</sup> Please refer Hsu (2016) for the details in model specification and the predictors.

<sup>37</sup> The final classification means using full sample information to decide good bear and bad bear market states.

<sup>38</sup> Chauvet and Piger (2008) is aimed to date business cycles. They require the probability of recession move from below to above 80% and remain above 80% for 3 consecutive months before a recession phase is identified. The peak date for this recession is then the month prior to this recession phase, while the trough is the month prior to three consecutive months of probabilities below 20%.

<sup>39</sup> According to Accuracy measure, the optimal cut-off satisfies:  $C^* = \arg \max_{c \in [0,1]} J(c)$ , where  $J(c)$  denotes the Youden Index, defined as  $J = Se(c) + Sp(c) - 1$ .

identified under the conventional binary class model, which is binary logit model. Table 2.2, Panel (a) presents the results, where left panel shows the results under multinomial logit model, and right panel shows that under binary logit model. The identified peak month and trough month of each phase is listed on column (1), (2), (6), and (7). Column (3) indicates the type of bear market. Column (4) and (8) indicate whether an identified bear market is overlapping with a NBER recession. “\*” indicates the bear market is overlapping with a NBER recession. Under multinomial logit model, there are 6 good bear and 4 bad bear market phases. Among good bear markets, only 2 out of 6 are accompanied with NBER recessions, whereas 3 out of 4 bad bear markets are accompanied with NBER recessions. On the other hand, under the binary logit model, only 4 out of 8 identified bear markets are accompanied with NBER recessions. Hence, in Panel (b), under multinomial logit model, given the signal of bad bear market, the probability of occurring NBER recession is 75%, while under the binary logit model, given an identified bear market, the probability is only 50%. This result implies that, even under out-of-sample examination, by classifying bear markets into two types, the bad bear market is a more reliable signal of economic recessions than the traditional identified bear market.

However, in literature, many studies investigate the model that target and forecast NBER recession directly. It would be interesting to compare the reliability of signaling recessions between using bad bear markets or forecasting recessions directly. I therefore use the same predictors and the same forecasting model (static binary logit mode with one period ahead forecasting horizon) with the NBER recession as the dependent variable to conduct an out-of-sample forecasting exercise. Since NBER usually announce recession turning points with a substantial delay, Kauppi and Saikkonen (2008) and Nyberg (2010) assume that the recession indicator is announced with 9 months delay, whereas Chauvet and Potter (2005), Kauppi (2008),

and Ng (2012) advocate the use of a 3 months delay. Here, I take the delay of NBER announcements into consideration by estimating model parameters using data up to time  $t - d - 1$ , where  $d = 0, 3, 6, \text{ and } 9$  and then using data at time  $t - 1$  to construct predictions. Therefore, I estimate 4 scenarios which assume the practitioner's understanding of the states of economy is no late, 3 months, 6 months, or 9 months late. Using the same method of identifying contraction phases, Table 2.3 presents the identified recession phases under 4 scenarios in Panel (a) to (d), respectively. The column "overlap with Rec" indicates whether an identified recession phase is overlapping with a NBER recession. The bottom line shows the probability of a NBER recessions realized given a forecasted recession phase. Consistent with intuition, among 4 scenarios, the assumption of no late intelligence in knowing economy states gives the highest reliability of signaling NBER recessions. The longer the delay, the less reliable the forecast is. Surprisingly, even under the assumption of no late intelligence in knowing economy state, the probability of correctly signaling a NBER recession is only 67% which is lower than that of an identified bad bear market, 75%. However, this only implies a simple static binary logit model which predicts recessions directly can't provide more robust signals than a predicted bad bear market. The comparison with more sophisticated forecasting models proposed in literature is beyond the discussion of this study.

## **2.4 Performances in providing timely information of business cycles**

In section 2.3, I discuss the reliabilities of signals from various forecasting models. In this section, I further investigate the timeliness of these signals under the out-of-sample exercise. Specifically, I consider the peak and trough turning points of predicted contraction phases (i.e., bad bear markets in multinomial logit model; bear markets in binary logit model; recession

phases in binary logit model targeting recessions directly) that successfully predict NBER recessions. Because of the specific rule of identifying a turning point, the dates that a peak or a trough turning point become available are 3 months after the turning points realized. I therefore compare the timing of peak and trough turning points among various forecasting models and the time when they are available with the NBER announcement dates. I focus on the last five NBER recessions where the announcement dates are available. Table 2.4 presents the identified peak turning points of bad bear markets under the multinomial logit model and bear markets under the binary logit model. Column (1) and (2) show the peak dates of bad bear markets, whereas column (3) and (4) show the peak dates of bear markets. Among the last 5 recessions, bad bear markets only signal 3 NBER recessions, while bear markets under the binary logit model signal 4 recessions. About the timeliness, though the peak dates of bad bear markets are 2.7 months on average latter than NBER declared peak dates, these peak turning points are available 2.7 months on average before NBER announcements. For the bear markets under the binary logit model, peak dates are 1.25 months on average ahead of NBER peak dates and are available 6.8 months on average before NBER announcements. This result implies both the peak turning points of bad bear markets under the multinomial logit model and bear markets under the binary logit model are pretty close to the NBER declared peak turning points, and both models can provide early warning signals for the beginning of NBER recessions before NBER announcements.

Turn to the binary logit model that forecast recession directly. Table 2.5 presents the results. With 4 scenarios about practitioners' understanding of economy states, column (1), (3), (5), and (7) report the peak dates of identified recession phases that overlapping with last 5 NBER recessions. Column (2), (4), (6), and (8) report the time when these peak dates are available. On

average, all the identified peak dates under 4 scenarios are ahead of NBER peak dates and also available before NBER peak dates announcements. This implies using the binary logit model to forecast recessions directly can provide early warning signals of NBER recessions under the simulated out-of-sample test. The identified peak dates under the scenario of no lag and 6 months lags in knowing economy states are closet to peak dates declared by NBER. On average, the forecasting model can signal the start of a recession 6.5 to 8.3 months prior to the NBER announcement.

In addition to the peak turning points, knowing the troughs of business cycles in advance is also critical for policy makers and investors<sup>40</sup>. Table 2.6 reports the identified troughs of bad bear markets under the multinomial logit model and of bear markets under the binary logit model. Column (1) and (2) show the trough dates of bad bear markets, whereas column (3) and (4) show the trough dates of bear markets. Under the multinomial logit model, the identified trough of bad bear market is 5 months on average latter than the trough of NBER recession, while the timing of recognizing a trough turning points is 6.3 months on average before a NBER announcement. For the bear market under the binary logit model, the identified trough is 2.8 months on average latter than the trough of a NBER recession, while the timing of identifying a trough points is 10.5 months prior to a NBER announcement. This table implies both bad bear markets (under the multinomial logit model) and bear markets (under the binary logit model) can provide early signals of the ending of NBER recessions before NBER announcements, though the identified trough turning points are behind NBER turning points.

Turning to the binary logit model that forecast recession directly. Table 2.7, column (1), (3), (5), and (7) report the trough dates of identified recession phases overlapping with NBER

---

<sup>40</sup> Siegel (1991) finds about two thirds of the gain in timing business cycle by switching investment portfolios is the result of predicting the end of the recession in advance.

recessions. Column (2), (4), (6), and (8) report the time when these trough dates are available. On average, all the identified trough points under 4 scenarios are behind NBER declared trough dates. Especially, the trough identified under the scenario of no lag intelligence in knowing economy states is latest on average, which is surprising. However, these trough points are still available 4 to 9 months on average ahead of NBER announcements, which means the binary logit models that forecast NBER recessions directly can still provide timely information of the end of NBER recessions.

To summarize the performances of signaling NBER recessions over various forecasting models, for peak turning points, all the models discussed in this chapter (i.e., bad bear markets under the multinomial logit model, bear markets under the binary logit model, and predicted recessions under the binary logit model targeting recession directly) can provide early warning signals of the starts of NBER recessions before NBER official announcements. The peak dates of bear markets identified under binary logit model are on average prior and closest to the NBER declared peak turning points. For trough turning points, even though the trough points identified under all models are on average behind the NBER declared trough points<sup>41</sup>, all these models still show the superiority in timely signaling the ends of recessions over NBER announcements. Among all the models, the trough points of bear markets under the binary logit model are still on average closest to the NBER declared trough turning points. Results from section 2.3 and section 2.4 imply that the bad bear market is a more reliable signal of economic recession but its turning points is not as close to NBER declared turning points as the bear market under the binary logit model. Figure 2.1 and Figure 2.2 present the out-of-sample bear market probabilities and the

---

<sup>41</sup> Actually, except the recession, 2001/03-2001/11, most of the trough points identified under all models are ahead of or very shortly behind of NBER declared trough points.

identified bad bear market phases under the multinomial logit model and bear market phases under the binary logit model respectively. Figure 2.3, Panel (a) to (d) present the out-of-sample recession probabilities and identified recession phases under 4 scenarios.

## 2.5 Conclusion

This study extends the discussion of chapter 1 which distinguishes stock bear markets further into good bear and bad bear markets. In chapter 1, I find the distinct interactions of two types of bear markets with the real economy. This chapter further investigates the reliability and timeliness of using bad bear markets to signal economic recessions and compares the performance with using conventional bear markets as signals under the binary logit model or using the binary logit model to predict recessions directly. With respect to the reliability of signaling NBER recessions, bad bear markets are more reliable with very few false alarms than other forecasting models under either in-sample or out-of-sample tests. As respect to improving timeliness over NBER announcements, on average, all the forecasting models can timely signal both the beginning (peak turning points) and the end (trough turning points) of recessions prior to NBER announcements. Under out-of-sample test, the identified peak and trough turning points of bad bear markets do not occur ahead of turning points declared by NBER, while the turning points of conventional bear markets (under the binary logit model) are preceding most of time and on average closet to the NBER turning points. However, with only 0.67% (in-sample) and 50% (out-of-sample) probability that an alarm is true (true positive), the conventional bear market is actually not a reliable tool to detect an economic recession, which is consistent with

Paul Samuelson's famous word<sup>42</sup> even after 47 years! Overall, both in-sample and out-of-sample results support that bad bear markets are more reliable to signal NBER recessions than the conventional bear markets or the model that forecasts recessions directly. Bad bear markets also provide timely information about business cycles over NBER announcements.

---

<sup>42</sup> "The stock market has predicted nine out of the last five recessions." - Paul Samuelson, 1966.

## 2.6 References

- Barro, R. and Ursua, J. F. (2009). "Stock-market crashes and depressions", *NBER working paper* No. 14760.
- Chauvet, M. and Piger, J. (2008). "A comparison of the real-time performance of business cycle dating methods", *Journal of Business and Economic Statistics*, 26 (1), 42-49.
- Chauvet, M. and Potter, S. (2005). "Forecasting recessions using the yield curve", *Journal of forecasting*, 24, 77-103.
- 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.
- Hsu, W. (2016). "Predicting and Capitalizing on Two Types of Stock Bear Markets in the U.S." *working paper*.
- Samuelson, P. (1996). "Science and Stocks" *Newsweek*, September 19, p.92.
- Kauppi, H. (2008). "Yield curve based probit models for forecasting U.S. recessions: stability and dynamics." *Helsinki Center of Economic Research Discussion Paper*, 221, 1-36.
- Kauppi, H., Saikkonen, P. (2008). "Predicting U.S. recessions with dynamic binary response models." *The Review of Economics and Statistics*, 90 (4), 777-791.
- Ng, E. C. (2013). "Forecasting U.S. recessions with various risk factors and dynamic probit models", *Journal of Macroeconomics*, 34, 112-125.
- Nyberg, H. (2010). "Dynamic probit models and financial variables in recession forecasting", *Journal of forecasting*, 29, 215-230.

Table 2.1 In-sample stock bear market classification and business cycles, Jan. 1967 to Dec. 2013.

In-sample		1967M1- 2013M12						
Panel (a)								
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
NBER recessions		bear market (two types)		type	peak		trough	
Peak	trough	peak	trough		G lead(m)	B lead (m)	G lead(m)	B lead(m)
		Nov-68	Apr-69	G				
Dec-69	Nov-70	4/31/1969	Jun-70	B		8		5
Nov-73	Mar-75	Dec-72	Apr-74	G	11		11	
Nov-73	Mar-75	Apr-74	Sep-74	B		-5		6
		Dec-76	Feb-78	G				
Jul-81	Nov-82	Nov-80	Jul-82	B		8		4
		Jun-83	May-84	G				
		Aug-87	Nov-87	G				
Jul-90	Mar-91	May-90	Oct-90	B		2		5
Mar-01	Nov-01	Aug-00	Apr-02	B		7		-5
		Apr-02	Sep-02	G				
Dec-07	Jun-09	Oct-07	Feb-09	B		2		4
average					11	3.67	11	3.17
Panel(b)								
NBER recessions		bear market (conventional)		peak	trough			
peak	trough	peak	trough	lead (m)	lead (m)			
Dec-69	Nov-70	Nov-68	Jun-70	13	5			
Nov-73	Mar-75	Dec-72	Sep-74	11	6			
		Dec-76	Feb-78					
Jul-81	Nov-82	Nov-80	Jul-82	8	4			
		Jun-83	May-84					
		Aug-87	Nov-87					
Jul-90	Mar-91	May-90	Oct-90	2	5			
Mar-01	Nov-01	Aug-00	Sep-02	7	-10			
Dec-07	Jun-09	Oct-07	Feb-09	2	4			
average				7.17	2.33			

This table presents the in-sample relationship between stock bear market and NBER declared recessions. Panel (a) presents the resulting good bear and bad bear markets under Hsu (2016) classification and their timing relationship with NBER recessions; Panel (b) presents the relationship in the case of conventional bear markets (i.e., under binary classification).

Table 2.2 Out-of-sample identified bear markets phases and their relations with NBER declared recessions, Jan. 1977 to Dec. 2013.

Panel (a)									
multinomial logit model					binary logit model				
	(1)	(2)	(3)	(4)		(6)	(7)	(8)	
	bear market					bear market			
No.	peak	trough	type	overlap Rec	No.	peak	trough	overlap Rec	
1	Mar-77	Apr-78	G		1	Mar-77	Apr-78		
2	Jun-81	Apr-82	G	*	2	Jun-81	Aug-82	*	
3	Apr-82	Sep-82	B	*	3	Feb-84	Jan-85		
4	Feb-84	Jul-84	G		4	Oct-87	May-88		
5	Nov-87	Jul-88	G		5	Aug-90	Jan-91	*	
6	Mar-94	Aug-94	G		6	Oct-00	Apr-03	*	
7	Oct-00	Mar-01	G	*	7	Jan-08	May-09	*	
8	Mar-01	Apr-03	B	*	8	Aug-11	Jan-12		
9	Dec-07	Jul-09	B	*					
10	Sep-11	Jan-12	B						
Panel (b)									
			B	G					
		prob	75%	33%			prob	50%	

Panel (a), left panel shows the results under multinomial logit model, and right panel shows the results under binary logit model. The identified peak month and trough month of each phase is listed on column (1), (2), (6), and (7). Column (3) indicates the type of bear market. Column (4) and (8) indicate whether an identified bear market is overlapping with a NBER recession. Entry “\*” indicates the bear market is overlapping with a NBER recession. Panel (b) shows the probability of overlapping with a NBER recession, given a bear market phase identified under each model.

Table 2.3: Out-of-sample identified recession phases under binary logit model and their relation with NBER declared recessions, Jan. 1977 to Dec. 2013.

No.	Panel (a)			Panel (b)			Panel (c)			Panel (d)		
	lag=0		overlap	lag=3		overlap	lag=6		overlap	lag=9		overlap
	peak	trough	with Rec	peak	trough	with Rec	peak	trough	with Rec	peak	trough	with Rec
1	Nov-77	Feb-78		May-77	Feb-78		Feb-77	Mar-78		Feb-77	Mar-78	
2	Feb-80	Aug-80	*	May-80	Sep-80	*	Apr-81	Oct-82	*	Apr-81	Nov-82	*
3	Apr-81	Mar-83	*	Dec-80	Oct-82	*	Jan-84	Nov-84		Jan-84	Nov-84	
4	May-84	Aug-84		Feb-84	Feb-85		Aug-90	Dec-90	*	Oct-00	May-03	*
5	Oct-00	May-03	*	Oct-00	May-03	*	Oct-00	May-03	*	Dec-07	Nov-08	*
6	Dec-07	Oct-09	*	Dec-07	Jul-09	*	Dec-07	Jul-09	*	Sep-11	Jan-12	
7				Sep-11	Jan-12		Jul-10	Oct-10				
8							Sep-11	Jan-12				
prob			67%			57%			50%			50%

Panel (a) to (d) present the identified recession phases under 4 scenarios, which assume no delay, 3 months, 6 months, or 9 months delay in knowing economy states under the binary logit model that predict NBER recession with one period ahead forecasting horizon. Column ‘overlap with Rec’ indicates whether the identified recession phase is overlapping with a NBER recession. Entry “\*” indicates a NBER recession is realized. The bottom line shows the probability of a NBER recessions realized given a forecasted recession phase.

Table 2.4 Out-of-sample performances in timeliness (peak) of multinomial logit model and binary logit model over the last 5 NBER recessions, Jan. 1977- Dec. 2013.

NBER recession	NBER announcement peak	multinomial logit model		binary logit model		multinomial logit model		binary logit model	
		(1) bad bear peak date	(2) peak date available	(3) bear peak date	(4) peak date available	lead NBER peak dates (m)	lead NBER announcements (m)	lead NBER peak dates (m)	lead NBER announcements (m)
Jan-80	20-Mar-00	fail		fail					
Aug-81	6-Jan-82	Apr-82	Jul-82	Jun-81	Sep-81	-8	-6	2	4
Jul-90	25-Apr-91	fail		Aug-90	Nov-90			-1	5
Mar-01	26-Nov-01	Mar-01	Jun-01	Oct-00	Jan-01	0	5	5	10
Dec-07	1-Dec-08	Dec-07	Mar-08	Jan-08	Apr-08	0	9	-1	8
average						-2.67	2.67	1.25	6.75

This table presents the identified peak turning points of bad bear markets under multinomial logit model and bear markets under binary logit model. Column (1) and (2) show the peak dates of bad bear markets, whereas column (3) and (4) show the peak dates of bear markets.

Table 2.5 Out-of-sample performances in timeliness (peak) of binary logit model that predicts recessions directly over the last 5 NBER recessions, Jan. 1977- Dec. 2013.

NBER peak	NBER announcement	(1) lag=0 peak	(2) available	(3) lag=3 peak	(4) available	(5) lag=6 peak	(6) available	(7) lag=9 peak	(8) available
Jan-80	3-Jun-80	Feb-80	May-80	May-80	Aug-80				
Aug-81	6-Jan-82	Apr-81	Jul-81	Dec-80	Mar-81	Apr-81	Jul-81	Apr-81	Jul-81
Jul-90	25-Apr-91					Aug-90	Nov-90		
Mar-01	26-Nov-01	Oct-00	Jan-01	Oct-00	Jan-01	Oct-00	Jan-01	Oct-00	Jan-01
Dec-07	1-Dec-08	Dec-07	Mar-08	Dec-07	Mar-08	Dec-07	Mar-08	Dec-07	Mar-08
NBER peak	NBER announcement	lead NBER peak	lead NBER announcements	lead NBER peak	lead NBER announcements	lead NBER peak	lead NBER announcements	lead NBER peak	lead NBER announcements
Jan-80	3-Jun-80	-1	1	-4	-2				
Aug-81	6-Jan-82	3	6	7	10	3	6	3	6
Jul-90	25-Apr-91					-1	5		
Mar-01	26-Nov-01	5	10	5	10	5	10	5	10
Dec-07	1-Dec-08	0	9	0	9	0	9	0	9
	average	1.75	6.5	2	6.75	1.75	7.5	2.67	8.3

With 4 scenarios about practitioners' understanding of economy states, column (1), (3), (5), and (7) report the peak dates of identified recession phases that overlapping with last 5 NBER recessions. Column (2), (4), (6), and (8) report the time when these peak dates are available. The last bottom line shows the average leading period (month) between the peak dates of model identified recessions and the peak dates declared by NBER, as well as the leading periods (month) between the time when model identified peak days are available and NBER announcements.

Table 2.6 Out-of-sample performances in timeliness (trough) of multinomial logit model and binary logit model over the last 5 NBER recessions, Jan. 1977- Dec. 2013.

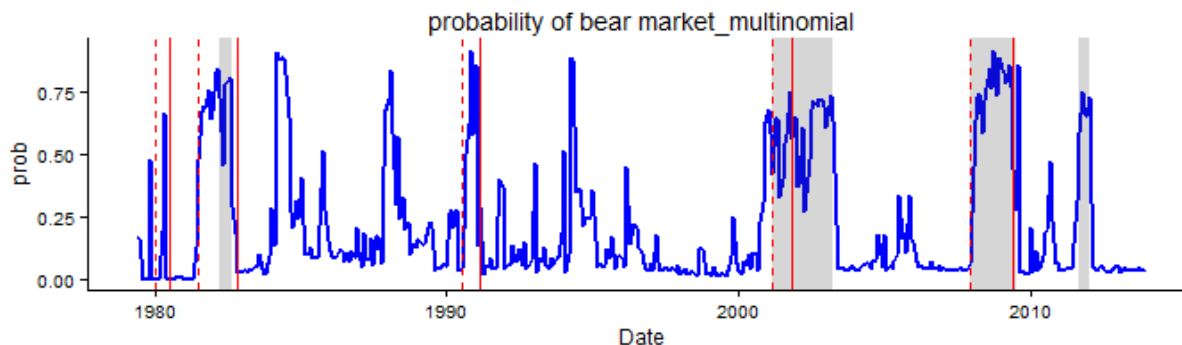
NBER recession	NBER announcement	multinomial logit model		binary logit model		multinomial logit model		binary logit model	
		(1)	(2)	(3)	(4)	lead NBER trough dates (m)	lead NBER announcements (m)	lead NBER trough dates (m)	lead NBER announcements (m)
trough	trough	trough	available	trough	available				
Jul-80	8-Jul-81								
Nov-82	8-Jul-83	Sep-82	Dec-82	Aug-82	Nov-82	2	7	3	8
Mar-91	22-Dec-92	fail		Jan-91	Apr-91			2	20
Nov-01	17-Jul-03	Apr-03	Jul-03	Apr-03	Jul-03	-17	0	-17	0
Jun-09	20-Oct-10	Jul-09	Oct-09	May-09	Aug-09	-1	12	1	14
					average	-5.33	6.33	-2.75	10.5

This table presents the identified trough turning points of bad bear markets under multinomial logit model and bear markets under binary logit model. Column (1) and (2) show the trough dates of bad bear markets, whereas column (3) and (4) show the trough dates of bear markets.

Table 2.7: Out-of-sample performances in timeliness (trough) of binary logit model that predicts recessions directly over the last 5 NBER recessions, Jan. 1977- Dec. 2013.

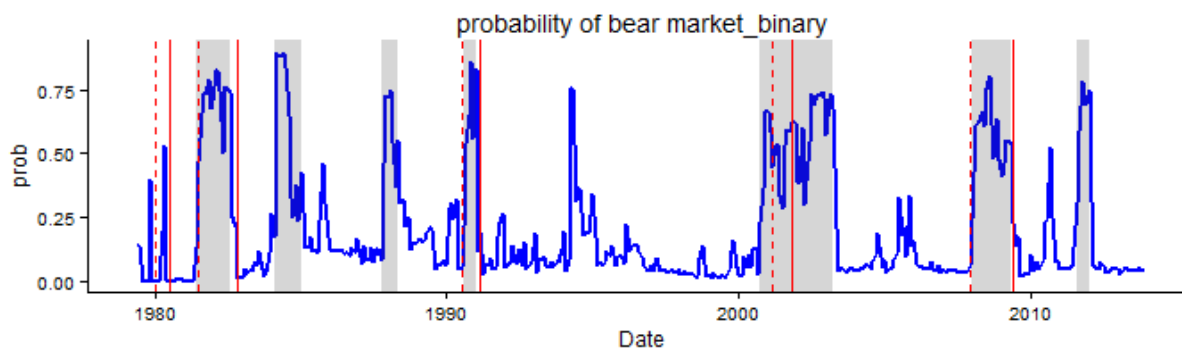
NBER trough	NBER announcement	(1) lag=0 trough	(2) available	(3) lag=3 trough	(4) available	(5) lag=6 trough	(6) available	(7) lag=9 trough	(8) available
Jul-80	8-Jul-81	Aug-80	Nov-80	Sep-80	Dec-80				
Nov-82	8-Jul-83	Mar-83	Jun-83	Oct-82	Jan-83	Oct-82	Jan-83	Nov-82	Feb-83
Mar-91	22-Dec-92					Dec-90	Mar-91		
Nov-01	17-Jul-03	May-03	Aug-03	May-03	Aug-03	May-03	Aug-03	May-03	Aug-03
Jun-09	20-Oct-10	Oct-09	Jan-10	Jul-09	Oct-09	Jul-09	Oct-09	Nov-08	Feb-09
NBER trough	NBER announcement	lead NBER trough	lead NBER announcement	lead NBER trough	lead NBER announcement	lead NBER trough	lead NBER announcement	lead NBER trough	lead NBER announcement
Jul-80	8-Jul-81	-4	8	-2	7				
Nov-82	8-Jul-83	-7	1	1	6	1	6	0	5
Mar-91	22-Dec-92					3	21		
Nov-01	17-Jul-03	-18	-1	-18	-1	-18	-1	-18	-1
Jun-09	20-Oct-10	-4	9	-1	12	-1	12	7	20
	average	-8.25	4.25	-5	6	-3.75	9.5	-3.67	8

Figure 2.1 Out-of-sample bear market probabilities and the identified bad bear market phases under the multinomial logit model



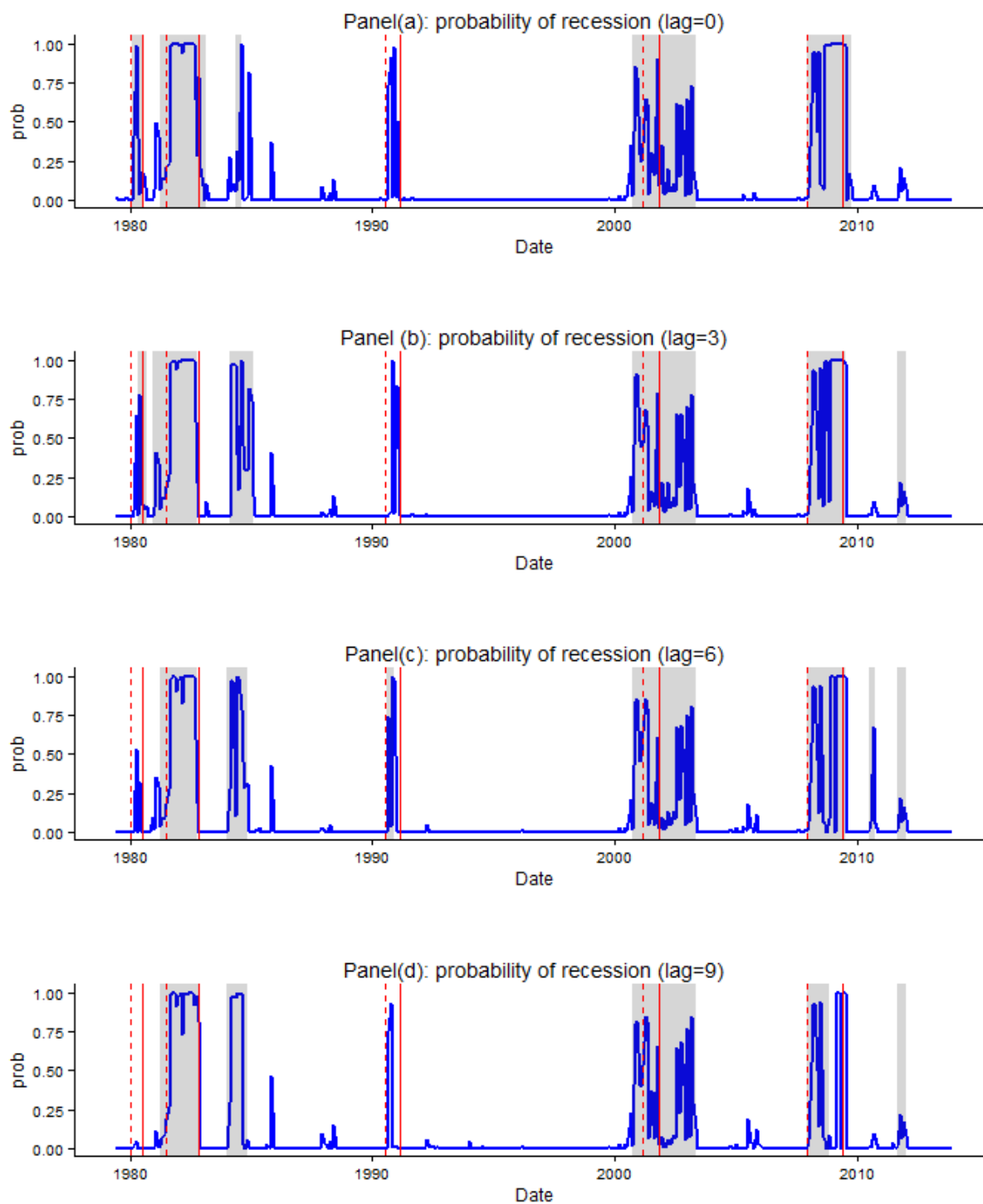
Blue line represents the out-of-sample bear market probabilities under multinomial logit model, from Jan. 1977 to Dec. 2013. The gray shaded bars indicate identified bad bear market phases. The red dash and solid vertical lines indicate the peak and trough turning points of last 5 NBER recessions respectively.

Figure 2.2 Out-of-sample bear market probabilities and identified bear market phases under the binary logit model



Blue line represents the out-of-sample bear market probabilities under the binary logit model, from Jan. 1977 to Dec. 2013. The gray shaded bars indicate identified bear market phases. The red dash and solid vertical lines indicate the peak and trough turning points of last 5 NBER recessions respectively.

Figure 2.3 Out-of-sample bear market probabilities and identified bear market phases under the multinomial logit model with 4 scenarios



Panel (a) to (d) represents the out-of-sample results under 4 different scenarios about the intelligence in knowing economy states. Blue line represents the out-of-sample recession probabilities under the binary logit model, from

Jan. 1977 to Dec. 2013. The gray shaded bars indicate identified recession phases. The red dash and solid vertical lines indicate the peak and trough turning points of last 5 NBER recessions respectively.

## Chapter 3

### Revisit the “Too Much Finance” in dynamic panel data analysis

#### 3.1 Introduction

The question of whether financial development leads to economic growth has invoked a large literature. Levine (2005) provides an extensive survey of the theoretical literature that describes how the service provided by the financial sector contributes to economic growth by: (1) producing ex-ante information about investment opportunities; (2) improving ex-post monitoring of investments and exerting corporate governances; (3) facilitating risk managements and diversifications; (4) mobilizing and pooling savings; and (5) easing the exchange of goods and services. Empirical studies that support a positive nexus between financial development and economic growth are ample. King and Levine (1993), pooling 77 countries during 1960-1989 periods, show that the degree of financial development helps to explain long-run growth. Levine and Zervo (1998) also support this evidence. Rousseau and Wachtel (2000, 2002), Beck, Levine, and Loayza (2000) along with Levine, Loayza, and Beck (2000) also find that measures directed toward financial sector developments do have a significant positive causal effect on growth in panel VAR estimates.

On the contrary, theories also show financial development can hurt growth. Specifically, financial development, by enhancing resources allocation and hence the returns to savings, may lower saving rates. If there are sufficiently large externalities associated with saving and investments, then financial development slows long-run growth<sup>43</sup>. Ram (1999) argues that the results pertaining to the finance-growth nexus are, at best, uncertain and ambiguous. Considering

---

<sup>43</sup> See Beck and Levine (2004)

both panel and time-series data for 8 Asian economics, Zhang (2003) shows that there was a significantly negative connection between banking development and economic growth during 1960-1999 period. Khan and Senhadji (2003) similarly demonstrate that certain banking development indicators become statistically insignificant when growth equations are estimated through the use of panels.

Despite the large literature in this area, the causal relation between financial development and economic growth is absent of a consensus of opinion. Khan and Senhadji (2003) explain the conflicting empirical results by the possible nonlinear causality between banking development and growth. Shen and Lee (2006) claim the relationship between financial development and growth may not be linear, but rather dependent on the economic conditions. They find the relationship between growth and bank development is better described as a weak inverse U-shape. By using different methods, Arcand, Berkes, and Panizza (2015) also suggest a nonlinear causality between financial depths and economic growth. They find there is a threshold on the positive effect of financial development on economic growth. Financial depth starts having negative effect on output growth when credits to the private sector reach 100% of GDP.

Empirically, studies in this literature use econometric techniques, such as cross-country, time-series, panel data, and firm-level studies to investigate the finance-growth nexus. King and Levine (1993), Levine (1997, 2003), and Levine and Zervos (1998) use econometric models in cross-country or time-series analysis. However, to get a causal relationship and control the unobserved heterogeneity effects across countries and time, panel data analysis are used in recent studies. The dynamic structure in modeling economic growth is standard in growth literature.

Previous studies also find economic growth can contribute to financial development<sup>44</sup>. Hence, a dynamic panel data model with a lagged dependent variable and other explanatory variables which may be affected by lagged feedbacks from the dependent variables is the model we want to examine. However, Nickell (1981) points out that the Fixed-Effect estimator in the dynamic panel model is biased and inconsistent for a dataset with large cross-section number ( $N$ ) and small time periods ( $T$ ). As  $T$  gets larger, the Fixed-Effect estimator becomes consistent. In finance-growth nexus literature, many studies use 3-year or 5-year spells to control the business cycle fluctuations and hence have relative short time periods with large cross-section country observations. To control the endogeneity from dynamic structure, feedback effects, and unobservable country effects, the difference GMM estimator - developed by Arellano and Bond (1991), and the system GMM estimator - developed by Arellano and Bover (1995), Blundell and Bond (1998) are used most often.

Blundell and Bond (1998) show that the difference GMM estimator can have very poor finite- sample properties in terms of bias and precision when the series is persistent (e.g. output or financial data) as the instruments are then weak predictors of endogenous changes<sup>45</sup>. They propose the use of extra moment conditions that rely on certain stationary conditions of the initial observation, as suggested by Arellano and Bover (1995). The resulting system GMM estimator has been shown to have better finite sample properties in terms of bias and root mean squared error (RMSE) than the difference GMM estimator in their study.

Rousseau and Wachtel (2000) make an important contribution to the literature by using panel data techniques, the difference GMM estimator, to access the relationship between stock

---

<sup>44</sup> Calderon and Liu (2003) find the two-way direction of granger causality between financial development and economic growth coexist. Law and Habibullah (2007) find real income per capita is a statistically significant deterrent of banking sector development and capital market development.

<sup>45</sup> When the series is very persistent, the lagged level is a weak instrument for its subsequent changes.

markets, banks, and economic growth. They find both banking and stock market developments explain subsequent growth. Beck and Levine (2004) use system GMM estimator and show that a system panel estimator that simultaneously uses both the difference panel data and the data from the original levels specification produces dramatic increases in both consistency and efficiency. They use data covering 40 countries with 5-year spell between 1976 and 1998 and find stock markets and banking development positively influence economic growth.

Though the difference and system GMM estimators can potentially correct the dynamic bias (Nickell, 1981), the proliferation of instruments can cause weak instrument problems that generate spurious findings in applications. Roodman (2009) shows the large instrument collection can over-fit endogenous variables even as it weakens the Hansen test of the instruments' joint validity. Though the system GMM estimator improve the estimation in both consistency and efficiency, Bun and Windmeijer (2010) and Bazzi and Clemens (2013) show the weak instrument bias can still exist in system GMM estimator in a dynamic panel data model. Specifically, the country-level output and financial data are characterized by highly persistent. The variance of country fixed effects is furthermore often expected to be high relative to the variance of idiosyncratic shocks. These characteristics combined may extend the weak instrument problem to the system GMM estimator and cause the bias increases with the ratio of country fixed effects variances to the idiosyncratic shock variances.

Moreover, to investigate the effect of financial development to economic growth, many studies add certain common controls to the regression, such as the square term of financial development, subsets of financial development, or product terms of financial development with other control variables, which results in high correlations among the regressors. Roodman (2008) points out this nearly multicollinearity can further amplify the weak instrument problems in the

GMM estimators and the estimation can be sensitive to the outliers. So, the presence of nearly multicollinearity is a sign of strong idiosyncrasies which could be mistaken for general patterns and is contradict to the purpose of using panel data analysis.

In this chapter, I reexamine the relationship between financial development and economic growth by using the same dataset of Arcand, Berkes, and Panizza (2015) but focus on the analysis in the dynamic panel data framework. Specifically, I use the system GMM estimator but carefully examine the weak instrument problems. I conduct a battery robustness tests that use the full instrument set, collapsed (full) instrument set, or restricted lags as instruments and examine their validity and strength. I further use the Half-Panel Jackknife Fixed-Effect estimator as an alternative method which has better performances in a set of Monte Carlo simulations, particularly in the case where the feedback effects and  $T$  is moderate. Last, being aware of the effects of outliers, I perform a series of outlier detection as robustness tests.

My empirical result doesn't show sufficient evidences to support a positive causal effect from financial development to economic growth, nor does it show a significant negative quadric causal effect. Instead, in some cases, I find non-robust negative effects of financial development on growth. This chapter contributes to the literature by addressing the potential weak instrument problems in GMM estimators ignored by many previous studies. I further confirm my findings by a novel method, Half-Panel Jackknife Fixed-Effect estimator, which can avoid the weak instrument issue and correct the endogenous bias from the dynamic panel model at the same time. I also take care the possible impacts from the outliers and conduct a series of robustness tests.

The rest of this chapter is as follows: Section 3.2 setups the empirical model, data, and econometric methods. Here, I use the system GMM estimators with multiple setting of

instrument sets to examine the finance-growth nexus in a dynamic panel data framework. Section 3.3 introduces an alternative method, Half-Panel Jackknife Fixed-Effect estimator. To compare the performances of different estimators in the finite samples, I conduct a set of Monte Carlo simulations with multiple scenarios. Section 3.4 contains a discussion of empirical findings. Section 3.5 conducts robustness tests for outliers. Section 3.6 concludes.

## 3.2 Empirical model, data, and methodology

### 3.2.1 Empirical model and data

The empirical model is based on Arcand, Berkes, and Panizza (2015) but we focus on the part of panel data analysis, which is similar to the model of Beck and Levine (2004). Following is the baseline model to examine the linkages between finance and growth:

$$g_{i,t} = \alpha_y y_{t-1} + X'_{i,t-1} \beta + \varphi_i + \eta_t + \varepsilon_{i,t}, \quad (3.1)$$

where  $g_{i,t}$  is 5-year average real per capita GDP growth (GROWTH),  $y_{i,t-1}$  is the log of initial GDP per capita (LGDP),  $X_{i,t-1}$  is a vector of growth determinants including the financial development measure of interests and other control variables at the beginning of 5-year spell;  $\varphi_i$  is an unobservable county fixed effect;  $\eta_t$  is a time fixed effect;  $\varepsilon_{i,t}$  is an idiosyncratic shock; and  $i$  and  $t$  represent county and time period (5-year), respectively.  $X_{i,t-1}$  are the beginning values at the 5-year spell which include: credits to the private sector (PC) which is the total credits to the private sector extended by deposit banks and other financial institutions over GDP; the log of average years of education (LEDU); the log of government consumption over GDP (LGC); the log of trade openness (LOPEN); and the log of inflation (LINF)<sup>46</sup>. The sample period is from

---

<sup>46</sup> The data is from Arcand, Berkes, and Panizza (2015). I thank authors for providing the data.

1961 to 2010. Arcand, Berkes, and Panizza (2015) find the “vanishing effects” of financial development is the result of its nonlinear effects. They add the square term of credits to the private sector (PC2) in the regression and find a negative effect on output growth when credits to the private sector reach 100% of GDP. Table 3.1 is the summary statistics of dataset. In this study, I first reexamine the regression without the square term (Eq.(3.1)) and then add the square of credits to the private sector (PC2) into the regression.

### 3.2.2 System GMM estimation

Due to the inclusion of lagged dependent variable (log of initial GDP per capita) on the right hand-side of Eq. (3.1), our model is a dynamic panel data model. More, the growth determinants could be affected by lagged feedback effects from the dependent variables, and so are weakly exogenous (or predetermined)<sup>47</sup>. I use the system GMM estimator to correct the biases from the lagged dependent variable (initial GDP per capita) and other weakly exogenous explanatory variables. Under the assumptions: (1) the error term,  $\varepsilon$ , is not serial correlated, and (2) the explanatory variables,  $X_{t-1}$ , are weakly exogenous, Arellano and Bond transform Eq. (3.1) into first-difference and propose the following moment conditions:

$$E[y_{i,t-s}\Delta\varepsilon_{i,t}] = 0 \text{ for } s \geq 2; \quad t = 3, \dots, T, \quad (3.2)$$

$$E[X_{i,t-s}\Delta\varepsilon_{i,t}] = 0 \text{ for } s \geq 2; \quad t = 3, \dots, T, \quad (3.3)$$

Alonso-Borrego and Arellano (1999) and Blundell and Bond (1998) show in the case of persistent explanatory variables, lagged levels are weak instruments for the difference equation.

In small samples, their Monte Carlo experiments show that the weak instruments can produce

---

<sup>47</sup> For example, Calderon and Liu (2003) find economic growth can positively affect financial developments. Law and Habibullah (2007) find real income per capita is a statistically significant positive deterrent of banking sector development and capital market developments. Beck and Levine (2004) use the same specification for the control variable  $X$ .

biased coefficients. Hence, I use an estimator that combines in a system the regression in differences with the regression in levels (Arellano and Bover, 1995, and Blundell and Bond, 1998). Specifically, this popular exploit an additional set of moment conditions:

$$E[(\varphi_i + \varepsilon_{i,t})\Delta y_{i,t-s}] = 0 \text{ for } s = 1; \quad t = 3, \dots, T, \quad (3.4)$$

$$E[(\varphi_i + \varepsilon_{i,t})\Delta X_{i,t-s}] = 0 \text{ for } s = 1; \quad t = 3, \dots, T. \quad (3.5)$$

These moment conditions are valid under joint mean stationary of the  $y_{i,t}$  and  $X_{i,t}$  processes. In theory, these moment conditions offer a credible identification strategy. Although generally the system GMM estimator is thought to be more robust than the difference GMM estimator, several recent studies (Hayakawa, 2009; Bun and Windmeijer, 2010; Bazzi and Clemens, 2013) show that this system estimator can also suffer from serious weak instrument biases.

### *3.2.3 Instrument sets and diagnostic tests for system GMM estimators*

To extend Arcand, Berkes, and Panizza (2015) and being aware of the prevalent weak instrument problem in the system GMM estimator in many dynamic panel data studies, I particularly try 3 different sets of instruments and conduct a variety of diagnostic tests to examine the validity and strength of instruments. In Arcand, Berkes, and Panizza (2015), they include all the available lags as instruments. Since the number of instruments increases quadratically with the sample period, except the full lags instrument set, I also consider the collapsed instrument set suggested by Roodman (2009), or follow Levine et al. (2000), using only one period lagging as instruments (restricted instrument set). I estimate regressions without and with the square of credits to the private sector (PC2) with all the specifications and diagnostics.

To test the validity of system GMM estimators, I first perform the Arellano-Bond AR (2) test. Since lagged repressors are used as instruments, an unbiased estimation requires the absence of second-order serial correlation in the error term (see Arellano and Bond, 1991). A p-value greater than 0.05 implies the absence of second-order autocorrelation. Second, to test the validity of the exclusion restrictions, I perform the Hansen  $J$ -test. Due to the system GMM, I also report the second test of the exclusion restrictions known as the difference-in-Hansen test which checks the validity of the additional exclusions restrictions that arise from the level equations of the system GMM model (Roodman, 2009).

To examine the strength of instruments, I conduct diagnostics for underidentification and weak instruments tests. To test for underidentification, I report p-values for a test of the null hypothesis that the structural equation is underidentified based on a Lagrange-Multiplier (LM) test using the rank-based  $rk$  statistic based on Kleibergen and Paap (2006). A rejection of the null implies the smallest canonical correlation between the endogenous variables and the instruments is nonzero. However, GMM estimators can still suffer from weak instrument biases even when the underidentification test is rejected, which makes it tenuous to test any meaningful hypothesis. Hence, I also report first-stage F statistics, Wald statistics based on Kleibergen and Paap (2006), which is generalized to non-independently and non-identically distributed errors. I report two-stage robust standard errors corrected for finite sample (Windmeijer, 2005) throughout.

### **3.3 Alternative estimation**

#### *3.3.1 Half-Panel Jackknife Fixed-Effect estimation*

As Roodman (2009) shows the system GMM estimations can have numerous instruments that results in overfitting the model and potentially leads to biased estimates. As advised in

Baltagi (2010) and Hsiao and Zhang (2015), reducing the number of instruments used is facing the trade-off between the reduction in biases and loss of efficiency. Up to date, to the best of my knowledge, there is no general guidance on the optimal number of instruments. Hence, I consider an alternative method, the Half-Panel Jackknife Fixed-Effect (HPJFE) estimator (Chudik, Pesaran, and Yang, 2016), which is particularly suitable for a dynamic panel model with weak-exogenous variables as regressors in their paper. Chudik et al. (2016) show that after the half-panel Jackknife procedure, the Nickell bias of order  $T^{-1}$  of the Fixed-Effect estimator can be effectively reduced to order  $T^{-2}$ , and a valid inference can be done when  $N/T^3 \rightarrow 0$ , as  $N, T \rightarrow \infty$  jointly. However, these conditions also mean the HPJFE estimator requires moderate  $T$  for a large  $N$  dataset. Estimation with too short  $T$  with large  $N$  dataset may result in a biased estimate. Assume the following dynamic model:

$$y_{i,t} = X'_{i,t}b + e_{i,t} = c + \lambda y_{i,t-1} + x'_{i,t}\beta + e_{i,t}, \quad (3.6)$$

subtracting the *sample* averages from equation (6) yields

$$\widetilde{y}_{i,t} = y_{i,t} - \bar{y}_i = \widetilde{X}'_{i,t}\widetilde{b} + \widetilde{e}_{i,t} = \lambda_y(y_{i,t-1} - \bar{y}_{i,-1}) + (x_{i,t} - \bar{x}_i)'\beta + (e_{i,t} - \bar{e}_i), \quad (3.7)$$

for  $i=1,2,\dots, N$  and  $t=1,\dots, T$ , where  $\widetilde{X}_{i,t} = (y_{i,t-1} - \bar{y}_{i,-1}, x'_{i,t} - \bar{x}_i)'$ ,  $\widetilde{b} = (\lambda_y, \beta)'$ .

The Half-Panel Jackknife Fixed-Effect estimator of  $\widetilde{b}$ <sup>48</sup>:

$$\widehat{\widetilde{b}}_{HPJFE} = 2\widehat{\widetilde{b}}_{FE} - \frac{1}{2}(\widehat{\widetilde{b}}_{a,FE} + \widehat{\widetilde{b}}_{b,FE}), \quad (3.8)$$

---

<sup>48</sup> Please see Chudik, Pesaran, and Yang (2016) for details in derivation and the asymptotic distribution of the HPJFE. They also provide a consistent estimator of its variance.

where  $\widehat{b}_{FE}$  is the Fixed-Effect estimator of Eq. (3.8) using the full sample of  $T$  time periods, and  $\widehat{b}_{a,FE}$  and  $\widehat{b}_{b,FE}$  are the Fixed-Effect estimators using the first  $T/2$  and the last  $T/2$  observations, respectively.

Before conducting the empirical estimations, to have a clear understanding of the finite sample properties of different estimators, I conduct a set of Monte Carlo simulations in the next section.

### 3.3.2 Monte Carlo simulations

To address the potential problems raising from the dynamic nature of economic growth and the feedback effects between current financial development and past economic growth, I setup a dynamic panel model with one weak exogenous regressor. I examine the performances of following estimators: (1) difference GMM; (2) system GMM; and (3) Half-Panel Jackknife Fixed-Effect (HPJFE); and (4) Fixed-Effect.

The data generating functions are as followed:

$$y_{i,t} = c_i + \lambda_y y_{i,t-1} + \beta x_{i,t} + u_{i,t} \quad (3.9)$$

$$= (1 - \lambda_y) \mu_{y,i} + \lambda_y y_{i,t-1} + (1 - \lambda_y) \delta x_{i,t} + u_{i,t},$$

$$x_{i,t} = (1 - \lambda_x) \mu_{x,i} + \lambda_x x_{i,t-1} + (1 - \lambda_x) k y_{i,t-1} + v_{i,t}, \quad (3.10)$$

for  $i=1,2,\dots,N$  and  $t= 1, 2, \dots,(T+100)$ , with the initial value  $y_1 = x_1 = 0$ . The first 100 observations ( $t=1,\dots, 100$ ) for  $x_{i,t}$  and  $y_{i,t}$  are discarded. The fixed effects and the idiosyncratic errors are generated as:

$$\mu_{xi} \sim IIDN(1,1), \quad \mu_{yi} = \mu_{xi} + \eta_{i,y}, \quad \eta_{yi} \sim IIDN(1,1),$$

$$\begin{aligned}
v_{i,t} &\sim \text{IIDN}(0, \sigma_{vi}^2), & \sigma_{vi}^2 &= \omega + 0.25\eta_{vi}^2, & \eta_{vi}^2 &\sim \text{IID } \chi^2(2), \\
u_{i,t} &\sim \text{IIDN}(0, \sigma_{ui}^2), & \sigma_{ui}^2 &= \omega + 0.25\eta_{ui}^2, & \eta_{ui}^2 &\sim \text{IID } \chi^2(2),
\end{aligned} \tag{3.11}$$

The setup allows the fixed effect in the  $y_{i,t}$  and  $x_{i,t}$  equations to be correlated, which in turn induces correlation between  $\mu_{yi}$  and  $x_{i,t}$ .

Considering the high persistence of economic growth, I set  $\lambda_y = 0.8$ , representing a high degree of persistence in  $y$ . I consider two values for the feedback coefficient, with  $k = 0.1, 0.3$ , representing a low and a medium degree of feedbacks, respectively. Through, I set  $\delta = 0.5$  and  $\lambda_x \in \{0, 0.1, \dots, 0.8\}$ , indicating the different persistence of  $x$  over time. The magnitude of  $\omega$ , which implicitly captures the ratio of the variance in idiosyncratic shock to the variance in country fixed effects, has fundamental effects on instrument strength. While the theoretical apparatus in Blundell and Bond (1998) presumes this variance ratio equal to one (i.e.,  $\omega = 0.5$ ), it's more likely that the ratio is less than one (i.e.,  $\omega < 0.5$ ) for typical applications in the empirical growth literature including the data used in this study<sup>49</sup>. The theoretical channel from relatively low variance in idiosyncratic shocks to the weak instruments is borne out in Bun and Windmeijer (2010). Here, I consider  $\omega = 0.25, 0.5, 4.5$ , indicating the ratio of the variance in idiosyncratic shock to the variance in country fixed effects is slightly lower than, equal to, or far larger than one. The observation is set as  $N = 100$ , and  $T = 10, 15$ . The number of replication is  $R = 1000$ . The parameter of interest is  $\beta = (1 - \lambda_y)\delta$ . Through all simulation sets, the difference and the system GMM estimations are conducted with full instrument sets that using all available lags of internal instruments and treating  $x$  as weakly exogenous (i.e., predetermined).

---

<sup>49</sup> Applying the Fixed-Effect estimation in our empirical baseline model, the ratio of variance in idiosyncratic shocks to the variance in the country fixed effects is 0.292, which is similar to Levine, Loayza, and Beck (2000), which has a ratio equal to 0.2.

The complete performance comparison in terms of biases and root mean square errors (RMSE) over all simulation sets are reported in Table 3.2 and Table 3.3, representing performances based on simulations with low feedback effect ( $k = 0.1$ ) and median feedback effect ( $k = 0.3$ ), respectively. Figure 3.1 shows estimates based on the simulation with low degree of feedback,  $k = 0.1$ , and  $\omega = 0.25, 0.5$  and  $4.5$  in the upper, middle, and lower panel, respectively. The simulation is based on 1000 repetitions where the data size is set as  $N = 100$ , and  $T = 10$ . The horizontal axis shows different assumed degree of persistence of  $x$  (i.e.,  $\lambda_x \in \{0, 0.1, \dots, 0.8\}$ ), and the vertical axis compares the estimated  $\hat{\beta}$  (black solid line) to the true  $\beta$  (red dotted line). The green dashed lines show the average 95% confident interval on  $\hat{\beta}$  across all repetitions. Figure 3.2 shows estimates based on the simulation with the same scenarios but with median feedback,  $k = 0.3$ . Figure 3.3 and Figure 3.4 report the performances of estimates with low and median feedbacks, respectively, but with  $T = 15$ .

In Figure 3.1, comparing the difference GMM estimators and the system GMM estimators, the performances of difference GMM estimators are poor with the biases and imprecisions increasing with the persistence of  $x$ . This is consistent with the finding of Blundell and Bond (1998). Comparing the HPJFE estimators and the Fixed-Effect estimators, the biases of the HPJFE estimators are slightly larger than that of the Fixed-Effect estimators when  $x$  becomes very persistent and having relatively large variances at all degree persistence of  $x$ . Among four estimators, the different GMM estimators apparently perform poorest (both in biases and in imprecision), and the HPJFE estimator and Fixed-Effect estimators have lower biases where the Fixed-Effect estimators have better precisions. Looking at the lower panel of Figure 3.1, when the variance in idiosyncratic shocks is much larger than the variance in country fixed effects, both two GMM estimators decrease their biases comparing to the upper panel (though the

difference GMM estimators are still very imprecise), whereas the HPJFE and Fixed-Effect estimators perform the same.

In Figure 3.2, where the degree of feedback is median ( $k = 0.3$ ), the difference GMM estimators still perform poorly with biases and imprecisions increasing with the persistence of  $x$ . Comparing the HPJFE estimators and Fixed-Effect estimators, different from Figure 3.1, the Fixed-Effect estimator have notably bigger biases in this case, which is consistent with Nickell (1981) that the bias resulting from the feedback effects increases with the degree of feedback.

Looking at Figure 3.3 and Figure 3.4, where  $T = 15$ , the variances of the HPJFE estimators decrease markedly, which is consistent with Chudik et al. (2016) that the HPJFE estimator has good finite sample properties when  $N$  is large and  $T$  is moderate.

Overall, the simulation results suggest the following for the estimation in a dynamic panel model with weak exogenous regressors, where the dataset is composed of large  $N$  and small  $T$  observations: (1) One should never consider the difference GMM estimator if the series is persistent. (2) When the feedback effect is moderate or even strong, the bias in the Fixed-Effect estimator is non-negligible. (3) When the variance in idiosyncratic shocks is smaller than the variance in cross-section fixed effects, the bias in the system GMM estimator becomes larger. (4) The HPJFE estimator has small biases no matter the feedback is weak or strong, but has large variances when  $T$  is very short. (5) When the  $N$  is large and  $T$  is moderate (or larger), the bias and the imprecision in the HPJFE estimator is reduced significantly.

### 3.4 Empirical results

The results of our baseline model is presented in Table 3.4, where the GDP growth (GROWTH) is regressed on beginning values over the 5-year spell of the following regressors: the log of GDP per capita (LGDP), credits to the private sector (PC); the log of average years of education (LEDU); the log of government consumption over GDP (LGC); the log of trade openness (LOPEN); and the log of inflation (LINF). Column (1) and column (2) represent the regressions of Pool OLS and the Fixed-Effect estimations.

In Pool OLS estimation, credits to the private sector (PC) has negative but insignificant effects. Under the Fixed-Effect estimation, the effect of credits to the private sector has significant negative effects. However, the Fixed-Effect estimate can be significantly biased in a dynamic panel model with weak exogenous regressors where the length of the panel  $T$  is small and  $N$  is large (Nickell, 1981) (i.e.,  $N = 133$ ,  $T = 10$  in this dataset.). From Wooldridge (2002), in a dynamic panel model with the lagged dependent variable,  $y$ , and only one weak exogenous variable,  $x$ , the bias from the Fixed-Effects estimator is in opposite sign of the correlation between current  $x$  and past  $y$ . From the literature, past economic growths should have positive effects on the financial development<sup>50</sup>, and therefore the coefficient of credits to the private sector maybe be negatively biased if it is the only weak exogenous regressor (other than the lagged dependent variable). However, in this empirical model, the sign and the magnitude of the bias in the estimated partial effect of credits to the private sector (PC) on economic growth will not only depend on the effect of past economic growth on current credits to the private sector, but also the effect of past economic growth on current values of any other included control variables.

---

<sup>50</sup> Calderon and Liu (2003) find economic growth can positively affect financial developments. Law and Habibullah (2007) find real income per capita is a statistically significant positive deterrent of banking sector development and capital market developments.

To control the endogenous biases from the dynamic panel model with weak exogenous variables, the system GMM estimations with full instruments, collapsed instruments (Roodman, 2009), and restricted instruments (i.e., only one period lagging) are used respectively. The results are presented in column (3), (4), and (5) of Table 3.4. Among the three system GMM estimations, the coefficients of credits to the private sectors are all negative but only the one with collapsed instruments (column (4)) is significant. Further examining the validity and the strength of instruments, the AR (2) tests are rejected in all three system GMM estimations indicating the absence of serial correlation in the error structures. Regarding the Hansen-*J* test for overall exclusion of the instrument set and the difference-in-Hansen test for the exclusion of the additional instruments in the level equations, the extremely high p-value under the full instrument set suggesting the possible overfitting problem. However, using the collapsed instrument set results in the rejects of both Hansen tests, casting doubts on the validity of the results in column (4). Using only one lag as instruments (column (5)), the model passes the two Hansen tests, but fail to reject Kleibergen-Paap (K-P) LM test for underidentification in the levels equation. The two K-P F statistics indicate weak instruments are likely to be an issue for all three system GMM estimators. Considering all diagnostics among the three system GMM estimators, the one with the restricted one period lagging instrument set (column (5)) shows better credibility which finds negative but insignificant effects of credits to the private sector on economic growth. For other control variables, the log of average years of education (LEDU) has positive impacts on economic growth in most of estimations; the log of government consumption over GDP (LGC) has negative impacts; the log of trade openness (LOPEN) has positive impacts; and the log of inflation (LINF) has negative impacts. All these estimates of control variables' coefficients are consistent with previous studies.

Being aware of the prevalence weak instrument problems in the system GMM estimators, I use an alternative method, Half-Panel Jackknife Fixed Effect (HPJFE) estimator, and present the results in column (6). One caveat for this estimator is that the short time sample ( $T = 10$ ) might be too short for the HPJFE to have a precise and unbiased estimate. Indeed, though the HPJFE estimate also get a negative coefficient of credits to the private sector, the coefficients of all regressors are insignificant.

Next, I examine the nonlinear causal relationship of financial development and economic growth by adding the square of credits to the private sector (PC2) in the regression. I estimate this model with the same estimations of Table 4 and present the results in Table 5. Among the three system GMM estimators, the coefficient of credits to the private sectors (PC) is only significantly negative under the collapsed instrument set, and the coefficients of the square of credits to the private sectors (PC2) are all insignificant with mixed signs. Checking the diagnostics for instruments, based on the two Hansen tests, the model with collapsed instruments still fails the exclusion restriction and hence is not reliable. Based on the two K-P LM tests, the estimation using only one lagging period (column (5)) is still more credible among the three system GMM estimators though it fails the underidentification test in levels equation and still suffer the potential weak instruments problems (i.e., the K-P F statistics are very small in both differences and levels equations.). Under this instrument set, the effect of credits to the private sector in level is positive whereas the effect in square term is negative, but none of them is significant which may be due to the fact of reducing efficiency by using less instruments. Column (6) reports the result of HPJFE estimator. The signs of estimated coefficients for the level and the square terms of credits to the private sector are the same with that of the system GMM estimator with restricted instruments, and again none of them is significant.

Overall, none of estimations from our baseline model and the model including the square of credits to the private sector provides sufficient evidences to support the positive causal effect from financial development to economic growth. Instead, in some cases, the significant negative effect of credits to the private sector in the level term is found, though the estimation itself might be biased due to the weak instrument issue. The nonlinear effect between finance-growth nexus is not supported either. In next section, I examine the robustness of this result. Specifically, I check the sensitivity to the outliers, use 5-year and 3-year spells data. Since the Pool OLS estimator and the Fixed-Effect estimator are biased in the dynamic panel model in theory, Nickell (1981), in the following robustness tests, only system GMM estimators with three different instrument sets and the Half-Panel Jackknife Fixed Effect estimator are used.

### **3.5 Robustness tests**

#### *3.5.1 Outliers*

In previous sections, to estimate the possible nonlinear causality between financial development and economic growth, I add the square of credits to the private sector (PC2) among the regressors. This causes a high correlation between the two explanatory variables (i.e., the correlation between PC and PC2 is 0.95.). As Roodman (2008) points out this nearly multicollinearity can impact not only the OLS estimator but also the GMM estimator by amplifying the weak instrument problems and causing the estimate sensitive to the outliers. To explore whether my results are driven by outliers, I first plot leverage versus residual squared plot in Figure 3.5, based on the Pool OLS regression in Table 3.4, column (1). This is a quick way of checking potential influential observations and outliers at the same time, and I find

countries, Bangladesh (BGD), China (CHN), Liberia (LBR), Iceland (ISL), Saudi Arabia (SAU), and Gabon (GAB), are most influential counties or outliers. Further, to know which country has larger impacts on the coefficients of credits to the private sector in the level and in the square terms, conditional on other control variables, I plot the partial-regressions of credit to the private sector in level (PC) and in square (PC2) terms at Figure 3.6 and Figure 3.7, which are based on the Pool OLS regression in Table 3.5, column (1). The two graphs show countries, Gabon (GAB), China (CHN), United States (USA), Iceland (ISL), appear to be influential. Therefore, I discard these 7 countries: GAB, CHN, ISL, LBR, SAU, IRL, and USA, from the dataset, and re-estimate all the estimations in Table 3.4 and Table 3.5 with the same specifications and report in Table 3.6, where Panel (A) is the estimation for the baseline model and Panel (B) is the estimation for model including the square term of credits to the private sector (PC2).

From the Table 3.6, Panel (A), the effect of credits to the private sector is still negative in all the estimations. Checking the diagnostics for the validity and the strength of instruments, the system GMM estimator with restricted only one period lagging instruments (column (3)) appears to be more credible, though the issue of weak instruments still exists. The HPJFE estimate is negative but insignificant. Table 3.6, Panel (B), among the three system GMM estimates, different from that of Table 3.5, all the estimated effects of credits to the private sector in level are negative but insignificant. As for the effect of the square term, the signs of estimates are mixed and all are insignificant. Under the HPJFE estimation, the signs of the estimates for the level and the square terms of credit to the private sector are the same with those of the table 5, column (6), and are insignificant as well. Based on Table 6, it shows discarding the potential outliers does not help the estimations to obtain a consistent and significant result. Next, I consider another robustness check by using 3-year spell dataset.

### 3.5.2 3-year growth spell

Given the 5-year growth spell only yields 10 observations for each country, I also work with 3-year growth spell which provides 17 observations per country. As it is stated in Chudik et al. (2016) (and is showed in the simulation experiments, section 3.1), the Half-Panel Jackknife Fixed Effect (HPJFE) estimator has better finite sample properties in a dynamic panel model when the sample size,  $T$  is moderate and  $N$  is large. Hence, using 3-year spell dataset should further strengths the consistency and efficiency of HPJFE estimator. Table 3.7 presents the estimation for the baseline model (Panel (A)), and the model including the square of credit to the private sector (Panel (B)), based on 3-year spell dataset. In Table 3.7, Panel (A), for the three system GMM estimations, most of the estimated coefficients have the same sign with that of in Table 3.6. Particularly, all the three system GMM estimators show the significant negative effects of credit to the private sector. However, the weak instrument issue implied by the vary low values of K-P F statistics is still unsolved and casting doubts on the results of all system GMM estimates. For the HPJFE estimates, though the sign of estimated coefficients do not change, the effects of trade openness (LOPEN) and inflation (LINF) become significant, which indicates having longer sample improves the efficiency of HPJFE estimator. From Table 3.7, Panel (B), among the three system GMM estimates, the effect of credits to the private sector in level is still negative and become significant only in the estimation with collapsed instruments. For the square terms (PC2), the results are mixed and none are significant. Last, for the HPJFE estimator, all the sign of estimates stay the same with that of the 5-year spell dataset, excepting the significant effects of the trade openness (LOPEN) and the inflation (LINF).

## 3.6 Conclusion

This chapter revisits the debate of finance-growth nexus. I ask whether financial development has causal effects on economic growth, and whether this relationship is nonlinear. I recognize the potential biases arising from the dynamic nature of economic growth and the feedback effects of economic growth on financial development, which would cause Nickell (1981) type biases in the Fixed-Effect estimator when the sample is a large  $N$  with small  $T$  dynamic panel data. Hence, in addition to the Fixed-Effect estimator, I consider the system GMM estimator (Blundell and Bond, 1998) and a novel method, Half-Panel Jackknife Fixed-effect estimator (HPJFE) (Chudik, Pesaran, and Yang, 2016) to correct the endogenous bias. I also conduct a set of Monte Carlo simulations to have a clear understanding of the finite sample performances between different estimators.

In the empirical part, I use the system GMM estimators with 3 different instrument sets and extensively examine the validity and the strength of the using instruments, which are ignored by many previous studies. In addition to using the HPJFE as an alternative estimator, I also consider the issue of outliers and use 3-year spell sample to increase the sample periods which could improve the consistency and efficiency of the HPJFE estimator. Overall, I don't find sufficient evidences to support a positive causal effects (if it's not negative), nor do I find a quadratic effects of financial development on economic growth.

Several factors might explain our results. First, the persistence of economic growth and financial data, and the relatively high variances of country fixed effects may extend the weak instrument problems to the system GMM estimator, which could bias the estimator to an uncertain direction. Second, the panel data used in this study is relatively short for the HPJFE estimator to have an efficient estimation. Having longer data might improve the precision of HPJFE estimators. On the other hand, the effect of financial development may depend on how it

has been accumulated, its composition, and its structure. The aggregate measure of credits to the private sector might provide too limited information to uncover a clear relationship. Further research is certainly needed to fully understand the link between financial development and economic growth.

### 3.7 References

- Alonso-Borrego, C., and Arellano, M., 1999. Systematically normalized instrumental variable estimation using panel data. *Journal of Business & Economic Statistics*, 17, 36-49.
- Arcand, J. L., Berkes, E., and Panizza, U., 2015. Too much finance? *Journal of Economic Growth*, 20, 105-148.
- Arellano, M., and Bond, S., 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies*, 58 (2), 277-297.
- Arellano, M., and Bover, O., 1995. Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, 68 (1), 29-51.
- Baltagi, B., 2010. *Econometric analysis of panel data*. John Wiley & Sons.
- Bazzi, S., and Clemens, M. A., 2013. Blunt instruments: Avoiding common pitfalls in identifying the causes of Economic growth. *American Economic Journal: Macroeconomics*, 2013, 5 (2), 152-186.
- Beck, T., Levine, R., and Loayza, N., 2000. Finance and sources of growth. *Journal of Financial Economics*, 58(1-2), 261-300.
- Beck, T., and Levine, R., 2004. Stock markets, banks, and growth: panel evidence. *Journal of Banking and Finance*, 28 (3), 423-442.
- Blundell, R., and Bond, S., 1998. Initial conditions and moment restrictions in dynamic panel data methods. *Journal of Econometrics*, 87 (1), 115-143.
- Bun, M. G., and Windmeijer, F., 2010. The weak instrument problem of the system GMM estimator in dynamic panel data models. *Econometrics Journal*, 13 (1), 95-126.
- Calderon, C., and Liu, L., 2003. The direction of causality between financial development and economic growth. *Journal of development economics*, 72, 321-334.
- Chudik, A., Pesaran, M. H., and Yang, J., 2016. Half-Panel Jackknife Fixed Effects Estimation of Panels with Weakly Exogenous Regressors. *Working Paper*, 2016.
- Hayakawa, K., 2009. A simple efficient instrumental variable estimator for panel AR (p) models when Both N and T are large. *Econometric Theory*, 25 (3), 873-890.
- Hsiao and Zhang (2015), IV, GMM or Likelihood approach to estimate dynamic panel models when either N or T or both are large. *Journal of Econometrics*, 187 (1), 312-322.
- Khan, M., and Senhadji, A., 2003. Financial development and Economic growth: an overview. *Journal of African Economics*, 12, 89-110.
- King, R. G., and Levine, R., 1993. Finance and growth: Schumpeter might be right. *The quarterly Journal of Economics*, 108(3), 717-37.

- Kleibergen, F., and Paap, R., 2006. Generalized reduced rank tests using the singular value Decomposition. *Journal of Econometrics*, 133 (1), 97-126.
- Law, S. H., and Habibullah, M. S., 2007. The determinants of financial development: institutions, openness and financial liberalization. *South African Journal of Economics*, 77 (1), 45-58.
- Levine, R., 1997. Financial development and economic growth: Views and agenda. *Journal of Economic Literature*, 35, 688-726.
- Levine, R., 2003. More on finance and growth: more finance, more growth? *Federal Reserve Bank of St. Louis Review*, 85(July), 31-46.
- Levine, R., 2005. Finance and growth: theory and evidence. In Aghion, P. and Durlauf, S. (Eds.), *Handbook of economic growth*, Chap. 12 (Vol. 1, pp. 865-34). Amsterdam: Elsevier
- Levine, R., Loayza, N., and Beck, T., 2000. Financial intermediation and growth: causality and causes. *Journal of Monetary and Economics*, 46(1), 31-77.
- Levine, R., and Zervos, S., 1998. Stock markets, banks, and economic growth. *The American Economic Review*, 88(3), 537-58.
- Nickell, S., 1981. Biases in dynamic models with fixed effects. *Econometrica*, 49, 1417-1426.
- Ram, R., 1999. Financial development and Economic growth: additional evidence. *Journal of Development Studies*, 35, 164-174.
- Roodman, D., 2008. Through the looking glass, and what OLS found there: on growth, foreign aid, and reverse causality, *working paper No. 137*, Center of Global Development.
- Roodman, D., 2009 (a). How to do xtabond2: An introduction to difference and system GMM in Stata, *Stata Journal*, 9 (1), 86-136.
- Roodman, D., 2009. Practitioners' corner: A note on the theme of too many instruments. *Oxford Bulletin of Economics and Statistics*, 71 (1), 135-158.
- Rousseau, P., and Wachtel, P., 2000. Equity markets and growth: Cross-country evidence on timing and outcomes, 1980-1995. *Journal of Banking and Finance*, 24, 1933-1957.
- Rousseau, P., and Wachtel, P., 2002. Inflation thresholds and the finance-growth nexus. *Journal of International money and Finance*, 21(6), 777-793.
- Shen, C. H., and Lee, C. C., 2006. Same financial development yet different economic growth-why? *Journal of Money, Credit, and Banking*, 38(7), 1907-1944.
- Windmeijer, F., 2005. A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of econometrics*, 126 (1), 25-51.
- Zhang, K., 2003. Does financial development promote economic: Economic growth in the East Asia? *China Journal of Finance*, 1, 1-10.

Table 3.1 Summary statistics: 5-year spell panel dataset

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
YEAR	2,420	1,984	14.36	1,961	2,006
PC	1,160	0.371	0.351	0.00477	2.698
LGDP	1,751	7.607	1.587	4.191	11.44
ID	2,420	121.5	69.87	1	242
GROWTH	1,791	1.917	3.793	-24.99	27.29
LGC	1,613	2.680	0.406	0.937	4.334
LOPEN	1,660	4.070	0.642	-0.0216	6.082
PC2	1,160	0.261	0.526	2.27e-05	7.277
LINF	1,848	2.418	1.308	-5.298	6.908
LEDU	1,410	2.133	0.751	0	3.274

## Data description and source

GROWTH	Annual percentage growth rate of GDP at market prices based on constant local currency. Aggregates are based on constant 2000 U.S. dollars. Source: World Bank World Development Indicators (WDI) (2011)
PC	Claims on private sector by deposit money banks and other financial institutions divided by GDP. Source: Beck et al. (November 2010 update) and Beck et al. (2000) when Beck et al. (2010) has missing data (PC2 is the square of PC)
EDUC	Average years of schooling of males and females above 25 years of age (the regressions use the log of this variable). Source: Barro and Lee (2010)
GC	General government final consumption expenditures a percentage of GDP (LGC is the log of GD). Source: WDI (2011)
OPEN	Trade openness (Calculated as exports plus imports divided by GDP) (LOPEN is the log of OPEN). Source: WDI
INFL	Inflation as measured by the consumer price index (annual %). Observations which inflation is less than -10% and then set to zero all the observations for which inflation takes on negative value and apply the inverse hyperbolic sine transformation. Source: WDI 2011

Table 3.2 Monte Carlo Simulation sets:  $k = 0.1$ 

$\omega$	Bias T=10			Bias T=15			RMSE T=10			RMSE T=15		
	0.25	0.5	4.5	0.25	0.5	4.5	0.25	0.5	4.5	0.25	0.5	4.5
$\lambda_x$ difference GMM												
0	-2.03	-2	-1.78	-1.17	-1.13	-0.98	5.02	5.01	4.86	3.57	3.56	3.47
0.1	-2.29	-2.26	-1.96	-1.28	-1.28	-1.09	5.36	5.35	5.13	3.73	3.72	3.62
0.2	-2.6	-2.56	-2.16	-1.41	-1.42	-1.19	5.75	5.72	5.4	3.86	3.9	3.69
0.3	-2.98	-2.93	-2.37	-1.7	-1.7	-1.35	6.19	6.15	5.67	4.1	4.04	3.85
0.4	-3.45	-3.38	-2.59	-1.94	-1.97	-1.45	6.72	6.66	5.95	4.27	4.36	3.91
0.5	-4.04	-3.93	-2.79	-2.26	-2.23	-1.61	7.37	7.27	6.22	4.54	4.6	4.11
0.6	-4.77	-4.6	-2.96	-2.59	-2.55	-1.74	8.18	8.01	6.47	4.85	4.84	4.15
0.7	-5.61	-5.32	-3.02	-3.05	-2.95	-1.74	9.16	8.87	6.67	5.35	5.2	4.21
0.8	-6.27	-5.81	-2.84	-3.28	-3.03	-1.56	10.15	9.67	6.87	5.59	5.42	4.17
system GMM												
0	1.64	1.32	0.09	1.91	1.73	0.42	3.8	3.73	3.59	3.43	3.35	3.05
0.1	1.69	1.37	0.11	1.97	1.68	0.37	3.88	3.8	3.65	3.45	3.34	3.03
0.2	1.73	1.41	0.14	1.98	1.68	0.43	3.93	3.85	3.68	3.51	3.39	3.07
0.3	1.76	1.44	0.17	1.88	1.74	0.39	3.95	3.87	3.67	3.44	3.35	3.04
0.4	1.77	1.45	0.21	1.79	1.63	0.42	3.95	3.86	3.62	3.38	3.39	3.08
0.5	1.77	1.45	0.26	1.77	1.52	0.4	3.9	3.8	3.54	3.34	3.32	3.02
0.6	1.73	1.43	0.3	1.58	1.41	0.31	3.8	3.7	3.43	3.25	3.11	2.92
0.7	1.65	1.35	0.31	1.36	1.25	0.25	3.64	3.53	3.27	3.04	2.98	2.77
0.8	1.47	1.2	0.32	1.19	1	0.26	3.38	3.26	3.05	2.87	2.8	2.54
HPJFE												
0	0.17	0.17	0.17	0.23	0.23	0.23	4.46	4.46	4.46	3.09	3.09	3.09
0.1	0.21	0.21	0.21	0.26	0.26	0.26	4.63	4.63	4.63	3.17	3.17	3.17
0.2	0.27	0.27	0.27	0.3	0.3	0.3	4.78	4.78	4.78	3.23	3.23	3.23
0.3	0.34	0.34	0.34	0.36	0.36	0.36	4.91	4.91	4.91	3.28	3.28	3.28
0.4	0.44	0.44	0.44	0.43	0.43	0.43	5.02	5.02	5.02	3.31	3.31	3.31
0.5	0.6	0.6	0.6	0.52	0.52	0.52	5.11	5.11	5.11	3.33	3.33	3.33
0.6	0.82	0.82	0.82	0.66	0.66	0.66	5.19	5.19	5.19	3.33	3.33	3.33
0.7	1.17	1.17	1.17	0.86	0.86	0.86	5.27	5.27	5.27	3.31	3.31	3.31
0.8	1.7	1.7	1.7	1.17	1.17	1.17	5.38	5.38	5.38	3.31	3.31	3.31
Fixed- Effect												
0	-1.05	-1.05	-1.05	-0.46	-0.46	-0.46	3.49	3.49	3.49	2.74	2.74	2.74
0.1	-1.03	-1.03	-1.03	-0.45	-0.45	-0.45	3.58	3.58	3.58	2.79	2.79	2.79
0.2	-1	-1	-1	-0.42	-0.42	-0.42	3.63	3.63	3.63	2.81	2.81	2.81
0.3	-0.94	-0.94	-0.94	-0.37	-0.37	-0.37	3.66	3.66	3.66	2.82	2.82	2.82
0.4	-0.83	-0.83	-0.83	-0.29	-0.29	-0.29	3.66	3.66	3.66	2.8	2.8	2.8
0.5	-0.67	-0.67	-0.67	-0.17	-0.17	-0.17	3.61	3.61	3.61	2.76	2.76	2.76
0.6	-0.42	-0.42	-0.42	0.02	0.02	0.02	3.53	3.53	3.53	2.69	2.69	2.69
0.7	-0.04	-0.04	-0.04	0.31	0.31	0.31	3.42	3.42	3.42	2.61	2.61	2.61
0.8	0.54	0.54	0.54	0.78	0.78	0.78	3.35	3.35	3.35	2.58	2.58	2.58

Table 3.3 Monte Carlo Simulation sets:  $k = 0.3$ 

$\omega$	Bias T=10			Bias T=15			RMSE T=10			RMSE T=15		
	0.25	0.5	4.5	0.25	0.5	4.5	0.25	0.5	4.5	0.25	0.5	4.5
$\lambda_x$ difference GMM												
0	-2.15	-2.13	-1.92	-1.21	-1.17	-1.05	5.05	5.05	4.91	3.58	3.57	3.49
0.1	-2.5	-2.47	-2.18	-1.38	-1.38	-1.16	5.44	5.43	5.21	3.75	3.78	3.66
0.2	-2.91	-2.88	-2.47	-1.7	-1.64	-1.31	5.89	5.87	5.54	3.95	3.96	3.75
0.3	-3.43	-3.39	-2.79	-1.94	-1.96	-1.59	6.42	6.39	5.88	4.25	4.21	3.9
0.4	-4.07	-4.01	-3.15	-2.35	-2.27	-1.86	7.07	7.01	6.24	4.54	4.47	4.1
0.5	-4.88	-4.78	-3.54	-2.78	-2.67	-2.11	7.9	7.79	6.64	4.87	4.75	4.32
0.6	-5.89	-5.72	-3.92	-3.33	-3.23	-2.29	8.97	8.78	7.06	5.35	5.32	4.47
0.7	-7.07	-6.78	-4.21	-3.93	-3.88	-2.56	10.27	9.96	7.45	5.9	5.88	4.68
0.8	-8.03	-7.56	-4.19	-4.45	-4.21	-2.46	11.56	11.06	7.78	6.48	6.3	4.73
system GMM												
0	1.52	1.26	0.1	1.87	1.66	0.47	3.76	3.71	3.6	3.4	3.32	3.03
0.1	1.58	1.31	0.12	1.88	1.69	0.51	3.83	3.79	3.66	3.4	3.35	3.13
0.2	1.64	1.36	0.14	1.9	1.68	0.5	3.89	3.84	3.69	3.44	3.39	3.13
0.3	1.68	1.4	0.17	1.83	1.67	0.46	3.92	3.86	3.68	3.35	3.36	3.1
0.4	1.71	1.43	0.19	1.77	1.58	0.4	3.91	3.85	3.63	3.37	3.25	2.99
0.5	1.73	1.45	0.24	1.76	1.48	0.36	3.87	3.8	3.55	3.26	3.23	3.04
0.6	1.72	1.44	0.28	1.66	1.4	0.33	3.78	3.69	3.43	3.23	3.09	2.87
0.7	1.68	1.4	0.31	1.47	1.27	0.26	3.63	3.53	3.28	3.1	3	2.83
0.8	1.56	1.29	0.33	1.23	1.06	0.37	3.38	3.27	3.06	2.84	2.78	2.62
HPJFE												
0	0.12	0.12	0.12	0.22	0.22	0.22	4.45	4.45	4.45	3.1	3.1	3.1
0.1	0	0	0	0.19	0.19	0.19	4.62	4.62	4.62	3.17	3.17	3.17
0.2	-0.14	-0.14	-0.14	0.15	0.15	0.15	4.78	4.78	4.78	3.23	3.23	3.23
0.3	-0.28	-0.28	-0.28	0.11	0.11	0.11	4.91	4.91	4.91	3.27	3.27	3.27
0.4	-0.4	-0.4	-0.4	0.07	0.07	0.07	5.02	5.02	5.02	3.29	3.29	3.29
0.5	-0.47	-0.47	-0.47	0.04	0.04	0.04	5.11	5.11	5.11	3.3	3.3	3.3
0.6	-0.41	-0.41	-0.41	0.06	0.06	0.06	5.15	5.15	5.15	3.27	3.27	3.27
0.7	-0.15	-0.15	-0.15	0.16	0.16	0.16	5.15	5.15	5.15	3.22	3.22	3.22
0.8	0.47	0.47	0.47	0.46	0.46	0.46	5.15	5.15	5.15	3.15	3.15	3.15
Fixed- Effect												
0	-1.11	-1.11	-1.11	-0.5	-0.5	-0.5	3.5	3.5	3.5	2.74	2.74	2.74
0.1	-1.27	-1.27	-1.27	-0.61	-0.61	-0.61	3.64	3.64	3.64	2.81	2.81	2.81
0.2	-1.43	-1.43	-1.43	-0.73	-0.73	-0.73	3.77	3.77	3.77	2.87	2.87	2.87
0.3	-1.59	-1.59	-1.59	-0.84	-0.84	-0.84	3.87	3.87	3.87	2.91	2.91	2.91
0.4	-1.7	-1.7	-1.7	-0.94	-0.94	-0.94	3.94	3.94	3.94	2.93	2.93	2.93
0.5	-1.73	-1.73	-1.73	-0.98	-0.98	-0.98	3.95	3.95	3.95	2.91	2.91	2.91
0.6	-1.64	-1.64	-1.64	-0.93	-0.93	-0.93	3.87	3.87	3.87	2.84	2.84	2.84
0.7	-1.33	-1.33	-1.33	-0.71	-0.71	-0.71	3.67	3.67	3.67	2.69	2.69	2.69
0.8	-0.65	-0.65	-0.65	-0.19	-0.19	-0.19	3.39	3.39	3.39	2.48	2.48	2.48

Table 3.4 Dynamic Panel data estimations with 5-year spell dataset: Baseline model

Dep. Growth	Pool OLS	fixed-effect	system GMM			HPJFE
	(1)	(2)	full	collapsed	restricted (lag=1)	(6)
LGDP(t-1)	-0.118 (0.128)	-3.537*** (0.549)	-0.406 (0.313)	0.101 (0.557)	-0.0709 (0.411)	-1.143 (2.799)
PC(t-1)	-0.440 (0.325)	-1.795*** (0.442)	-0.599 (0.636)	-2.606** (1.138)	-1.042 (0.934)	-0.933 (3.258)
LEDU(t-1)	1.288*** (0.277)	-1.801** (0.703)	2.148*** (0.547)	1.828* (0.958)	1.961*** (0.736)	-1.030 (2.056)
LGC(t-1)	-1.132*** (0.242)	-1.763*** (0.468)	-2.074*** (0.616)	-1.512 (1.245)	-3.128*** (0.816)	-2.679 (2.119)
LOPEN(t-1)	0.411** (0.165)	2.426*** (0.517)	1.789*** (0.446)	6.643*** (1.376)	2.626*** (0.602)	2.749 (2.430)
LINF(t-1)	-0.116 (0.0834)	-0.201* (0.106)	-0.281** (0.127)	-0.462* (0.241)	-0.336** (0.143)	-0.477 (0.443)
country FE		Yes	Yes	Yes	Yes	Yes
time FE		Yes	Yes	Yes	Yes	
N.obs	917	917	917	917	917	868
N.country			133	133	133	130
N. iv in system GMM			334	70	118	
AR(1) pval			0	0	0	
AR(2) pval			0.081	0.077	0.066	
Hansen Test Pval			1	0.024	0.381	
Diff In Hansen Test Pval			1	0.004	0.736	
difference eq:						
K-P LM stat.						
pval			0.29	0.0397	0.023	
KP Wald F stat			2.773	1.63	1.827	
levels eq:						
K-P LM stat.						
pval			0.1875	0.0697	0.1875	
KP Wald F stat			1.329	0.543	1.329	

This table reports the results of a set of panel regressions aimed at estimating the causal effect of credits to the private sector on economic growth using 5-year non-overlapping panel data. Column (1) and (2) are the regressions estimated by Pool OLS and Fixed-Effect estimations. Column (3)-(5) are the regressions estimated by the system GMM with full instrument set, collapsed instrument set, and restricted one period lagging instrument set, respectively. Column (6) represents the regression estimated by Half-Panel Jackknife Fixed-Effect (HPJFE) estimation. The set of controls includes time fixed effects, country fixed effects, and the lag of: log of GDP per capita (LGDP); the level of credits to the private sector (PC); the log of average years of education (LEDU); the log of government consumption over GDP (LGC); the log of trade openness (LOPEN); and the log of inflation (LINF). The bottom panel of the table reports the diagnostics for the system GMM estimators. Robust (Windmeijer) standard errors for the system GMM estimators are reported in parentheses. \*\*\* indicates p-value < 0.01, \*\* indicates p-value < 0.05; and \* indicates p-value < 0.1.

Table 3.5 Dynamic Panel data estimation with 5-year spell dataset: Model including the square term

Dep. Growth	Pool OLS	fixed-effect	system GMM			HPJFE
	(1)	(2)	full	collapsed	restricted (lag=1)	(6)
LGDP(t-1)	-0.194 (0.131)	-3.625*** (0.541)	-0.330 (0.316)	0.298 (0.548)	-0.00409 (0.370)	-1.351 (3.327)
PC(t-1)	2.068** (0.899)	-0.532 (1.034)	1.580 (1.880)	-4.725** (2.391)	0.129 (1.688)	3.089 (9.049)
PC2(t-1)	-1.519*** (0.462)	-0.660 (0.469)	-1.186 (0.952)	1.104 (0.883)	-0.530 (0.750)	-2.584 (4.521)
LEDU(t-1)	1.236*** (0.273)	-1.803*** (0.684)	1.969*** (0.715)	1.784* (0.911)	1.725** (0.755)	-1.146 (2.400)
LGC(t-1)	-1.184*** (0.244)	-1.789*** (0.467)	-2.164*** (0.631)	-1.629 (1.411)	-3.248*** (0.853)	-2.800 (2.514)
LOPEN(t-1)	0.368** (0.166)	2.379*** (0.507)	1.314*** (0.448)	6.641*** (1.501)	2.297*** (0.571)	2.700 (2.809)
LINF(t-1)	-0.0809 (0.0816)	-0.192* (0.105)	-0.258** (0.120)	-0.459* (0.248)	-0.282* (0.161)	-0.456 (0.516)
country FE		Yes	Yes	Yes	Yes	Yes
time FE		Yes	Yes	Yes	Yes	
N.obs	917	917	917	917	917	868
N.country			133	133	133	130
N. iv in system GMM			388	80	136	
AR(1) pval			0	0	0	
AR(2) pval			0.113	0.069	0.101	
Hansen Test Pval			1	0.052	0.524	
Diff In Hansen Test Pval			1	0.013	0.877	
difference eq:						
K-P LM stat. pval			0.1609	0.0205	0.0315	
KP Wald F stat			4.735	1.644	1.74	
levels eq:						
K-P LM stat. pval			0.1457	0.0857	0.1457	
KP Wald F stat			1.636	0.415	1.636	

This table reports the results of a set of panel regressions aimed at estimating the causal effect of credits to the private sector on economic growth using 5-year non-overlapping panel data. Column (1) and (2) are the regressions estimated by Pool OLS and Fixed-Effect estimations. Column (3)-(5) are the regressions estimated by the system GMM with full instrument set, collapsed instrument set, and restricted one period lagging instrument set, respectively. Column (6) represents the regression estimated by Half-Panel Jackknife Fixed-Effect (HPJFE) estimation. The set of controls includes time fixed effects, country fixed effects, and the lag of: log of GDP per capita (LGDP); the level of credits to the private sector (PC); the square of credits to the private sector (PC2); the log of average years of education (LEDU); the log of government consumption over GDP (LGC); the log of trade

openness (LOPEN); and the log of inflation (LINF). The bottom panel of the table reports the diagnostics for the system GMM estimators. Robust (Windmeijer) standard errors for the system GMM estimators are reported in parentheses. \*\*\* indicates p-value < 0.01, \*\* indicates p-value < 0.05; and \* indicates p-value < 0.1.

Table 3.6 Dynamic Panel data estimations with 5-year spell dataset excluding the outliers

Dep. Growth	Panel (A)				Panel (B)			
	system GMM			HPJFE	system GMM			HPJFE
	full	collapsed	restricted (lag=1)		full	collapsed	restricted (lag=1)	
(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	
LGDP(t-1)	-0.406 (0.313)	0.101 (0.557)	-0.0709 (0.411)	-1.143 (2.799)	-0.0600 (0.338)	-0.0842 (0.518)	0.385 (0.358)	-1.551 (4.256)
PC(t-1)	-0.599 (0.636)	-2.606** (1.138)	-1.042 (0.934)	-0.933 (3.258)	-0.267 (1.709)	-3.167 (2.522)	-1.230 (2.075)	2.589 (12.02)
PC2(t-1)					-0.576 (0.951)	0.691 (1.231)	-0.384 (1.004)	-2.297 (6.671)
LEDU(t-1)	2.148*** (0.547)	1.828* (0.958)	1.961*** (0.736)	-1.030 (2.056)	1.574** (0.639)	2.320** (0.941)	1.290* (0.756)	-1.057 (2.983)
LGC(t-1)	-2.074*** (0.616)	-1.512 (1.245)	-3.128*** (0.816)	-2.679 (2.119)	-2.038*** (0.694)	-2.555*** (1.100)	-2.793*** (0.788)	-2.605 (3.013)
LOPEN(t-1)	1.789*** (0.446)	6.643*** (1.376)	2.626*** (0.602)	2.749 (2.430)	1.375** (0.550)	4.648*** (0.942)	2.058*** (0.633)	2.428 (3.317)
LINF(t-1)	-0.281** (0.127)	-0.462* (0.241)	-0.336** (0.143)	-0.477 (0.443)	-0.318** (0.141)	-0.633*** (0.234)	-0.292* (0.160)	-0.477 (0.627)
country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
time FE	Yes	Yes	Yes		Yes	Yes	Yes	
N.obs		867	867		867	867	867	820
N.country		126	126		126	126	126	123
N. iv in system GMM					388	80	136	
AR(1) pval	0	0	0		0	0	0	
AR(2) pval	0.058	0.037	0.057		0.053	0.033	0.069	
Hansen Test Pval	1	0.027	0.412		1	0.091	0.685	
Diff In Hansen Test Pval	1	0.017	0.763		1	0.053	0.938	
difference eq: K-P LM stat. pval	0.1212	0.0361	0.0054		0.124	0.0355	0.0057	
KP Wald F stat	6.542	1.496	2.077		5.142	1.493	2.059	
levels eq: K-P LM stat. pval	0.1808	0.2355	0.1801		0.1842	0.096	0.1842	
KP Wald F stat	1.852	0.407	1.852		1.836	0.388	1.836	

This table reports the results of a set of panel regressions aimed at estimating the causal effects of credits to the private sector on economic growth using 5-year non-overlapping panel data excluding countries: GAB, CHN, ISL, LBR, SAU, IRL, and USA. Column (1)-(3) are the regressions estimated by the system GMM with full instrument set, collapsed instrument set, and restricted one period lagging instrument set, respectively. Column (4) represents

the regression estimated by the Half-Panel Jackknife Fixed-Effect (HPJFE) estimation. The set of controls includes time fixed effects, country fixed effects, and the lag of: log of GDP per capita (LGDP); the level of credits to the private sector (PC); the square of credits to the private sector (PC2); the log of average years of education (LEDU); the log of government consumption over GDP (LGC); the log of trade openness (LOPEN); and the log of inflation (LINF). The bottom panel of the table reports the diagnostics for the system GMM estimators. Robust (Windmeijer) standard errors for the system GMM estimators are reported in parentheses. \*\*\* indicates p-value < 0.01, \*\* indicates p-value < 0.05; and \* indicates p-value < 0.1.

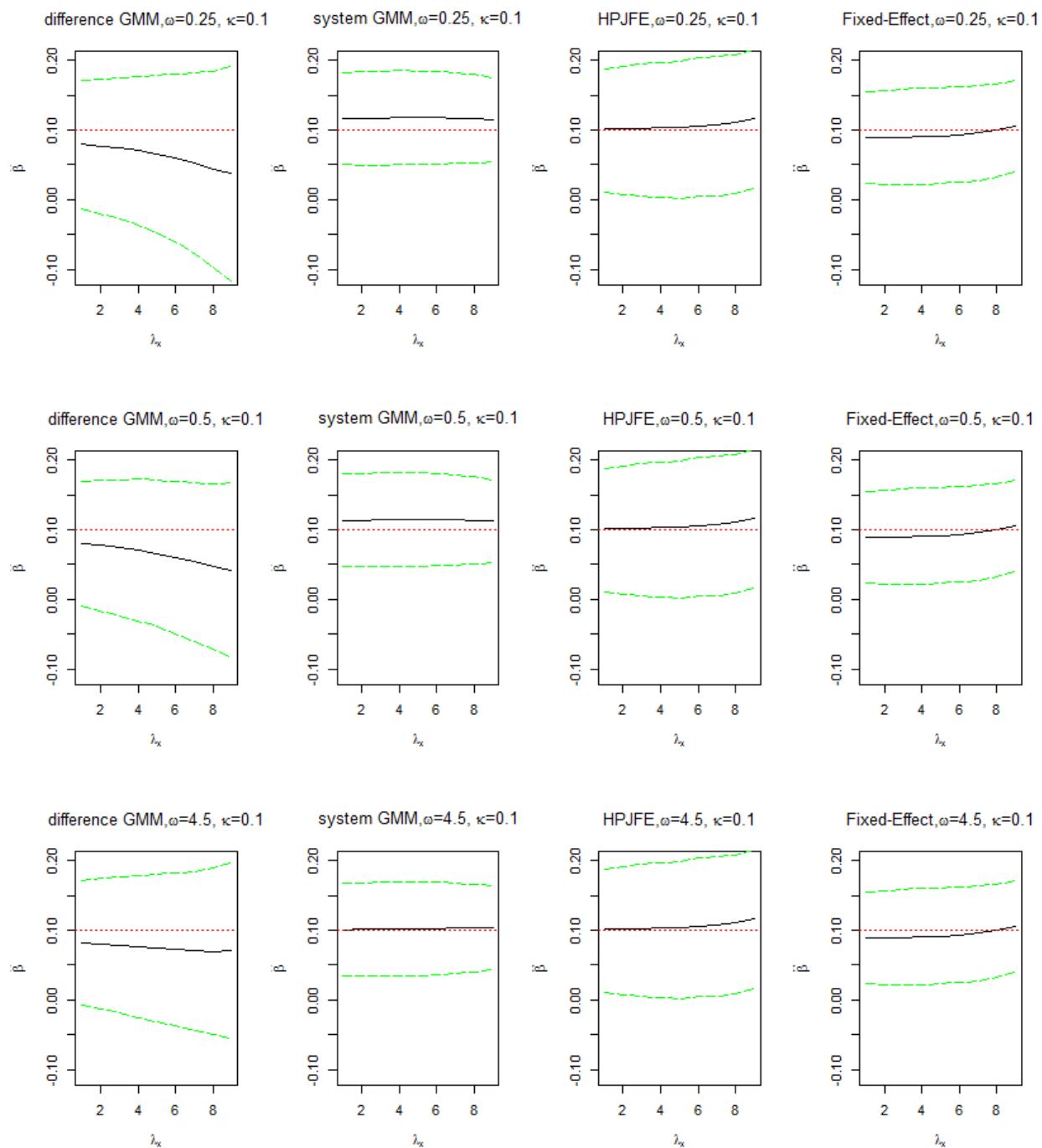
Table 3.7 Dynamic Panel data estimations with 3-year spell dataset excluding the outliers

Dep. Growth	Panel (A)				Panel (B)			
	system GMM			HPJFE	system GMM			HPJFE
	full	collapsed	restricted (lag=1)		full	collapsed	restricted (lag=1)	
(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	
LGDP(t-1)	0.135 (0.378)	0.359 (0.771)	0.525 (0.376)	-1.228 (1.619)	-0.0529 (0.409)	0.609 (0.691)	0.673* (0.367)	-1.384 (1.833)
PC(t-1)	-1.880* (0.971)	-3.088** (1.255)	-2.583*** (0.730)	-0.919 (1.959)	-0.199 (2.361)	-7.130** (3.114)	-2.713 (2.216)	2.487 (5.097)
PC2(t-1)					-0.606 (1.174)	2.509 (1.570)	0.314 (1.042)	-2.105 (2.753)
LEDU(t-1)	0.509 (0.524)	0.381 (1.153)	0.576 (0.641)	-0.551 (0.937)	0.872* (0.468)	0.00129 (0.989)	0.206 (0.645)	-0.624 (1.049)
LGC(t-1)	-1.584** (0.699)	-2.169* (1.306)	-3.119*** (0.717)	-1.691 (1.250)	-1.932*** (0.714)	-2.157 (1.349)	-2.944*** (0.674)	-1.812 (1.414)
LOPEN(t-1)	1.903*** (0.563)	6.674*** (1.296)	1.931*** (0.612)	2.679** (1.340)	1.654** (0.792)	7.037*** (1.316)	1.899*** (0.655)	2.689* (1.497)
LINF(t-1)	-0.250** (0.126)	-0.549*** (0.201)	-0.327** (0.135)	-0.585** (0.233)	-0.258** (0.115)	-0.557*** (0.180)	-0.254* (0.139)	-0.570** (0.259)
country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
time FE	Yes	Yes	Yes		Yes	Yes	Yes	
N.obs	1331	1331	1331		1331	1331	1331	
N.country	126	126	126		126	126	126	
N. iv in system GMM					912	128	220	
AR(1) pval	0	0	0		0	0	0	
AR(2) pval	0.77	0.876	0.977		0.765	0.862	0.909	
Hansen Test Pval	1	0.166	0.999		1	0.343	1	
Diff In Hansen Test Pval	1	0.133	1		1	0.869	1	
difference eq: K-P LM stat. pval	0.3043	0.0005	0.0029		0.3894	0.0004	0.021	
KP Wald F stat	32.636	2.625	1.895			2.53	1.664	
levels eq: K-P LM stat. pval	0.7838	0.1217	0.7838		0.8396	0.0581	0.8396	
KP Wald F stat	1.321	0.421	1.321		1.414	0.55	1.414	

This table reports the results of a set of panel regressions aimed at estimating the causal effects of credits to the private sector on economic growth using 3-year non-overlapping panel data excluding countries: GAB, CHN, ISL, LBR, SAU, IRL, and USA. Column (1)-(3) are the regressions estimated by the system GMM with full instrument

set, collapsed instrument set, and restricted one period lagging instrument set, respectively. Column (4) represents the regression estimated by the Half-Panel Jackknife Fixed-Effect (HPJFE) estimation. The set of controls includes time fixed effects, country fixed effects, and the lag of: log of GDP per capita (LGDP); the level of credits to the private sector (PC); the square of credits to the private sector (PC2); the log of average years of education (LEDU); the log of government consumption over GDP (LGC); the log of trade openness (LOPEN); and the log of inflation (LINF). The bottom panel of the table reports the diagnostics for the system GMM estimators. Robust (Windmeijer) standard errors for the system GMM estimators are reported in parentheses. \*\*\* indicates p-value < 0.01, \*\* indicates p-value < 0.05; and \* indicates p-value < 0.1.

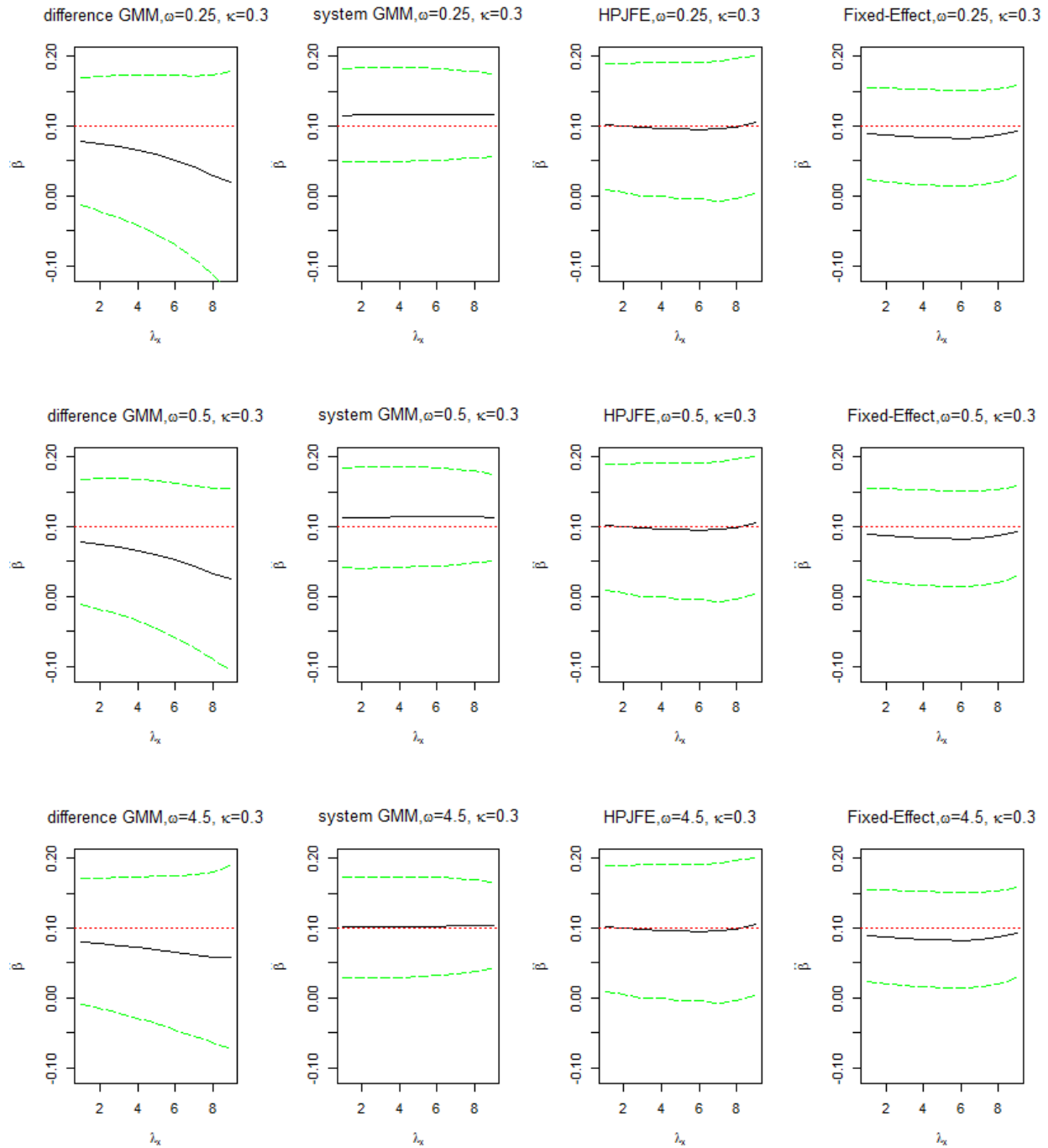
Figure 3.1 Estimates based on the simulations with  $k = 0.1$ ;  $\omega = 0.25, 0.5, 4.5$ ;  $\lambda_x \in \{0, 0.1, 0.2, \dots, 0.8\}$ ;  $T=10$ .



The graph show parameter estimates and 95% confident intervals from simulations of the model (9) (10) (11) based on 1000 draws of a sample size  $N=100$ ,  $T=10$ . The Upper panel represents estimates based on the simulation with  $\lambda_y = 0.8$ ;  $\delta = 0.5$ ;  $k = 0.1$ ;  $\omega = 0.25$ . The middle panel represents estimates based on the simulation with  $\lambda_y = 0.8$ ;  $\delta = 0.5$ ;  $k = 0.1$ ;  $\omega = 0.5$ . The lower panel represents estimates based on the simulation with  $\lambda_y = 0.8$ ;  $\delta = 0.5$ ;  $k = 0.1$ ;  $\omega = 4.5$ . Red dot line indicates the true value  $\beta = 0.1$ . The black solid lines are the estimates under four

estimators: difference GMM; system GMM; HPJFE; and Fixed-Effect. The green dash lines represent the 95% confident interval.

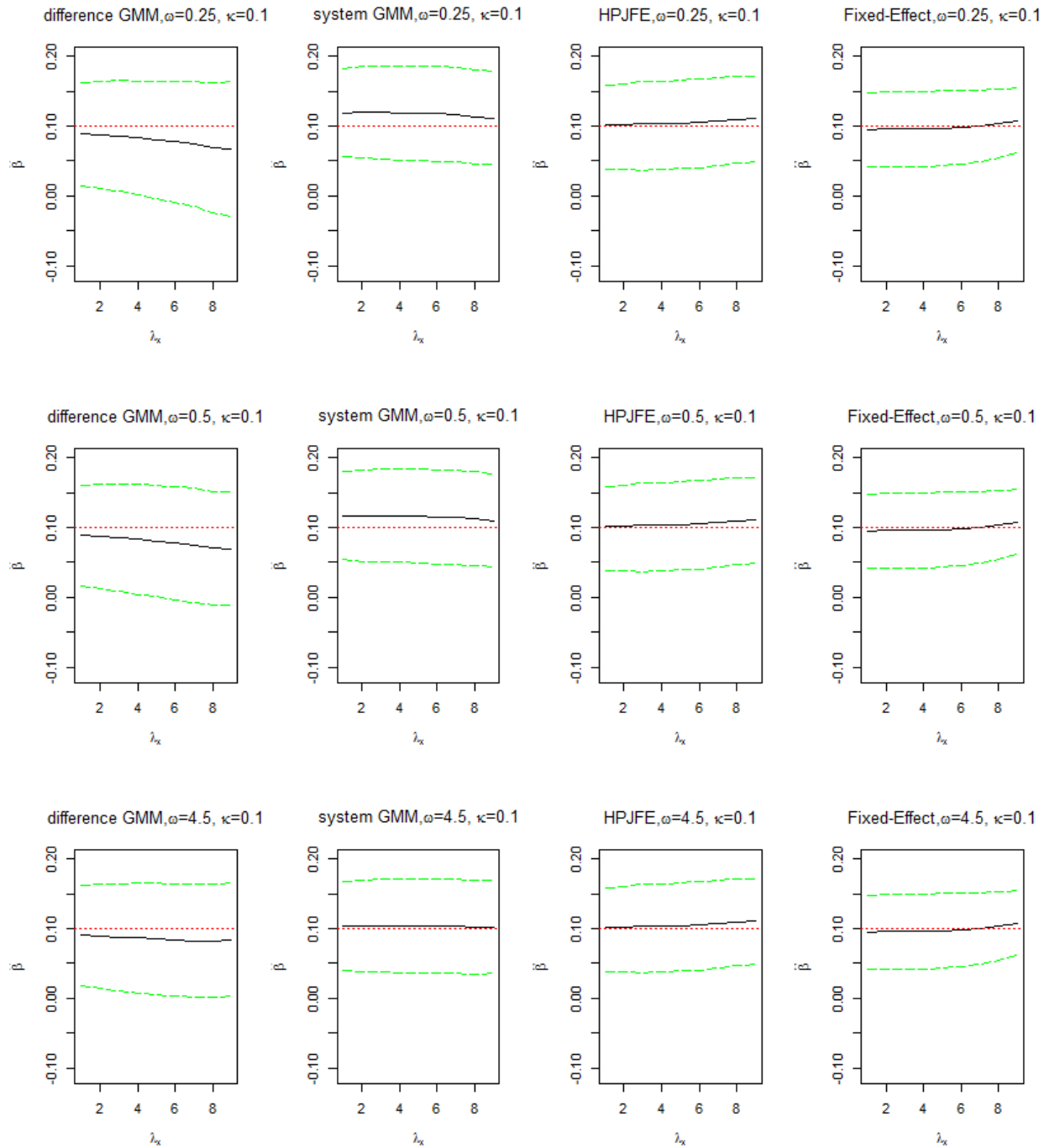
Figure 3.2 Estimates based on the simulations with  $k = 0.3$ ;  $\omega = 0.25, 0.5, 4.5$ ;  $\lambda_x \in \{0, 0.1, 0.2, \dots, 0.8\}$ ;  $T=10$ .



The graph show parameter estimates and 95% confident intervals from simulations of the model (9) (10) (11) based on 1000 draws of a sample size  $N=100$ ,  $T=10$ . The Upper panel represents estimates based on the simulation with  $\lambda_y = 0.8$ ;  $\delta = 0.5$ ;  $k = 0.3$ ;  $\omega = 0.25$ . The middle panel represents estimates based on the simulation with  $\lambda_y = 0.8$ ;  $\delta = 0.5$ ;  $k = 0.3$ ;  $\omega = 0.5$ . The lower panel represents estimates based on the simulation with  $\lambda_y = 0.8$ ;  $\delta = 0.5$ ;  $k = 0.3$ ;  $\omega = 4.5$ . Red dot line indicates the true value  $\beta = 0.1$ . The black solid lines are the estimates under four

estimators: difference GMM; system GMM; HPJFE; and Fixed-Effect. The green dash lines represent the 95% confident interval.

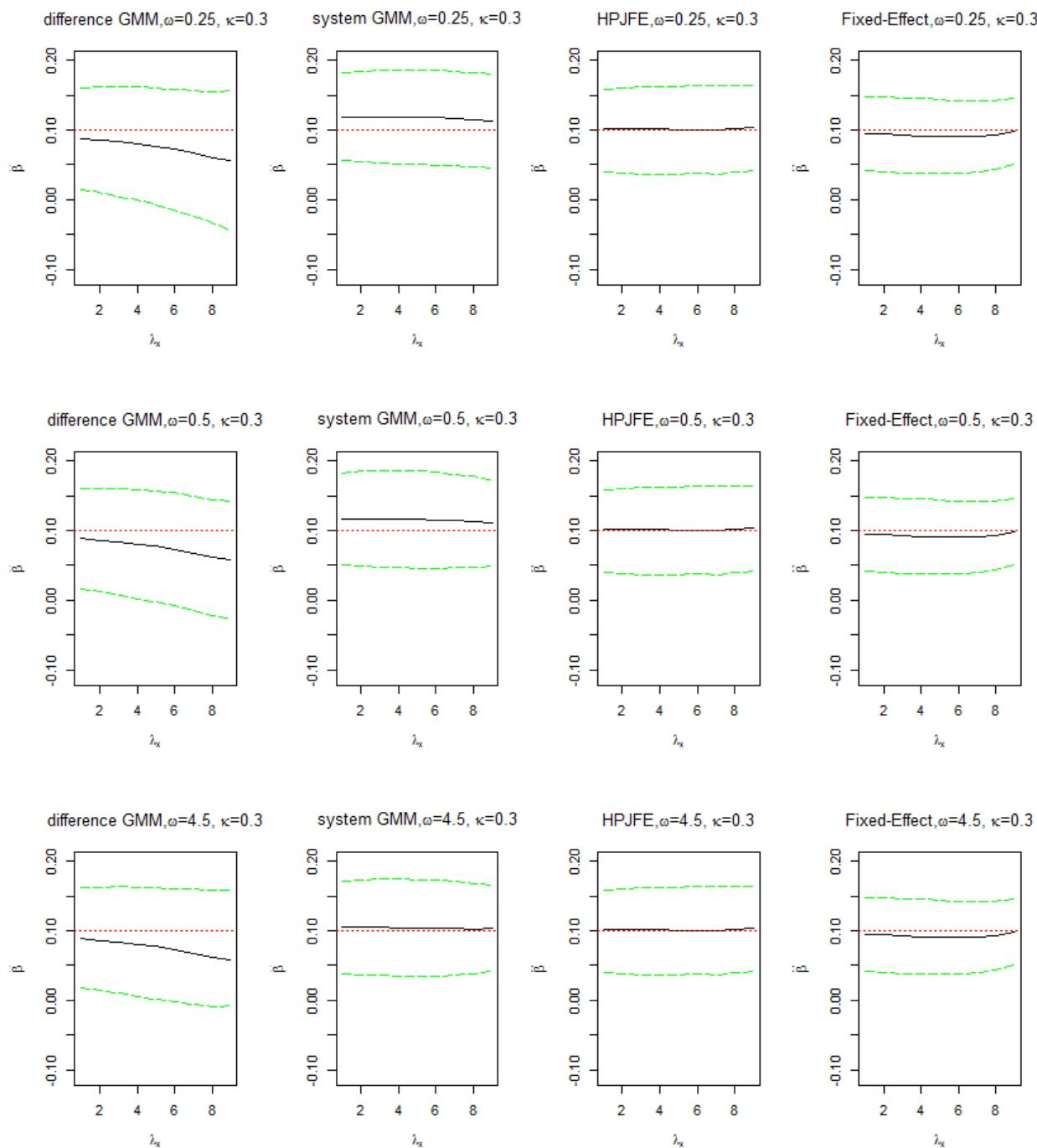
Figure 3.3 Estimates based on the simulations with  $k = 0.1$ ;  $\omega = 0.25, 0.5, 4.5$ ;  $\lambda_x \in \{0, 0.1, 0.2, \dots, 0.8\}$ ;  $T=15$ .



The graph show parameter estimates and 95% confident intervals from simulations of the model (9) (10) (11) based on 1000 draws of a sample size  $N=100$ ,  $T=15$ . The Upper panel represents estimates based on the simulation with  $\lambda_y = 0.8$ ;  $\delta = 0.5$ ;  $k = 0.1$ ;  $\omega = 0.25$ . The middle panel represents estimates based on the simulation with  $\lambda_y = 0.8$ ;  $\delta = 0.5$ ;  $k = 0.1$ ;  $\omega = 0.5$ . The lower panel represents estimates based on the simulation with  $\lambda_y = 0.8$ ;  $\delta = 0.5$ ;  $k = 0.1$ ;  $\omega = 4.5$ . Red dot line indicates the true value  $\beta = 0.1$ . The black solid lines are the estimates under four

estimators: difference GMM; system GMM; HPJFE; and Fixed-Effect. The green dash lines represent the 95% confident interval.

Figure 3.4 Estimates based on the simulations with  $k = 0.3$ ;  $\omega = 0.25, 0.5, 4.5$ ;  $\lambda_x \in \{0, 0.1, 0.2, \dots, 0.8\}$ ;  $T=15$ .



The graph shows parameter estimates and 95% confidence intervals from simulations of the model (9) (10) (11) based on 1000 draws of a sample size  $N=100$ ,  $T=15$ . The upper panel represents estimates based on the simulation with  $\lambda_y = 0.8$ ;  $\delta = 0.5$ ;  $k = 0.3$ ;  $\omega = 0.25$ . The middle panel represents estimates based on the simulation with  $\lambda_y = 0.8$ ;  $\delta = 0.5$ ;  $k = 0.3$ ;  $\omega = 0.5$ . The lower panel represents estimates based on the simulation with  $\lambda_y = 0.8$ ;  $\delta = 0.5$ ;  $k = 0.3$ ;  $\omega = 4.5$ . Red dot line indicates the true value  $\beta = 0.1$ . The black solid lines are the estimates under four

estimators: difference GMM; system GMM; HPJFE; and Fixed-Effect. The green dash lines represent the 95% confident interval.

Figure 3.5 Leverage plot based on the Pool OLS estimates of column (1), Table 3.4

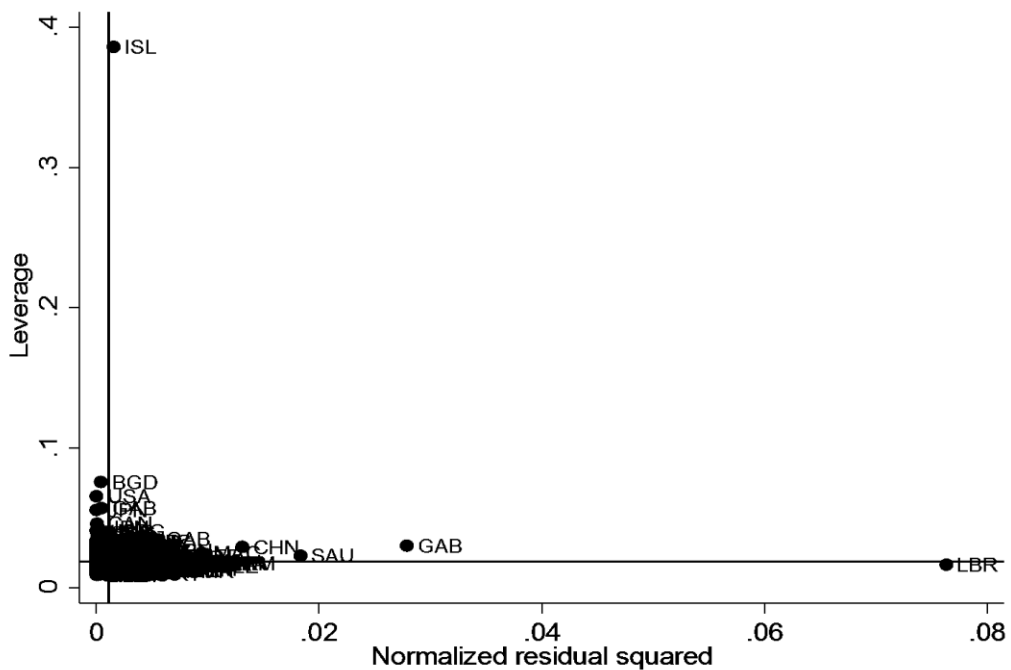


Figure 3.6 Partial correlations between credits to the private sector (PC) and growth, based on column (1), Table 3.5

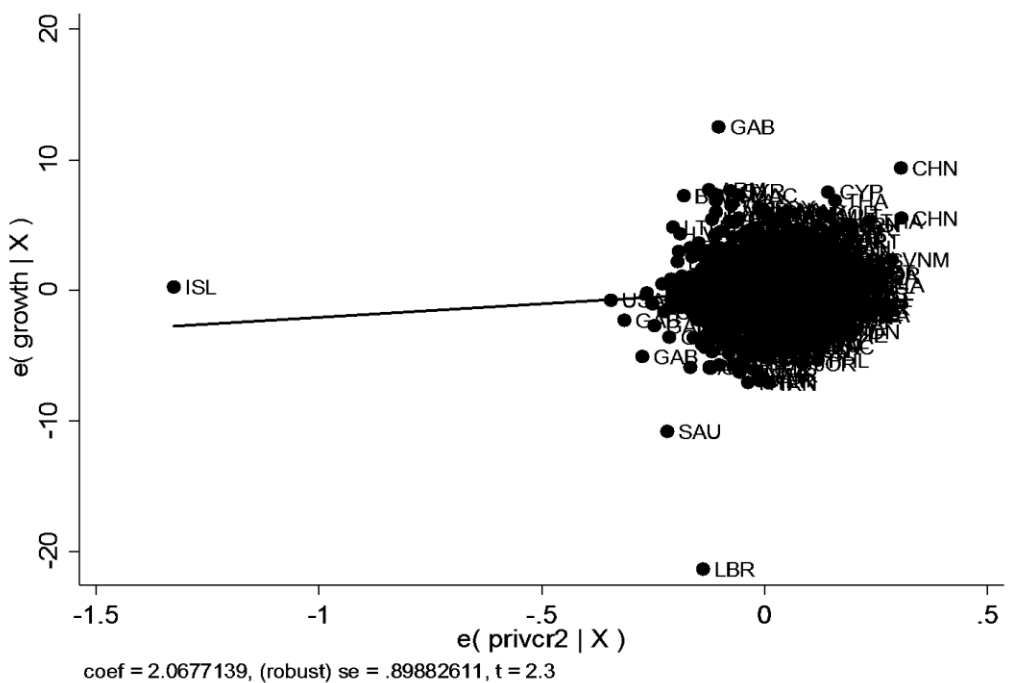


Figure 3.7 Partial correlations between the square of credits to the private sector (PC2) and growth, based on column (1), Table 3.5.

