

BOFIT Discussion Papers

2 • 2011

Marco Lo Duca and Tuomas Peltonen

Macro-financial vulnerabilities and future financial stress - Assessing systemic risks and predicting systemic events



EUROJÄRJESTELMÄ
EUROSYSTEMET

Bank of Finland, BOFIT
Institute for Economies in Transition

BOFIT Discussion Papers
Editor-in-Chief Aaron Mehrotra

BOFIT Discussion Papers 2/2011
2.3.2011

Marco Lo Duca and Tuomas Peltonen:
Macro-financial vulnerabilities and future financial stress - Assessing
systemic risks and predicting systemic events

ISBN 978-952- 462-697-2
ISSN 1456-5889
(online)

Suomen Pankki
Helsinki 2011

Contents

Abstract	4
Tiivistelmä	5
1 Introduction.....	6
2 Measuring financial stress and identifying systemic events	9
3 Data.....	14
4 Empirical analysis.....	16
5 Conclusions.....	26
References.....	28

All opinions expressed are those of the authors and do not necessarily reflect the views of the Bank of Finland.

Marco Lo Duca¹ and Tuomas Peltonen

Macro-financial vulnerabilities and future financial stress - Assessing systemic risks and predicting systemic events²

Abstract

This paper develops a framework for assessing systemic risks and for predicting (out-of-sample) systemic events, i.e. periods of extreme financial instability with potential real costs. We test the ability of a wide range of “stand alone” and composite indicators in predicting systemic events and evaluate them by taking into account policy makers’ preferences between false alarms and missing signals. Our results highlight the importance of considering jointly various indicators in a multivariate framework. We find that taking into account jointly domestic and global macro-financial vulnerabilities greatly improves the performance of discrete choice models in forecasting systemic events. Our framework shows a good out-of-sample performance in predicting the last financial crisis. Finally, our model would have issued an early warning signal for the United States in 2006Q2, 5 quarters before the emergence of money markets tensions in August 2007.

JEL Codes: E44, E58, F01, F37, G01.

Keywords: Early warning indicators, asset price booms and busts, financial stress, macro-prudential policies.

¹ Corresponding author: Marco Lo Duca, International Policy Analysis Division, European Central Bank, Kaiserstrasse 29, D-60311, Frankfurt am Main. Email: marco.lo_duca@ecb.europa.eu tel. +49 69 1344 5690.

² The authors want to thank Daniela Bragoli, Arjana Brezigar-Masten, Carmen Broto, Alexander Chudik, Mardi Dungey, Carsten Detken, Michael Fidora, Marcel Fratzscher, Philipp Hartmann, Jean Imbs, Gilles Noblet, Livio Stracca, and participants to the European Economic Association Annual Meeting 2010, INFINITY 2010 conference, the 2010 Workshop of the Eurosystem and Latin American Central, the 2010 ECB workshop "A Global Dimension on Early Warning Models and Macroprudential Analysis", the 2010, 8th ESCB Workshop on Emerging Markets, the 2011 Bank of Korea and BIS conference on “Macroprudential Regulation and Policy” for useful comments and discussions. All remaining errors are of our own. The views presented in the paper are those of the authors and do not necessarily represent the views of the European Central Bank or the Eurosystem.

Marco Lo Duca and Tuomas Peltonen

Macro-financial vulnerabilities and future financial stress - Assessing systemic risks and predicting systemic events

Tiivistelmä

Tässä tutkimuksessa luodaan menetelmä, jonka avulla voidaan arvioida sellaisia systeemisiä riskejä ja ennustaa sellaisia systeemisiä tapahtumia, joihin liittyy merkittävää finanssimarkkinoiden epävakautta ja mahdollisesti suuria reaalitalouden kustannuksia. Tutkimuksessa testataan lukuisten yksittäisten ja komposiitti-indikaattoreiden kykyä ennustaa systeemisiä tapahtumia. Näiden indikaattorien arvioinnissa otetaan huomioon päätöksentekijöiden preferenssit väärin hälytysten ja tapahtumien huomiotta jättämisen välillä. Tulokset korostavat monien eri indikaattorien yhtäaikaisen huomioonoton tärkeyttä monesta muuttujista koostuvassa mallissa. Tulosten mukaan kotimaisten ja globaalien makrorahoitusherkkyysien yhtäaikainen huomioonotto parantaa huomattavasti diskreetin valinnan mallien kykyä ennustaa systeemisiä tapahtumia. Tutkimuksessa luotu menetelmä ennustaa viimeaikaisen finanssikriisin hyvin. Malli olisi antanut varhaisen varoitusmerkin Yhdysvalloille vuoden 2006 toisella neljänneksellä eli viisi neljännestä ennen rahamarkkinajännitteiden ilmaantumista elokuussa 2007.

Avainsanat: varhaiset varoitusmerkit, varallisuushintojen nousut ja laskut, rahoitusmarkkinoiden jännitteet, vakauspoliittika.

1 Introduction

The current financial turmoil has demonstrated the importance of understanding and measuring systemic risks and predicting systemic events, i.e. events *when financial instability becomes so widespread that it impairs the functioning of the financial system to the extent that economic growth and welfare suffer materially*.³

Borio and Lowe (2002, 2004) show that widespread financial distress typically arises from the unwinding of financial imbalances that build up disguised by benign economic conditions, such as periods of stable and low inflation. Using annual data for 34 countries for the period 1960-1999, they show that sustained rapid credit growth combined with large increases in asset prices increased the probability of episodes of financial instability. Recently, Cardarelli et al. (2011), using data for 17 main advanced economies, show that the likelihood that stress in the financial system causes more severe economic downturns is higher when stress is preceded by the building up of balance sheet vulnerabilities in the form of a rapid expansion of credit, and a run-up in house prices. Moreover, in a paper closely related to our study, Misina and Tkacz (2009) investigate whether credit and asset price movements can help to predict financial stress in Canada by using linear and non-linear threshold models. According to their findings, business credit emerges as an important leading indicator among all variables considered in their study.

The paper builds upon the above studies and contributes to the financial crisis literature by developing a unified framework for assessing systemic risks, stemming from domestic and global macro-financial vulnerabilities, and for predicting (out of sample) systemic events i.e. periods of extreme financial instability with potential real costs. Within this framework it is possible to assess the relative importance of the factors contributing to the probability of a systemic event. It is also possible to identify potential vulnerabilities on the basis of a scenario analysis of the evolution in the domestic and global macro-financial environment.

We extend the existing literature on predicting financial crises in several ways. First, we identify past systemic events by using a composite index measuring the level of systemic tensions in the financial system of one country. This approach provides an objective criterion for the definition of the starting date of a crisis and it contrasts with the standard way of identifying crises that exploits qualitative information (see e.g. Laeven and Valencia, 2008). More specifically, we refine the approaches of the IMF (2009) and the ECB (2009a) by calculating a country specific Financial

³ See the definition of the concept of systemic risk in the ECB Financial Stability Review, December 2009.

Stress Index (FSI) with a robust method of aggregation. This reduces the revisions to the index due to the arrival of new information and, therefore, it makes the index more suitable for the use in econometric models. We then identify systemic events as episodes of extreme financial stress that has led to negative real economic consequences on average⁴. In this way, we avoid the selection bias that would occur if we chose only cases where extreme financial stress has always led to negative real economic consequences. The selection bias could emerge because a policy action (that we do not control for) might have prevented the negative economic outcome. Our approach to identify systemic events can be seen as an extension of Eichengreen *et al.* (1995), where an index of exchange market pressure is used to identify currency crises⁵. Compared to Eichengreen *et al.* (1995), our financial stress index is broader than the exchange market pressure index because it includes several market segments. This enables us to identify episodes that are truly systemic and not segment specific. In addition, by defining systemic events as episodes of extreme financial stress with potential real economic consequences, we focus on financial crises that are relevant for policy makers who want to avoid real costs. The real cost dimension is absent in Eichengreen *et al.* (1995), where a simple statistical rule is used to identify crisis periods.

Second, in predicting the identified systemic events, we evaluate the joint role of domestic and global vulnerabilities. This strategy encompasses traditional approaches that look only at the role of domestic vulnerabilities (Bussière and Fratzscher, 2006 and Berg, Borensztein and Pattillo, 2005) and other studies showing that global vulnerabilities are important determinants of domestic financial instability (Alessi and Detken, 2011). In addition, we also analyse the role of the interactions between domestic factors and the interplay of global developments with the domestic conditions.

Third, we evaluate both "stand alone" macroprudential indicators of vulnerabilities and composite indicators calculated using discrete choice models. The evaluation of the two categories of indicators is done with a common methodology that takes into account policy makers' preferences between issuing false alarms and missing systemic events (Demirguc-Kunt and Detragiache, 1999, Bussière and Fratzscher, 2008 and Alessi and Detken 2011). We are therefore able to test

⁴ According to the analysis in section 2, we identify an episode of extreme financial stress (or a systemic event) when the FSI crosses the 90th percentile of the country distribution. In section 2, we show that the 90th percentile threshold anticipates real costs on average.

⁵ The index is calculated as equal variance weighted average of exchange rate changes, interest rate changes, and reserve changes. Crises are defined as periods when the pressure index is at least two standard deviations above the mean.

whether a multivariate analysis based on discrete choice models outperforms stand alone indicators recently suggested by the literature.

Our empirical analysis covers a set of 28 emerging market and advanced economies with quarterly data between 1990 Q1 and 2009 Q4. Our results show that stand alone measures of asset price misalignments and credit booms are in general useful leading indicators of systemic events. Interestingly, global measures of credit expansion and asset price developments perform better than indicators of domestic fragilities⁶. Interactions between domestic variables as well as between global and domestic variables are among the best stand alone indicators. However, our results highlight the importance of considering jointly various indicators in a multivariate framework, as we find that discrete choice models outperform “stand alone” indicators in predicting systemic events. We find that taking into account jointly domestic and global macro-financial vulnerabilities greatly improves the performance of discrete choice models in forecasting systemic events. In addition, considering interactions between domestic and global macro-financial vulnerabilities further improves the performance of the model. Our framework displays a good out-of-sample performance in predicting the last financial crisis. Our model would have issued an early warning signal for the United States in 2006 Q2, 5 quarters before the emergence of the tensions in money markets in August 2007 that started the crisis. Our analysis show that both domestic (credit cycle and macro-overheating) and global factors (equity valuations and macro-overheating) were important determinants of systemic risk in the United States in the period before the crisis. Knowing the sources of systemic risk can guide the policy maker in choosing policy responses. Some risks can be mitigated by domestic policies, however, the importance of global factors as sources of systemic risk suggests that international cooperation and coordinated policy action are crucial to preserve global financial stability.

The remainder of the paper is organised as follows. Chapter 2 introduces the measure for financial stress and defines systemic events. Chapter 3 describes the data. Chapter 4 presents the empirical analysis, while Chapter 5 concludes.

⁶ These results are also supported by the conclusions of with Borio and Lowe (2002), Gerdesmeier et al. (2009) and Alessi and Detken (2011)

2 Measuring financial stress and identifying systemic events

Construction of the financial stressindex

To identify systemic events, we construct a *Financial Stress Index (FSI)* for each country in our sample, and evaluate at which levels it has, on average, had negative implications for the real economy. This approach, by capturing systemic tensions in the financial system of a country, provides an objective criterion for the definition of the starting date of a systemic event. Furthermore, it contrasts with the standard way of identifying crises that exploits qualitative information (see for example Laeven and Valencia, 2008).

Typically, when negative shocks, such as bursts of asset price bubbles, or banking, financial and currency crises hit the economy, it is possible to observe tensions in several market segments. The larger and broader the shock is (i.e. the more systemic the shock is), the higher the co-movement among variables reflecting tensions. Therefore, by aggregating variables measuring stress across markets segments, it is possible to create a Financial Stress Index that captures the start and the evolution of a crisis⁷.

Our FSI is a country-specific composite index, covering the main segments of the domestic financial market, and it consists of the following five components: (1) the spread of the 3-month interbank rate over the 3-month Government bill rate (Ind_1)⁸; (2) negative quarterly equity returns (multiplied by minus one, so that negative returns increase stress; positive returns are disregarded and set to 0) (Ind_2); (3) the realised volatility of the main equity index (Ind_3); (4) the realised volatility of the nominal effective exchange rate (Ind_4); and (5) the realised volatility of the yield on the 3-month Government bill (Ind_5)⁹.

Each component j of the index for country i at quarter t is transformed into an integer that ranges from 0 to 3 according to the country-specific quartile of the distribution the observation at quarter t belongs to ($q_{j,i,t}$). For example, a value for component j falling into the fourth quartile of

⁷ Several indices for the measurement of financial tensions have been developed in recent years, examples are Illing and Liu (2006) and Hakkio and Keeton (2009). The IMF (2008, 2009) presented the work by Cardarelli et al. (2009) and Balakrishnan et al. (2009), who constructed financial stress indices for a broad set of advanced and emerging economies. The ECB (2009a) presented a Financial Stress Index for the global economy, based on work by Fidora and Straub (2009). It should be highlighted that research on the measurement of financial stress and the construction of indices capturing systemic events is currently very active (see e.g. Hollo, Kremer and Lo Duca, 2010).

⁸ Or the spread among the interbank and the T-bill rates of the closest maturity to three months when the latter maturity is not available.

⁹ In the calculation of realised volatilities for equity, nominal effective exchange rate and Government bill rate, i.e. components (Ind_3) to (Ind_5), average daily absolute changes over a quarter were used.

the distribution would be transformed into “3”¹⁰. Note that each variable is measured in a way that higher values indicate higher stress levels, therefore higher values of the transformed variables indicate higher stress.

The Financial Stress Index is computed for country i at time t as a simple average of the transformed variables as follows:

$$FSI_{i,t} = \frac{\sum_{j=1}^5 q_{j,i,t}(Ind_{j,i,t})}{5} \quad (1)$$

Hollo, Kremer and Lo Duca (2010) show that the standardization method based on quartiles that we use is more robust than a standardization based on mean and variance, especially when the number of components of the index is small. More specifically, with the “quartile” standardization method, adding new observations to the sample produces only small revisions to the historical levels of the index (*ex post* stability). Large revisions of the historical levels of the index would complicate the analysis of the Financial Stress Index and its use in econometric models¹¹.

In calculating the Financial Stress Index, we face a trade off between the degree of precision of the index at the country level and the degree of homogeneity of the index across countries and time. For some countries with more developed financial systems, it would be possible to calculate a more detailed Financial Stress Index aggregating the information from several financial instruments and several markets segments. The set of instruments and segments is, however, limited for the emerging economies that are included in our sample. Since in our study the cross country dimension prevails, we give more importance to the homogeneity of the Financial Stress Index across countries. We believe this does not affect our results mainly for two reasons. First, once some crucial segments of the financial system are included in the index, adding components to it does not substantially change the “shape” of financial stress indices (Hollo, Kremer and Lo Duca, 2010). Second, our focus is on the detection of systemic events, i.e. we look only at extreme values of financial stress. Identified extreme values are robust to the composition of the FSI.

¹⁰ The only exception to this standardisation method is the indicator for negative stock market returns. As this indicator is most of the time equal to zero, the standardisation through quartiles leads often to a very volatile variable jumping from 0 to 3 directly. Therefore to standardise this variable we just divide this indicator by its maximum value over the sample. We then rescale the transformed indicator so that it ranges from 0 to 3.

¹¹ Hollo, Kremer and Lo Duca (2010) discuss advantages and disadvantages of approaches for the calculation of financial stress indices.

The financial stress indices for countries in the sample are plotted in Figures A1 and A2 in the Appendix. As it can be seen from the Figures, the FSIs capture well past episodes of high financial stress or crises, such as the Asian financial crisis in 1997, the Russian crisis in 98, the burst of the IT bubble in 2000-2001 and the last Global Financial Crisis. For many advanced economies, the Global Financial Crisis led to the highest level of financial stress since the start of the sample in 1990, while in many emerging economies the level of financial stress was higher during the Asian financial crisis, or during some country-specific crisis, such as the Russian crisis of 1998 or the crisis in Argentina of 2001.

Financial stress and the real economy

Financial instability and stress can impact economic activity through various channels. First, shocks that affect the creditworthiness of borrowers tend to strengthen the output fluctuations through the financial accelerator, as changes in the values of collateral impact the willingness of the financial system to provide credit to the economy (Bernanke and Gertler, 1995, and Bernanke et al., 1999, Kiyotaki and Moore, 1997). Second, factors that impact lenders' balance sheets can magnify economic downturns as if banks' capital is weakened, banks may become more reluctant to provide capital to the real sector or can even be forced to deleverage, leading to sharper economic downturns (Bernanke and Lown, 1991, Kashyap and Stein, 1995). Third, the development and structure of the financial system determine the degree of interconnection between real and financial sectors in the economy (IMF, 2006, Rajan and Zingales, 2003).

Recently, Cardarelli et al. (2011) show that out of 113 financial stress episodes identified for 17 main advanced economies since 1980, 29 were followed by an economic slowdown and an equal number by recessions. The remaining 55 financial stress episodes were not followed by an economic downturn. The authors find that the median cumulative output losses (relative to trend or until recovery) in downturns that followed financial stress episodes were about 2.8 percent of GDP for slowdowns and about 4.4 percent of GDP for recessions. Real costs of downturns anticipated by financial stress episodes are found to be significantly larger than the real costs of downturns that were not preceded by financial stress.

Policy makers' main concern regarding financial stress is that financial instability could become so widespread that it would impair the functioning of the financial system to the extent that economic growth and welfare suffer materially. Therefore, it is important to study the relationship

between the Financial Stress Index and measures of real economic activity, and to calibrate the thresholds for the FSI at which negative economic outcomes have occurred in the past¹². One way to do this is to analyse the relationship between the Financial Stress Index and the real GDP.

Figure 1 reports the median deviation (in percent) of the real GDP from its trend¹³ (output gap) for different percentiles of the distribution of the Financial Stress Index (two quarters ahead). As it can be seen from the Figure, levels of the Financial Stress Index above the 90th percentile of the country distribution of the index anticipate negative deviations of the real GDP from its trend (i.e. economic slowdowns or recessions).

(INSERT FIGURE 1 HERE)

Furthermore, Figure 2 shows the evolution of the deviation of the real GDP from its trend in the quarters after financial stress reached the last decile of the country distribution. It can be seen that the high level of stress anticipates a significant slowdown in economic activity that lasts up to 5 quarters. During the period while GDP remains below the trend, the cumulated costs range between 4 to 5% of the GDP.

(INSERT FIGURE 2 HERE)

These findings confirm that high levels of financial stress should be a concern for policy makers, as they could lead to a slowdown of the economy or even to a loss of the level of the real output. Thus, our focus will be on periods of extreme financial stress.

In the analysis, we focus on episodes of extreme financial stress that have often, i.e. on median cases, led in the past to negative consequences for the real economy. We define these episodes as **systemic events**. We focus on the level of stress that on average have led to negative real consequences to avoid a selection bias by choosing only cases where financial stress has for certainty led to negative real economic consequences. This is because a policy action might have prevented the negative economic outcome.

In our benchmark case we identify systemic events when the Financial Stress Index crosses the 90th percentile of the country distribution. We adopt this threshold because on average it anti-

¹² Hakkio and Keeton (2009) perform a similar analysis of the linkages between financial stress shocks and economic performance.

¹³ The trend is calculated using Hodrick-Prescott filter with the smoothing parameter set to 1600.

pates real consequences in terms of negative deviation of real GDP from trend, as suggested by Figure 1. Following this approach, we identify a set of 94 systemic events. We find the following starting dates for well known crisis episodes in the 1990s and 2000s: 1994 Q1 for Brazil, 1994 Q4 for the Mexican crisis; 1997 Q2 for the Asian crisis in Thailand, 1997 Q3 for Hong Kong and other main Asian countries, 1998 Q3 for the Russian crisis, 1999 Q1 for the Brazilian crisis; 2001 Q3 for the Argentinean crisis; 2007Q3 for the last financial crisis in the United States. In many cases, these episodes spread to several other economies. For example, after starting in 2007 Q3 with severe problems in money markets and volatility in other market segments in the United States and in the euro area, the latest crisis spread in successive waves across countries in 2008 Q1, 2008 Q3 and finally it reached emerging markets in 2008 Q4. Several of the episodes that we identify are also in the list of crises compiled by Laeven and Valencia (2008).

Our approach to identify systemic events can be seen as an extension of Eichengreen *et al.* (1995), where an index of exchange market pressure is used to identify currency crises¹⁴. Compared to Eichengreen *et al.* (1995), our Financial Stress Index is broader than the exchange market pressure index, because it includes several market segments. This enables us to identify episodes that are truly systemic and not segment specific. In addition, by defining crises as episodes of extreme financial stress with potential real economic consequences, we focus on events that are relevant for policy makers who want to avoid real costs. The real cost dimension is absent in Eichengreen *et al.* (1995), where a simple statistical rule is used to identify crisis periods.

Definition of the dependent variable

The objective of the study is predicting the occurrence of systemic events within a given time horizon that in our benchmark specification is set to 6 quarters¹⁵. To do this, we proceed in three steps:

First, we transform the Financial Stress Index into a binary variable that we call “systemic event”. The variable takes value 1 in the quarter when the FSI moves above the predefined threshold of the 90th percentile of the country distribution that anticipates real consequences on average, as suggested by Figure 1.

¹⁴ The index is calculated as equal variance weighted average of exchange rate changes, interest rate changes, and reserve changes. Crises are defined as periods when the pressure index is at least two standard deviations above the mean.

¹⁵ The time horizon of 6 quarters is chosen because within this time interval policy makers can adopt measures to prevent the materialisation of systemic events. Shorter time horizons are less relevant for policy making because the potential for effective pre-emptive actions is lower. For robustness, we anyway try time horizons of 2, 4 and 8 quarters. The results are discussed in the section on the robustness tests.

Second, we set the dependent variable to 1 in the 6 quarters preceding the systemic event and to 0 in all the other periods. The dependent variable mimics an ideal leading indicator that perfectly signals “systemic events” by “flashing” in the 6 quarters before the event.

Finally, we drop from the sample all the observations that are not informative about the transition from tranquil times to systemic events. It means that we drop the periods when financial stress remains above the predefined threshold that identifies systemic events. We also drop tranquil periods that are not longer than 6 quarters, as the short distance between the extreme stress episodes delimiting them suggests that we should not consider these periods as “normal”.¹⁶

3 Data

To assess the level of systemic risks and to predict systemic events, we construct indicators commonly used in the macroprudential literature to predict crises (Borio and Lowe, 2002 and 2004, and Alessi and Detken, 2011). These indicators capture the building up of vulnerabilities and imbalances both in the domestic and global economy. In this regard, we focus on asset price and credit developments, valuation levels and proxies for leverage in the economy. However, we also control for macroeconomic conditions with a broad set of indicators.

We build a comprehensive dataset of quarterly macro and financial data for period 1990 Q1 – 2009 Q4 for 28 countries, of which 10 advanced¹⁷ countries and 18 emerging economies¹⁷. The data is obtained either from Haver Analytics, Bloomberg and Datastream¹⁸.

¹⁶ Bussière and Fratzscher (2006) point out that including in the estimation of early warning models the period of economic recovery after a crisis produces the so called “post crisis bias”. In recovery periods, economic variables go through an adjustment process before reaching again the path they have during tranquil periods. The recovery period therefore should be excluded from the analysis as it is not informative of the path leading from the pre-crisis regime to the crisis. Bussière and Fratzscher address this issue by using a multinomial logit model with “three regimes” for the dependent variable (calm period, crisis and recovery). In our paper, as we drop periods in which stress is high, potentially we already disregard recovery periods, at least partially. However, we check the robustness of our results by dropping observations up to two quarters after the end of the stress periods to ensure that the post crisis bias is addressed. Only marginal gains in the performance of the model are obtained when dropping the additional two quarters.

¹⁷ The advanced countries are the following: Australia, Denmark, Euro area, Japan, New Zealand, Norway, Sweden, Switzerland, the United Kingdom, and the United States. The emerging economies are the following: Argentina, Brazil, China, Czech Republic, Hong Kong, Hungary, India, Indonesia, Malaysia, Mexico, the Philippines, Poland, Russia, Singapore, South Africa, Taiwan, Thailand and Turkey.

¹⁸ Credit and money variables are seasonally adjusted using X12 seasonal adjustment procedure, and all real variables are deflated using CPI price index

Table 1 summarises the core variables included in the study.

(INSERT TABLE 1 HERE)

Following the literature (e.g. Alessi and Detken, 2011), we test several transformations of the indicators, such as annual changes and deviations from moving averages or trends¹⁹. To proxy for global macro-financial imbalances and vulnerabilities, we calculate a set of global indicators by averaging the transformed variables for the following four countries or regions: the United States, euro area, Japan and the United Kingdom²⁰.

Starting from the core set of variables listed in Table 1, we calculate interactions between domestic variables, between international variables and between domestic and international variables. Interactions are calculated by multiplying the relevant variables and are aimed to capture the joint dynamic of two indicators. The first group of interactions measures the joint dynamics of asset price growth and asset valuations. They are calculated by multiplying all possible measures of asset price growth by all possible measures of asset valuation included in our core set of indicators reported in Table 1. We calculate these interactions both for domestic and global variables. The second group of interactions measures the joint dynamics of credit growth and leverage. They are calculated by multiplying all possible core indicators of credit growth by core indicators of leverage. Also these interactions are calculated both for domestic and global variables. In addition, we calculate other domestic interactions capturing the joint dynamic of asset prices and credit by multiplying indicators of asset price growth or valuation by indicators of credit growth or leverage. We finally calculate interactions between domestic and global variables by multiplying domestic indicators of asset price growth and valuation, and indicators of credit growth and leverage by their equivalent indicators measured at the global level. In total, our dataset includes more than 200 indicators.

Our analysis is conducted as much as possible in a real-time analysis fashion²¹. At each point in time, only information available to the policy makers up to that point in time is used. This implies that we take into account that certain variables, as for example GDP, are not available to the

¹⁹ We estimate the trend with the Hodrick-Prescott filter. Following Borio and Lowe (2004), we use two different values of the smoothing parameter, namely 1600 and 400,000.

²⁰ We also calculated global averages by using weighted GDP averages of all countries in our sample. We only report the results for the global variables calculated using the United States, euro area, Japan and the United Kingdom as the results are substantially the same.

²¹ The literature on early warning models deals with large datasets of macro data for several countries of which several are emerging markets. "Real time datasets" that contain information on the revisions of data after the first publication do not exist yet for several countries in our sample. Our analysis is therefore a real-time analysis in the sense that we take into account publication lags, as in other early warning models (Alessi and Detken, 2011).

policy makers in real time because of publication lags. To take into account publications lags, we used lagged variables. For GDP, money and credit related indicators the lag ranges from 1 to 2 quarters depending on the country.

The real time analysis also implies that de-trended variables are computed using only real time information. Therefore, we recursively calculate trends at each time t , using only the information available up to that moment.

4 Empirical analysis

In order to test the performance of different stand alone indicators of vulnerabilities and their joint performance in the discrete choice model framework, we evaluate the indicators on the basis of assumptions on policy makers' preferences between issuing false alarms and missing systemic events.

In doing so, we calculate optimal thresholds for policy intervention for both stand alone indicators of vulnerabilities and for the probabilities of systemic events estimated with discrete choice models.

The remaining of the section is organised in the following way. First, we describe the approach used to extract early warning signals from the indicators, while taking into account policy makers' preferences. Second, we report the empirical investigation using stand alone measures of financial fragilities. Third, we report the empirical investigation with the discrete choice models. Fourth, we run an out-of-sample forecast exercise. Finally, we discuss the robustness of our analysis.

Evaluation of signals and calculation of optimal threshold for the indicators

To find out which vulnerabilities are the best indicators of systemic risks and systemic events, and to calibrate the optimal threshold for policy action, we follow the approaches by Demirguc-Kunt and Detragiache (2000) and Alessi and Detken (2011)²². The optimal threshold for policy action for an indicator is the one that maximises a measure of utility (i.e. "usefulness") that takes into account policy maker preferences between Type I and Type II errors²³. Once the optimal threshold is found for each indicator on the basis of a set of preferences, the best performing indicator is the one that

²² Bussière and Fratzscher (2008) also address the issue of policy maker preferences in calibrating the optimal early warning thresholds and the timing of policy interventions.

²³ Normally, the threshold for an indicator is chosen based on some kind of information criteria, e.g. noise-to-signal ratio. Bussière and Fratzscher (2008) and Alessi and Detken (2011) highlight that this approach has several drawbacks.

maximises the measure of usefulness among all indicators. We discuss next how to calculate the measure of usefulness for an indicator for a given threshold and set of preferences.

As it is common in the signalling literature (Kaminsky, Lizondo and Reinhart, 1998), a signal is issued when the indicator is above the predefined threshold. Consequently, we can classify the outcomes according to the following schema:

	Systemic event within a given time horizon	No systemic event within a given time horizon
The indicator is above the threshold (Signal)	A (correct signals)	B (wrong signals)
The indicator is below the threshold (No Signal)	C (missing signals)	D (correct absence of signals)

Following Alessi and Detken (2011), we then define a loss function that depends on the preferences of the policy maker between Type I and Type II errors:

$$L(\mu) = \mu (C / (A + C)) + (1 - \mu) (B / (B + D)) \quad (2)$$

The term $C / (A + C)$ is the share of systemic events that have not been signalled (i.e. the share of missing signals or Type I errors), while $B / (B + D)$ is the share of tranquil (normal) periods that were incorrectly signalled as systemic events (i.e. the share of false alarms or Type II errors).

The parameter μ describes the relative preference of the policy maker between Type I and Type II errors. For a value of $\mu = 0.5$, the policy maker is equally concerned about Type I and Type II errors. The policy maker is less concerned of missing signals when $\mu < 0.5$. Conversely, the policy maker is less concerned of issuing wrong signals when $\mu > 0.5$.

If the policy maker disregards the signal given by the indicator (i.e. either she assumes that a signal is never issued or that the signal is always issued) she faces a loss equal to $\text{Min} [\mu, 1 - \mu]$.

Thus, an indicator is “useful” for the policy maker if the loss obtained by ignoring the indicator is higher than the loss obtained by taking it into consideration. It is possible to define the usefulness U in the following way:

$$U = \text{Min} [\mu, 1 - \mu] - L(\mu) \quad (3)$$

The measure of usefulness U is computed for each indicator and for each threshold (and for different set of preferences). For a given set of preferences, the best threshold for an indicator is

the one that achieves the highest score in terms of U among the tested thresholds. The best indicator is the one that achieves the highest U among all the indicators.

At this stage it is important to clarify (i) how thresholds are selected and (ii) the assumptions on the parameter μ describing policy maker preferences.

We express thresholds as percentiles of the distribution of the indicators by country. This procedure generates country-specific cut-off levels for each indicator. Thus, our approach lies between those of Borio and Lowe (2004) and Alessi and Detken (2011). The former test the predictive power of constant cut-off levels across time and countries. The latter express thresholds at time t as percentiles of the distribution of the indicators by country up to time t , therefore the cut-off levels are country and time dependent. This approach is the ideal choice for real time analysis, as only the information available to policy makers in real time is used. In our paper, we have to depart from this approach as the length of our data sample does not allow us to compute percentiles in real time. However, we adopt the “pure” real time approach, when we evaluate the out-of-sample performance of our indicators in predicting the last Global Financial Crisis.

Regarding policy makers’ preferences, in our benchmark analysis we take the point of view of a policy maker who is equally concerned of issuing false alarms and missing systemic events, i.e. we assume that $\mu = 0.5$. This could be considered the point of view of a neutral external observer who does not want to commit any mistakes and is only concerned of correctly calling a systemic event. The point of view of local policy makers or international institutions in charge of giving policy recommendations could be different, as the costs of missing systemic events and issuing false alarms are different (e.g. through reputational costs or real costs). It is likely that the last financial crisis increased the concerns of policy makers of missing systemic events. However, it is difficult to assess whether policy makers could be assumed to be relatively more concerned of missing crises versus issuing false alarms²⁴.

Stand alone indicators of vulnerability

In the following, we test the predictive power of all the indicators that we have in our dataset that were described in Section 3. The dataset includes several domestic and global stand alone indicators of vulnerabilities inspired by the early warning system and macroprudential literature, and based on asset price (equity and property prices), credit (credit and monetary aggregates) and macro (GDP,

²⁴ For a more comprehensive discussion of the issue, see Bussière and Fratzscher (2008) and Alessi and Detken (2011).

inflation, government deficit, current account deficit) developments. We evaluate the performance of the different indicators according to the evaluation method discussed in the previous section.

Table 2 reports the best performing 5 global indicators (upper part) and domestic ones (lower part), as well as some statistics to assess the efficiency of the indicators in predicting systemic events over a horizon of 6 quarters, under the assumption that preferences are balanced between issuing false alarms and missing signals ($\mu=0.5$). More specifically, Table 2 reports usefulness U , the noise-to-signal ratio (NtSr), the percentage of systemic events predicted by the indicator (%predicted), the probability of a systemic event conditional to a signal (Cond Prob) and the difference between the conditional and the unconditional probability of a systemic event (Prob Diff)²⁵.

(INSERT TABLE 2 HERE)

The following observations can be made regarding stand alone indicators:

- Overall, the majority of indicators have a larger than zero “usefulness”, which means that the neutral observer would benefit from using the indicators rather than ignoring them²⁶.
- The best stand alone indicator among all is a global indicator, namely the percentage deviation of the ratio equity market capitalisation to GDP from the trend. This is in contrast with other studies that find that the deviation from trend of the ratio credit to GDP is the best indicator (Alessi and Detken, 2011 for example). However, according to our results, the deviation from trend of the global credit to GDP ratio ranks as second best stand alone indicator. In general, the performance of indicators based on equity prices is very similar to the performance of indicators based on credit aggregates.
- Credit indicators dominate indicators for monetary aggregates, as the latter do not appear among the top indicators. This confirms the finding of the literature that credit is a better predictor of financial crisis/stress than money aggregates (see e.g. Alessi and Detken 2011, Borio and Lowe 2004, or Schularick and Taylor 2009).
- Global indicators perform better than domestic indicators (in line with Alessi and Detken, 2011). The top 5 global indicators are the best performers among all indicators, while the first domestic indicator ranks only seventh among all the indicators.
- Interactions among indicators are important. The interaction between real equity prices growth and equity valuations (price/earning ratio) at the global level is among the top 5 global indicators. Also, among the domestic factors the interaction between growth in real equity prices and valuation ratios (price/earning ratio) ranks the second best.

²⁵ As in Kaminsky et al. (2008) the efficiency measures are calculated in the following way: the noise to signal ratio (NtSr) is the ratio between false signals as a proportion of periods in which false signals could have been issued and good signals as a proportion of periods in which good signals could have been issued (i.e. $NtSr = (B/(B+D))/(A/(A+C))$). The percentage of crisis predicted by the indicator (%predicted) is simply the ratio between good signals and the number of periods in which good signals could have been issued ($\% \text{ predicted} = A/(A+C)$). The probability of a crisis conditional on a signal (Cond Prob) is the ratio between good signals and the total number of signals issued ($\text{Cond Prob} = A/(A+B)$). Finally the difference between the conditional and the unconditional probability of a crisis (Prob Diff) is calculated as $\text{Cond Prob} - (A + C) / (A + B + C + D)$.

²⁶ The full set of results for stand alone indicators is available from the authors on request.

- “Overheating” at the global level (i.e. the percentage deviation from trend of the real GDP at the global level), figures among the top indicators.

Discrete choice models – a logit model

Next, we use a logit model to jointly estimate the impact of multiple vulnerability indicators to the probability of a systemic event. Furthermore, we calculate the optimal thresholds for policy intervention, by using the approach that takes into account policy makers’ preferences, described at the beginning of the Chapter. This approach departs from the standard practise used in the early warning models, where arbitrary thresholds for the estimated probabilities are used to signal incoming systemic events. The identification of early warning thresholds also helps the policy maker to judge in real time, whether the probabilities are elevated.

The benchmark specification of the logit model is the following:

$$\text{Prob}_{i,t}[\text{Dep}_{i,t} = 1] = \frac{e^{X_{it}\beta}}{1 + e^{X_{it}\beta}} \quad (4)$$

where $\text{Prob}_{i,t}[\text{dep}_{i,t} = 1]$ is the probability of a systemic event for a country i at time t within the next $h=[2,4,6,8]$ quarters, and X_{it} is the set of macro-financial vulnerabilities observed in country i at time t . As described in Section 2, we set our dependent variable $\text{dep}_{i,t}$ to 1 in the h quarters preceding the systemic event, and to 0 in all the other periods. The dependent variable mimics an ideal leading indicator that perfectly signals “systemic events” by “flashing” in the h quarters before the event. In our benchmark specification $h = 6$.

The country specific probability of a systemic event, i.e. systemic risk, is a function of financial vulnerabilities that are shown to perform well as stand alone indicators for predicting crises, according to the analysis carried out in the previous section. In our benchmark model, the explanatory variables are grouped into three main sets, namely the domestic, the global and the interactions between domestic and global factors.

The first set consists of variables that measure domestic conditions. It includes growth in domestic asset prices (equity) and bank credit, asset price valuation levels, and the level of leverage in the economy. In our benchmark specification, growth in equity prices and bank credit are measured by the real (net of inflation) annual growth of the local MSCI equity index and of the amount

outstanding of credit granted to the private sector. Asset price valuations are measured by the deviation of the ratio equity market capitalisation to GDP from its trend, while leverage is measured as the deviation of the ratio private credit to GDP from its trend²⁷. The domestic block of variables also includes the interaction between asset price developments and valuation levels, as well as the interaction between credit growth and leverage. The interactions are computed by the product of the two relevant variables. Finally, domestic macroeconomic environment is controlled for with the following variables: annual real GDP growth, annual CPI inflation, current account deficit in percentage of GDP, and government deficit in percentage of GDP.

The second set of explanatory variables aims at capturing the global macro-financial environment. These variables are included because from the recent literature on macroprudential indicators (Alessi and Detken, 2011) and from our empirical analysis of stand alone indicators of vulnerabilities, it emerges that global factors have a significant influence on domestic financial stability. Similarly to the domestic set of variables, we include growth in global asset prices and bank credit, global asset price valuation levels, and the global level of leverage to the model. In addition, the set of explanatory variables also includes the interaction between global asset price developments and valuation levels, as well as the interaction between global credit growth and leverage. Finally, global macroeconomic conditions are captured by real GDP growth and inflation.

The third set of explanatory variables includes the interplay between domestic and global indicators of vulnerabilities, computed as the product between the relevant domestic and international variables. The introduction of this group of variables captures additional fragilities that emerge when the overheating of the domestic economy coincides with the vulnerabilities in the global conditions.

In the robustness section, we evaluate our results by changing the specification of the benchmark model and the variables used to measure the different fragilities.

Regarding the estimation strategy, due data limitations, we pool the information of our unbalanced panel, and assume that the constant c and the slope coefficients β of the logit model do not

²⁷ Trends are computed with the Hodrick-Prescott filter setting the smoothing parameter λ to 400,000. Regarding, equity valuations it would be optimal to use price earning ratios, however time series for these data are not available since 1990 for a large portion of our set of countries. Therefore, we opted to use the ratio equity market capitalisation to GDP as a proxy for valuations after de-trending the ratio to correct for the non-stationarity due the progress in developing local stock markets. Regarding leverage, the deviation from the trend of the ratio private credit to GDP is a commonly used measure of leverage (Borio and Lowe, 2002), against the background of the lack of uniform coverage across countries and time of data on leverage of financial intermediaries, households and corporations.

change across time and countries. The appropriateness of a pooled approach is discussed by Fuertes and Kalotychou (2006) and Davis and Karim (2008).

To take into account country specific fixed effects and potential cross country differences in the scale of the regressors, as well as to avoid that our results are affected by large outliers, we follow the method by Berg, Borensztein and Pattillo (2005), and measure variables in country specific percentile terms.

Table 3 reports the estimated coefficients for the benchmark model (column 5), that includes the set of explanatory variables described above, as well as statistics for alternative models that are used for comparison (columns 1-4). The alternative models are:

- “Currency crisis” model: it includes explanatory variables often used in the currency crises literature, as for example the real exchange rate, macro conditions and credit growth (see table 3).
- “Macroprudential” model: it adds equity price growth and valuation to the set of explanatory variables of the “Currency crisis” model²⁸.
- “Domestic” model: it includes the explanatory variables of the “Macroprudential” model with the addition of (i) the general government deficit, (ii) the interaction between equity growth and equity valuation, and (iii) the interaction between credit growth and leverage.
- “Domestic and international (no interactions)” model: it includes the explanatory variables of the “Domestic” model with the exclusion of the interactions terms. It also includes global growth and inflation, as well as global credit growth and leverage, and global equity growth and valuation.

Table 3 also includes the estimated marginal effects of the independent variables in the benchmark model (column 6) and the estimated coefficients for two models that use data only for the 18 emerging markets included in our sample (columns 7 and 8).

As the main objective of the paper is to evaluate the performance of the models in predicting systemic events according to the framework that takes into account policy maker preferences, we draw the attention of the reader only on a few features that emerge from Table 3²⁹.

First, the fit of the benchmark model (column 5) is better than the fit of any alternative model (with the exclusion of the models that include emerging markets only). Second, the information criteria support the choice of the benchmark model and the inclusion of interaction terms³⁰. Third, in the benchmark model, domestic factors, as well as global factors and the interaction be-

²⁸ However, it excludes the real exchange rate due to data limitations.

²⁹ In addition, due to non-linearities and data transformations, the interpretation of the estimated coefficients is not straightforward.

³⁰ The AIC criterion for the models is the following: “Macroprudential” AIC=1235.8; “Domestic” AIC=1228.2; “Domestic and international (no interactions)” AIC=1041.2; “Benchmark” AIC=986.454. Models for emerging markets: “Domestic and international (no interactions)” AIC=623.9; “Benchmark” AIC=522.0.

tween domestic and global ones are statistically significant and have, in most cases, the expected signs. Fourth, the estimated coefficients (and marginal effects) capturing the impact of global variables as well as those capturing the interaction between domestic and global variables are, in most cases, larger in the model for emerging markets only. This suggests that the determinants of systemic risks are the same in emerging and advanced economies. The main difference between emerging markets and advanced economies is the relative importance of the different factors, with emerging economies being more exposed to global factors³¹.

We now turn to the evaluation of the performance of the models in predicting systemic events. The selection of the best model is done in the following way: once the probability of financial stress is estimated, we use the approach by Alessi and Detken (2011) to evaluate whether the policy maker can extract “useful” signals from it. Thus, we find the thresholds for the estimated probability that maximises the “U” statistic for each model (for the given preference parameter $\mu=0.5$). The best model is the one that achieves the highest usefulness U score for the given preference parameter.

(INSERT TABLE 4 HERE)

Table 4 reports the performance statistics for the above models. The main results are the following:

- First, all models have larger than zero usefulness score U that means that the models provide statistical gains for policy makers who are equally concerned of issuing false alarms and missing systemic events.
- Second, all the models except the “Currency crisis” model outperform the best stand alone indicators (see Table 2) in terms of their usefulness. The “Currency crisis” model, however, would rank third among stand alone indicators.
- Third, the benchmark model including global variables and all the interactions clearly outperforms the other models. The benchmark model successfully predicts more than 80% of the systemic events. Furthermore, the difference between the unconditional probability and probability of a systemic event conditional to observing a signal from the model is almost 40%.

To sum up, our results highlight that analysing multiple signals from various sources of vulnerabilities in a multivariate framework, such as the discrete choice models, are more comprehensive tools than stand alone indicators to assist policy makers in evaluating systemic risks and predicting systemic events. Furthermore, it is crucial to take into consideration both domestic and international sources of vulnerabilities as well as their interaction.

³¹ This conclusion is supported by Dungey et al. (2010).

We turn now to the evaluation of the out-of-sample performance of the logit models.

Out-of-sample performance of the models

We evaluate the out-of-sample performance of the Logit models over the evaluation period 2005 Q2 to 2007 Q2 (8 quarters) in the following way³²:

- 1) We recursively estimate the model at each quarter t in the evaluation period using the information that would have been available in real time from the beginning of the sample (1990 q1) to quarter t .
- 2) We collect the real time signals from the model over the evaluation period (assuming the benchmark scenario of a forecast horizon of 6 quarters and policy preference parameter of $\mu=0.5$).
- 3) We compute *ex post* the number of missed signals and false alarms issued by the model over the evaluation period and compute the measure of usefulness U described in the previous chapter.
- 4) We rank the models according to the usefulness parameter (U).

This approach provides a new, structured way to assess the out-of-sample performance of the models.

Table 5 summarises the results of the out-of-sample evaluation, which indicate that the benchmark model and the three alternative models would have been useful tools for policy makers in predicting the last financial crisis. As it was the case with in-sample predictions, the benchmark model that incorporates both domestic and global variables as well as their interactions outperforms by far the other models.

(INSERT TABLE 4 HERE)

Figure 3 shows the out-of-sample performance of the benchmark model for the United States for the last financial crisis. It shows that the probability of a systemic event within 6 quarters in the United States was close to the early warning threshold already in 2006 q1 and crossed it in 2006 q2. According to our Financial Stress Index, the switch from the tranquil period to the extreme financial stress period occurred in 2007 Q3, when the tensions in the money markets emerged and spread to other segments of the financial system. Therefore, our benchmark model is able to correctly anticipate the systemic event with a lead of 5 quarters. Furthermore, in the 5 quarters preceding the crisis, it flags that the systemic risks are elevated and a systemic event could be imminent.

(INSERT FIGURE 3 HERE)

³² We choose 2005Q2 as starting point for our real time evaluation of the model in order to have long enough time series for all the countries for expressing regressors in country specific percentiles. We stop the evaluation period in 2007Q2 as in 2007Q3 the last financial crisis started in the US.

Source of systemic risk and policy analysis

To choose the correct policy response, policy makers need to understand the factors contributing to the predicted systemic risks. Regarding the above example of the United States, Figure 4 shows the contribution of domestic and global macro-financial conditions to the probability of a systemic event as a percent share in 2006 Q2, and for comparison in 2006 Q4. As it is evident from the Figure, strong growth, as well as, buoyant equity markets occurring jointly at the domestic and global level (captured by the interaction between domestic and global factors) had the largest relative contribution to systemic risk (around 20.3%) in 2006 Q2. The other contributors to systemic risks were global equity market (18.7%); global macro conditions (17.7%), domestic credit cycle (12.9%), domestic equity market (9.4%) and global credit cycle (5.0%)³³. Knowing the sources of systemic risk can guide the policy maker in choosing policy responses. Some risks can be mitigated by domestic policies, however, the importance of global factors as sources of systemic risk suggests that international cooperation and coordinated policy action are important to preserve financial stability.

(INSERT FIGURE 4 HERE)

Robustness checks

In order to ensure the robustness of the results, we conducted the following robustness tests on discrete choice models:

- **Definitions of vulnerabilities.** We tested various definitions and transformations of the vulnerabilities (see Table 1). Overall, the results from the alternative models were qualitatively the same as in the benchmark model. Moreover, they always had relatively high positive values of the usefulness parameter U , compared to the stand alone indicators of vulnerabilities. For instance, regarding asset valuations, we used price/earnings ratios as it is common in the literature, and obtained similar results to the benchmark model. However, in this case, due to availability of data, our sample size was reduced and the analysis was not possible for all the countries in the sample.
- **Contagion effects.** To capture contagion effects we add the current average level of financial stress at the global level or, alternatively, in the region, to the set of explanatory variables. Adding contagion variables does not improve the performance of the benchmark model in predicting systemic events.
- **Role of capital inflows.** Adding net capital inflows to the set of explanatory does not increase the performance of the model. The impact of capital inflows is indirectly captured by domestic asset price and credit dynamics.

³³ Potentially, it would be possible to decompose further systemic risk and calculate the contribution of each of the variables included in the logit model.

- **Forecasting horizon.** We test the following forecast horizons for predicting systemic events: 2 quarters, 4 quarters, 6 quarters (benchmark) and 8 quarters. Overall, the performance of the model is relatively robust across forecasting horizons (see Table A1 in the Appendix). The best performance is achieved on average over the 8 quarter period, followed by the 6 quarter period. Normally, over the 4 and 2 quarter periods, the model performance slightly decreases³⁴.
- **Policy markers' preferences.** In our benchmark analysis we assume that the policy maker has the preferences of a neutral observer (she is equally concerned of Type I and Type II errors). If we change this assumption (see Table A2 in the Appendix), we find that overall policy makers would benefit from the signals of the models as their usefulness score is positive, especially when the policy maker's preferences are close to the balanced preferences (i.e. either $\mu=0.4$ or $\mu=0.6$).
- **Post crisis bias.** We test whether our results are affected by a post crisis bias. Bussière and Fratzscher (2006) point out that including in the estimation of early warning models the economic recovery period after a crisis produces so called "post crisis bias". In recovery periods, economic variables go through an adjustment process before reaching again the path they have during tranquil periods. The recovery period, therefore, should be excluded from the analysis as it is not informative of the path leading from the pre-crisis regime to the crisis. Bussière and Fratzscher address this issue by using a multinomial logit model with "three regimes" for the dependent variable (calm period, crisis and recovery). In this paper, as we exclude from the estimation sample the periods in which financial stress is high following the switch from tranquil regime to an extreme stress regime, we at least partially disregard some periods of economic recovery. However, we check the robustness of our results by excluding observations up to two quarters after the end of the stress periods to ensure that the post crisis bias is addressed. Only marginal gains in the performance of the model are obtained when dropping the additional two quarters from the sample.

5 Conclusions

This paper contributes to the financial crisis literature by developing a unified framework for assessing systemic risks, stemming from domestic and global macro-financial vulnerabilities, and for predicting (out-of-sample) systemic events i.e. periods of extreme financial instability with potential real costs.

We extend the existing literature on predicting financial crises in several ways. First, we identify past systemic events by using a composite index measuring the level of systemic tensions in the financial system of one country. Second, in predicting the identified systemic events, we evaluate the joint role of domestic and global vulnerabilities. In addition, we also analyse the role of the interactions between domestic factors and the interplay of global developments with the domes-

³⁴ The performance also decreases for time horizons longer than 8 quarters.

tic conditions. Third, we evaluate both "stand alone" macroprudential indicators of vulnerabilities and composite indicators calculated using discrete choice models. The evaluation of the indicators is done with a common methodology that takes into account policy makers' preferences (Demirgüç-Kunt and Detragiache, 1999, Bussière and Fratzscher, 2008 and Alessi and Detken 2011).

Our empirical analysis covers a set of 28 emerging market and advanced economies with quarterly data since 1990. Our results show that stand alone measures of asset price misalignments and credit booms are in general useful leading indicators of systemic events. Interestingly, global measures of credit expansion and asset price developments perform better than indicators of domestic fragilities³⁵. Interactions between domestic variables as well as between global and domestic variables are among the best stand alone indicators. However, our results highlight the importance of considering jointly various indicators in a multivariate framework, as we find that discrete choice models outperform "stand alone" indicators in predicting systemic events. We find that taking into account jointly domestic and global macro-financial vulnerabilities greatly improves the performance of discrete choice models in forecasting systemic events. In addition, considering interactions between domestic and global macro-financial vulnerabilities further improves the performance of the model.

Our framework displays a good out-of-sample performance in predicting the Global Financial Crisis. Our model would have issued an early warning signal for the United States in 2006 Q2, 5 quarters before the emergence of the tensions in money markets that started the crisis in August 2007. Our analysis shows that both domestic (credit cycle and macro-overheating) and global factors (equity valuations and macro-overheating) were important determinants of systemic risk in the United States in the period before the crisis. Knowing the sources of systemic risk can guide the policy maker in choosing policy responses. Some risks can be mitigated by domestic policies, however, the importance of global factors as sources of systemic risk suggests that international cooperation and coordinated policy action are crucial to preserve global financial stability.

In principle, the framework can also be used to identify potential vulnerabilities on the basis of a scenario analysis of the evolution of the domestic and global macro-financial environment.

³⁵ These results are also supported by the conclusions of with Borio and Lowe (2002), Gerdesmeier et al. (2009) and Alessi and Detken (2011)

References

- Alessi and Detken (2011). “Quasi real time early warning indicators for costly asset price boom/bust cycles: A role for global liquidity”. *European Journal of Political Economy*, Vol. 27, issue I.
- Balakrishnan, Danninger, Elekdag and Tytell (2009). “The Transmission of Financial Stress from Advanced to Emerging Economies”. IMF Working Paper WP/09/133.
- Bernanke and Gertler (1995). “Inside the Black Box: The Credit Channel of Monetary Policy Transmission”. *Journal of Economic Perspectives*, Vol. 9 (Autumn), pp. 27–48.
- , and Gilchrist (1999). “The Financial Accelerator in a Quantitative Business Cycle Framework”. *Handbook of Macroeconomics*, Vol. 1C, pp. 1341–93 (Elsevier: Amsterdam).
- Bernanke and Lown (1991). “The Credit Crunch”. *Brookings Papers on Economic Activity*, Vol. 22 (1991–2), pp. 205–48.
- Bordo, Dueker and Wheelock. (2000). “Aggregate Price Shocks and Financial Instability: An Historical Analysis”. NBER Working Paper No. 7652.
- Borio and Lowe (2002). “Asset Prices, Financial and Monetary Stability: Exploring the Nexus”. BIS Working Papers, No. 114.
- (2004). “Securing Sustainable Price Stability: Should Credit Come Back from the Wilderness?”. BIS Working Papers, No. 157.
- Borio and Drehmann (2009). “Assessing the Risk of Banking Crises – Revisited”. *BIS Quarterly Review*, March.
- Bussière and Fratzscher (2006). “Towards a new early warning system of financial crises”. *Journal of International Money and Finance*, Vol. 25(6), pp. 953-973.
- (2008). “Low probability, high impact: Policy making and extreme events”. *Journal of Policy Modeling* 30 (2008), pp. 111-121.
- Cardarelli, Elekdag, and Lall (2011). “Financial stress and economic contractions”. *Journal of Financial Stability*, Vol. 7.
- Davis and Karim (2008). “Comparing Early Warning Systems for Banking Crisis”. *Journal of Financial Stability*, Vol. 4, Issue II.
- Demirguc-Kunt and Detragiache (2000). “Monitoring Banking Sector Fragility. A Multivariate Logit”. *World Bank Economic Review*, Vol, 14, Issue 2.
- Demirguc-Kunt and Detragiache (2005). “Cross Countries Empirical Studies of Banking Distress: A survey”. IMF Working Paper, WP/05/96.
- Dungey, Fry, Gonzalez-Hermosillo and Martin (2010). “Are All Crises Alike?”. Oxford University Press, New York.
- ECB (2009a). “Global Index for Financial Turbulence”, a box in Financial Stability Review, December 2009.
- (2009b). “The concept of systemic risk”, a special feature in Financial Stability Review, December 2009.

- Eichengreen, Rose and Wyplosz (1996). "Contagious currency crises: first tests". *Scandinavian Journal of Economics*, Vol. 98, Nr. 4.
- Fidora and Straub (2009). "Global Index for Financial Turbulence". ECB mimeo.
- Fuertes and Kalotychou (2006). "Early Warning System for Sovereign Debt Crisis: the role of heterogeneity". *Computational Statistics and Data Analysis*, Vol. 5, pp. 1420 – 1441.
- Gerdesmeier, Roffia and Reimers (2009). "Asset price misalignments and the role of money and credit". ECB Working Paper No. 1068.
- Hakkio and Keeton (2009). "Financial Stress: What is it, How can it be measured and Why does it matter?" Federal Reserve Bank of Kansas City Economic Review, Second Quarter 2009.
- Hollo, Kremer and Lo Duca (2010). "CISS – A composite indicator of systemic stress in the financial system". ECB Working Paper Series, forthcoming.
- Illing and Liu (2006). "Measuring financial stress in a developed country: An application to Canada". *Journal of Financial Stability* 2, 243 – 65.
- IMF, 2006. *World Economic Outlook, September 2006: How Do Financial Systems Affect Economic Cycles?* World Economic and Financial Surveys.
- , 2008. *World Economic Outlook, October 2008: Financial Stress and Economic Downturns*. World Economic and Financial Surveys.
- , 2009. *World Economic Outlook, April 2009: How Linkages Fuel the Fire: The Transmission of Financial Stress from Advanced to Emerging Economies*. World Economic and Financial Surveys.
- Kaminsky, Lizondo and Reinhart (1998). "Leading Indicators of Currency Crises". *IMF Staff Papers*, Vol. 45, No. 1, pp. 1 – 48.
- Kashyap and Stein (1995). "The Impact of Monetary Policy on Bank Balance Sheets". *Carnegie-Rochester Conference Series on Public Policy*, Vol. 42 (June), pp. 151–95.
- Kiyotaki and Moore (1997). "Credit Cycles". *Journal of Political Economy*, Vol. 105 (April), pp. 211–48.
- Laeven and Valencia (2008). "Systemic Banking Crises: A New Database". IMF Working Paper WP/08/224.
- Misina and Tkacz (2009). "Credit, Asset Prices, and Financial Stress". *International Journal of Central Banking* Vol. 5, No. 4, pp. 95-122.
- Rajan and Zingales (2003). "Banks and Markets: the Changing Character of European Finance". CEPR Discussion Papers, No. 3865.
- Schularick and Taylor (2009). "Credit Booms Gone Bust: Monetary Policy, Leverage Cycles and Financial Crises, 1870-2008". NBER Working Paper 15512.

Table 1 List of variables

Description	Domestic Variables						Global Average
	(1)	(2) Deviation from Moving Average (Short)	(3) Deviation from Moving Average (Long)	(4) Annual Change	(5) Deviation from Trend (short)	(6) Deviation from Trend (Long)	(7)
ratio money to GDP		x	x		x	x	x
ratio m2 to GDP		x	x		x	x	x
real effective exchange rate		x	x	x	x	x	
nominal effective exchange rate		x	x	x	x	x	
real GDP		x	x	x	x		x
consumer price index		x	x	x	x		x
ratio credit to the private sector to GDP		x	x		x	x	x
real money				x			x
real m2				x			x
real house prices				x			x
real equity prices (MSCI based)				x	x	x	x
real credit to the private sector to GDP				x			x
general government debt (% of GDP)	x						
general government deficit (% of GDP)	x						
current account deficit (% of GDP)	x						
price earning ratios	x				x	x	x
stock market capitalisation over GDP					x	x	x

Notes: The table lists the core set of variables used in the empirical analysis; other transformations of the variables (i.e. interactions among them) that are used in the analysis are described in the main text. The first column reports the description of the variable; column (1) indicates whether the level of the original variable is used in the analysis; columns (2) and (3) indicate whether percentage deviations from short (8 quarters) and long (20 quarters) moving averages of the variable are used; column (4) indicates whether the annual percentage change of the original variable is used in the analysis; columns (5) and (6) indicate whether percentage deviations from Hodrick-Prescott trends of the variable are used; the “short” (“long”) Hodrick-Prescott trend is computed with the smoothing parameter λ set to 1600 (400000) following Borio and Lowe (2004); The last column (7) indicates whether global averages have been computed for all the transformations of the variables listed in columns (1) to (6).

Table 2 Performance of stand alone indicators of vulnerabilities ($\mu=0.5$ and forecasting horizon 6 quarters)

Global variables						
Indicator	Threshold (percentile)	U	NtSr	%Predicted	Cond Prob	Prob Diff
Percentage deviation of the ratio equity market capitalisation to GDP from Hodrick-Prescott ($\lambda=400000$) trend	55	0.21	0.45	76.91%	48.44%	18.72%
Percentage deviation of the ratio private credit to GDP from Hodrick-Prescott ($\lambda=1600$) trend	55	0.21	0.43	73.09%	49.36%	19.63%
Percentage deviation of real GDP from Hodrick-Prescott ($\lambda=1600$) trend	60	0.16	0.50	64.69%	45.93%	16.21%
Interaction between real equity prices (percentage deviation from Hodrick-Prescott ($\lambda=400000$) trend) and price earning ratios	67	0.16	0.45	57.22%	48.80%	18.99%
Interaction between international and domestic real credit growth	63	0.16	0.48	61.04%	44.32%	16.47%
Domestic variables						
Indicator	Threshold (percentile)	U	NtSr	%Predicted	Cond Prob	Prob Diff
Percentage deviation of the ratio equity market capitalisation to GDP from Hodrick-Prescott ($\lambda=400000$) trend	68	0.15	0.46	54.22%	48.67%	18.44%
Interaction between real equity prices (percentage deviation from Hodrick-Prescott ($\lambda=400000$) trend) and price earning ratios	71	0.14	0.44	50.21%	49.68%	19.52%
Percentage deviation of real equity prices from Hodrick-Prescott ($\lambda=400000$) trend	55	0.13	0.60	66.04%	42.38%	11.92%
Percentage deviation of the ratio private credit to GDP from Hodrick-Prescott ($\lambda=400000$) trend	55	0.12	0.63	66.34%	38.73%	10.28%
Percentage deviation of the private credit GDP ratio from 20 quarter moving average	79	0.12	0.35	37.54%	51.52%	24.49%

Notes: The Table reports the top 5 indicators for both the global (upper part) and domestic (lower part) category, and the optimal threshold in terms of percentile of the country distribution of the indicator at which the indicator issues a signal. Thresholds are calculated for $\mu=0.5$ (neutral observer). The Table also reports in columns the following measures to assess the efficiency of indicators: usefulness "U" (see formula 3); the noise to signal ratio (NtSr) i.e. the ratio between false signals as a proportion of periods in which false signals could have been issued and good signals as a proportion of periods in which good signals could have been issued ($NtSr = (B/(B+D))/(A/(A+C))$); the percentage of crisis predicted by the indicator (%predicted) i.e. the ratio between good signals and the number of periods in which good signals could have been issued (% predicted = $A/(A+C)$); the probability of a crisis conditional on a signal (Cond Prob) i.e. the ratio between good signals and the total number of signals issued (Cond Prob = $A/(A+B)$); the difference between the conditional and the unconditional probability of a jump (Prob Diff) i.e. $Cond Prob - (A + C) / (A + B + C + D)$.

Table 3 Estimation results of the Logit models ($\mu=0.5$ and forecasting horizon 6 quarters)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Currency crisis model	Macro Prudential model	Domestic model	Domestic and international (no interactions)	Benchmark	Benchmark (marginal effects)	Domestic and international (no interactions only EMEs)	Benchmark (only EMEs)
Domestic variables	real GDP growth	0.0152 ***	0.0069 **	0.0047	0.0057	0.0066 *	0.0008 *	-0.0002
	inflation	0.0027	0.0082 ***	0.0070 **	0.0050	0.0061	0.0008 *	0.0164 ***
	current account deficit	0.0012	0.0060 **	0.0070 ***	0.0079 **	0.0075 *	0.0009 **	0.0118 ***
	general government deficit			-0.0073 **	-0.0032	-0.0011	-0.0001	-0.0040
	real effective exchange rate overvaluation	0.0014						-0.0018
	real equity growth		0.0049 *	0.0015	0.0089 ***	0.0188 ***	0.0023 ***	0.0146 ***
	equity valuation		0.0291 ***	0.0266 ***	0.0142 ***	0.0119 ***	0.0015 ***	0.0011
	(A1) equity interaction (growth&valuation)			0.0063 **		0.0045	0.0006	0.0012
	real credit growth	0.0107 ***	0.0011	-0.0011	0.0013	-0.0068	-0.0008	0.0043
	leverage	0.0195 ***	0.0222 ***	0.0160 ***	0.0180 ***	0.0105 *	0.0013 *	0.0201 ***
	(B1) credit interaction (growth&leverage)			0.0032		0.0135 ***	0.0017 ***	0.0362 ***
interaction leverage&valuation			0.0067 **					
interaction equity&credit growth			0.0019					
Global variables	real GDP growth			0.0015	0.0004	0.0000	-0.0003	0.0003
	inflation			0.0170 ***	0.0303 ***	0.0038 ***	0.0182 ***	0.0685 ***
	real equity growth			-0.0014	-0.0123 ***	-0.0015 ***	-0.0062	-0.0177 ***
	equity valuation			0.0396 ***	0.0347 ***	0.0043 ***	0.0403 ***	0.0434 ***
	(A2) equity interaction (growth&valuation)				0.0133 ***	0.0017 ***		-0.0021
	real credit growth			0.0146 ***	0.0093	0.0012	0.0266 ***	0.0004
	leverage			0.0032	0.0415 ***	0.0052 ***	0.0017	0.1195 ***
(B2) credit interaction (growth&leverage)				-0.0430 ***	-0.0054 ***		-0.1292 ***	
Interaction between domestic and global variables	interaction domestic&international leverage				0.0078 **	0.0010 **		0.0103 *
	interaction domestic&international valuations				-0.0029	-0.0004		-0.0065
	interaction domestic&international credit growth				-0.0074 **	-0.0009 **		-0.0020
	interaction domestic&international equity growth				0.0168 ***	0.0021 ***		0.0355 ***
	A1 x A2				-0.0124 **	-0.0015 **		-0.0181 **
B1 x B2				-0.0008	-0.0001		0.0020	
Constant	-3.7562 ***	-5.1805 ***	-4.8874 ***	-8.4201 ***	-9.2450 ***		-9.2059 ***	-12.0821 ***
Number of countries	28	28	28	28	28		17	17
Number of observations	902	1275	1275	1275	1275		745	745
Pseudo R-squared	0.1278	0.1903	0.2036	0.3398	0.3894		0.3191	0.4585

Notes: Robust standard errors have been used in the estimation. *** denotes statistical significance at 1% level, ** at 5% level and * at 10% level.

Table 4 Performance of Logit models ($\mu=0.5$ and forecasting horizon 6 quarters)

Model	Threshold (percentile)	U	NtSr	%Predicted	Cond Prob	Prob Diff
Benchmark (EME only)	65	0.33	0.22	84.73%	63.24%	35.99%
Benchmark	68	0.32	0.20	80.95%	65.83%	37.83%
Benchmark (no interactions)	58	0.31	0.31	88.80%	55.52%	27.52%
Benchmark (no interactions - EME only)	68	0.30	0.24	78.33%	61.39%	34.14%
Domestic model	67	0.26	0.28	71.71%	57.79%	29.79%
Macro Prudential	63	0.24	0.34	74.23%	53.32%	25.32%
Currency crisis	69	0.19	0.38	60.33%	49.16%	22.33%

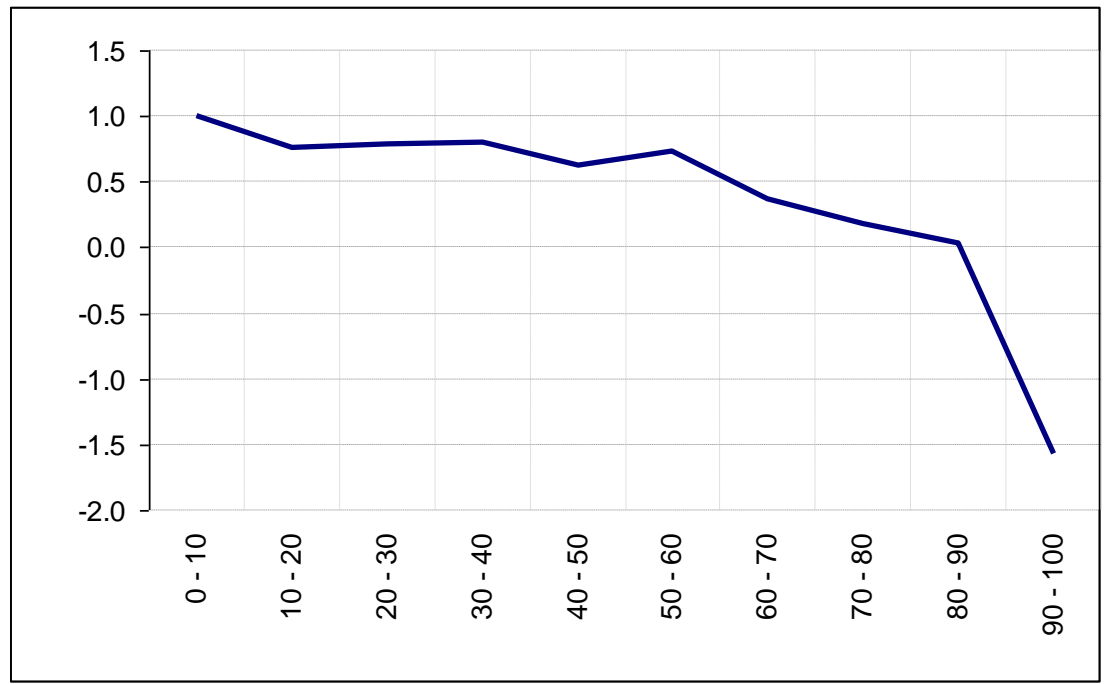
Notes: See notes to Table 2.

Table 5 Out-of-sample performance of Logit models ($\mu=0.5$ and forecasting horizon 6 quarters)

Model	U	NtSr	% Predicted	Cond Prob	Prob Diff
Benchmark (EMEs only)	0.22	0.55	97.50%	39.39%	13.08%
Benchmark	0.18	0.57	83.91%	48.34%	13.54%
Benchmark (no interactions)	0.15	0.64	85.06%	45.40%	10.60%
Domestic Model	0.12	0.67	72.41%	44.37%	9.57%
Macro Prudential	0.10	0.75	77.01%	41.61%	6.81%
Benchmark (no interactions - only EMEs)	0.09	0.81	100.00%	30.53%	4.22%
Currency Crisis	0.06	0.84	77.01%	38.95%	4.15%

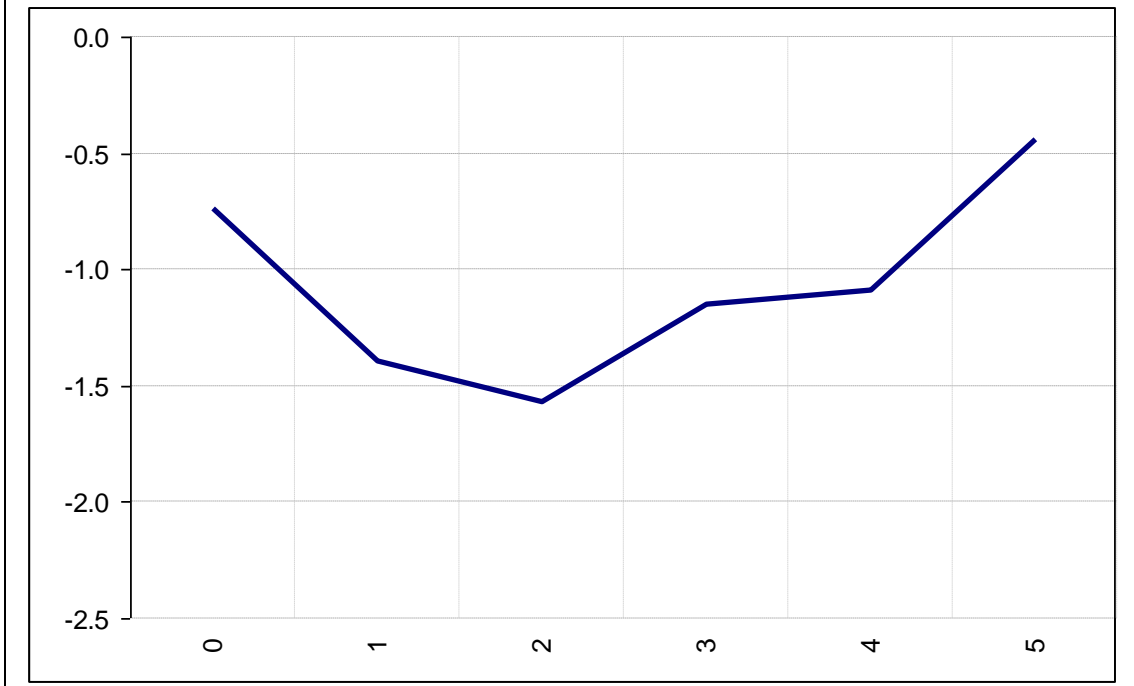
Notes: See notes to Table

Figure 1 Level of the financial stress index at quarter t and deviation of real GDP from trend at quarter $t+2$



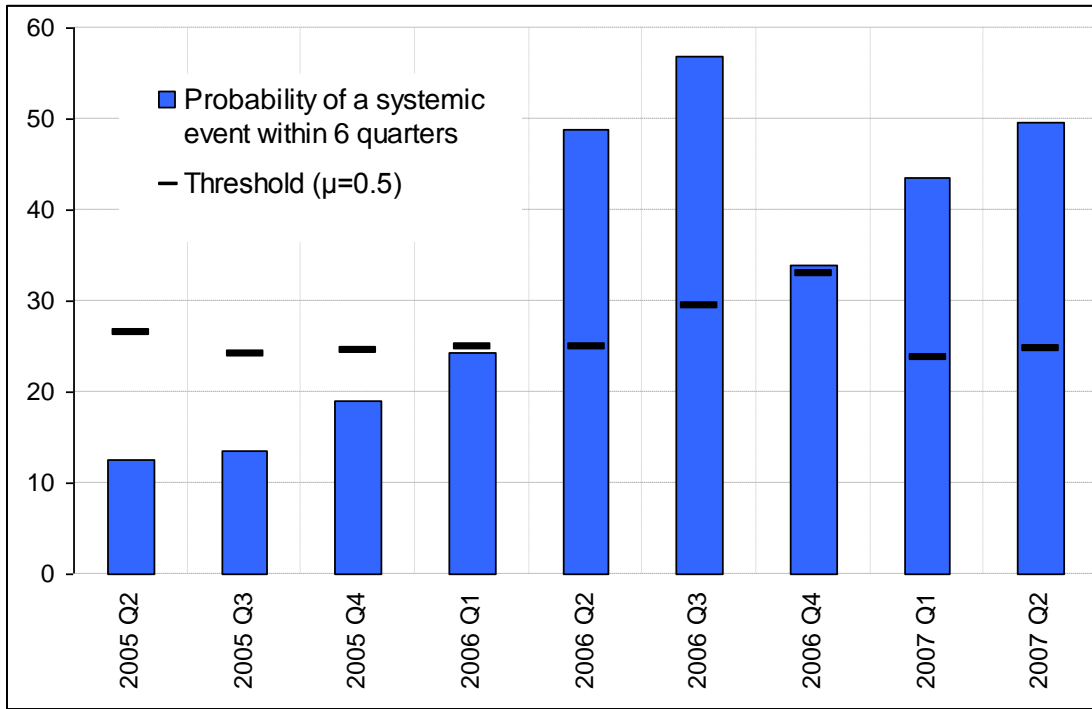
Note: The X-axis represents the percentile of the country distribution of the Financial Stress Index at time t , while the Y-axis represents real GDP deviation from its trend at time $t+2$, measured in percent of the trend.

Figure 2 Median real GDP loss in the quarters following extreme financial stress



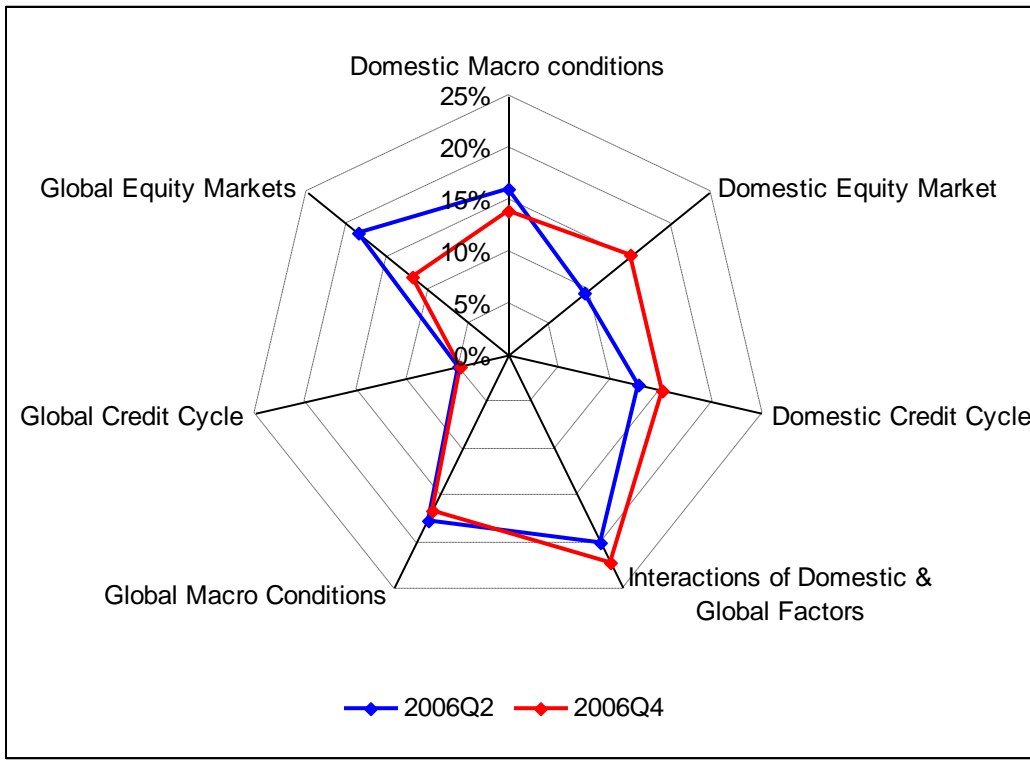
Note: Extreme financial stress is an episode when financial stress crosses the 90th percentile of the country distribution. The X-axis represents time (in quarters), while the Y-axis represents real GDP deviation from its trend, measured in percent of the trend.

Figure 3 Predicting the 2008/2009 financial crisis in the United States.
Out-of-sample performance of the benchmark logit model in the period 2005q2 – 2007q2



Note: The X-axis represents time (in quarters), while the Y-axis represents the probability of a systemic event within the next 6 quarters (threshold optimised for $\mu=0.5$). The probability is the output of the benchmark logit model.

Figure 4 Predicting the 2008/2009 financial crisis in the United States. Factors contributing to systemic risk (% share) in 2006q2 and 2006q4



Note: Domestic and global macro conditions include the following variables: real GDP growth, CPI inflation, current account deficit and General Government Deficit. Domestic and global equity market conditions include equity price growth, equity price level (misalignment) and their respective interaction terms. Domestic and global credit cycle include credit growth, proxy for leverage (deviation from trend credit to GDP ratio) and their respective interactions. Interactions of domestic and global factors include the interactions of domestic and global credit and equity price growth rates and misalignments.

Appendix

Figure A1 Financial Stress Index for emerging economies in the sample

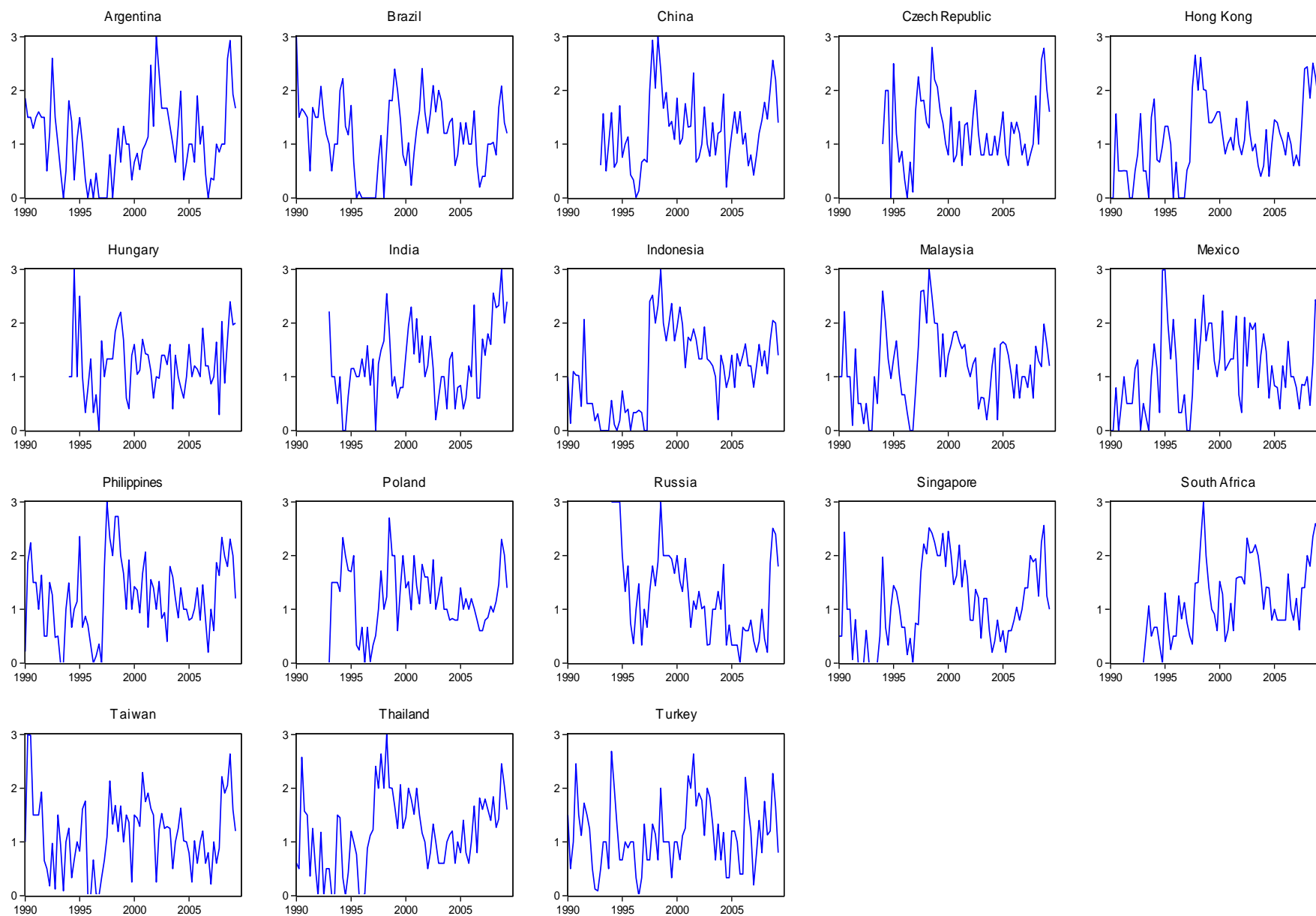


Figure A2 Financial Stress Index for advanced economies in the sample

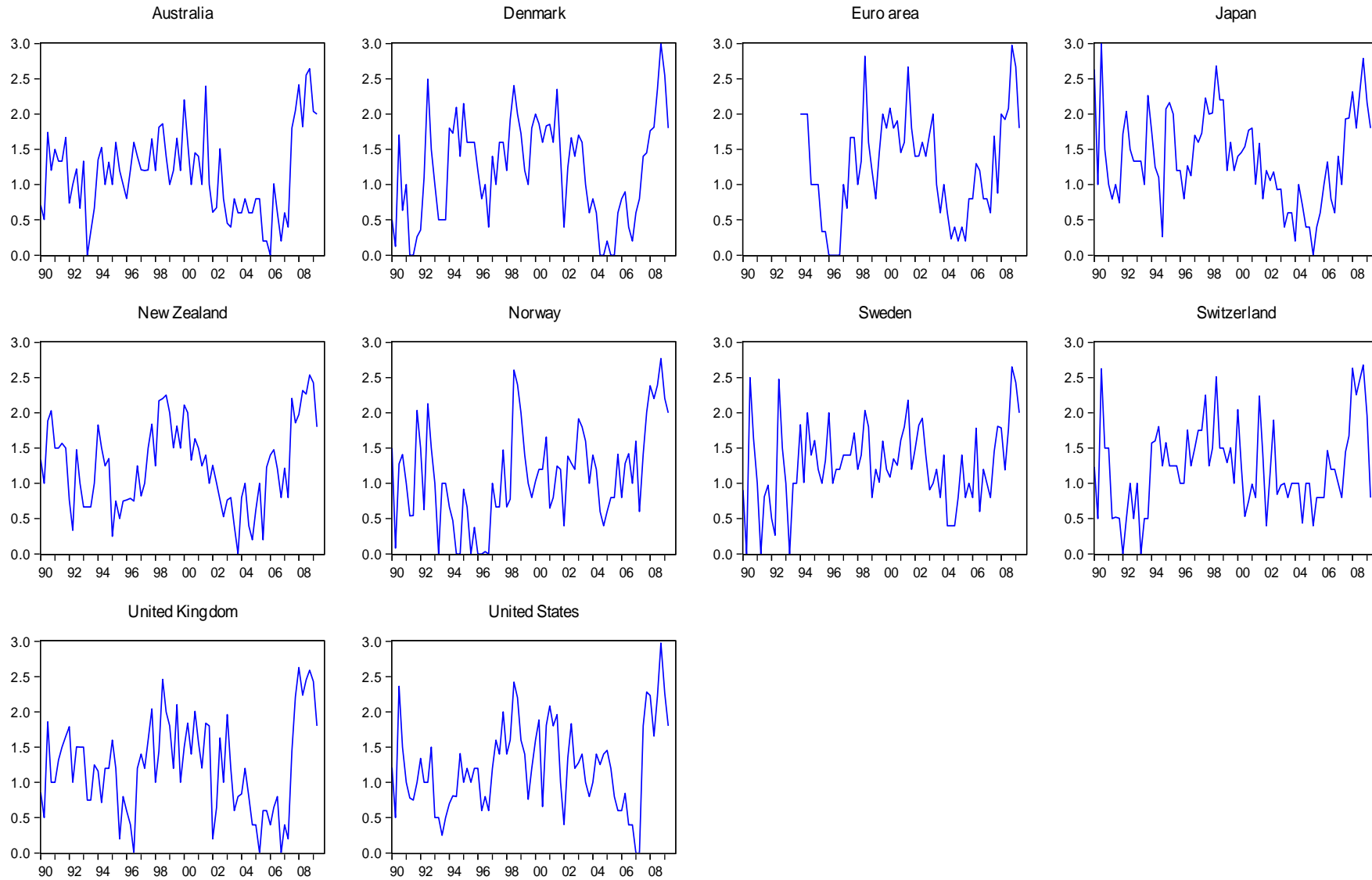


Table A1 Performance of the benchmark Logit model over different forecasting horizons ($\mu = 0.5$)

Model	Forecasting Horizon	Threshold (percentile)	U	NtSr	%Predicted	Cond Prob	Prob Diff
Benchmark	8 quarters	62	0.34	0.19	84.43%	74.76%	38.99%
Benchmark	6 quarters	68	0.32	0.20	80.95%	65.83%	37.83%
Benchmark	4 quarters	67	0.30	0.29	83.54%	45.21%	26.15%
Benchmark	2 quarters	74	0.29	0.28	80.80%	28.21%	18.41%

Notes: See notes to Table 2

Table A2 Performance of the benchmark Logit model using different values for the parameter μ (forecasting horizon 6 quarters)

Model	μ	Threshold (percentile)	U	NtSr	%Predicted	Cond Prob	Prob Diff
Benchmark	0.5	68	0.32	0.20	80.95%	65.83%	37.83%
Benchmark	0.4	69	0.23	0.19	79.83%	67.06%	39.06%
Benchmark	0.6	65	0.22	0.23	83.47%	62.74%	34.74%
Benchmark	0.7	53	0.14	0.37	91.88%	51.57%	23.57%
Benchmark	0.3	72	0.13	0.17	74.51%	69.27%	41.27%
Benchmark	0.8	53	0.07	0.37	91.88%	51.57%	23.57%
Benchmark	0.2	81	0.06	0.11	57.14%	77.27%	49.27%

Earlier BOFIT Discussion Papers

For a complete list of Discussion Papers published by BOFIT, see
www.bof.fi/bofit

- 2010
- No 1 Anatoly Peresetsky: Bank cost efficiency in Kazakhstan and Russia
 - No 2 Laurent Weill: Do Islamic banks have greater market power?
 - No 3 Zuzana Fungáčová, Laura Solanko and Laurent Weill: Market power in the Russian banking industry
 - No 4 Allen N. Berger, Iftekhar Hasan and Mingming Zhou: The effects of focus versus diversification on bank performance: Evidence from Chinese banks
 - No 5 William Pyle and Laura Solanko: The composition and interests of Russia's business lobbies: A test of Olson's "encompassing organization" hypothesis
 - No 6 Yu-Fu Chen, Michael Funke and Nicole Glanemann: Off-the-record target zones: Theory with an application to Hong Kong's currency board
 - No 7 Vladimir Sokolov: Bi-currency versus single-currency targeting: Lessons from the Russian experience
 - No 8 Alexei Karas, William Pyle and Koen Schoors: The effect of deposit insurance on market discipline: Evidence from a natural experiment on deposit flows
 - No 9 Allen N. Berger, Iftekhar Hasan, Iikka Korhonen, Mingming Zhou: Does diversification increase or decrease bank risk and performance? Evidence on diversification and the risk-return tradeoff in banking
 - No 10 Aaron Mehrotra and José R. Sánchez-Fung: China's monetary policy and the exchange rate
 - No 11 Michael Funke and Hao Yu: The emergence and spatial distribution of Chinese seaport cities
 - No 12 Alexey A. Ponomarenko and Sergey A. Vlasov: Russian fiscal policy during the financial crisis
 - No 13 Aaron Mehrotra and Alexey A. Ponomarenko: Wealth effects and Russian money demand
 - No 14 Asel Isakova: Currency substitution in the economies of Central Asia: How much does it cost?
 - No 15 Eric Girardin and Konstantin A. Kholodilin: How helpful are spatial effects in forecasting the growth of Chinese provinces?
 - No 16 Christophe J. Godlewski, Zuzana Fungáčová and Laurent Weill: Stock market reaction to debt financing arrangements in Russia
 - No 17 Zuzana Fungáčová, Laurent Weill, Mingming Zhou: Bank capital, liquidity creation and deposit insurance
 - No 18 Tuuli Koivu: Monetary policy, asset prices and consumption in China
 - No 19 Michael Funke and Michael Paetz: What can an open-economy DSGE model tell us about Hong Kong's housing market?
 - No 20 Pierre Pessarossi, Christophe J. Godlewski and Laurent Weill: Foreign bank lending and information asymmetries in China
- 2011
- No 1 Aaron Mehrotra and Jenni Pääkkönen: Comparing China's GDP statistics with coincident indicators
 - No 2 Marco Lo Duca and Tuomas Peltonen: Macro-financial vulnerabilities and future financial stress - Assessing systemic risks and predicting systemic events

Bank of Finland
BOFIT – Institute for Economies in Transition
PO Box 160
FIN-00101 Helsinki

 + 358 10 831 2268

bofit@bof.fi

<http://www.bof.fi/bofit>