

PREDICTING CRISIS: THIS TIME IS DIFFERENT (?)

A Master's Thesis

by

TUĞBA SAĞLAMDEMİR

Department of
Economics

İhsan Doğramacı Bilkent University
Ankara

September 2012

To my family...

PREDICTING CRISIS: THIS TIME IS DIFFERENT (?)

Graduate School of Economics and Social Sciences
of
İhsan Doğramacı Bilkent University

A Master's Thesis

by

TUĞBA SAĞLAMDEMİR

In Partial Fulfillment of the Requirements For the Degree
of
MASTER OF ARTS

in

THE DEPARTMENT OF
ECONOMICS
İHSAN DOĞRAMACI BILKENT UNIVERSITY
ANKARA

September 2012

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Arts in Economics.

Assoc. Prof. Dr. Selin Sayek Böke

Supervisor

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Arts in Economics.

Assoc. Prof. Dr. Fatma Taşkın

Second Reader

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Arts in Economics.

Assoc. Prof. Dr. Levent Akdeniz

Examining Committee Member

Approval of the Graduate School of Economics and Social Sciences

Prof. Dr. Erdal Erel

Director

ABSTRACT

PREDICTING CRISIS: THIS TIME IS DIFFERENT (?)

SAĞLAMDEMİR, Tuğba

M.A., Department of Economics

Supervisor: Assoc. Prof. Dr. Selin Sayek Böke

September 2012

This thesis aims to predict the currency, banking and debt crises and more specifically investigates the effect of the housing sector on the crisis prediction. This study is not only constructing a crisis prediction method, which uses the previous literatures data set, but also proposing a new one including the housing market data, and comparing the performances of the two in order to measure the impact of the housing market on the prediction power. The data are taken from World Bank, IMF, OECD and Eurostat and cover the years between 1999 and 2010. Multinomial logistic regression is used for crisis prediction. As an advantage, it prevents the ‘post-crisis bias’ problem; by this way the robustness of the analysis is also improved. The multinomial logistic regression is run for two different time windows ‘ $t-1, t, t+1$ ’ and ‘ $t, t+1, t+2$ ’ as t denoting the current year. By inclusion of the housing market data, the prediction power is increased from 60% to 100% in the case of ‘ $t-1, t, t+1$ ’. The model sends no false alarms in this case. For the case of ‘ $t, t+1, t+2$ ’, the in-sample prediction power is improved from 68% to 95% and the false alarm ratio is reduced from 6.6% to 3%. The-out-of-sample predictive performance of the system in ‘ $t, t+1, t+2$ ’ is improved from 33% to 60% by the inclusion of the housing market data. Due to the restrictions of the data set, out-of-sample analysis could not be performed for the case of ‘ $t-1, t, t+1$ ’. The proposed crisis prediction method succeeds in predicting

the crises of 2000s by using housing sector data. The impact of the housing sector in predicting crisis is clearly shown. Also, it is shown that chosen time window for the multinomial logistic regression in predicting crisis can lead to variations on the predicted results.

Key words: Predicting crisis, Currency Crisis, Banking Crisis, Debt Crisis, Post-crisis bias, Multinomial Logistic Regression, The Housing Market

ÖZET

KRİZ TAHMİNLEME: BU DEFA FARKLI (?)

SAĞLAMDEMİR, Tuğba

Yüksek Lisans, Ekonomi Bölümü

Tez Yöneticisi: Doç. Dr. Selin Sayek Böke

Eylül 2012

Bu tez çalışması, para, bankacılık ve borç krizlerini tahminlemeyi ve daha özelde, konut piyasası verilerinin kriz tahminleme üzerindeki etkilerini araştırmayı amaçlamaktadır. Bu çalışma, sadece geçmiş çalışmalarda kullanılan verilerden oluşan bir kriz tahminleme metodu kurgulamamakta, aynı zamanda konut piyasası verilerinin kriz tahminlemedeki etkilerini araştırmak için bunları da içeren yeni bir kriz tahminleme metodu da önermektedir. Çalışmada kullanılan veriler, Dünya Bankası, Uluslararası Para Fonu (IMF), Ekonomik Kalkınma ve İşbirliği Örgütü (OECD) ve Eurostat veritabanlarından alınmıştır ve 1999 - 2010 yıllarıarasını kapsamaktadır. Kriz tahminlemede ‘Çok Terimli Lojistik Regresyon’ tekniği kullanılmıştır. Bir avantaj olarak, bu yöntem, ”kriz sonrası sapma” problemini önlemektedir. Analizin sağlamlığı, bu şekilde geliştirilmektedir. Regresyon analizi, t ’nin şimdiki yılı belirttiği durumda, ‘ $t-1, t, t+1$ ’ ve ‘ $t, t+1, t+2$ ’ şeklinde iki farklı zaman penceresi için koşturulmuştur. Konut piyasası verilerinin eklenmesiyle, ‘ $t-1, t, t+1$ ’ penceresindeki tahminleme gücü %60’tan %100’e yükseltilmiştir. Bu durumda, sistem herhangi bir yanlış uyarı vermemektedir. ‘ $t, t+1, t+2$ ’ penceresi içinse, sistemin örneklem içi tahminleme gücü %68’den %95’e çıkarılmış ve yanlış uyarı oranı %6.6’dan %3’e düşürülmüştür. ‘ $t, t+1, t+2$ ’ penceresi için, sistemin örneklem dışı tahminleme gücü konut piyasası verilerinin dahil edilmesiyle, %33’ten %60’a yükseltilmiştir. Veri kümesindeki bazı kısıtlamalar nedeniyle, ‘ $t-1, t, t+1$ ’ penceresi için örneklem dışı

performans analizi yapılamamıştır. Önerilen kriz tahminleme metodu, konut piyasası verileri ile 2000’li yıllardaki krizleri tahminlemeyi başarmıştır. Konut piyasası verilerinin kriz tahminlemedeki etkisi açık şekilde gösterilmiştir. Ayrıca, çok terimli regresyon analizi ile kriz tahminlemesi yapılırken, seçilen zaman pencerelerinin tahminleme sonuçlarında değişikliklere neden olabileceği de gösterilmiştir.

Anahtar Kelimeler: Kriz tahminleme, Para Krizleri, Bankacılık Krizleri, Borç Krizleri, Kriz Sonrası Sapma, Çok Terimli Regresyon, Konut Sektörü

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisor Selin Sayek Böke for her supervision and invaluable guidance. I will always be indebted to her for her continuous support, patience and encouragement during my studies.

I also would like to thank to Refet Gürkaynak for his continuous support and encouragement throughout my graduate study.

I am grateful to my examining committee members, Fatma Taşkın and Levent Akdeniz for their helpful comments and suggestions. Also, I would like to thank Erinç Yeldan and ÇağrıSağlam for their helpful comments and suggestions during my graduate study.

I would like to thank to Bilkent University and TUBITAK for their financial support.

I owe my special thanks to my family for their unconditional love, support, patience and encouragement throughout my life.

I am thankful to the big brothers: M. Orkun Sağlamdemir and Ömer Yetik for their patience, understanding and support throughout my study. It is very pleasing to know that, it is independent from where I am, I will never walk alone.

I am thankful to Özge Filiz Yağcıbaşı, for being my fellow who shows me that distance becomes meaningless when her friendship is the matter, it is a chance to have a sister like her.

I am thankful to Fulya Özcan and Seçil Yıldırım not only for sharing the time of my graduate study, but also for their friendship, which will last forever. They are one of the most important gains of this graduate study.

Last, but not least, I would like to thank to Maria Nawandish and Çiğdem Akbulut for their friendships and continuous support.

TABLE OF CONTENTS

ABSTRACT	iii
ÖZET	v
ACKNOWLEDGEMENTS	vii
TABLE OF CONTENTS	x
LIST OF TABLES	xi
LIST OF FIGURES	xii
CHAPTER 1: INTRODUCTION	1
CHAPTER 2: LITERATURE SURVEY	6
CHAPTER 3: THE HOUSING MARKET	11
CHAPTER 4: DEFINITIONS, DATA AND METHODOLOGY	13
4.1 Definitions and Data	13
4.2 Methodology	20
CHAPTER 5: ANALYSIS	26
5.1 Event Study Analysis	26
5.2 Empirical Analysis	43
5.2.1 Empirical results for the time window: ‘t-1, t, t+1’	46

5.2.2 Empirical results for the time window: ‘t, t+1, t+2’	55
CHAPTER 6: CONCLUSION	66
BIBLIOGRAPHY	70
APPENDIX A:FURTHER DETAILS	76
A.1 Definitions Data and Methodology	76

LIST OF TABLES

4.1	Data set classified according to categories	15
4.2	Countries used as data source	17
5.1	Regime definition for the multinomial logit model for t-1, t, t+1 . . .	46
5.2	Results of the multinomial logit regression for t-1, t, t+1	48
5.3	Performance of the model without housing data for t-1, t, t+1	52
5.4	Performance of the model with housing data for t-1, t, t+1	53
5.5	Crises correctly called with the model with housing data, t-1, t, t+1 .	54
5.6	Performance of the model without housing data on the reduced data set for t-1, t, t+1	55
5.7	Regime definition for the multinomial logit model for t, t+1, t+2 . .	55
5.8	Results of the multinomial logit regression for t, t+1, t+2	57
5.9	Performance of the model without housing data for t, t+1, t+2 . . .	61
5.10	Performance of the model with housing data for t, t+1, t+2	62
5.11	Crises correctly called with the model with housing data, t, t+1, t+2	63
5.12	Performance of the model without housing data on the reduced data set for t, t+1, t+2	64
A.1	Data set classified according to source	76

LIST OF FIGURES

1.1	Housing Prices for some developed countries	3
5.1	Event Study Analysis for time window ‘t-1, t, t+1’ - I	31
5.2	Event Study Analysis for time window ‘t-1, t, t+1’ - II	32
5.3	Event Study Analysis for time window ‘t-1, t, t+1’ - III	33
5.4	Event Study Analysis for time window ‘t, t+1, t+2’ - I	39
5.5	Event Study Analysis for time window ‘t, t+1, t+2’ - II	40
5.6	Event Study Analysis for time window ‘t, t+1, t+2’ - III	41
5.7	Event Study Analysis for time window ‘t, t+1, t+2’ - IV	42

CHAPTER 1

INTRODUCTION

This hypothesis of this study is formed on the supposed impact of the housing sector on predicting crisis. This question is valuable for the following reasons: Crises are events that have been re-occurring since the 14th century and policy makers could not prevent outbreak of crises from that time. This is why crisis prediction is crucial for the whole economy. In this study, the housing sector is analyzed as an important indicator for predicting crisis; since most of the crises seem to have the common characteristics of having started in the housing markets. Due to all these reasons above, the housing market implication on predicting crisis will analyzed in this thesis.

Economic crisis, as being one of the common problems of the world, has a history that dates back to 14th century in England [Reinhart and Rogoff(2008b)]. Since the crisis creates a destructive impact on the economy, both policy makers and academics search for policies to avoid the crisis. If the economy consisted of mechanisms, which followed the same rules and did not change until the policy makers made any regulations, they would achieve their aim. However, the reality is different. There are many actors in the system and each of them makes different contributions. Due to this fact, the measures, which are taken to avoid the crisis, do not work.

Once it is accepted that the economic crises are situations which the economies cannot escape from, the optimum solution turns out to become predicting the coming crisis before it hits the economy, in order to protect the whole economic system as much as possible. As a consequence, the requirement for predicting crisis before its outbreak arises.

There exist quite a large number of studies on predicting crisis in the literature. They generally use financial and macroeconomic data to predict crisis and work only on one type of crisis on its own. Some of them work on banking crises, some on currency crises, and some on debt crises. Only Babecký et al (2012) try to find the early warning indicators for currency, banking and debt crisis [Babecký(2012)]. However, none of the previous studies both try to predict all types of crises and use housing sector data among with the financial and macroeconomic variables except Rose et al (2011).

The previous studies generally emphasize the similar indicators, which come from macroeconomic and financial data. Also, more or less, they usually apply the same techniques, which are being used from the very beginning of the early warning system studies. Every analysis with the aim of constructing an early warning system is a step to achieve higher prediction power; however, since economy has a very dynamic structure, both the prediction techniques and the data to be used in the estimation process have to evolve through time to improve the ability of catching the alterations in the system. In that respect, importance of the housing market on economic crises is analyzed in this thesis, and an early warning system is constructed by taking the effects of the housing sector data into consideration.

In this thesis, all types of crisis (either currency, banking or debt) are aimed to be predicted by combining the indicators used in the previous studies in the literature with the housing sector variables. Since the housing sector is the common problematic part of the economy for the countries which have economic crises after 2000s, it is

worth to search for its impact on predicting crisis. Rose et al (2011) is the first study, which uses ‘percentage change in real estate prices’ as an indicator of housing market data in crisis prediction. Although the authors introduce the housing market data into the system, they fail to predict the crisis for 107 countries. As it is seen from the following figure, the housing prices followed upward trend for some developed countries such as Australia, Canada, United Kingdom and United States:

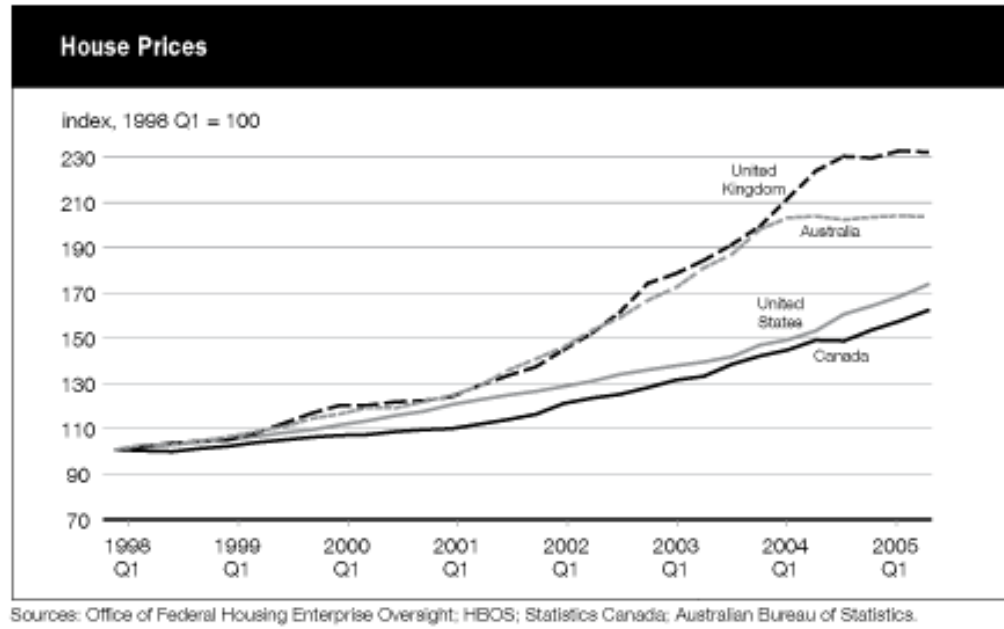


Figure 1.1: Housing Prices for some developed countries

In this study, the analysis on the crisis prediction is separated into two cases: In the first case, the crisis is predicted by running a multinomial logistic regression over a time window of ‘ $t-1$, t , $t+1$ ’ where ‘ t ’ denotes the current year of investigation, ‘ $t-1$ ’ is the year before the current one and ‘ $t+1$ ’ is the year after it. Hence, given that a crisis has occurred at time t , then the behavior of the variables within $t-1$ to $t+1$ is analyzed. That is, the values both before and after the crisis are taken into consideration. In the second case, again the multinomial logistic regression technique is used. However, the time window is now changed to ‘ t , $t+1$, $t+2$ ’. In this case,

the analyzed time window covers the current year and two consecutive years after it. Hence, the behavior of the variables before the crisis is not taken into consideration but the focus is on their current and future values.

For both time windows of multinomial logistic regression, the estimation is constructed from two different stages. In the first stage, the variables which are used in the prior crisis prediction studies are used in the multinomial logistic regressions for both time windows to show the prediction powers of these variables in predicting the crises that arose after 2000s. Then, in the second stage these prior studies' variables are combined with the housing market data set, which indicates the effects of housing sector. This combined data set is used in the multinomial logistic regressions for both time windows to show the prediction powers of these variables. By comparing the results of first and second stage analyses, it is possible to see the difference, which arises due the inclusion of the housing sector data.

In this thesis, the following questions are answered: Do housing market data make any improvements in predicting crisis? Does the prediction power differ depending on the time window of the multinomial logistic regression? Which time window is more suitable to be used in crisis prediction to attain a higher prediction power?

Applying the procedure described above, the empirical results show that the housing sector data, make an improvement in predicting crisis. The prediction power changes depending on the time window of the multinomial logistic regression. On the other hand, since the publicly available housing data set is restricted, it does not allow to make an out-of-sample performance analysis for the time window of 't-1, t, t+1'. Consequently, it is not possible to make a comparison between the crisis prediction powers of the two different time windows.

This thesis is organized as follows: Chapter 2 analyzes the literature. Chapter 3 explains the importance and impact of the housing market on the economy. Chapter

4 gives details on the data and the methodology. Chapter 5 reviews the event study and empirical analysis results. Chapter 6 concludes.

CHAPTER 2

LITERATURE SURVEY

The economy is a complete system, which consists of different dynamics. As the countries economies are getting integrated in terms of financial dependence on each other, the economic decisions which are taken by one country turns out to be an important indicator for all economies over the world. As a result of this, this system needs to be controlled in order to protect it against an outbreak of a crisis. The policy makers in each country try to control and arrange any variation for every indicator, which may affect the other components and cause the whole system to collapse, by utilizing the tools and mechanisms they have in hand.

Although there are many regulatory mechanisms, which arrange and supervise the system, unexpected changes still do appear. These changes sometimes become so uncontrollable that they drive the economy into crisis. These crises may be the result of some already existing dynamics or may arise due to a new dynamic, which has recently been introduced to the system and has become crucial in an ongoing basis on the economic conditions.

As a matter of fact, the economic crisis history is as old as the history of economy. Since there are many indicator factors constituting the whole economy - and new indicators get involved in this system continuously - it is impossible to think of an

economy, which never suffers from an economic crisis. On the other hand, the most appropriate solution for reducing the destructive impact of a crisis on the economy is to predict the coming of it and take the required precautions so as to protect the economy as much as possible. This indeed is the motivation behind designing an Early Warning System, aim of which is to predict a crisis before it hits the economy.

There are different types of economic crises as currency crisis, banking crisis and debt crisis. The policy makers generally tried to construct Early Warning Systems for each type of crisis on its own. After mid 90s, the currency crisis turned out to become a common problem of the economic systems. This crisis type is analyzed by Kaminsky et al (1998) and in this analysis, a new non-parametric approach is constructed, which is also called as the signal approach [Kaminsky and Reinhart(1998)]. This approach is prominent in the field of early warning systems, whose prediction power is 70% in the in-sample analysis. By this study, it has been shown that it was possible to predict a crisis with a non-parametric approach. After this analysis, the early warning system literature has been introduced a new term as false signal, standing for the cases in which the system warns about a coming crisis but there is no upcoming crisis indeed.

Then researchers sought for new estimation techniques to construct early warning systems. Berg et al (1999a) use the same data set and crisis definition as Kaminsky et al to work on the currency crisis [Berg and Pattillo(1999a)]. However, their estimation technique is the probit approach for designing the early warning system, where the dependent variable takes the value of one for the case of a coming crisis and zero for all other cases. The probit approach is more practical than the signal approach since it allows testing the statistical significance and coefficient constancy overtime and countries.

In addition to the probit approach Demirguc-Kunt et al (1997) introduce a new estimation tool [Demirguc-Kunt(1997)]. They try to find the main reason behind the

banking crisis by working on both developing and developed countries. They apply multivariate logit model to identify the determinants of the banking crisis.

Since the important thing is not only constructing the early warning system but also predicting the crisis with the highest success ratio, the literature seek for new techniques, new data sets and indicators to increase their prediction power. Binomial logit is another technique used but its results suffer from post-crisis bias which is described in Section 4.2.

Bussiere et al (2002) construct an early warning system aiming for predicting currency crisis by using the estimation technique of multinomial logistic regression [Bussiere and Fratzscher(2002)]. Since they criticize the prediction power of the binomial logit model, their analysis consists of both binomial and multinomial logistic regression models. The data set they use contains 20 open economies and spans the years from 1993 to 2001. According to their estimations, the binomial logistic regression predicts the crisis entry periods with a success ratio of 56,9% and multinomial logistic regression estimates the same periods with 65,5% success. After their contribution, the multinomial logistic regression technique replaces the binomial logistic regression in the early warning system literature.

Although the early warning system techniques are generally parametric, there are some analyses, which compare the prediction power of the different methods such as the one done by Peltonen (2006) for predicting currency crisis [Peltonen(2006)]. In this study, two early warning systems are constructed by using two different approaches: probit approach, which is parametric, and artificial neural network (ANN) model as a non-parametric approach. Their data set comprises eight exogenous indicators, and the time interval spans the period between 1980 and 2001 for 24 emerging economies. The main contribution of this paper is that, it compares the prediction power of the probit model with that of the Artificial Neural Network approach for in-sample and out-of-sample performance in predicting currency crisis. Then it is shown

that the ANN approach outperforms the probit model for predicting currency crisis regarding the in-sample performances, but both methods' out-of-sample performances are weak.

Although most of the early warning systems are designed to predict the currency crisis, after 2000s, with the financial liberalization and globalization, the reason of the economic crises do change with the structure of the economy. As a consequence of this change, different types of crises emerge and attract the literature's attention. Manasse et al (2003) constructed an early warning system with the aim of predicting debt crisis [Manasse and Roubini(2003)]. Their data set consists of 47 markets and the time interval spans years from 1970 to 2002. In this analysis, both a non-parametric approach - Classification and Regression Tree (CART) - and a parametric approach binomial logistic regression - are applied and their prediction powers are compared. According to the results, binomial logistic regression predicts 74% of the crises where CART predicts 89% of the crisis entries. On the other hand, CART sends more false alarms than the binomial logistic regression. No out-of-sample analyses are applied.

The debt crisis is analyzed by many other researchers such as Bruner et al (1987) and Roubini et al (2005). Roubini et al (2005) apply CART as well, using 47 countries' data, which belongs to the period of 1970 to 2002 [Roubini(2005)]. 10 exogenous variables are used for the estimation process and the prediction power of this model is 85% for the in-sample performance and 35% for the out-of-sample performance.

To estimate the debt crises, more techniques and crisis definitions are used with the aim of increasing the estimation power. One of them is constructed by Ciarlone et al (2005) [Ciarlone and Trebeschi(2005)]. In this analysis, they use multinomial logistic regression to predict the crisis. Also, they run a binomial logistic regression to make a comparison between the estimation powers of these two models. By using 28 countries' data, which also span the years from 1980 to 2002, they find that, with the binomial logistic regression's in-sample prediction power of the model is 72,5%. For the same

data set, the in-sample prediction power increases to 76% with the multinomial logistic regression. They do not make an out-of-sample performance analysis for the models. One year later, again Ciarlone et al (2006) construct a new early warning system to predict debt crisis [Ciarlone and Trebeschi(2006)]. The debt crisis definition and multinomial logistic regression construction is different compared to their previous analysis. This analysis uses the same data set, but their prediction power is 78% for the in-sample performance and 70% for the out-of-sample performance.

Rose et al (2009, 2011) construct an early warning system, which uses multiple-indicators multiple-causes (MIMIC) model for the estimation for 107 countries, to predict the 2008 financial crisis [Rose(2009), Rose(2011)]. They analyze 60 potential causes of this crisis but they accept that their early warning system could not predict the crisis.

Although, Rose et al (2009, 2011) could not predict this type of crisis by using housing sector data, the searched hypothesis question of the impact of housing sector on predicting crisis is still important. As it will be argued in the following section, the indicators of the housing market are important on predicting crisis as crisis indicators.

CHAPTER 3

THE HOUSING MARKET

In this study, the housing sector is analyzed as a potential reason for the economic crisis and constitutes the most important part of the crisis prediction. There are lots of independent reasons denoting the importance of the housing sector, which led to the question of whether it could be a cause for the crisis or not and motivated this research.

Although the housing prices generally change with the overall price level in the economy, the situation is a little bit different compared to the usual behavior, as explained by Andre (2010) as follows: During the 2001 recession, house prices got disconnected from the business cycle. Low interest rates and mortgage innovations fueled housing demand, which in turn led to an expansion in the housing prices. This housing price expansion had three main aspects as: Real house price has soared, the expansion has been exceptionally long, most of the industrialized countries experienced house price boom simultaneously. In reaction to this rise in the housing prices, the residential investment increased [Andre(2010)].

Although the number of the houses increased, their price level did not show any reduction and this increased the house owners' wealth. The level of wealth is a determinant factor of the consumption level for the overall economy, as shown by

Case et al (2001). Their study shows that, the changes in the national wealth are associated with the changes in consumption. In addition, they pointed out that, the tendency to consume out of stock market wealth is different from the tendency to consume out of housing wealth [Case(2001)]. The different impact of the housing wealth and stock market wealth on consumption is compared by Benjamin et al (2004) for the US economy. They reach the result that, the decline in the stock market during 2000 and 2001 had a limited impact on aggregate demand because of an offsetting real estate wealth effect [Benjamin(2004)].

Goodhart et al (2008) accept the link between aggregate consumption and housing sector as a determining factor for monetary variables and macroeconomics. They found out that, there is a multidirectional link between house prices, monetary variables and the macro economy. Moreover, this link between house prices and monetary variables is found to be stronger over a more recent sub-sample from 1985 till 2006. The housing boom is not only important for the housing sector itself but it also affects the financial sector, monetary sector and credit sector [Goodhart(2008)].

Furthermore, the real estate boom is accepted as a source of fragility by Buiter et al (2009) since it channeled investment away from more productive areas into unproductive residential construction [Buiter(2009)].

This framework forms the basis of my thesis. It suggests that housing market developments influence crisis. This also forms the basis for the analysis by Rose et al (2009, 2011). But they could not predict crisis [Rose(2011)].

Although the importance of the housing sector is pointed out by many studies in the literature, the impact of this sector on the crisis has not been shown. Hence, the aim of this study is to analyze the effects of the housing sector on the occurrence of economic crisis, and it is expected to find an improvement to the prediction power of the early warning systems by using the housing sector data.

CHAPTER 4

DEFINITIONS, DATA AND METHODOLOGY

4.1 Definitions and Data

In predicting crisis, there are many important steps in the construction part and the first one of them is defining the crisis conditions. The next step after categorizing the conditions is the prediction of the coming of a crisis. At this point, the most important thing is choosing the right indicator variables. The main goal of crisis prediction is to construct a model, which is capable of catching upcoming crises with the minimum possible number of misses. Therefore, the system needs some precise threshold, which helps in charting out. There are some basic questions answers of which help in leading the variation of the model. These questions are as follows:

- What is the definition of the crisis?
- Which countries constitute the research area?
- What is the time interval and which explanatory variables are in use?

By answering these questions, the general outline will be designed and then it will be easy to construct the model by using these answers. Basically, there are three kinds of economic crises as banking crisis, currency crisis and debt crisis. They differ from each other in terms of some basic causes and results but their general effect is the same: lowering nations' welfare. Motivated by this fact, the system proposed in this study aimed to predict all these three types of crises. A brief definition for each type of crisis can be given as follows [Reinhart and Rogoff Online Resources]:

• **Currency crisis:** The economic situation is defined to be a currency crisis if the annual inflation rate is 20 percent or higher and the annual depreciation versus the United States dollar is 15 percent or more.

• **Banking Crisis:** If one or both of the following two conditions hold, the economy is said to have a banking crisis:

- (i) Bank runs that lead to the closure, merging or takeover by the public sector of one or more financial institutions.
- (ii) If there are no runs, the closure, merging, takeover or large scale government assistance of an important financial institution (or group of institutions) that marks the start of similar outcomes for other financial institutions.

• **Debt Crisis:** It is identified in the case of a failure to meet a principal or interest payment on the due date (or within the specified grace period). The episodes also include instances where rescheduled debt is ultimately extinguished in terms less favorable than the original obligation. In addition to this condition, the situations of banks being forced to freeze their deposits or forcible conversions of such deposits from dollar to local currency are considered to be debt crisis conditions as well. These conditions for different types of crises cases have been defined in accordance with the previous studies of Reinhart and Rogoff to identify currency, banking and debt crises [Reinhart and Rogoff Online Resources].

The data set is formed from 111 exogenous variables, which is mainly divided into two main branches. The first group represents the prior studies variables, which were collected to predict currency, banking and debt crises. The second part of the data set consists of housing market data. This second group of variables has not been used in the prediction of the crises until now. The overall data set classified according to their categories are given in the following table:

Table 4.1: Data set classified according to categories

Source	Set of Variables
Capital Account	Net open position in the foreign exchange to capital ratio, FDI, total reserve growth, FDI to GDP
Debt Profile	Household debt to GDP, short term debt to international reserves, domestic credit to private sector, interest payment on total external debt, total external debt stocks, short term external debt stocks, short term debt to total reserves, short term debt to total external debt, interest payments on short term external debt, central government debt as percentage of GDP, private non-guaranteed external debt stocks, public and publicly guaranteed external debt stocks, bank non-performing loans to total gross loans, total debt service percent of exports, interest payments on long term external debt, total external debt to GDP, external debt to exports, international reserve to total external debt, short term debt to GDP, total external debt to total reserves, short term debt international reserve growth, real domestic credit growth, interest payments on external debt to international reserve, interest payment on short term debt to GDP
Current Account	Exchange rate, export, import, current account balance, export growth rate, import growth, current account to GDP, Deviations of real exchange rate from trend
International Variables	Foreign exchange reserves, use of IMF credit, portfolio equity net inflows, net ODA received
Financial Liberalization	Deposit rate, bank liquid reserve to bank asset ratio, domestic credit provided by banking sector percent of GDP, deposit insurance, interest rate spread, risk premium on lending, S&P global equity indices (annual percentage change), stocks traded (total value), stocks traded turnover ratio, international reserve growth

Source	Set of Variables
Other Financial Variables	Money supply, return on equity, liquid asset to total asset, treasury bill rate, liquid asset to short term liabilities, non performing loans to total gross loans, return on asset, sectoral distribution of total loans(deposit takers), sectoral distribution of total loans(residents), bank capital to asset ratio, inflation volatility, change in terms of trade, the ratio of M2 reserve to international reserve
Real Sector	Industrial production, GDP, unemployment rate, GDP growth, trade in services percent of GDP, real interest rate, inflation, GDP per capita, gross saving percent of GDP
Institutional variables	Capital adequacy ratio, degree of openness to international trade, financial requirement to total reserve
Fiscal Variables	Short term interest rates of government securities and government bonds, medium- log term government securities and government bonds, government revenue excluding grants percent of GDP, government expense percent of GDP, tax revenue percent of GDP, fiscal surplus to GDP
Housing Sector	Residential real estate prices, residential real estate loans to total loans, commercial real estate loans to total loans, house prices percentage change over previous period, house price ratio to rent ratio, house price to income ratio, household wealth and indebtedness as a percentage of nominal disposable income, housing expense in percent of GDP, building permits, gross residential loans, house price percentage change, housing completions, housing starts, net residential loans, nominal house price indices (2000=100), nominal house prices annual percentage change, number of housing transactions, real gross fixed investment in housing annual change, representative interest rates on new mortgage loans, residential mortgage debt to GDP ratio, total dwelling stock, total outstanding residential loans, rate of change of housing price to GDP growth, the ratio of real gross fixed investment in housing annual change to GDP growth, the ratio of representative interest rate on new mortgage loans to real interest rate

Table 4.2: Countries used as data source

Countries
Austria, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Netherlands, Poland, Portugal, Romania, Slovakia, Spain, Sweden, United Kingdom, Iceland, Norway, Russia, Turkey, Ukraine and United States

While grouping the above variables, the categorization is done according to prior studies [Kaminsky and Reinhart(1998)]. This data set includes the period between 1999 and 2010 for the countries given in Table A.1:

Housing sector and its components are among those factors, weights of which started overriding the effects of other variables in causing crises after 90s. Also their impact on the economy gradually increase [Goodhart(2008)]. Since the aim is constructing a framework to study the role of the housing market, which is capable of predicting different crisis, related factors identified by previous studies plus housing market data will be included in the analysis.

Multinomial logit regression is the methodology, which is used in constructing the crisis prediction framework. In this model, the system needs to use sufficiently enough data to make prediction.

As mentioned above, with the inclusion of housing variables, the analysis narrows down in terms of both the countries and time interval in both cases of time windows used ('t-1, t, t+1' case and 't, t+1, t+2' case). The number of observations used in the analysis decreased from 1942 observations to 362 observations for 't-1, t, t+1' case and to 332 observations in 't, t+1, t+2' case when housing market indicators are added into the analysis.

The dependent variable of the regression indicates the state of the economy identified for three different cases as pre-crisis, crisis and tranquil periods. The dependent variable is generated over a crisis indicator variable, which takes the value of 1 for

the years with crises and value of 0 otherwise. (The details of the construction of time windows and the dependent variables are given in Section 5.2). These periods of crisis, pre-crisis and tranquility are created in accordance with the previous studies in the crisis predicting literature [Glick(1999), Manasse(2009), Laeven(2008), Reinhart and Rogoff Online Resources].

While running the estimations, the approach proposed by Ciarlone et al (2005, 2006) is followed. In the first step separate multinomial logistic regressions have been run for each single variable on its own to check their statistical significance. The main estimation is run by only including those, which were proven to be significant at this first stage. In the estimation, some groups of these variables had exhibited similar properties in terms of their effects on the economy. As a result of these similarities, it was possible to create some sub-groups from these variables in accordance with the prior studies for currency, banking and debt crises [Berg and Pattillo(1999b), Bussiere and Fratzscher(2002), Kaminsky and Reinhart(1998), Demirguc-Kunt(1997), Manasse and Roubini(2003), Ciarlone and Trebeschi(2005), Ciarlone and Trebeschi(2006)].

After these groups were constructed, the multinomial logit regressions were run for each group for predicting crises. After this, for every group top three best performers, showing the highest variation passing through from tranquil period to pre-crisis period in terms of odds ratio, were selected for the next step. For some of these groups, the mlogit regression could not converge to a solution because of either insufficient number of observations or concavity problems. Such groups have been further divided into smaller sub-groups until a successful mlogit run could be achieved. This method of creating groups and selecting best performers for the next step continued up until reaching the final working combination of variables whose odds ratios showed the highest deviation while passing from tranquil period into the pre-crisis period with in 95% confidence interval.

Finally, all sub-groups' best performers were put together to create the set of main independent variables to be used in the model for predicting crisis. In the analysis part the prediction was done according to two different time windows as 't-1, t, t+1' and 't, t+1, t+2' where t denotes the current year of concern. Since the estimation of these two models depended on different time windows, the independent variables, which were used in these models, are not same.

The final group, which is used in the case of time window 't-1, t, t+1' consists of the following variables: The ratio of FDI to GDP, total reserve growth, the ratio of domestic credit to private sector (% of GDP), trade in services (% of GDP), the ratio of domestic credit provided by banking sector (% of GDP), export growth, GDP growth, real domestic credit growth, bank capital to asset ratio, S&P global equity indices (annual % change), tax revenue (% of GDP), real interest rate, deposit rate, interest rate spread, representative interest rate on new mortgage loans, total outstanding residential loans, net residential loans.

The final group, which is used in the case of time window 't, t+1, t+2' is formed from the following variables: International reserve growth, the ratio of current account to GDP, total reserve growth, the ratio of FDI to GDP, change in terms of trade, domestic credit to private sector (% of GDP), domestic credit provided by banking sector (% of GDP), real domestic credit growth, import growth rate, degree of openness, export growth rate, S&P global equity indices (annual % change), tax revenue (% of GDP) , bank capital to asset ratio, interest rate of government securities & government bonds, bank liquid reserves to bank asset ratio, central government total debt (% of GDP), residential mortgage debt to GDP ratio, real gross fixed investment in housing annual change, total dwelling stock, housing completions, representative interest rate on new mortgage loans, nominal house prices annual percentage change, change of housing price to GDP growth.

4.2 Methodology

There are different kinds of estimation techniques in the early warning systems. These techniques are divided into two main groups as parametric and non-parametric techniques. Since the variables are assumed to be statistically independent, using a parametric estimation method is much more suitable for constructing an early warning system.

Logistic regression is one of the parametric techniques, which has many advantages as follows:

- Logistic regression allows properties of a linear regression model to be exploited.
- The logistic regression value can vary between $-\infty$ and $+\infty$. Although, the model coefficients' value change $-\infty$ and $+\infty$, the probability remains 0 and 1, by this way it gets easy to make interpretation and analysis depending on the data.
- The logit model can directly affect the odds ratio, as an advantage of this property, the changes in the model can be totally reflected to the ratio [Online Resources PennState University].

Logistic regression has two branches as binomial logistic regression and multinomial logistic regression. Binomial regression function analyzes the economy by separating it into crisis and adjustment periods. This method has been used in many studies to predict crisis [Ciarlone and Trebeschi(2005), Manasse and Roubini(2003), Davis and Karim(2008b)]. Although the models in these papers predicted many crises, they were ridden by the post crisis bias.

Post-crisis bias means that, while predicting a crisis, the independent variables are analyzed depending on their values during and directly after a crisis. However, the aim for constructing early warning systems is predicting crises before they hit

the economy. The most suitable way of doing this is comparing the variation of the variables before a crisis compared to their values during tranquil periods, when their values are sustainable. But a binomial logit model compares the pre-crisis observation with that in both the tranquil periods and crisis/adjustment periods. That is, the binomial logit model does not differentiate the tranquil period from an adjustment period. This induces an important bias, as the variation of the independent variables is very different during tranquil times as compared to crisis/recovery period.

There are two ways to overcome this post-crisis bias. The first one is dropping all crisis/adjustment observations from the model and the second way is using a discrete dependent variable, which gives more than two outcomes: multinomial logistic regression [Bussiere and Fratzscher(2002)].

The setup of the multinomial regression model is the same as that in logistic regression; the key difference is that, the dependent variable of logistic regression is formed from two outcomes whereas the dependent variable of multinomial logistic regression consists of more than two possible outcomes.

In order to explain the logistic regression, logistic function has to be introduced as:

$$\pi(x) = \frac{e^{(\beta_0 + \beta_1 X_1 + e)}}{e^{(\beta_0 + \beta_1 X_1 + e)} + 1} = \frac{1}{e^{-(\beta_0 + \beta_1 X_1 + e)} + 1}, \quad (4.1)$$

$$g(x) = \ln \frac{\pi(x)}{1 - \pi(x)} = \beta_0 + \beta_1 X_1 + e, \quad (4.2)$$

$$\pi(x)/1 - \pi(x) = e^{\beta_0 + \beta_1 X_1 + e} \quad (4.3)$$

The logistic function takes values from negative infinity to positive infinity and the output value changes between 0 and 1.

In these equations, $g(x)$ represents the logit function of given predictor X. ln present the natural logarithm, $\pi(x)$ gives the probability of being in a case, β_0 is the intercept from the linear regression equation. $\beta_1 X_1$ is the regression coefficient multiplied by some value of the predictor, and base e means the exponential function. The e in the linear regression equation stands for the error term.

To apply a logistic regression model, a series of N observed data points is needed. Each data point i , consists of a set of M explanatory variables from $x_{1,i}$ to $x_{M,i}$ and associated dependent variable Y_i .

In this model, it is assumed that the dependent variable Y is a random variable distributed according to the Bernoulli distribution. Each outcome of the dependent variable is determined by an unobserved probability p_i , which is special to the outcome itself but also connected to the explanatory variables as well. The overall picture for the logistic regression can be explained by the following equations [Greene(2003)]:

$$\begin{aligned}
Y_i | x_{1,i}, \dots, x_{m,i} &\sim \text{Bernoulli}(p_i) \\
\mathbb{E}[Y_i | x_{1,i}, \dots, x_{m,i}] &= p_i \\
Pr(Y_i | x_{1,i}, \dots, x_{m,i}) &= \begin{cases} p_i, & \text{if } Y_i = 1 \\ 1 - p_i & \text{if } Y_i = 0 \end{cases} \quad (4.4) \\
Pr(Y_i | x_{1,i}, \dots, x_{m,i}) &= p_i^{Y_i} (1 - p_i)^{(1-Y_i)}
\end{aligned}$$

The logistic regression can be designed by modeling the probability value of p_i using a linear predictor function. Hence, p_i will be a linear combination of the explanatory variables and a set of regression coefficients that are specific to the model.

The predictor function $f(i)$ for a data point i will be in the following form:

$$f(i) = \beta_0 + \beta_1 x_{1,i} + \dots + \beta_M x_{m,i} \quad (4.5)$$

$\beta_0, \dots, \beta_M$ are the regression coefficients and each gives the relative impact of a particular explanatory variable on the dependent variable. If,

- the regression coefficients $\beta_0, \beta_1, \dots, \beta_k$ are grouped into a single vector $\boldsymbol{\beta}$ of size $k + 1$;
- for each data point I , an additional explanatory variable $x_{0,i}$ is added, with a fixed value of 1, similar to the intercept coefficient β_0 ;
- the explanatory variables $x_{0,i}, x_{1,i}, \dots, x_{k,i}$ are grouped into a single vector $\mathbf{X}_i$ of size $k + 1$

then the linear predictor function turns into

$$f(i) = \boldsymbol{\beta} \cdot \mathbf{X}_i \quad (4.6)$$

The basic setup of the multinomial logistic regression is similar to that of the logistic regression. However, the dependent variable for a multinomial logistic regression is a categorical variable that is, it has more than two discrete possible outcomes. For multinomial logistic regression the probability of observation i having the outcome k is given by the linear predictor function $f(k, i)$ as following:

$$f(k, i) = \beta_{0,k} + \beta_{1,k}x_{1,i} + \beta_{2,k}x_{2,i} + \dots + \beta_{M,k}x_{M,i} \quad (4.7)$$

$\beta_{m,k}$ is the regression coefficient that relates the m^{th} explanatory variable with the k_{th} dependent variable outcome. As with the same in the logistic regression function, the predictor function can be written as;

$$f(i) = \boldsymbol{\beta}_k \cdot \mathbf{X}_i \quad (4.8)$$

$\boldsymbol{\beta}_k$ is the set of regression coefficients related with outcome k and $\mathbf{x}_i$ is the set explanatory variables related with observation i .

To clarify the multinomial logistic model, one can assume that the multinomial logistic regression for a dependent variable with K different possible outcomes is like running a series of K-1 independent binomial logistic regressions in which one outcome is chosen as the base and the rest K-1 outcomes are separately regressed relative to the base outcome. In the case of the last outcome K being selected as the base, the probabilities would be calculated as follows:

$$\begin{aligned}
\ln \frac{Pr(Y_i = 1)}{Pr(Y_i = K)} &= \beta_1 \cdot \mathbf{X}_i \\
\ln \frac{Pr(Y_i = 2)}{Pr(Y_i = K)} &= \beta_2 \cdot \mathbf{X}_i \\
&\dots \\
\ln \frac{Pr(Y_i = K-1)}{Pr(Y_i = K)} &= \beta_{K-1} \cdot \mathbf{X}_i
\end{aligned} \tag{4.9}$$

Separate set of regression coefficients are introduced, one for each possible outcome. After exponentiation of both sides of the equations and solving, the resulting probabilities would be:

$$\begin{aligned}
Pr(Y_i = 1) &= Pr(Y_i = K) e^{\beta_1 \cdot \mathbf{X}_i} \\
Pr(Y_i = 2) &= Pr(Y_i = K) e^{\beta_2 \cdot \mathbf{X}_i} \\
&\dots \\
Pr(Y_i = K-1) &= Pr(Y_i = K) e^{\beta_{K-1} \cdot \mathbf{X}_i}
\end{aligned} \tag{4.10}$$

Since the sum of the K probabilities is equal to 1, the probability of the base outcome will be:

$$Pr(Y_i = K) = \frac{1}{1 + \sum_{k=1}^{K-1} e^{\beta_k \cdot \mathbf{X}_i}} \tag{4.11}$$

The rest of the probabilities can be found as follows:

$$\begin{aligned}
Pr(Y_i = 1) &= \frac{e^{\beta_1 \cdot \mathbf{X}_i}}{1 + \sum_{k=1}^{K-1} e^{\beta'_k \cdot \mathbf{X}_i}} \\
Pr(Y_i = 2) &= \frac{e^{\beta_2 \cdot \mathbf{X}_i}}{1 + \sum_{k=1}^{K-1} e^{\beta'_k \cdot \mathbf{X}_i}} \\
&\dots \\
Pr(Y_i = K - 1) &= \frac{e^{\beta_{K-1} \cdot \mathbf{X}_i}}{1 + \sum_{k=1}^{K-1} e^{\beta'_k \cdot \mathbf{X}_i}}
\end{aligned} \tag{4.12}$$

The multinomial logistic regression model has an important assumption, which is the independence of irrelevant alternatives. The independence of irrelevant alternatives assumption states that, the odds of preferring one class to another do not depend on the presence or absence of other irrelevant alternatives. This condition makes it possible to model the choice of K alternatives as a set of K-1 independent binary choices, in which one alternative is chosen as the base outcome and the other K-1 alternatives are compared against it.

In this study, the dependent variable, Y, for the multinomial logit model has 3 different discrete outcomes as 0, 1 and 2 as 0 corresponding to tranquil period, 1 corresponding to the pre-crisis period and 2 corresponding to the state of being in crisis (may be mentioned adjustment period as well). In such a case, the calculated probabilities for the possible outcomes can be given as follows:

$$\begin{aligned}
Pr(Y_{i,t} = 0) &= \frac{1}{1 + e^{X_{i,t}\beta_1} + e^{X_{i,t}\beta_2}} \\
Pr(Y_{i,t} = 1) &= \frac{e^{X_{i,t}\beta_1}}{1 + e^{X_{i,t}\beta_1} + e^{X_{i,t}\beta_2}} \\
Pr(Y_{i,t} = 2) &= \frac{e^{X_{i,t}\beta_2}}{1 + e^{X_{i,t}\beta_1} + e^{X_{i,t}\beta_2}}
\end{aligned} \tag{4.13}$$

CHAPTER 5

ANALYSIS

5.1 Event Study Analysis

The analysis started with 111 exogenous variables and the number of variables decreased to 17 in the case of time window ‘t-1, t, t+1’ and to 24 in the case of time windows ‘t, t+1, t+2’ after a series of groupings and selecting the best three performers from all groups. The groupings have been done according to the related literature [Berg and Pattillo(1999b), Bussiere and Fratzscher(2002), Kaminsky and Reinhart(1998), Demirguc-Kunt(1997), Manasse and Roubini(2003), Ciarlone and Trebeschi(2005)]. Although, the literature was successful in predicting more than a half of the crises, it is still necessary to check the behavior of each exogenous variable before, during and after the crises periods in order to make sure that they are the indicator variables, variations of which signals the coming of a crisis.

By this analysis, the best performers’ behaviors, which are selected by the grouping procedure, are analyzed for a 7-year time interval. If the variables behave differently between the tranquil period and pre-crisis period, it will prove that the variables are suitable to use in predicting crisis. If they follow the same path in both tranquil and

pre-crisis period, they have to be dropped from the set [Ciarlone and Trebeschi(2005), Ciarlone and Trebeschi(2006)]

To analyze the behavior of the exogenous variables the event study analysis will be used, which is a technical tool that helps to show the variations in the behavior of the variables. The multinomial logistic regression is applied for two different time windows of three consecutive years as ‘t-1, t, t+1’ and ‘t, t+1, t+2’ where t denotes the current year. As a result of this, two separate event study analyses have been done for the exogenous variables.

The case of time window ‘t-1, t, t+1’ is given in Figure 5.1 to Figure 5.3 and the case of time window ‘t, t+1, t+2’ is given in Figure 5.4 to Figure 5.7. In the graphs, the bold horizontal line presents the average value of the variable during tranquil periods, and the solid lines stand for the average value of the variable in the interval around the crisis periods, with 3 years before and 3 years after the crisis, among with a 95% confidence interval denoted by the two dotted lines. The time interval for the event study starts from ‘t-3’ and ends at ‘t+3’. The indicator ‘t’ stands for the years in which the crisis indicator variable (which is designed to capture banking, currency and debt crises) is equal to 1 for the first time, which signals the start of a crisis. The event study analysis of the case of ‘t-1, t, t+1’ suggests the following:

- Real GDP growth rate is not only significantly lower during pre-crisis years than the tranquil periods’ average, but it follows the same worsening path during a crisis period as well and it maintains the same downward motion until ‘t+1’. This period involves the shifting from an average of 4 percent to minus 1 percent. Then the recovery period begins.
- Domestic credit to private sector (% of GDP) diverges from the tranquil periods’ average from the beginning of the pre-crisis period; it reaches its peak level just before the crisis and returns to the tranquil average at the end of the analyzed time period ‘t+3’.

- Trade in services (% of GDP) begins with 28 percent, which is above the tranquil periods' average of 25 percent. Then it deteriorates just before the crisis, at 't-1', and it continues this slow-down until the end of the time interval, where the variable reaches 20 percent.
- Real interest rate starts with 2 units, follows a rapid increase starting from 't-1', this continues during the crisis period, and it reaches to 6 units, which is above the tranquil periods' average of 3.8. After 't+1' the real interest rate falls down towards the tranquil periods' average.
- Domestic credit provided by banking sector (% of GDP) starts above the tranquil periods' average and it continues increasing. It reaches to its peak level, which is nearly 150, at the time of 't-1' and follows a downtrend and it finishes the time interval below its starting point.
- Export growth rate starts falling down at 't-1'. It falls down from 0.15 and reaches to its minimum, which is -0.05, at 't+1'. Then it attains its average level near the end of the time interval.
- Total reserve growth starts below its tranquil periods' average of 0.1. It falls below its tranquil periods' average at the start of the crisis, 't', and then the recovery procedure begins and it reaches to 0.25 at 't+3'.
- FDI to GDP starts with the tranquil periods' average of 0.2 and it starts falling down by 't-1' and it maintains this throughout the crisis and the recovery starts a while later than the crisis period, 't'. But it remains below its starting point at the end of the time interval.
- Deposit rate starts with the tranquil periods' average and then it starts increasing at 't-3'. It follows this trend until 't+1' and then its rate of increase gets augmented until 't+2'. Then it attains a downward slope towards its tranquil

periods' average. But at the end of the interval, the deposit rate is still above the tranquil periods' average.

- Interest rate spread starts with a downward move until 't-1'. Then it begins increasing at 't-1' and this continues until 't'. It experiences a slight decrease starting with the crisis. Its value remains the same until 't+3', where it is still above the tranquil periods' average.
- S&P global equity indices (annual % change) starts below the tranquil periods' average. It increases during the period from 't-3' to 't-2' and then, approaching the crisis, a noticeable decline starts which continues until the crisis hits the economy. Then, it attains an upward trend moving above its tranquil periods' average. However, the instability continues and the variable reaches the end of the interval with a value, which is way below its tranquil average.
- Bank capital to asset ratio starts with a value of 7.2 percent ratio, which is below its tranquil periods' average of 7.6. It follows a downward up until 't+1'. Then, the recovery begins and the ratio reaches to its tranquil periods' average back again at the time of 't+2'.
- Tax revenue starts decreasing from a value of 22 percent at 't-3' until 't'. Just before the crisis hits the economy at the time of 't', the tax revenue is equal to the its tranquil periods' average of ~ 20 . Then it maintains it downward trend down to the value of 14 percent.
- The residential loans start above the tranquil periods' average by approximately 30000 units and then it decreases to the tranquil periods' average by the crisis period 't'. It begins increasing at the end of 't+2'.
- Representative interest rate is under the tranquil periods' average between 't-3' and 't-1'. It follows a slightly downward trend during this time interval. After

this falling period it keeps increasing up until 't+1'. However, it experiences a sharp decline and reaches to its minimum, at 't+2', 2 years after the crisis hits the economy.

- Total outstanding residential loans follow a sustainable trend with a very slight slope from 't-3' to 't+1', with a value of approximately 500000. After that, it attains a rather higher slope and ends up at around 1000000 where the tranquil periods' average is around 400000.
- Real domestic credit growth exhibits a sharp decline starting at 't-1' by the value of ~ 0.09 and it maintains this downward slope until it falls to 0.03 at 't+1'. Despite the slight increase seen between 't+1' and 't+2', the value of the variable reaches to its minimum of ~ 0.02 by 't+3'.

All variables, which are chosen as the best performers, varied through tranquil period to pre-crisis period, so the multinomial logistic regression will include all variables mentioned above for 't-1, t, t+1'.

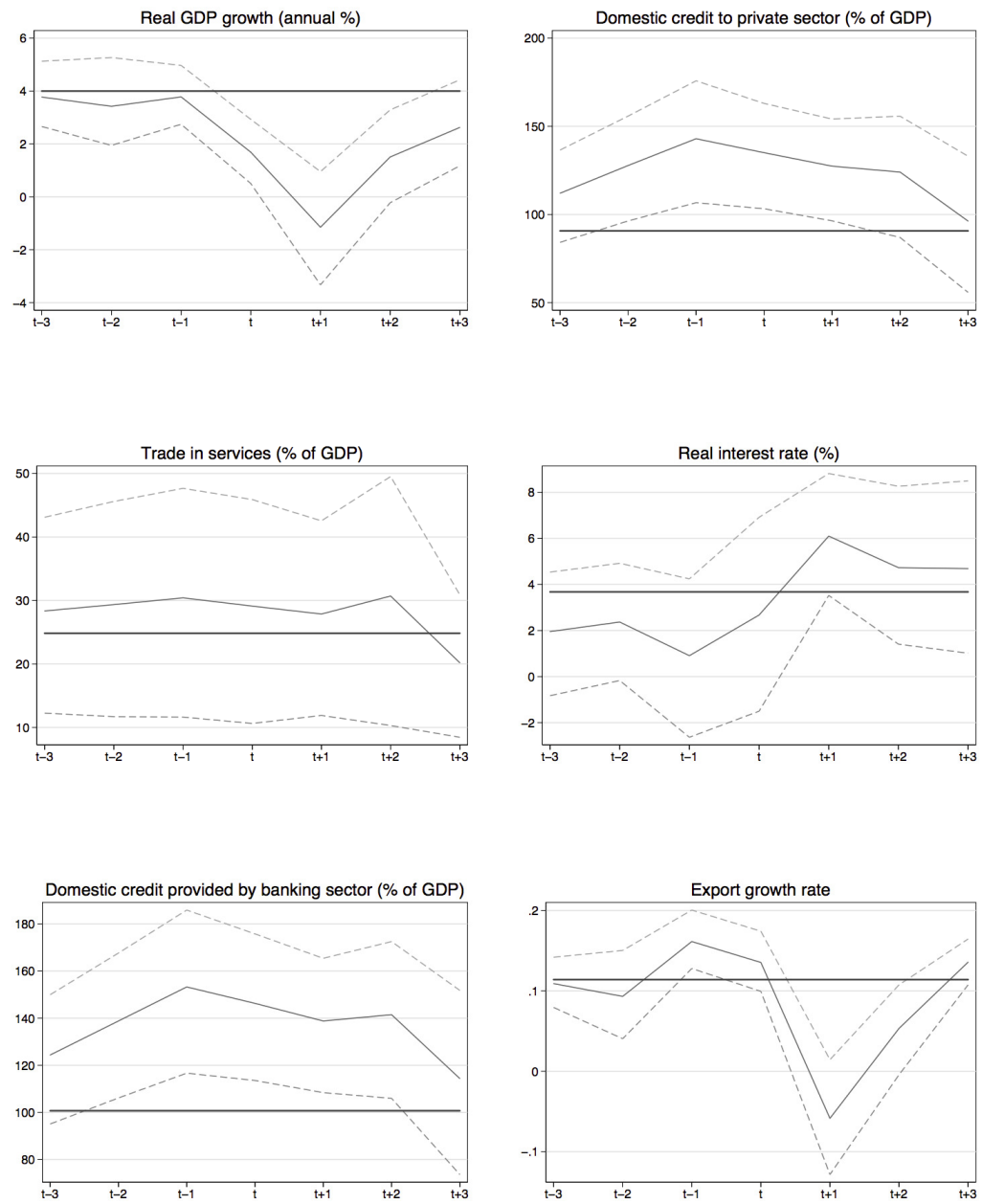


Figure 5.1: Event Study Analysis for time window ' $t-1, t, t+1$ ' - I

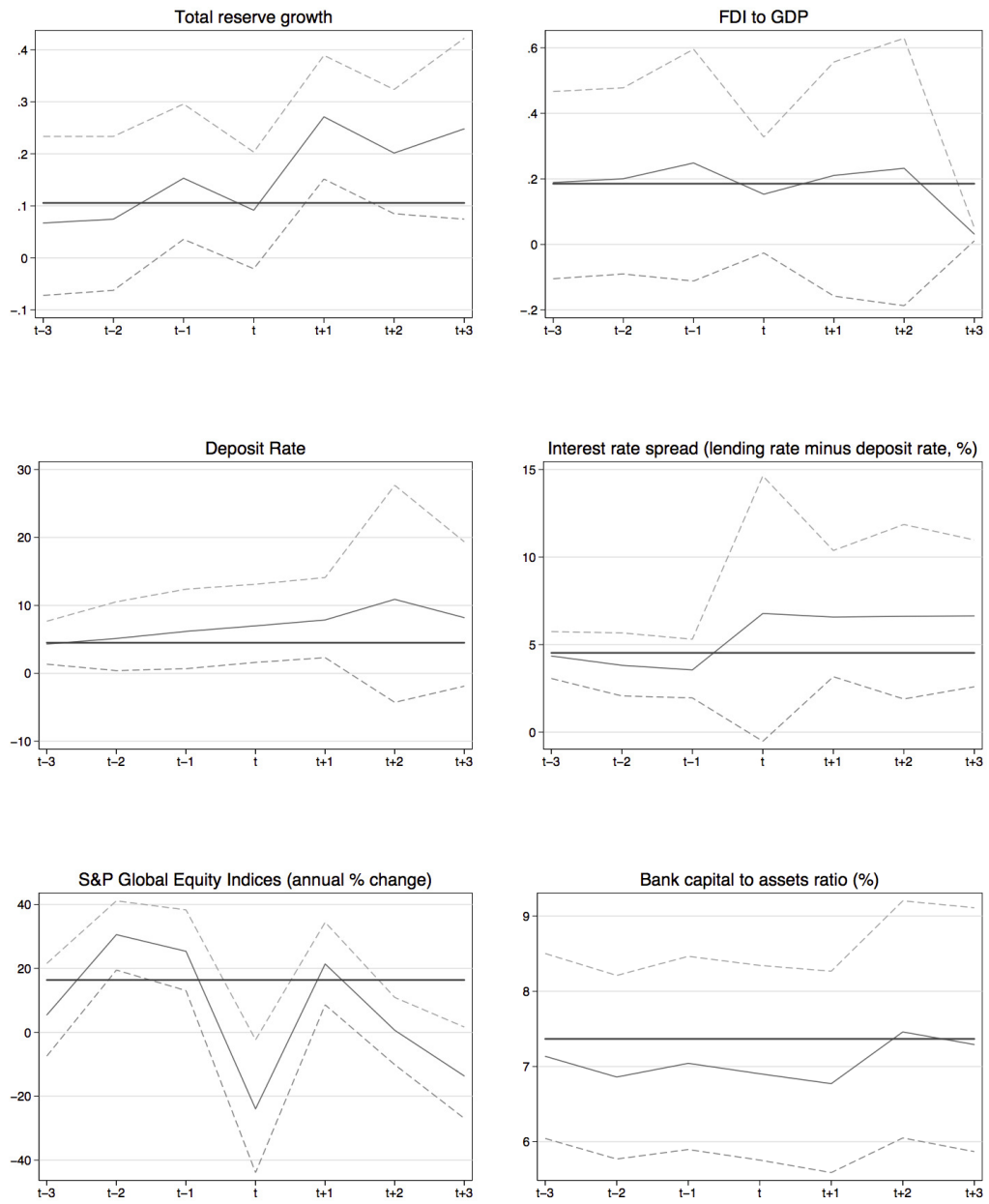


Figure 5.2: Event Study Analysis for time window ' $t-1, t, t+1$ ' - II

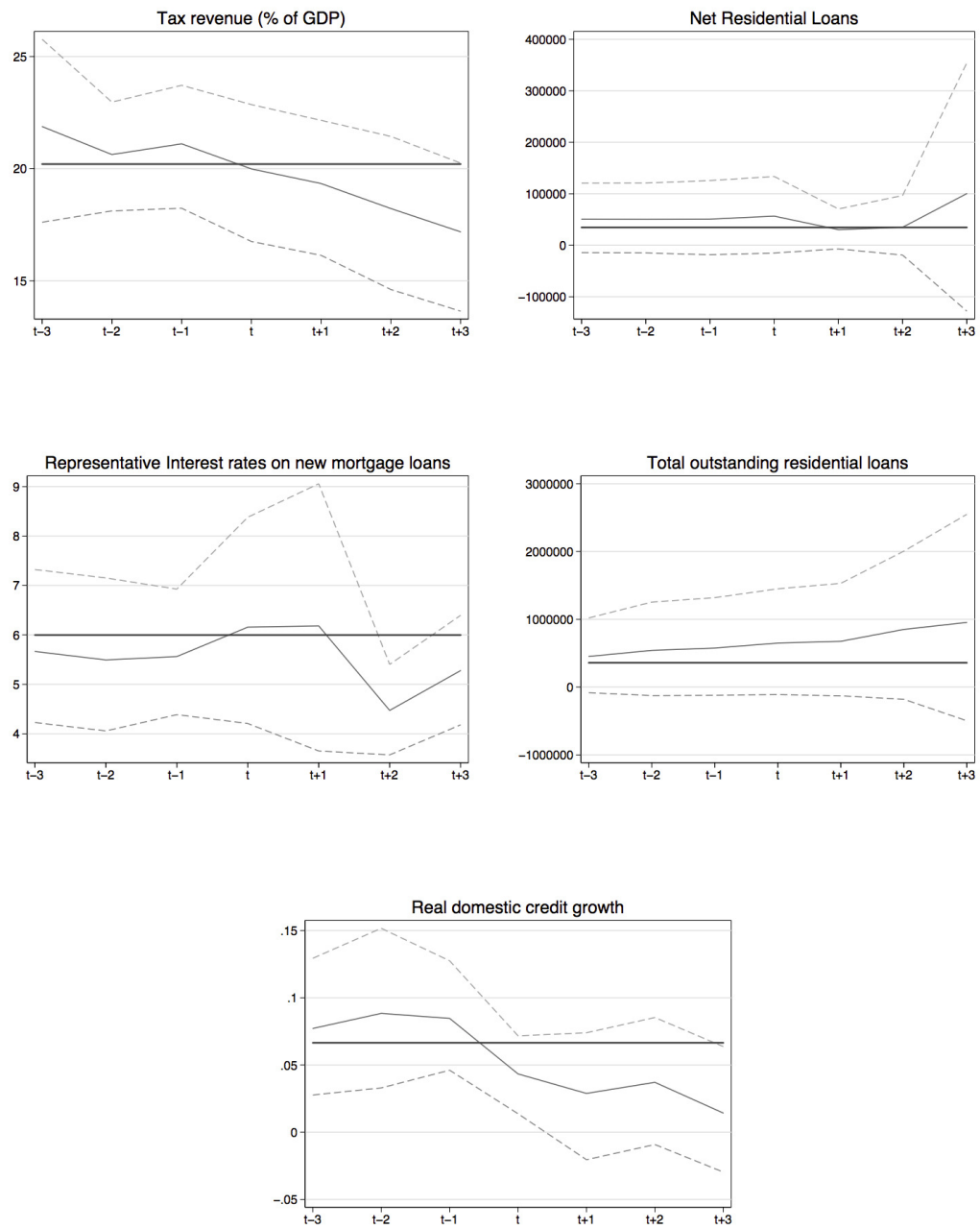


Figure 5.3: Event Study Analysis for time window ' $t-1, t, t+1$ ' - III

The case of ‘t-1, t, t+1’ and the case of ‘t, t+1, t+2’ are different from each other regarding the definitions of the pre-crisis, crisis and tranquil periods. The explanatory variables used in these models and their impacts on predicting crises are different from each other as well.

As a result of this, the event study analysis for the case of ‘t, t+1, t+2’ suggests the following:

- Domestic credit to private sector starts above its tranquil periods’ average and even exhibits a rapid increase between ‘t-3’ and ‘t-1’. After this point, a negative movement starts and decline sharpens after ‘t+2’ and the value approaches to its tranquil mean at ‘t+3’.
- Bank liquid reserve to bank asset ratio starts already below its tranquil periods’ value and decreases until ‘t+1’. Then starts a sharp increase, which takes the value even higher than the tranquil mean at the end of the interval.
- Domestic credit provided by banking sector begins on 118, which is above the tranquil period’s average of ~ 95 . Then it maintains an upward trend between ‘t-3’ and ‘t-1’. After that, with a slight decrease starting at ‘t-1’, it maintains a rather flat behavior until it experiences a sharp decrease by ‘t+2’ which takes the value closer to its tranquil mean at ‘t+3’.
- Degree of openness start with 0.65, which is below the tranquil period’s average. It has a downward trend in the period between ‘t-1’ and ‘t+1’. Then it enters into a recovery period, however at ‘t+3’, it is still below its starting value of 0.65, which is also below its tranquil periods’ average.
- Export growth rate begins below the tranquil periods’ average, and then it has a sharp fall between ‘t-1’ and ‘t+1’, in which it falls way below the tranquil average of 0.12 down to 0. Then it enters the recovery period, and finishes at 0.9 but it cannot reach to the tranquil periods’ average.

- Total reserve growth begins with ~ 0.14 which is slightly above the tranquil periods' average of 0.12. Then, it starts decreasing at 't-1', just the year before the crisis. It maintains this downward trend until the crisis, 't' and then it enters into a recovery and attains the value of 0.25 at 't+3', three years after the crisis.
- FDI to GDP starts at a point at which it is slightly above the tranquil periods' average. Then it starts decreasing by the year just before the crisis and reaches to its minimum at 't', the start of the crisis. Then it keeps increasing during the period from 't' to 't+2'. However, the instability continues and it almost becomes 0 at 't+3', which is below its tranquil mean.
- Central government debt begins above the tranquil periods' average. When the crisis hits the economy at 't', it starts increasing from 50 percent to 60 percent within the year following the crisis. By 't+3', it is still above the tranquil periods' average.
- S&P global equity indices falls from nearly 30 percent to minus 10 percent in the pre-crisis period. Then the recovery starts and in this period, it increases to 10 percent, which is a little bit below its tranquil periods' average. Then, it experiences a sharp fall again.
- Bank capital to asset ratio decreases in the period between 't-1' and 't+1'. Then it gets into the recovery period and returns to and exceeds its average level of 7.6 between 't+2' and 't+3'.
- Tax revenue starts at about 21, which is above the tranquil periods' average of 20. Then it enters into a downfall by 't-1', the year just before the crisis. It falls by an amount of nearly 5 percent until 't+3'. It finishes at 14 percent, which is far below the tranquil periods' average.

- Interest rate of government securities and government bonds is generally below the tranquil periods' average, although it approaches to the average between 't-2' and 't+1'. Then it enters into a falling period and starts increasing again 2 years after the crisis, but it is still below its average value at 't+3'.
- Housing completions decrease by 50000 units in the period between 't-1' and 't+1'. After this period, it attains an upward trend.
- Nominal house prices annual percentage change reacts to the upcoming crisis with a very sharp fall from nearly 14 percent to 4 percent until the crisis hits the economy. After the crash, it enters into the recovery period but after the crisis, it is seen that the impact is still effective on the nominal house prices annual percentage change because it still cannot reach to its tranquil periods' average at 't+3'.
- Real gross fixed investment in housing responds to the crisis earlier than its crash. It decreases from 7 percent to minus 9 percent during the period between 't-2' and 't+2'. The downward trend continues during the adjustment period and a strong recovery is observed starting from 't+2', two years after the crisis.
- Representative interest rates on new mortgage loans increases in the pre-crisis period up until the crisis hits the economy. Then, by the beginning of the recovery period it falls below the tranquil periods' average by reaching the minimum at 't+2' and returns from that point to try attaining the tranquil mean again.
- Residential mortgage debt to GDP ratio diverges from the tranquil period's average. The residential mortgage debt to GDP ratio starts from nearly 38 percent at 't-3' and at the end of 't+2' it becomes 58 percent. After 't+2', a recovery is observed and the value starts decreasing to approach to the tranquil mean, however, at 't+3', it is still nearly at 39 percent, which is higher than

the tranquil average of 25 percent. This signifies that impact of crisis is still effective on the variable even 3 years after the crisis.

- Total dwelling stock decreases by nearly 2000 units between the pre-crisis and the crisis period. After the recovery period, there occurs a huge rise in the amount of dwelling stock, approximately by 10000 units until the end of the recovery period.
- Import growth rate is also affected from the crisis. Its tranquil periods' average is equal to 0.12. There is a sharp decline during the period between 't-1' and 't+1' from a value of 0.19 to 0. Only then, the recovery begins, and the value reaches to its tranquil mean again at 't+3'.
- International reserve growth follows a stable but very slight downward trend, which is always lower than the tranquil periods' average. After the crisis bursts, it starts following an upward trend till the end of the period it slightly exceeds the pre-crisis period's value.
- Current account to GDP variable's tranquil periods' average is -0.01 and it is worsening from the very beginning of the pre-crisis period. It falls down to -0.03 at 't-1'. After a recovery period with some fluctuations, it manages to reach to its average value at 't+3' again.
- Real domestic credit growth's tranquil periods' average is 0.06. The downward motion starts at 't-2', however, it is much more significant in the period between 't-1' to 't+1' where the value decreases from ~ 0.09 to 0 moving below the tranquil average. For the rest of the recovery period, the fluctuations continue for making it possible to attain the average value again.

- Change in terms of trade is generally close to its tranquil mean in the pre-crisis period. However, in the recovery times, there occur some significant fluctuations in the variable due to the effects of the crisis.
- Rate of change of housing price to GDP growth ratio responds the crisis by falling down in the period between 't-3' and 't+1', with a decrease from 11 percent to -4 percent. Then in the adjustment period between 't+2' and 't+3', it reaches to its tranquil mean, which is about 2.

All variables, which are chosen as the best performers, varied through tranquil period to pre-crisis period, so the multinomial logistic regression will include all variables mentioned above for 't, t+1, t+2'.

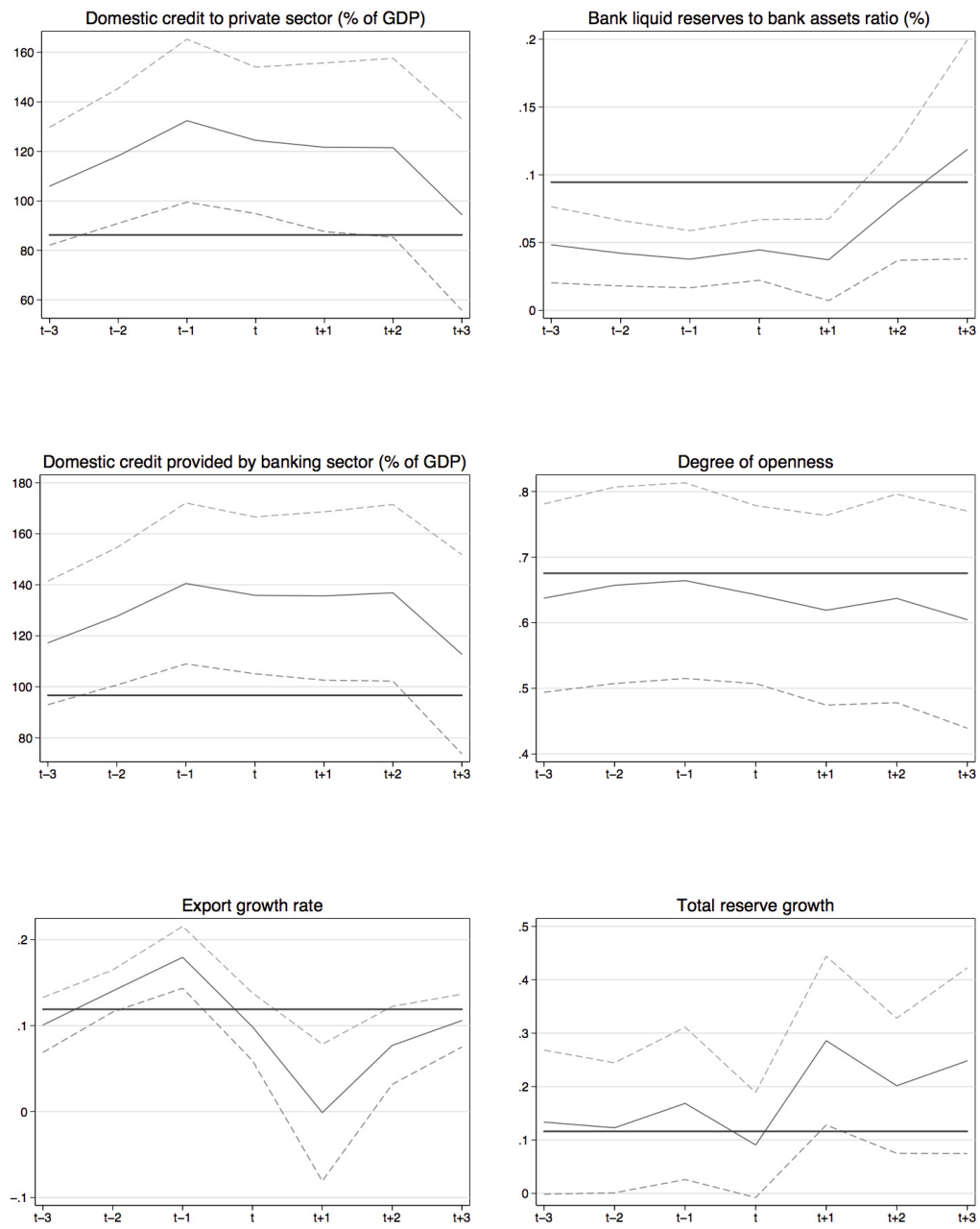


Figure 5.4: Event Study Analysis for time window ' $t, t+1, t+2$ ' - I

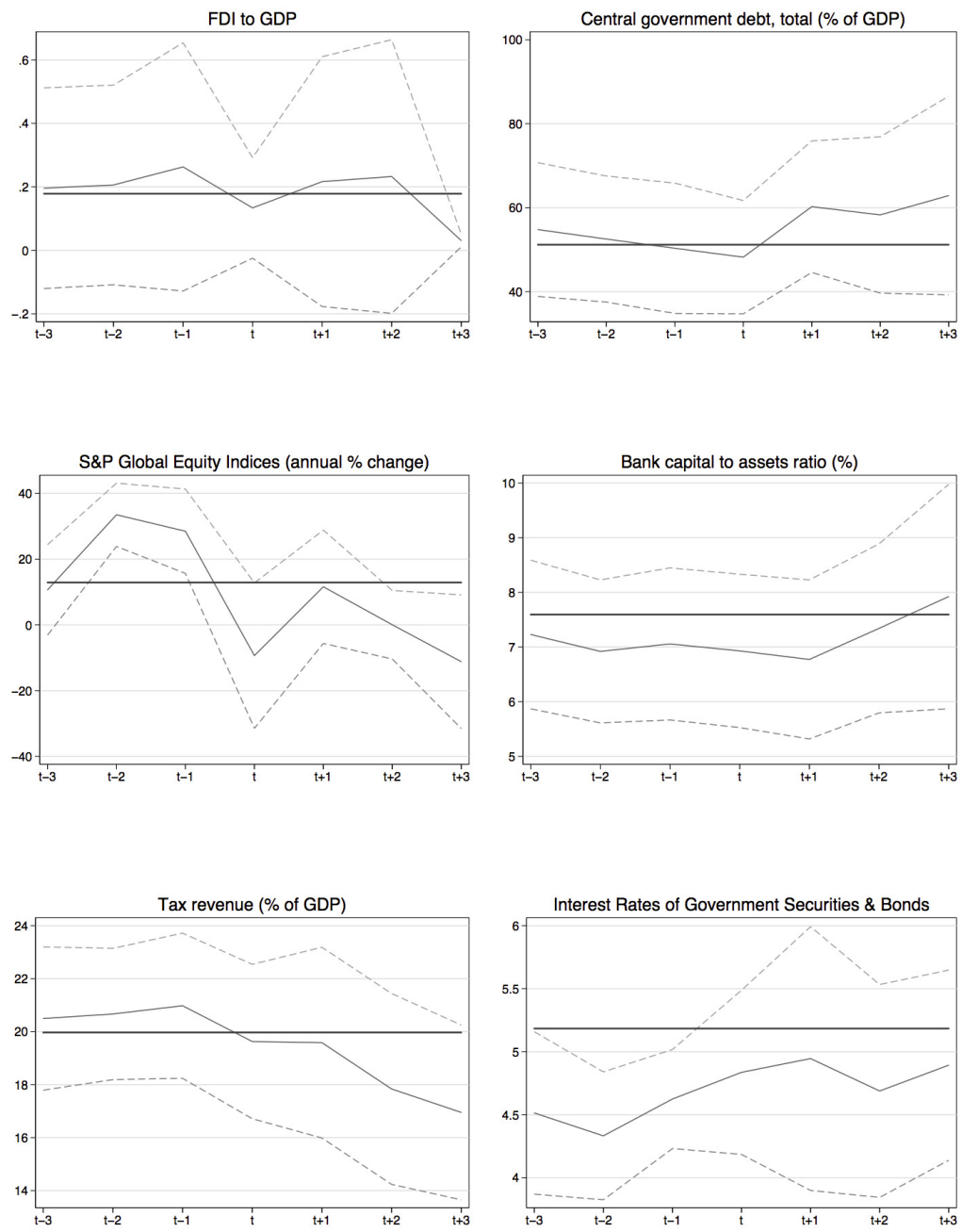


Figure 5.5: Event Study Analysis for time window ' $t, t+1, t+2$ ' - II

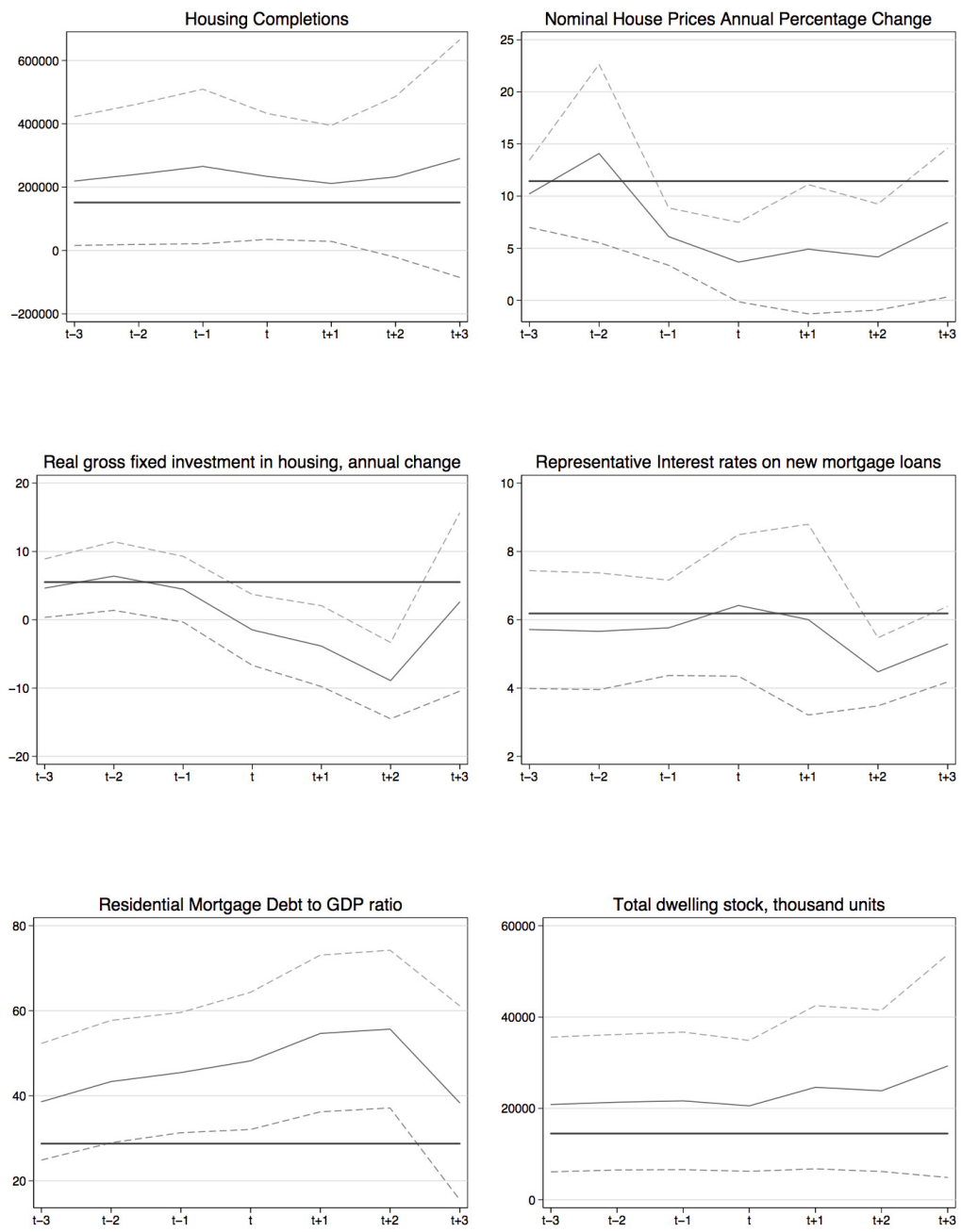


Figure 5.6: Event Study Analysis for time window ' $t, t+1, t+2$ ' - III

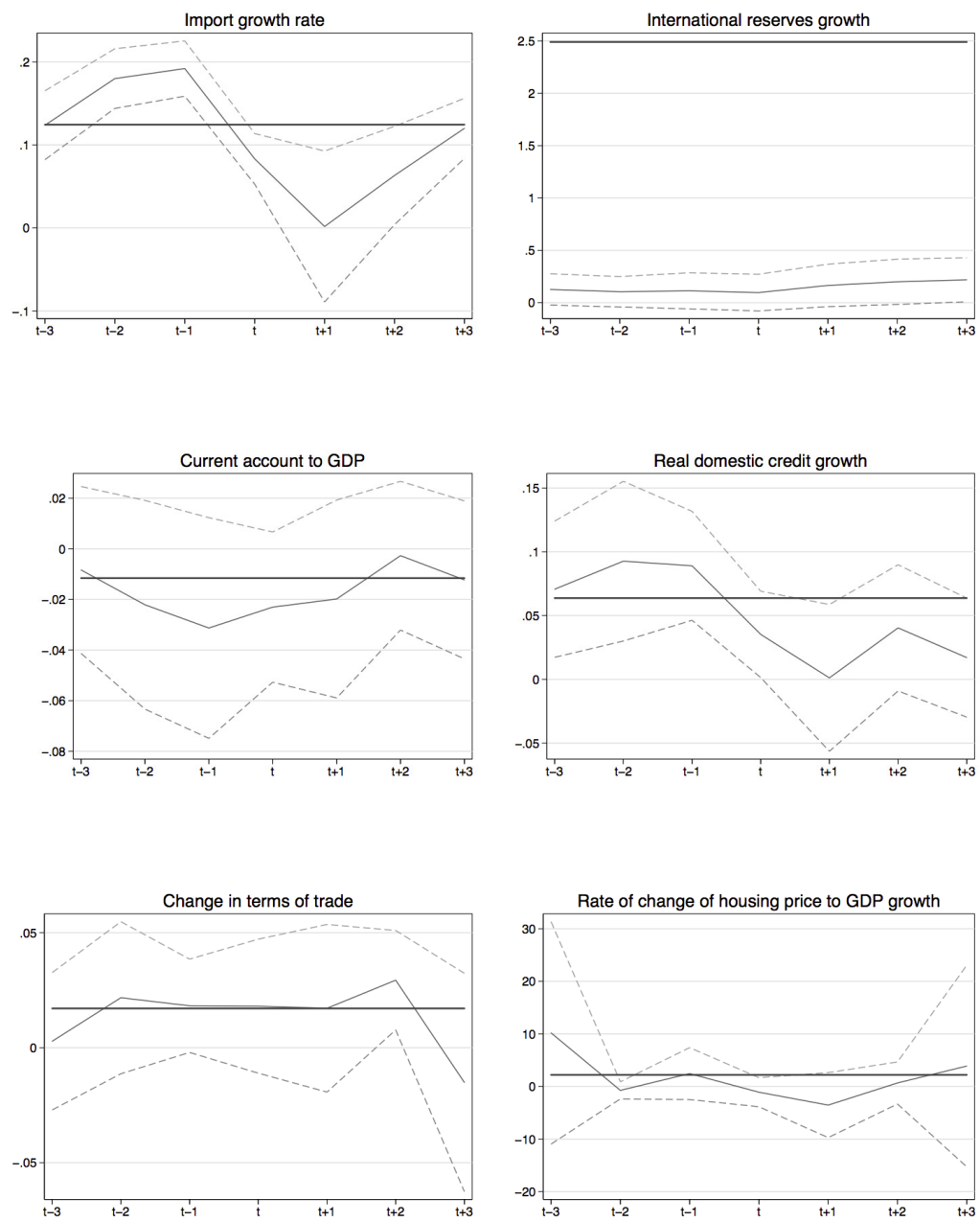


Figure 5.7: Event Study Analysis for time window 't, t+1, t+2' - IV

5.2 Empirical Analysis

In choosing which variables are more appropriate to be used for crisis prediction, the multinomial logistic regression, which was firstly applied by Bussiere and Fratzscher (2002) to constitute a new approach to predict crisis, was used as an estimation tool [Bussiere and Fratzscher(2002)]. In applying this approach, the time interval is divided into three sub-periods as a tranquil period, a pre-crisis period and an adjustment period. In the tranquil period, the economy does not have a risk of crisis and the economic variables follow a predictable path. In the pre-crisis period, the economic variables start giving some signals about the coming crisis by deviating from their average path. In the adjustment period, the economy has already been hit by the crisis and measures are being taken for the recovery. Because of this, the economic variables start approaching to their tranquil regime values again and they follow a more predictable path.

This analysis differs from the literature by the fact that it focuses on the impact of the housing market on the crisis prediction. It uses two different techniques - in terms of the time windows they use for prediction - and investigates the prediction powers of these different techniques. It aims predicting not only one but all three types of crises as banking, currency and debt.

In the multinomial logistic regression, the probabilities for a country being in each of the mentioned three different periods is calculated as the following:

$$\begin{aligned} Pr(Y_{i,t} = 0) &= \frac{1}{1 + e^{X_{i,t}\beta_1} + e^{X_{i,t}\beta_2}} \\ Pr(Y_{i,t} = 1) &= \frac{e^{X_{i,t}\beta_1}}{1 + e^{X_{i,t}\beta_1} + e^{X_{i,t}\beta_2}} \\ Pr(Y_{i,t} = 2) &= \frac{e^{X_{i,t}\beta_2}}{1 + e^{X_{i,t}\beta_1} + e^{X_{i,t}\beta_2}} \end{aligned} \tag{5.1}$$

When Y is equal to 0, it implies that, the economy is in the tranquil period. If Y is equal to 1, it means that the economy is in the pre-crisis and in the case of Y being equal to 2, the economy is in the adjustment regime.

Here, β_1 and β_2 indicate the marginal impact of a change in the explanatory variables on the probability of being in pre-crisis or adjustment period relative to the probability of being in the tranquil period respectively as shown below:

$$\begin{aligned} \frac{Pr(Y_{i,t} = 1)}{Pr(Y_{i,t} = 0)} &= e^{X_{i,t}\beta_1} \\ \frac{Pr(Y_{i,t} = 2)}{Pr(Y_{i,t} = 0)} &= e^{X_{i,t}\beta_2} \end{aligned} \tag{5.2}$$

The multinomial logistic regression consists of three steps and while constructing these steps the methodology of Ciarlone et al was followed [Ciarlone and Trebeschi(2005), Ciarlone and Trebeschi(2006)]:

- First of all, the multinomial logistic regression was run for each of the 111 variables independently from each other. Although it is a fact that the multinomial logistic regression is meaningful if both the relation between the variables and their deviations are taken into consideration, the aim of this procedure was controlling the Wald statistics and the sign of the coefficient of each variable before beginning the group analysis. This step is helpful for determining the significance and the impact of the variables on the estimation. Indeed, there occurred some problematic cases in choosing the variables for constructing the model. There were some variables (such as fiscal surplus to GDP ratio), which were statistically significant and their predictive powers were very high. However, it has been realized in the final step that those variables would destroy the validity of the constructed model by enlarging its confidence interval and destroying the Wald statistics. As a result, this makes it impossible to construct

a set of variables which is suitable to run a multinomial logistic regression to predict crisis by using such variables. Hence, at the end of first stage, that kind of variables were dropped from the data set.

- The second stage includes the grouping of the variables, which come from a similar structure and whose deviations from their trend create similar results on the economy. These variables had already satisfied the requirements of the first stage, which means that they are significant in the 95% confidence interval and they satisfy the Wald statistics. Also they are suitable to form a group to run a multinomial logistic regression on. While grouping the variables, the groups were formed depending on previous studies and data sources [Glick(1999), Manasse(2009), Laeven(2008), Reinhart and Rogoff Online Resources]. The variables, which have already been used in previous studies, constituted the first main group, and the variables representing the housing market data constituted the second one. This grouping technique was preferred because the aim was to measure the impact of the housing market data on the prediction power of the early warning system.
- The best performers of each group have been selected according to their significance on the transition from the tranquil to pre-crisis periods. Then, a general multinomial logistic regression has been run over this final group of variables, which successfully have made their way through the previous steps by satisfying the mentioned requirements.

Since both the construction and the results in terms of prediction power are different for the two time windows, ‘t-1, t, t+1’ and ‘t, t+1, t+2’ cases will be analyzed separately in the following sub-sections.

Table 5.1: Regime definition for the multinomial logit model for $t-1$, t , $t+1$

States of the binomial crisis indicator			Regime in the multinomial
<i>At time $t-1$</i>	<i>At time t</i>	<i>At time $t+1$</i>	<i>Model at time t</i>
1	0	0	Tranquil ($Y = 0$)
0	0	0	
0	0	1	Pre-crisis ($Y = 1$)
1	1	0	Adjustment ($Y = 2$)
1	1	1	
0	1	0	
1	0	1	
0	1	1	

5.2.1 Empirical results for the time window: ‘ $t-1$, t , $t+1$ ’

As mentioned earlier, a crisis indicator variable has been defined for identifying the years with a crisis. This variable takes the value of 1 for the years with crisis and the value of 0 otherwise. The identification of the state of the economy as either tranquil, pre-crisis or adjustment is done according to the values that the crisis indicator variable takes in a time window of three consecutive years. For the case of ‘ $t-1$, t , $t+1$ ’, this window consists of the current year, the year before and the year after it. Given these explanations, there exist eight different combinations of values that the crisis indicator may take in a 3-year window. These combinations and their categorization as either tranquil, pre-crisis or adjustment period are given in Table 5.1.

As an example, assume that the current year ‘ t ’ is 2001. If it is known that a country had a crisis in 2001 and 2002 but none in 2000, the crisis indicator variable will take the values of 0, 1 and 1 consecutively for the 3-year time window of 2000-2002. This combination corresponds to the last row of Table 5.1. Hence, in the data set, the value of the dependent variable Y will become 2 for this given country in year 2001, corresponding to an adjustment period observation. Consequently, the crisis indicator variable, which has only two values as 0 and 1 is converted into a

categorical variable Y , with three different values as 0, 1 and 2. Hence, it becomes possible to apply a multinomial logistic regression by using this variable Y as the dependent variable.

As can be seen from the table, the pre-crisis period ($Y = 1$) is defined by a condition where the economy has not been in a crisis over the last two years but it has an economic problem in the year ahead. The tranquil period ($Y = 0$) is identified as the case in which the economy either has no crisis in the 3-year period at all, or had only one crisis at time $t-1$, the year before the current one. The rest of the cases are identified as adjustment periods.

The multinomial logistic model is forecasted by assuming the tranquil period as the benchmark (the base outcome). In the final regression, the model is constructed by 17 macroeconomic and financial variables, behaviors of which have been analyzed in the event study analysis. The entire sample includes the period from 1999 to 2011. The final regression has been run over 362 observations. It satisfies the 95% confidence interval and the threshold value used for the probabilities is 50% for determining tranquil, pre-crisis and adjustment periods.

The final multinomial logistic regression includes the following variables with each being given a short abbreviation:

- *FDItoGDP*: The ratio of FDI to GDP,
- *TotResGrwth*: Total reserve growth,
- *DmstcCrdttoPrvtSec*: The ratio of domestic credit to private sector (% of GDP),
- *TrdinServ*: Trade in services (% of GDP),
- *DmstcCrdtBnkSec*: The ratio of domestic credit provided by banking sector (% of GDP),
- *ExprtGrwth*: Export growth,
- *GDPGrwth*: GDP growth,
- *RealDmstcCrdtGrwth*: Real domestic credit growth,

- *BnkCaptoAsstRat*: Bank capital to asset ratio,
- *SPGlobEqInd*: S&P global equity indices (annual % change),
- *TaxRev*: Tax revenue (% of GDP),
- *RealIntrRat*: Real interest rate,
- *DpstRat*: Deposit rate,
- *IntrRatSprd*: Interest rate spread,
- *ReprIntrRatMortLoan*: Representative interest rate on new mortgage loans,
- *TotOutstdResidlLoan*: Total outstanding residential loans
- *NetResidLoan*: Net residential loans.

The following table summarizes the final estimation results of the general multinomial logistic regression (Table 5.2):

Table 5.2: Results of the multinomial logit regression for t-1, t, t+1

Variables	Coeff.	Std. Err.	Z-stat.	95% Conf. Intrvl	
<i>Pre-crisis (Y = 1)</i>					
FDItoGDP	-202.4851	10.22642	-19.80	-222.5285	-182.4416
TotResGrwth	-99.41258	3.612384	-27.52	-106.4927	-92.33244
DmstcCrdttoPrvtSec	6.795125	0.178791	38.01	6.444701	7.145548
TrdinServ	-0.4409177	0.1042907	-4.23	-0.6453238	-0.2365116
DmstcCrdtBnkSec	-5.035147	0.1806483	-27.87	-5.389211	-4.681082
ExprtGrwth	586.0838	23.32143	25.13	540.3747	631.793
GDPGrwth	12.40456	0.4107437	30.20	11.59952	13.2096
RealDmstcCrdtGrwth	-32.86714	11.66812	-2.82	-55.73623	-9.998036
BnkCaptoAsstRat	-9.483088	1.773855	-5.35	-12.95978	-6.006396
SPGlobEqInd	0.7620294	0.0265112	28.74	0.7100684	0.8139905
TaxRev	8.319374	0.4405457	18.88	7.455921	9.182828
RealIntrRat	-0.5358567	0.1505403	-3.56	-0.8309103	-0.240803
DpstRat	9.150635	0.7062367	12.96	7.766436	10.53483
IntrRatSprd	57.93009	1.292595	44.82	55.39665	60.46353
ReprIntrRatMortLoan	-6.008512	0.7048006	-8.53	-7.389896	-4.627128
TotOutstdResidlLoan	0.0000438	7.45E-06	5.88	0.0000292	0.0000584
NetResidLoan	0.0029007	0.0000483	60.06	0.0028061	0.0029954
"Const. "	-698.1562	13.34334	-52.32	-724.3086	-672.0037

In the first part of the Table 5.2, the determinant indicator for the Early warning system is the coefficient part of the analysis, which gives the marginal impact of every component on predicting crisis. The coefficients in the first part of the table represent the estimation result for β_1 , which gives information about the likelihood of the economy's entering into a crisis in the following year against the likelihood of its staying in the tranquil period.

This multinomial logistic regression is significant in the 95% confidence interval and the threshold value used for the probabilities in determining the tranquil, pre-crisis and adjustment periods is 50%.

Depending on the estimation results, exogenous indicators, which changes from tranquil period to pre-crisis period, are interpreted depending on their odds ratios, as follows:

- FDI to GDP ratio and real domestic credit growth affect entering into crisis very much as, if there is a decrease in these values, the probability of entering into the crisis increases. Also if these values do not increase, the country in crisis has trouble in exiting from the ongoing crisis.
- Total reserve growth's marginal effect on entering into a crisis is also important as, if there is a reduction in the total reserve of the countries, this will increase the probability of entering into a crisis and if there is a rise in the reserve amount of a country, this will ease its exit from the crisis.
- Domestic credit to private sector behaves different from the previous indicators as if there is an increase in the value of this variable, it will increase the probability of entering into the crisis and a decrease in this value indicates the probability of being in adjustment period.
- Trade in services and bank capital to asset ratio behave against the effect of domestic credit to private sector. Their decrease increases the chances of entering

into a crisis and if there is a recovery in these values, this will indicate being in the recovery period rather than the tranquil.

- Domestic credit provided by banking sector is a variable which is expected to behave similar to the domestic credit to private sector, but their effect are opposite; a reduction in the domestic credit provided by banking sector increases the probability of entering into a crisis and if this value recovers, this indicates a recovery from the crisis.
- Export growth rate and GDP growth rate affect the entering into and exiting from the crisis in the following way: if there is an increase in the export growth rate and GDP growth rate, they will increase the probability of entering into a crisis and a fall in these rates indicates recovery.
- Representative interest rates on new mortgage loans and real interest rate's marginal effect on entering into crisis change in the same manner as, if there is a reduction in these variables, this means that these variations are signals of a coming crisis.
- S&P global equity indices and tax revenue behave in the similar way on entering and exiting from the crisis, as the variation in these two variable accelerate the entering into the crisis, and they are also important in exiting from the crisis.
- Deposit rate and interest rate spread behave in the same manner in determining the probability of entering into and exiting from the crisis. It seems that, they both have positive marginal effects on the probability of entering into a crisis. This may be related to the concretionary and expansionary impact of these ratios on the general economy.
- Total outstanding residential loans indicates the effects of the housing markets. It seems to have a positive marginal effect on the probability of entering into

a crisis. Also its decrease will be an indicator of a recovery period rather than tranquility.

- Net residential loans represent the variations of the housing market and it has a positive marginal effect on the probability of entering into a crisis. The reason may be that, the countries whose residential loan amount is excessive are more fragile compared to the other countries. This fragility increases the probability of entering into a crisis.

Predictive Ability

If accurate and harmonious variables are chosen and used in the estimation, crisis prediction becomes an easy work but the success of predicting crisis determine the contribution of the work into the literature

The success of a crisis prediction depends on different factors. Defining a threshold level the determining factor of which cases are to be identified as crisis and which are not. Then, it is expected from the crisis prediction mechanism to give signal for the cases of crisis, probabilities of which are above the defined threshold level. Choosing a correct threshold level is very critical. If a rather lower threshold level is chosen, there may be some cases for which the early warning system gives signals of an upcoming crisis, although no such threat exists, that is some periods of tranquility or recovery may be falsely alarmed as a pre-crisis period. These types of cases are identified as Type 1 error cases. On the other hand, if the threshold level is set too high than required, then there may occur some pre-crisis periods, which are missed by the early warning system by not giving any signals. This will result in some pre-crisis periods to be considered as either tranquil or adjustment. This will obviously avoid taking the necessary actions required in a pre-crisis period and will severely impact the economy.

Hence, the threshold value should be selected by taking these requirements into consideration. For this analysis, in order to be on the safe side, the threshold value to be used has been defined as 50% for both the probability of being in pre-crisis and in recovery periods compared to tranquility.

While making the prediction the aim is to measure the effect of the housing market variables in crisis prediction performance. In the first, general multinomial logistic regression model without the housing market variables, the best performers, which are also used in the previous studies' data sets, are the following: The ratio of FDI to GDP, total reserve growth, the ratio of domestic credit to private sector (% of GDP), trade in services (% of GDP), the ratio of domestic credit provided by banking sector (% of GDP), export growth, GDP growth, real domestic credit growth, bank capital to asset ratio, S&P global equity indices (annual % change), tax revenue (% of GDP), real interest rate, deposit rate and interest rate spread.

The in sample performance of this early warning system, consisting of these mentioned variables, is given in the following table (Table 5.3):

Table 5.3: Performance of the model without housing data for t-1, t, t+1

<i>State</i>	<i>Correctly called</i>	<i>Not correctly called</i>	<i>Total</i>
<i>Tranquil</i>	107	3	110
<i>Pre-crisis</i>	3	2	5
<i>Adjustment</i>	10	3	13

This multinomial logistic regression uses a probability threshold value of 50% in predicting the tranquil, pre-crisis and adjustment periods. It correctly predicts the 60% of the in-sample crises, one year before they hit the economies. Two crises are missed by this model: 2005 France economic crisis and 2005 Ireland economic crisis.

The model does not send a false alarm, as there exists no adjustment or tranquil period observations for which a pre-crisis alarm is generated. In this model's observation set, there are 110 tranquil period observations ($Y = 0$) and the early warning

system predicts 107 of them in correctly. (110 cases for which $Y = 0$, and 107 cases out of them, for which both $Y = 0$ and $P_0 \geq 0.5$) Missed pre-crisis observations are considered as tranquil period observations in this model, which implies that although there is a crisis, the model accepts and estimates them as tranquil periods since although $Y = 1$, $P_0 \geq 0.5$. Also in this model's observation set, there are 13 adjustment period observations and the early warning system predicts 10 of them correctly. (13 cases in which $Y = 2$, for 10 out of them both $Y = 2$ and $P_2 \geq 0.5$)

After predicting crisis by using the best performers of the prior studies , a second multinomial logistic regression model is constructed by introducing the following housing market variables into the first best performers' data set to show the impact of this sector on predicting crisis. The best performers of the housing data are the following: Representative interest rates on new mortgage loans, total outstanding residential loans, and net residential loans. After combining both data sets the in-sample performance of multinomial logistic regression in predicting crisis is given in the following table (Table 5.4):

Table 5.4: Performance of the model with housing data for t-1, t, t+1

<i>State</i>	<i>Correctly called</i>	<i>Not correctly called</i>	<i>Total</i>
<i>Tranquil</i>	84	0	84
<i>Pre-crisis</i>	5	0	5
<i>Adjustment</i>	11	0	11

This new model predicts 100% of the in-sample crises one year before they hit the economies. The list of caught crises are given in Table 5.5.

The model does not send a false alarm, as there exists no adjustment or tranquil period observation for which a pre-crisis alarm is generated. In this model's observation set, there are 84 tranquil period observations ($Y = 0$) and the early warning system predicts 84 of them correctly. (84 cases where $Y = 0$, for all of them

Table 5.5: Crises correctly called with the model with housing data, t-1, t, t+1

<i>Country</i>	<i>Years</i>
France	2005
Ireland	2005
Latvia	2008
Netherlands	2008
Ukraine	2008

$P_0 \geq 0.5$) Also in this model's observation set, there are 11 adjustment period observations and the early warning system predicts 11 of them correctly. (11 cases for which $Y = 2$, for all of them $P_2 \geq 0.5$ as well.)

As obviously seen from the results, after taking housing market variables into account as crisis indicators, the model does not miss any crises. In addition to this, housing market variables also increase the model's success in predicting the tranquil and adjustment periods. As it is given, the model predicts all three cases with 100% success.

The publicly available housing market data span only the period from 1999 to 2010. Moreover, as given in Table 5.1 only one out of eight different cases are identified as pre-crisis ($Y = 1$) for the time window of 't-1, t, t+1'. As a result of these facts, the number of observation with $Y = 1$ is not large enough to be able to make an out-of-sample performance analysis for the time window of 't-1, t, t+1'. Only after enlarging the data set with the upcoming years' data, it will become possible to make such an out-of-sample performance analysis for these variables and this time window.

It has to be noted that, inclusion of the housing sector variables into the model decreases the total number of valid observations in the data set from 128 to 100 as can be seen from Tables 5.3 and 5.4. Hence, as a further means for the quantitative comparison of the two models, the first model constructed by the best performers of the prior studies has also been invoked in this reduced data set to evaluate its predictive performance. The results of this run are given in Table 5.6

Table 5.6: Performance of the model without housing data on the reduced data set for $t-1$, t , $t+1$

<i>State</i>	<i>Correctly called</i>	<i>Not correctly called</i>	<i>Total</i>
<i>Tranquil</i>	81	3	84
<i>Pre-crisis</i>	3	2	5
<i>Adjustment</i>	8	3	11

5.2.2 Empirical results for the time window: ‘ t , $t+1$, $t+2$ ’

For the case of ‘ $t-1$, t , $t+1$ ’, the identification of the state of the economy as either tranquil, pre-crisis or adjustment is again done according to the values that the crisis indicator variable takes in a time window of three consecutive years. However, for this case, the time window consists of the current year and consecutive 2 years following the current one. Again there exist eight different combinations of values that the crisis indicator may take in a 3-year window. These combinations and their categorization as either tranquil, pre-crisis or adjustment period are given in Table 5.7.

Table 5.7: Regime definition for the multinomial logit model for t , $t+1$, $t+2$

States of the binomial crisis indicator			Regime in the multinomial
<i>At time t</i>	<i>At time $t+1$</i>	<i>At time $t+2$</i>	<i>Model at time t</i>
0	0	0	Tranquil ($Y = 0$)
1	0	0	
0	0	1	Pre-crisis ($Y = 1$)
0	1	0	
0	1	1	
1	1	0	Adjustment ($Y = 2$)
1	1	1	
1	0	1	

Same as the previous one, tranquil period is denoted by $Y = 0$, pre-crisis is denoted by $Y = 1$ and the adjustment period is denoted by $Y = 2$. The cases corresponding to each period is given as follows: The cases in which there is no crisis in any of the

three years, or there is a crisis at time t , but no crises in $t+1$ and $t+2$ are considered as ‘tranquil periods’. The cases in which there is no crisis at time t , but there exists a crisis either in $t+1$ or $t+2$ or both are considered as periods of ‘pre-crisis’. The rest of the eight combinations are considered as adjustment or recovery periods, which means that a crisis has already hit the economy and the necessary adjustments are being done. The multinomial logistic model is forecasted by assuming the tranquil period as the benchmark (the base outcome). In the final regression, the model is constructed by 24 macroeconomic and financial variables, behaviors of which have been analyzed in the event study analysis. The entire sample includes the period from 1999 to 2011. The final regression has been run over 332 observations. The final multinomial logistic regression includes the following variables:

- *IntResGrwth*: International reserve growth,
- *CurrAccGDP*: The ratio of current account to GDP,
- *TotResGrwth*: Total reserve growth,
- *FDItoGDP*: The ratio of FDI to GDP,
- *ChngTrd*: Change in terms of trade,
- *DmstcCrdttoPrvtSec*: Domestic credit to private sector (% of GDP),
- *DmstcCrdtBnkSec*: Domestic credit provided by banking sector (% of GDP),
- *RealDmstcCrdtGrwth*: Real domestic credit growth,
- *ImprtGrwth*: Import growth rate,
- *DegOpen*: Degree of openness,
- *ExprtGrwth*: Export growth rate,
- *SPGlobEqInd*: S&P global equity indices (annual % change),
- *TaxRev*: Tax revenue (% of GDP),
- *BnkCaptoAsstRat*: Bank capital to asset ratio,
- *IntrRatGovSec*: Interest rates of government securities & government bonds,
- *BnkLiqtoBnkAsst*: Bank liquid reserves to bank asset ratio,

- *GovTotDbt*: Central government total debt (% of GDP),
- *ResidMortDbtGDP*: Residential mortgage debt to GDP ratio,
- *RealFixInvHous*: Real gross fixed investment in housing annual change,
- *TotDwlStck*: Total dwelling stock,
- *HousComp*: Housing completions,
- *ReprIntrRatMortLoan*: Representative interest rate on new mortgage loans,
- *NomHousPrc*: Nominal house prices annual percentage change,
- *ChngHousPrcGDP*: Change of housing price to GDP growth.

Table 5.8 summarizes the final estimation results of the general multinomial logistic regression:

Table 5.8: Results of the multinomial logit regression for t, t+1, t+2

Variables	Coeff.	Std. Err.	Z-stat.	95% Conf. Intrvl	
<i>Pre-crisis (Y = 1)</i>					
IntResGrwth	-18.40241	11.48476	-1.60	-40.91213	4.107301
CurrAccGDP	-68.86347	74.88574	-0.92	-215.6368	77.90988
TotResGrwth	12.33979	13.37741	0.92	-13.87945	38.55903
FDItoGDP	-24.93837	17.3483	-1.44	-58.94043	9.063678
ChngTrd	10.9874	13.36763	0.82	-15.21268	37.18748
DmstcCrdttoPrvtSec	-0.0735938	0.2600634	-0.28	-0.5833087	0.4361211
DmstcCrdtBnkSec	-0.0242062	0.285955	-0.08	-0.5846676	0.5362553
RealDmstcCrdtGrwth	3.649001	20.92005	0.17	-37.35354	44.65154
ImpprtGrwth	-16.97865	63.21763	-0.27	-140.8829	106.9256
DegOpen	-10.01837	8.704804	-1.15	-27.07947	7.04273
ExprtGrwth	67.23506	85.76387	0.78	-100.859	235.3291
SPGlobEqInd	0.054558	0.0356393	1.53	-0.0152937	0.1244096
TaxRev	1.562956	0.8609468	1.82	-0.1244692	3.25038
BnkCaptoAsstRat	-8.311859	4.285097	-1.94	-16.7105	0.0867775
IntrRatGovSec	-3.907184	6.418303	-0.61	-16.48683	8.672458
BnkLiqtoBnkAsst	-255.6601	260.3078	-0.98	-765.854	254.5339
GovTotDbt	-0.1772111	0.1603141	-1.11	-0.4914209	0.1369988
ResidMortDbtGDP	-0.1376594	0.2006145	-0.69	-0.5308565	0.2555378
RealFixInvHous	0.382227	0.4609537	0.83	-0.5212257	1.28568
TotDwlStck	-0.0000291	0.0001736	-0.17	-0.0003695	0.0003112
HousComp	0.0000299	0.0000217	1.38	-0.0000127	0.0000724
ReprIntrRatMortLoan	5.869147	5.460862	1.07	-4.833946	16.57224

Variables	Coeff.	Std. Err.	Z-stat.	95% Conf. Intrvl
NomHousPrc	-0.1079562	0.3518175	-0.31	-0.7975057 0.5815934
ChngHousPrcGDP	-0.3186199	0.6975067	-0.46	-1.685708 1.048468
Const.	28.94143	24.21257	1.20	-18.51433 76.39718

The interpretations of the coefficients β_1 and β_2 of the multinomial logistic regression are the same as those given for the time window of ‘t-1, t, t+1’.

According to the estimation results, the following interpretations can be made:

- International reserve growth, FDI to GDP and domestic credit provided by banking sector (% of GDP) behave in similar way in entering into the crisis. If there is a high reduction in these values, the probability of entering into the crisis increases and in the adjustment phase, a decline in these values indicates that the adjustment phase is still going on.
- Bank liquid reserves to bank assets ratio, bank capital to assets ratio and domestic credit to private sector’s marginal contributions are as follows: If there is a fall in this ratios, the probability of entering into crisis increases. This situation may arise due to the fact that this situation makes the economy more fragile and this makes it open to speculative attacks, which may end up with a crisis.
- Interest rates of government securities and government bonds, total central government debt and current account to GDP affect both the probability of being in the pre-crisis and that of staying in the adjustment phase. If there is a decrease in these variables, this gives a signal about the coming of a crisis and if there is an improvement in these variables, this indicates the economy’s approaching to the tranquil regime again.
- Import growth rate and degree of openness are among the important contributors of this early warning system. A small reduction in these indicators gives

a signal about the coming crisis. If there is a raise in these variables, it will indicate that the recovery period is still going on.

- Export growth rate, tax revenue (% of GDP) and change in terms of trade are the variables, which affect each other. As a result of this, their contributions to the early warning system are similar. If there is a reduction in these variables, it will show that there is a slowing down in the economy and this will be a signal about the coming crisis. These indicators represent the summary of the economic activities as they give information about the trade and the fiscal position of the government. However, in the recovery period, if there is an improvement in these variables, it will indicate being in the recovery period.
- Real domestic credit growth and S&P global equity indices give information about the country's financial position. Their marginal contributions on entering into crisis are positive, and if there is a fall in these variables, it indicates a slowdown in the economy, so it gives signal about the coming crisis.
- Total reserve growth rate's marginal contribution is as follows: if the reserve ratio worsens, the economy enters into the crisis and if it becomes better, this indicates recovery in the economy.
- Residential mortgage debts to GDP ratio, real gross fixed investment in housing and total dwelling stock move in the same direction. In the case of a decrease in these variables the probability of a crisis increases. On the contrary, an increase in these variables after the crisis indicates a recovery period.
- Housing completions and representative interest rates on new mortgage loans give information about the household's future expectations. If there is a decrease in these values, it indicates an upcoming slowing down. And in recovery

periods, if these variables' variations are positive, this will give information about the exiting from the crisis.

- Nominal house prices annual percentage change and rate of change of housing price to GDP growth give information about the household's purchasing power and also the general situation of the economy. If the household's purchasing power is in a better position, it shows that the economic conditions are sustainable, which means there is no risk of a crisis. But if there is a decrease in these ratios, it indicates the contraction period of the economy as giving a signal about the upcoming crisis.

Predictive Ability

The success of an early warning system is measured by its prediction performance. Since the aim of this study is to measure the impact of the housing market variables on the predicting performance, the analysis consists of two data sets. The first data set is composed of the previous studies' independent variables. (These are the best performers, which have been selected from the sub-groups as explained previously in the text.). The first set contains the following variables: International reserve growth, the ratio of current account to GDP, total reserve growth, the ratio of FDI to GDP, change in terms of trade, domestic credit to private sector (% of GDP), domestic credit provided by banking sector (% of GDP), real domestic credit growth, import growth rate, degree of openness, export growth rate, S&P global equity indices (annual % change), tax revenue (% of GDP), bank capital to asset ratio, interest rate of government securities & government bonds, bank liquid reserves to bank asset ratio, central government total debt (% of GDP). The probability threshold values used in determining the tranquil, pre-crisis and adjustment periods are same and they are all equal to 50%.

The in sample performance of this first early warning system, consisting of these mentioned variables, is given in the following table (Table 5.9):

Table 5.9: Performance of the model without housing data for t , $t+1$, $t+2$

<i>State</i>	<i>Correctly called</i>	<i>Not correctly called</i>	<i>Total</i>
<i>Tranquil</i>	80	13	93
<i>Pre-crisis</i>	24	11	35
<i>Adjustment</i>	22	5	27

This model predicts 68% of all the pre-crisis observations available in the set. Eleven pre-crisis conditions are missed by this model: Austria 2007, France 2004, Germany 2007, Greece 2004, Ireland 2004, Italy 2004 and 2007, Spain 2007, Sweden 2007, United States 2001 and 2005.

9 of these missed pre-crisis cases are called as tranquil period observations by this early warning system, which implies that although there is a pre-crisis condition, the model accepts and estimates it as tranquil period. (for 9 cases, $Y = 1$ but $P_0 \geq 0.5$).

In this model's observation set, there are 93 tranquil period observations ($Y = 0$) and the early warning system predicts 80 of them correctly. (93 cases for which $Y = 0$, for 80 out of them, $Y = 0$ & $P_0 \geq 0.5$). The Early warning system missed 13 tranquil period observations, 4 of them are classified as adjustment period observations and 8 of them are classified as pre-crisis period observations.

Also in this model's observation set, there are 27 adjustment period observations and the early warning system predicts 22 of them correctly. (27 cases for which $Y = 2$, for 22 cases out of them both $Y = 2$ & $P_2 \geq 0.5$). The missing 3 adjustment period observations are classified as tranquil periods.

The model sends false alarms, which means that although the economy is either in the tranquil or adjustment period, the early warning system gives a signal about an upcoming crisis. There are 8 tranquil period observations, for which false signals

of an upcoming crisis are given. (8 cases for which $Y = 0$ & $P_1 \geq 0.5$). The ratio of the false alarms generated by this model is 6.67% (8 cases out of 120).

Here, it has to be noted that as obviously seen from Table 5.7, in the case of ‘t, t+1, t+2’ it is possible to predict a crisis either one or two-years in advance. Hence, for this time window, not all of the pre-crisis conditions correspond to a ‘unique crisis’ each. For some of the crises, not one but two pre-crisis calls do exist. Consequently, if the number of correctly called unique crises is considered, the success ratio of this model is 94%, that is, it catches 94% of the unique crises in the set.

After predicting crisis by using the best performers of the prior studies, a second multinomial logistic regression model is constructed by introducing the following housing market variables into the first best performers’ data set to show the impact of this sector on predicting crisis. The best performers of the housing data are the following: Residential mortgage debt to GDP ratio, real gross fixed investment in housing annual change, total dwelling stock, housing completions, representative interest rate on new mortgage loans, nominal house prices annual percentage change, change of housing price to GDP growth. After combining both data sets the in-sample performance of the multinomial logistic regression in predicting crisis is given in the following table (Table 5.10):

Table 5.10: Performance of the model with housing data for t, t+1, t+2

<i>State</i>	<i>Correctly called</i>	<i>Not correctly called</i>	<i>Total</i>
<i>Tranquil</i>	45	2	47
<i>Pre-crisis</i>	19	1	20
<i>Adjustment</i>	19	0	19

This new model predicts 95% of all the pre-crisis conditions available in the observation set. Only one pre-crisis observation is missed by this early warning system: Italy 2004. The multinomial logistic regression model classifies this missing pre-crisis case as a tranquil period observation.

In this model's observation set, there are 47 tranquil period observations ($Y = 0$) and the early warning system predicts 45 of them correctly. (47 cases for which $Y = 0$, for 45 out of them, $Y = 0$ & $P_0 \geq 0.5$). The Early warning system missed 2 tranquil period observations; both of them are classified as pre-crisis observations (2 cases for which $Y = 0$ & $P_1 \geq 0.5$).

Also in this model's time interval, there are 19 adjustment period observations and the early warning system predicts all 19 of them correctly. (19 cases for which $Y = 2$, for all of them $Y = 2$ & $P_2 \geq 0.5$).

This system also gives 2 false signals. Although these periods are tranquil periods, the model accepts them as pre- crisis periods and gives false signals. (2 cases for which $Y = 0$ & $P_1 \geq 0.5$) . Hence, the ratio of the false alarm ratio of this model is 3.33% (2 cases out of 66).

As mentioned earlier, for the time window of 't, t+1, t+2', a crisis can be predicted one-year or two-years in advance. Hence, for this model, the 20 pre-crisis conditions available in the set correspond to a set of 11 unique crises. This new model with the housing market variables successfully calls all of these unique 11 crises as listed in Table 5.11.

Table 5.11: Crises correctly called with the model with housing data, t, t+1, t+2

<i>Country</i>	<i>Years</i>
Denmark	2008
Germany	2005, 2008
Ireland	2005, 2007
Italy	2005
Netherlands	2008
Spain	2008
Sweden	2008
United States	2002, 2007

It has to be noted that, for this time window, inclusion of the housing sector variables into the model decreases the total number of valid observations in the data set from 155 to 86 as can be seen from Tables 5.9 and 5.10. Hence, as a further means for the quantitative comparison of the two models, the first model constructed by the best performers of the prior studies has also been invoked in this reduced data set to evaluate its predictive performance. The results of this run are given in Table 5.12

Table 5.12: Performance of the model without housing data on the reduced data set for t , $t+1$, $t+2$

<i>State</i>	<i>Correctly called</i>	<i>Not correctly called</i>	<i>Total</i>
<i>Tranquil</i>	39	8	47
<i>Pre-crisis</i>	13	7	20
<i>Adjustment</i>	16	3	19

Regarding the number of correctly called pre-crisis conditions, introduction of the housing market variables dramatically improves the performance of the system from 68% to 95%. In addition to this, the housing market variables also increase the prediction success of the tranquil periods and the adjustment periods as well. Hence, the overall predictive ability of the system is significantly improved.

Using this time window of ' t , $t+1$, $t+2$ ', it becomes possible to make an out-of-sample performance analysis to check the validity of the results. As mentioned earlier the publicly available housing market data spans the period from 1999 to 2010 and this puts a restriction on the number of available observations. Furthermore, in order to be able to make an out-of-sample analysis, the data set has been divided into two parts: First part as being up-to 2007 and the second part being from 2007 and on. The first part has been used for the in-sample performance analysis and the model has been constructed over it, and the second part has been used as the out-of-sample set. For this analysis, 4 variables from the original model needed to be further dropped since the model could not converge to a result due to the reduced number of

in-sample observations and the comparatively large number of missing values in these 4 variables: Domestic credit to private sector (% of GDP), domestic credit provided by banking sector (% of GDP), real domestic credit growth and S&P Global equity indices.

For the model constructed by the independent variables of the previous studies, the out-of-sample performance is 27% regarding the number of correctly called pre-crisis conditions. After introducing the housing market variables, the out-of-sample performance improves from 27% to 60%, again regarding the number of correctly called pre-crisis conditions. In terms of the correctly called unique crises, the out-of-sample performance of the model with the housing variables is 60%, which is far above the 33% success ratio of the old model without the housing market variables. This out-of-sample performance will definitely get improved if the missing values of the dropped 4 variables are provided and these variables are added back into the model again.

CHAPTER 6

CONCLUSION

This thesis constructs a multinomial logistic regression model to predict crisis by using both macro economic data, financial market data and housing market data. The literature on the early warning systems focuses on currency crisis, banking crisis and debt crisis. Although there are several studies in the literature, which aims to predict crises, none of them could manage to predict the crises of 2000s. The aim of this study is to predict crisis by using two different time windows by utilizing the technique of multinomial logistic regressions. The main hypothesis of this study is to measure the impact of the housing sector on predicting crisis. The findings in this study indicate that the housing sector includes essential indicators for predicting crisis.

The data used in multinomial logistic regressions are taken from the World Bank, IMF, OECD and Eurostat, and spans the period between 1999 and 2010 for 29 different countries. The data set consists of two main categories: First group includes the data from the collection of previous studies' macro economic and financial sector data set. The second part of the data set includes the housing sector data, which is the contribution of this study to the crisis prediction literature. The estimation started with 111 exogenous variables, they were divided to groups according to the similarity

of their effects on the economy, and then the best performers of each group have been selected to constitute the final multinomial logistic regression data set. Since there are two different time windows that affect the criteria of choosing appropriate variables, the selected variables for the final data set are different for these two different time windows.

Initially, a model is constructed from the variables of the previous studies' data set without adding the housing sector data by using multinomial logistic regression. Then, the housing data set is introduced and a new multinomial logistic regression model is formed, which uses both the previous studies' data and the housing sector data.

The results show that, for the time window of 't-1, t, t+1', the housing data improves the in-sample prediction power of the multinomial logistic regression model for crisis entries from 60% to 100%. The model produces no false alarms in this case. For the time window of 't, t+1, t+2', the housing data improves the in-sample prediction power of the multinomial logistic regression model for crisis entries from 94% to 100% regarding the number of correctly called unique crises. If the number of correctly called pre-crisis periods are considered, the prediction power of the system is improved from 68% to 95%. The false alarm ratio of this model is reduced from 6.6% to 3% after inclusion of the housing data.

The out-of-sample performance of the model is estimated for the time window of 't, t+1, t+2' as well. With the housing data set, the multinomial logistic regression model's out-of-sample performance improved from 33% to 60% proving that the model not only makes accurate predictions for the period available in the in-sample data set but also it is accurate in its predictions about the future. The housing data set is not suitable for estimating the out-of-sample performance for the time window of 't-1, t, t+1', but it will be possible after the housing sector data for the years in the near future are collected.

This study differs from the prior studies from many perspectives: This multinomial logistic regression model is constructed to predict not only one, but all types of crises as currency crisis, banking crisis and debt crisis. Its data set includes the housing sector variables, which has not been widely used in the crisis prediction literature up to now. Moreover, the analysis part of this estimation is constructed from two different time windows and new definitions are provided for the pre-crisis and adjustment periods. By this way the crisis prediction power was aimed to be increased.

In this study, it has been found that, the housing sector is an important factor contributing to economic crisis. Hence, this sector has to be taken into account in designing crisis prediction systems. The improvement in prediction power is evident in both time window analyses.

However, it has to be noted that the size of the data set is not large enough for being able to compare the performances of the two different time windows since, as described in Section 5.2, by construction the number of observations categorized as pre-crisis in the case of ‘ $t-1, t, t+1$ ’ is only one-third of those categorized as pre-crisis for the case of ‘ $t, t+1, t+2$ ’. Hence, this puts a further restriction on the data set, which is already restricted by the availability of publicly available housing market data. Hence, in order to be able to make a appropriate comparison between the performances of the different time windows, a housing data set, which spans a much larger period of time is required.

On the other hand, despite these shortcomings; the findings of this study point to some early evidence about the important role played by the housing sector in leading to economic crisis.

As a future work, the data set might be enlarged especially at the housing market variables side in order to make it possible to use more observations for both in-sample and out-of-sample analyses. By this way, the differences in the performances of the two different time windows, if any, might become much more apparent. Furthermore,

enlarging the time window used, and trying to use different time-windows might contribute to the predictive power of the system.

BIBLIOGRAPHY

- [Andre(2010)] Andre, C., “A birds eye view of OECD housing markets”, *OECD Economics Department Working Papers*, No. 746, January, 2010.
- [Babecký(2012)] Babecký, J., Havránek T., Matějů J., Rusnák M., Smídková K., Vaiscek B., “Banking, Debt, and Currency Crises: Early Warning Indicators for Developed Countries”, *IES Working Paper 20/2012*. IES FSV. Charles University, 2012.
- [Beckmann(2003)] Beckmann, D., Menkhoff, L. , “Early Warning Systems: Lessons from New Approaches”, 2003.
- [Benjamin(2004)] Benjamin, J. D., Chinloy, P., Jud, D. D., “Real Estate Versus Financial Wealth in Consumption”, *Journal of Real Estate Finance and Economics*, *Kluwer Academic Publishers*, Vol. 29:3, pp.341-354, 2004.
- [Berg and Pattillo(1999a)] Berg, A., Pattillo, C., “Predicting currency crises: the indicators approach and an alternative”, *Journal of International Money and Finance*, 18:4, pp. 561 - 586, August, 1999.
- [Berg and Pattillo(1999b)] Berg, A., Pattillo, C., “Are currency crises predictable? A test”, *IMF Staff Papers*, 46:2, pp. 107 - 138, June 1999.
- [Berg and Pattillo(1999c)] Berg, A., Pattillo, C., “What caused the Asian crises: An early warning system approach”, *Economic Notes*, Review of Banking, Finance and Monetary Economics, Volume 28, Issue 3, pages 285-334, November 1999.
- [Bruner(1987)] Bruner, R.F., Simms, J. M., “The International Debt Crisis and Bank Security Returns in 1982”, *Journal of Money, Credit and Banking*, Vol. 19, No. 1, pp. 46-55, February, 1987.
- [Brutti(2011)] Brutti Brutti, F., “Sovereign defaults and liquidity crises”, *Elsevier Journal of International Economics*, No. 84, pp.65-72, 2011.
- [Buiter(2009)] Buiter, Willem, “Fiscal Expansions in Submerging Markets; the Case of the USA and the UK”, *Financial Times Online*, 2009.

- [Bussiere and Fratzscher(2002)] Bussiere, M. , Fratzscher, M. , “Towards a new early warning system of financial crises”, *European Central Bank Working Paper* No. 145, 2002
- [Bussiere and Fratzscher(2006)] Bussiere, M. , Fratzscher, M. , “Towards a new early warning system of financial crises”, *Elsevier Journal of International Money and Finance*, No. 25, pp. 953-973, 2006.
- [Candelon[2010] Candelon Candelon, B., Palm, F. C., “Banking and Debt Crises in Europe: The Dangerous Liaisons”, *De Economist* , Vol. 158, No. 1, pp.81-99, 2010.
- [Case(2001)] Case, K. E., Quigley, J. M. and Shiller, R. J., “Comparing Wealth Effects: The stock market vs. the Housing market”, *NBER Working Paper* No. 8606, NBER, 2001.
- [Ciarlone and Trebeschi(2005)] Ciarlone, A., Trebeschi, G., “Designing an early warning system for debt crises”, *Elsevier Emerging Markets Review*, No. 6, pp.376-395, 2005.
- [Ciarlone and Trebeschi(2006)] Ciarlone, A., Trebeschi, G., “A multinomial approach to early warning system for debt crises”, *Bank of Italy, Economic Research and International Relations Area, Temi di discussione (Economic working papers)*, No. 588, 2006.
- [Darvas(2011)] Darvas, Z., Ferry, J. P., Sapir, A., “A comprehensive approach to the euro-area debt crisis”, *Bruegel Policy Brief*, Bruegel Working Paper, No. 05, 2011.
- [Davis and Karim(2008a)] Davis, E. P. , Karim, D., “Comparing early warning systems for banking crises”, *Elsevier Journal of Financial Stability*, No. 4, pp. 89120, 2008
- [Davis and Karim(2008b)] Davis, E. P. , Karim, D. , “Could Early Warning Systems Have Helped To Predict the Sub-Prime Crisis?”, *National Institute Economic Review*, No. 206: 35, 2008.
- [Demirguc-Kunt(1997)] Demirguc-Kunt, A., Detragiache, E., “The Determinants of Banking Crises”, *Policy Research Working Paper Series*, No. 1828, September, 1997.
- [Drechel(2010)] Drechsel, K., Scheufele, R., “Should We Trust in Leading Indicators? Evidence from the Recent Recession”, *IWH Discussion Papers*, Vol. 10, Halle Institute for Economic Research, 2010.
- [Edison(2002)] Edison, H. J. , “Do indicators of financial crises work? An evaluation of an early warning system”, *International Journal of Finance and Economics*, No. 8, pp. 11-53, December 2002.

- [Eichengreen(1994)] Eichengreen, B. , Rose, A. , Wyplosz, C. , “Speculative attacks on pegged exchange rates: An empirical exploration with special reference to the European monetary system”, *NBER Working Paper Series* No. 4898, Cambridge, National Bureau of Economic Research, 1994.
- [Eichengreen(1995)] Eichengreen, B. , Rose, A. , Wyplosz, C. , “Exchange market mayhem: The antecedents and aftermath of speculative attacks”, *Economic Policy*, pp. 249 - 312, October 1995.
- [Eichengreen(1996)] Eichengreen, B. , Rose, A. , Wyplosz, C. , “Contagious currency crises”, *NBER Working Paper Series* No. 5681, Cambridge, National Bureau of Economic Research, 1996.
- [Eichengreen(2011)] Eichengreen, B. et al, “Rethinking Central Banking”, *Committee on International Economic Policy and Reform*, September 2011.
- [Eng(2011)] Eng, M., “Book Review: ‘This Time is Different: Eight Centuries of Financial Folly’ by Reinhart and Rogoff”, *Journal of Emerging Markets*, pp.68-74, 2011.
- [Featherstone(2011)] Featherstone, K., “The Greek Sovereign Debt Crisis and EMU: A Failing State in a Skewed Regime”, *Journal of Common Market Studies*, Vol. 49, No. 2, pp.193-217, 2011.
- [Fioramanti(2008)] Fioramanti, M., “Predicting sovereign debt crises using artificial neural networks: A comparative approach”, *Elsevier Journal of Financial Stability*, No. 4, pp.149-164, 2008.
- [Flood(1984)] Flood, R.P., Garber, P. M., “Collapsing exchange rate regimes: some linear examples”, *Journal of International Economics*, No. 17, pp. 113, 1984.
- [Frankel(1996)] Frankel, J., Rose, A., “Currency crashes in emerging markets: An empirical treatment”, *International Finance Discussion Papers*, No. 534, Board of Governors of the Federal Reserve System, January 1996.
- [Fuertes and Kalotychou(2006)] Fuertes, A. M., Kalotychou, E., “Early warning systems for sovereign debt crises: The role of heterogeneity”, *Elsevier Computational Statistics and Data Analysis*, No. 51, pp.1420-1441, 2006.
- [Fuertes and Kalotychou(2007)] Fuertes, A. M., Kalotychou, E., “Optimal design of early warning systems for sovereign debt”, *Elsevier International Journal of Forecasting*, No. 23, pp.85-100, 2007.
- [Glick(1999)] Glick, R., Hutchison M., “Banking and Current Crises: How Common Are Twins?”, *Proceedings of Federal Reserve Bank of San Francisco*, Issue: September, 1999.

- [Goldstein(2000)] Goldstein, M. , Kaminsky, G., Reinhart, C. , “Assessing financial vulnerability: An early warning system for emerging markets”, *Institute for International Economics*, ISBN paper 0-88132-237-7, Washington DC., June 2000
- [Goodhart(2008)] Goodhart, C., Hofmann, B., “House Prices, Money, Credit and the Macroeconomy”, *European Central Bank Working Paper Series*, No: 888, April, 2008.
- [Graciela(2004)] Graciela L. Kaminsky, Carmen M. Reinhart, Carlos A. Végh, “When It Rains, It Pours: Procyclical Capital Flows and Macroeconomic Policies”, *NBER Macroeconomics Annual*, Vol. 19, 2004.
- [Greene(2003)] Greene, W.H., “Econometric analysis”, Prentice Hall , 5th edition, NJ, 2003.
- [IMF Database] IMF Online Database.
- [Kaminsky and Reinhart(1998)] Kaminsky, G. , Lizondo, S. , Reinhart, C. , “Leading indicators of currency crises”, *International Monetary Fund Staff Papers*, Vol. 45, No. 1, pp. 1-48, 1998
- [Kaminsky and Reinhart(1999)] Kaminsky, G. L., Reinhart, C. M., “The twin crises: The causes of banking and balance of payments problems”, *American Economic Review*, 89:3, 473 - 500, June 1999.
- [Kaminsky and Reinhart(2000)] Kaminsky, G. L., Reinhart, C. M., “On crises, contagion, and confusion”, *Elsevier Journal of International Economics*, No. 51, pp.145-168, 2000.
- [Laeven(2008)] Laeven, L., Valencia, F., “Systemic Banking Crises: A New Database, *IMF Working Paper Series*, No: 224, November, 2008.
- [Lestano(2003)] Lestano, J. J. , Kuper, G., “Indicators of financial crises do work! An early-warning system for six Asian countries”, 2003.
- [Manasse(2009)] Manasse, P. , Roubini, N. , “Rules of thumb for sovereign debt crises”, *Journal of International Economics*, No. 78, pp.192205, 2009.
- [Manasse and Roubini(2003)] Manasse, P., Roubini, N., Schimmelpfennig, A., “Predicting Sovereign Debt Crises”, *IMF Working Paper Series*, Working Paper 03/221, November 2003.
- [Masson(2000)] Masson, P. , Borensztein, E., Berg, A. , Milesi-Ferretti, G. M., Pattillo, C. , “Anticipating balance of payments crises: The role of early warning systems”, *International Monetary Fund Occasional Papers*, No. 186, 2000.
- [Nersisyan(2011)] Nersisyan, Y., Wray, L. R., “Review : ‘This Time is Different: Eight Centuries of Financial Folly’ by Reinhart and Rogoff”, *Challenge* , Vol. 54, No. 1, pp.113-120, January-February 2011.

- [Obstfeld(1986)] Obstfeld, M., “Rational and self-fulfilling balance-of-payments crises”, *American Economic Review*, Vol. 76, No. 1, pp. 72-81, 1986.
- [Parra(2008)] Parra, S. N. , “Who saw sovereign debt crises coming?”, *OECD Development Centre*, Working Paper No. 274, November 2008.
- [Parra(2009)] Parra, S. N. , “Sovereign Debt Crises and Early Warning Indicators: The Role of the Primary Bond Market”, *OECD Development Centre Policy Insights*, No. 89, February 2009.
- [Peltonen(2006)] Peltonen, T. A., “Are Emerging Market Currency Crises Predictable? - A Test”, *European Central Bank Working Paper Series*, No. 571, January, 2006.
- [Qian(2010)] Qian, R., Reinhart, C. M., Rogoff, K. S., “On Graduation From Default, Inflation and Banking Crisis: Elusive or Illusion?”, *NBER Working Paper Series*, Working Paper 16168, July 2010.
- [Reinhart(2009)] Reinhart, C. M., “The Economic and Fiscal Consequences of Financial Crises”, *Munich Personal RePEc Archive*, No. 13025, January 2009.
- [Reinhart and Felton(2008)] Reinhart, C. M., Felton, A., “The First Global Financial Crisis of the 21st Century”, *Munich Personal RePEc Archive*, No. 11862, December 2008.
- [Reinhart and Rogoff(2008a)] Reinhart, C. M., Rogoff, K. S., “Is the 2007 U.S. Subprime Financial Crisis So Different? An International Historical Comparison”, *NBER Working Paper Series*, Working Paper 13761, January 2008.
- [Reinhart and Rogoff(2008b)] Reinhart, C. M., Rogoff, K. S., “This Time is Different: A Panoramic View of Eight Centuries of Financial Crises”, *NBER Working Paper Series*, Working Paper 13882, March 2008.
- [Reinhart and Rogoff(2009)] Reinhart, C. M., Rogoff, K. S., “This Time It’s Different: Eight Centuries of Financial Folly-Chapter 1”, *Munich Personal RePEc Archive*, No. 17452, September 2009.
- [Reinhart and Rogoff(2010a)] Reinhart, C. M., Rogoff, K. S., “From Financial Crash to Debt Crisis”, *NBER Working Paper Series*, Working Paper 15795, March 2010.
- [Reinhart and Rogoff(2010b)] Reinhart, C. M., Rogoff, K., “Why we should expect low growth amid debt”, *Financial Times (FT.com)*, January 2010.
- [Reinhart and Rogoff(2010c)] Reinhart, C. M., Rogoff, K., “Debt and growth revisited”, *Center for Economic Policy Research: VOX*, August 2010.
- [Reinhart and Rogoff(2010d)] Reinhart, C. M., Rogoff, K. S., “Growth in a Time of Debt”, *NBER Working Paper Series*, Working Paper 15639, January 2010.

- [Reinhart and Rogoff(2011a)] Reinhart, C. M., Rogoff, K. S., “A Decade of Debt”, *NBER Working Paper Series*, Working Paper 16827, February 2011.
- [Reinhart and Rogoff(2011b)] Reinhart, C. M., Rogoff, K. S., “The Forgotten History of Domestic Debt”, *The Economic Journal*, No. 121, May 2011.
- [Reinhart and Rogoff Online Resources] <http://www.reinhartandrogoff.com>
- [Rose(2009)] Rose, A. K. , Spiegel, M. M., “Cross-Country Causes and Consequences of the 2008 Crisis: Early Warning”, *CEPR Discussion Paper*, No. 7354, 2009
- [Rose(2011)] Rose, A. K., Spiegel, M. M., “Cross-Country Causes and Consequences of the Crisis: An Update”, *Federal Reserve Bank of San Francisco Working Paper Series*, January 2011
- [Roubini(2005)] Manasse, P. , Roubini, N. , “Rules of thumb for sovereign debt crises”, *IMF Working Paper*, Working paper 05/42, March 2005.
- [Online Resources PennState University] <https://onlinecourses.science.psu.edu/stat507/node/18>

APPENDIX A

FURTHER DETAILS

A.1 Definitions Data and Methodology

Table A.1: Data set classified according to source

Source	Set of Variables
IMF	Exchange rate, Money supply, Export, Import, Industrial production, Return on equity, Current account, Foreign exchange reserves, GDP, Liquid asset to total asset, Capital adequacy ratio, Treasury bill rate, Residential real estate prices, Residential real estate loans to total loans, Commercial real estate loans to total loans, Deposit rate, Household debt to GDP, Interest rates of government securities and government bonds, Interest rates government securities and government bonds medium-long term, Interest rates government securities and government bonds short term, Liquid asset to short term liabilities, Net open position in the foreign exchange to capital, Non performing loans to total gross loans, Return on asset, Sectorial distribution of total loans (deposit takers), Sectorial distribution of total loans (residents), Unemployment rate

Source	Set of Variables
World Bank	GDP growth, Short term debt to international reserves, Domestic credit to private sector, Trade in services percent of GDP, Real interest rate, Inflation, Government revenue excluding grants percent of GDP, Government expense percent of GDP, Bank liquid reserve to bank asset ratio, Domestic credit provided by banking sector percent of GDP, GDP per capita, Deposit insurance, Interest payment on total external debt, Total external debt stocks, Short term external debt stocks, Short term debt to total reserves, Short term debt to total external debt, Interest payments on short term external debt, FDI, Use of IMF credit, Central government debt as percentage of GDP, Private non-guaranteed external debt stocks, Public and publicly guaranteed external debt stocks, Bank non-performing loans to total gross loans, Interest rate spread, Portfolio equity net inflows, Risk premium on lending, S&P global equity indices (annual percentage change), Stocks traded (total value), Stocks traded turnover ratio, Gross saving percent of GDP, Net ODA received, Total debt service percent of exports, Bank capital to asset ratio, Tax revenue percent of GDP, Interest payments on long term external debt
OECD	House prices percentage change over previous period, House price ratio to rent ratio, House price to income ratio, Household wealth and indebtedness as a percentage of nominal disposable income, Housing expense in percent of GDP
Eurostat	Building permits, Gross residential loans, House price percentage change, Housing completions, Housing starts, Net residential loans, Nominal house price indices (2000=100), Nominal house prices annual percentage change, Number of housing transactions, Real gross fixed investment in housing annual change, Representative interest rates on new mortgage loans, Residential mortgage debt to GDP ratio, Total dwelling stock, Total outstanding residential loans

Source	Set of Variables
Calculated	Degree of openness to international trade, Export growth rate, The ratio of total external debt to GDP, Total external debt to exports, The ratio of international reserve to total external debt, Short term debt to GDP, Total external debt to total reserves, Total reserve growth, FDI to GDP, Financial requirement to total reserve, Inflation volatility, Import growth, International reserve growth, Current account to GDP, Short term debt to international reserve growth, Real domestic credit growth, interest payments on external debt to international reserve, Change in terms of trade, Fiscal surplus to GDP, Interest payment on short term debt to GDP, Deviations of real exchange rate from trend, Rate of change of housing price to GDP growth, The ratio of real gross fixed investment in housing annual change to GDP growth, The ratio of representative interest rate on new mortgage loans to real interest rate, The ratio of M2 reserve to international reserve