The First Global Crisis of the Twenty-first Century: Was it Predictable?

Saktinil Roy  
Athabasca University  
Email: saktinil@athabascau.ca  

February 25th, 2009

I am grateful to Carmen Reinhart and Kenneth Rogoff for providing a large part of the database utilized for this paper. I would like to thank Paul McNeilis for his comments on a different but similar paper. Robert Shiller’s lecture on “Irrational Exuberance” at the University of Memphis as well as his speech in the panel discussion on financial crisis in the 2009 AEA meetings in San Francisco helped develop some of the ideas expressed in this paper. All errors are my own.

Financial support for this research from the Centre for Innovative Management of Athabasca University is gratefully acknowledged.
Abstract

Some authors find similarities between the US subprime crisis and other historical banking crises. Most notable are Reinhart and Rogoff (2008a, 2008b). They demonstrate that the US subprime crisis and historical banking crises in other countries share strikingly common patterns in the run-up of GDP growth, asset prices, accumulation of debt, current account deficits and capital flow. Thus, they conclude, certain changes in the behavior of these economic indicators are possible “precursors of crisis”. The present paper maintains that similarities across banking crises are best interpreted with robust statistical tests; on the other hand, an out-of-sample forecasting exercise must determine if the changes observed in an economic indicator signal a crisis.

The issues are addressed by answering three related questions: i) Could we predict the US crisis based on data on other historical crises? ii) How well in advance? iii) What indicators and which historical country experiences could give the best predictions? It is found that with a suitable probit model specification the recent US crisis could be predicted consistently starting in the year of 2003, but with data on a smaller group of historical crises and a smaller set of indicators (current account, public debt, and real estate price) than in Reinhart and Rogoff (2008a). At the same time, the prediction assumes introduction of an indicator that accounts for the relationship between growth of income inequalities and incidence of crisis. Thus, as with some other historical crises, protracted increases of income inequality in the United States contributed significantly. I conclude that a crisis – like the recent one in the United States – is best predicted by looking at both similarities and dissimilarities across historical cases.

JEL: C2, E6, and F3

Keywords: subprime, crisis, prediction, indicator, logit, probit, panel
I. Introduction

The US subprime crisis, by now a global crisis, is best described as “unprecedented” since the chronic days of the Great Depressions. However, some authors suggest that despite its scale and far reaching impacts, the lessons from financial crises that ever since roiled several other parts of the world could foretell the story. This paper develops an early warning system for financial crisis and confirms this proposition.

Of course, the crisis was predicted (most notably, Shiller, 2005). It was shown that long term domestic economic factors, such as growth of income and population, were unable to explain the quick surges in the US housing market. And, similar housing booms in history that were associated with such public exuberance also imploded very similarly.

Along the same line of research, in recent times some other authors have noted marked similarities between the US subprime crisis and other historical banking crises. Most notable are Reinhart and Rogoff (2008a, 2008b). They (2008a)\(^1\) discuss a few crisis indicators and demonstrate patterns that are strikingly common to the recent US case and the systemic bank-centered crises of the advanced economies. The indicators are real estate real price, real equity price, ratio of current account balance to GDP, real GDP per capita, and public debt. They interpret some specific changes in the behavior of these indicators as possible precursors of crisis.

---

\(^1\) Reinhart and Rogoff (2008b) arrive at the same conclusion with an extended database that goes back to 1800 and includes a core sample of sixty-six countries. In addition, they find that banking crises have been incidental equally in high income countries as in middle-to-low income countries.
This paper journeys from the pillar founded by these authors. It maintains that apparent similarities across crises are best interpreted with robust statistical tests. In the same vein, only an out-of-sample forecasting exercise must determine if a certain change in the behavior of an economic indicator signals a crisis. Accordingly, three related questions are asked: i) Could we predict the recent US crisis based on data on other historical crises? ii) How well in advance? iii) What indicators and which historical country experiences could give the best predictions? In view of the protracted increase in income inequality in the United States over the last three decades, I also ask if the historical crisis episodes are better understood with similar phenomena and hence, if the recent US experience could be predicted with any greater confidence.

To address the issues, in the first round of exercise a panel probit regression is performed with yearly historical data on the same set of indicators and countries as in Reinhart and Rogoff (2008a). Then, the specification is used to test if the US subprime crisis can be predicted out of sample starting in 2003. It is found that if an alarm for a crisis is issued when the probability of a crisis exceeds 50% then the past financial crises of advanced countries could predict the 2007-2008 US case starting in 2003, but also would have generated a false alarm thereby suggesting that there would be a financial crisis sometime between 2001 and 2004. On the other hand, if rather an alarm must be issued when the probability of a crisis exceeds 75% then the crisis could be predicted starting only in 2005. Further, not all indicators considered by Reinhart and Rogoff (2008a) are found to be significant. In particular, public debt doesn’t help explain the financial crises in the past. Limiting the within-sample group of countries only to the five (Spain, Norway,
Finland, Sweden, and Japan) that had suffered the biggest crises improves prediction accuracy. But again, neither all indicators are significant, nor is the subprime crisis fully predicted.

In the final exercise, a panel probit model is specified with data on a different sub-group of the same countries considered by Reinhart and Rogoff (2008a). This group consists of Finland (1991), Italy (1990), Japan (1992), New Zealand (1987), the United Kingdom (1995), and the United States (1984), where the crisis periods are in parentheses. This final specification has robust goodness of fit measures and high within sample prediction accuracy. But, most importantly, it predicts the US sub-prime failure starting in 2003 and does not produce any false alarm. All projected probabilities of crisis are 75% or higher. But, much like what we saw in the previous two exercises, only a subset of the indicators considered by these authors are significant. These are percent of current account balance in GDP, public debt and real estate real price. Notably, however, an additional indicator, growth difference between average productivity and real wage rate that I introduce as a measure of “growth of income inequality”, is also significant.

I conclude that the US subprime crisis could be predicted consistently starting in as early as 2003. This was possible based on a common set of indicators for some advanced countries that suffered bank-centered financial crisis in the postwar period. The accomplishment, however, is by accounting also for the growing income inequality in the United States and other advanced nations over several years before the respective crisis episodes.

The finding that income inequality explains and predicts the crisis is consonant with Krugman (2007) who suggests that growing income inequality under the current tax structure in
the United States contributed to the desperation of middle-class subprime borrowers, and then, after the housing bubble burst, to the increase in defaults. It is also consistent with the observation by Shiller (2007) that the housing boom and the subsequent fall in prices were most pronounced in the low and middle price tiers. Significance of income inequality and public debt conform to the historical data on countries with which the US 2007-2008 crisis is predicted. Britain, Finland, Japan, Italy, and the United States (savings and loan crisis in the mid 80’s) had experienced rise in income inequality and expansion of public debt over a period of four to five years preceding the respective crisis episodes.

The significance of the other indicators, real estate real price and ratio of current account balance to GDP, confirm the conclusions by Reinhart and Rogoff (2008a, 2008b). The authors emphasize persistent current account deficit and “bonanza” in capital inflow as crisis precursors. Further, these authors and Shiller (2008) suggest that a prior housing bubble has been too common to systemic banking crises across the world. At the same time, the fact that house prices help predict the recent US crisis lend credence to Shiller (2008) and Krugman (2008): like other systemic financial crises in history, the subprime fiasco was caused by irrational exuberance. Before the housing market started going down around 2006, the market participants as well as the policy makers in the United States and other countries persistently believed in ever-rising house prices, which however was never true in history.

In the rest of the paper first the nature and causes of the US subprime crisis is discussed with reference to the historical financial crises. Then, after briefly discussing the literature on
prediction of financial crisis and reviewing the method adopted by Reinhart and Rogoff (2008a), the empirical method central to this paper is developed and the results are reported.

II. Probable Causes of the US Subprime Crisis

The subprime crisis and virtual collapse of the US financial sector have naturally stimulated a flurry of investigations (see Gallegati et al., 2008; Krugman, 2008; Laeven et al, 2008; Lim, 2008; Mizen, 2008; Reinhart and Rogoff, 2008a, 2008b; Schnabl and Hoffman, 2008; Mills and Kiff, 2007; Shiller, 2007; Watson, 2008; Wray, 2007). While the crisis incidence has been viewed and explained in varied ways, Reinhart and Rogoff (2008a, 2008b) succinctly capture some of the most plausible explanations. They note that unfettered financial liberalization, together with a vast inflow of foreign capital and a rat-race to invest in a new, apparently credible technology -- much like what happened as a prelude to a plethora of other crises in history -- sparked the first global crisis of the twenty first century. The “new technology” in the recent U.S. case showed up in the form of complex loan packaging that was thought to be hedging risks across financial markets. However, these non-transparent mixes deceived the investors and market experts. Moreover, when house prices dropped and defaults on sub-prime loans increased these bundles permeated the effects leading to a nation-wide credit crunch. This has an emphasis in Gallegati et al. (2008) where under deteriorating economic conditions the interdependencies in real and financial assets increase the risk of contagion and reduce the ability of risk sharing.
Nevertheless, very little of the literature presents a rigorous analysis of the other end of the subprime market, *viz.* the subprime borrowers. What made such a borrower and then made her default? Most studies depict a sub-prime borrower, if any, as going above her expenditure and repayment capabilities to enjoy costly home-ownership and other luxuries. The motivation, in other words, was to catch up with the Joneses even though it meant crossing the limits of income.

But the comprehensive study by Warren and Tyagi (2005) on the surge of U.S. bankruptcies paints a different picture. The study is further confirmed and cited by Krugman (2007) who discusses the adverse effects of rising income and social inequalities on the US housing markets. By 2005, they find, the number of families filing bankruptcy was five times the number in the early 80’s. This has been primarily due to the fact that today’s average middle-class family, even after two members earning, spend about 75% of the family income on necessities, whereas a generation ago this spending was only 54% of the income of an average family which typically had only one earner. A large part of the increase in the share of basic expenditure is due to rise in the prices of home-mortgages. But families bought such costly mortgages in a desperate drive to find good school-districts, so that they could provide their children meaningful education, a good social life, and a secured future. The findings of Iacoviello (2008) are indicative of a similar fact. He develops a model with heterogeneous agents that mimic the US time series of income distribution from 1963 through 2003, and finds that long run increase in household debt can be explained only by increasing income gaps. In light of the analysis by Warren and Tyagi (2005), Krugman (2007) concludes that increase in the number
of mortgage debts, which the borrowers were not sure of if they could manage safely, has been partly due to the current tax structure in the United States that raised income and social inequalities over the last three decades.

Protracted increases of income inequality and costs of living for the low and middle income earners in the United States since the mid 70’s have been examined thoroughly (D’iaz-Gim´enez et al., 2002; Eckstein and Nagypal, 2004; Weinberg and Steelman, 2005; Bryan and Martinez, 2008). Bryan and Martinez (2008) analyze the US earnings distribution data since the 1960’s and find that during the period from 1975 to 2002 only the top 10 percent of the labor income distribution increased by more than the overall earning per worker. And, in fact, between 1975 and 1997 labor income in the 10th percentile shrank by about 7 percent.

*Figure 1*, which plots the US average labor productivity and real wage rate of the manufacturing sector on the same scale, presents a similar account. It shows that from 1950 through 1978 both grew almost at the same rate -- but starting in 1978 real wage rate followed a downward trend, while average productivity grew at a slightly higher rate from the mid 80’s. This means that since the late 70’s not only the relative share of the wage earners in GDP declined, but also that they actually experienced shrinkage of real earnings for every hour worked. Further, and may be most relevant in the current context, starting from 2002-2003, the gap between average productivity and real wage rate has been widening at a faster rate, while the real wage rate itself has been falling. Data indicate that it is also around the same time that the GDP share of US government debt started increasing drastically.
Data on other advanced countries, especially Britain, Finland, Japan and Italy, suggest that such dual increases in income inequality and public debt over a certain period preceding a crisis incidence are not special to the US experience. Japan passed through the trauma of a “lost decade” following the housing and stock bubbles in the early 90’s. But it also experienced a sharp 33% increase of its GDP share of government debt from 1987 to 1991. Over the same

---

2 Sources:

i) Money wage rate (manufacturing sector) and CPI, *International Financial Statistics*.

period the country’s average labor productivity increased by 7%, whereas its real wage rate increased by only 1.7%.

While income inequality might have contributed to increase in subprime borrowings in the United States, at the same time the borrowers as well as those who invested in subprime assets were spurred by the belief that house prices were to rise indefinitely. As Shiller (2008) delineates, much like what happened during the technology bubble of the 1990’s, it was precisely “irrational exuberance” that germinated and spread as a social contagion process. Such exuberance and the related speculative bubble grew via feedback loops: house prices rose, which made people believe that it was going to rise further and also increased the speed of contagion of the same belief, which in turn actually caused further rise in prices. Similar over-optimisms were observed in Japan before the stock and real estate bubbles burst in 1992 and also in Sweden and Finland before the financial crises in 1991.

But Shiller (2008) also notes that the housing boom in the United States was lot more pronounced in the low-price tier than in the high-price tier.3 Furthermore, after the bubble burst the sharpest fall in house prices was observed in the low-price market. He takes “subprime phenomenon” as the most plausible explanation. The subprime loans expanded at an accelerated pace since 2001, but particularly among the low income buyers for purchasing low-priced houses. Similarly, the biggest drop in prices in the low-price tier since 2006 is explained by the greatest number of defaults and foreclosures among the low income borrowers.

3 He examines the San Francisco housing market, but asserts that the same difference may be observed in many other cities.
The essential point to be noted here is that this observation provides additional support to the conclusions by Warren and Tyagi (2005) and Krugman (2007). While the entire US economy had nourished a wishful thinking of ever-rising home prices, the run-up was most noticeable among the low income groups. If the prices were going to rise, for a low income family buying a house had a two-fold justification: “buy it now or never” and “secure the future with an asset that will increase in value”. In other words, their current financial risks in the face of growing inequalities prompted trading even more risks for a large future expected gains. Of course, the work by Friedman and Savage (1948) in the tradition of expected utility theory offers an explanation for such a behavior. The theory suggests that people might become risk lovers when they perceive a probable lift from a low-income group to a high-income group. It however presumes that agents can figure out and hence know the exact probabilities of different outcomes. The finding in psychological literature that human beings fail to understand the complexity of risks and tend to be overconfident (see Ross, 1987 and Shiller, 2003) not only challenges this underlying assumption but also provides an alternative explanation for such observed behavior. Finally, the opaque nature of the subprime loans discussed earlier in this section further weakened their ability to understand the extent to which risks would increase.

The discussion so far suggests that while crisis indicators adopted by Reinhart and Rogoff (2008a) might help explain and predict the US subprime crisis, we need to examine whether financial crises in the past had any relation with “income inequality” and if this can be taken as one other crisis indicator.
III. The Literature on Financial Crisis Prediction

Once the probable causes of the US financial crisis have been identified one has to run statistically meaningful tests on whether all or most of these also contributed to the historical crises. This will then allow us to examine if the recent US crisis could be predicted with the historical experiences. For this purpose the choice of an appropriate model is made by reviewing the related literature. It is seen that in parallel to the theoretical constructs on financial crisis, much of recent literature has focused on developing early warning systems. This area has received renewed interests due to pervasive currency and banking crashes across the world\(^4\). It presents two important strands of research on prediction of financial crisis: the indicators approach and the approach based on probit and logit models.

Kaminsky, Lizondo, and Reinhart (1998) and Kaminsky and Reinhart (1998, 1999) develop the indicators approach, in which the evolution of a set of monthly indicators are monitored individually and a signal to a crisis is issued whenever an indicator exceeds an optimum threshold value. The optimal thresholds, calculated as percentiles with respect to each indicator and country, are determined by minimizing the ‘noise-to-signal’ ratio (ratio of ‘false’ alarms to ‘true’ alarms) across countries. Edison (2003) analyzes and extends this approach to apply to an individual country case.

Outside the indicators approach, several other authors (see Berg, Candelon, and Urbain, 2008; Berg and Pattillo, 1998, 1999a, 1999b; Eichengreen, Rose and Wyplosz, 2003a, 2003b; Frankel and Rose, 1996; Kamin and Babson, 1999; Reagle and Salvatore, 2000; Chionis and Liargovas, 2003) adopt logit or probit models to analyze and forecast financial crisis (specifically, currency crisis). In this approach, given panel data on a set of variables for several countries, probability of a crisis occurring within a certain period in a country is estimated – and, accordingly, an alarm is generated when the estimated or predicted probability crosses a threshold.

One obvious advantage of the probit/logit based method is that with the assumption of a standard probability distribution of the parameters, suitable statistical tests can be performed to examine if an indicator significantly contributes to the probability of a crisis. Further, unlike in the indicators approach, in logit and probit models the relevant information contained in the indicators are combined to better understand the incidence of a crisis. Some studies suggest that a warning system based on the indicators approach may produce less satisfactory results out of sample than the probit/logit based approach (see, for example, Alvarez-Plata and Schrooten, 2004; Berg and Pattillo, 1999a). This paper therefore adopts the probit/logit based approach for estimating the probability of a crisis.
IV. The Analysis by Reinhart & Rogoff (2008a)

A. The Analysis

As outlined in the introductory section, apart from examining if the US subprime crisis could be predicted, one other, but very related, motivation for the paper is to run statistical tests on the conclusions drawn by Reinhart and Rogoff (2008a). Hence, before discussing the estimation and forecasting exercises and the related results, here I present a brief review of the data and methodology adopted by these authors. Henceforth, the authors are dubbed as RR.

The group of countries, the crisis episodes, and the leading indicators considered by the authors are given below.

*The “Big Five” Crises:


*Other Postwar Bank-centered Financial Crises:


*Indicators

Real Housing Prices, Real Equity Prices, Ratio of Current Account Balance to GDP, Growth of Real GDP per Capita, Public Debt

In a comparison between the US subprime crisis and the average across all other advanced economy crises, RR demonstrates similar trends for each indicator over four years prior to the crisis episodes. Furthermore, they show that the similarities, in general, are more prominent when the comparison is done with only the big five crises. With respect to the indicators, they
find that persistent increase in current account deficit, rise in real estate prices, and increase in public debt in particular are common prior to the crisis episodes. Therefore, they conclude, these changes are unmistakably crisis precursors.

B. Limitations of the Analysis

A comparison of the US crisis with the average of historical crises in terms of leading indicators provides useful insights into the underlying similarities. Nevertheless, despite its appeal the method has its own limitations for any actual prediction of a crisis incidence. The relevant questions to ask: Could these similarities foretell the recent US experience? How many years in advance?

In other words, whether or not a crisis in the future can be foreseen with the historical data and if a certain change in the behavior of an indicator should be accepted as a “precursor of crisis” are issues that must be settled only in practice – specifically with an actual out-of-sample forecasting exercise. The RR analysis may not be adequate for accomplishing this task (admittedly, the authors don’t claim as such). First, as the literature on crisis prediction suggests, a comparison of trends only over a period of four years is hardly enough to understand the similarities (or dissimilarities) in the underlying dynamics that triggered the crises. If we assume that these four years actually foreshadow a crisis episode⁵, then it is also important to know if there were similarities in trends over a period that go farther back in history. This enables a distinction between the periods that are tranquil and the periods that are expected to signal a

---

⁵ In the literature, generally two years are taken as signalling a crisis.
crisis, which by far is an important criterion for “predicting” a crisis with any success. For example, going back to the year of 2003 if we had asked whether the 2007-2008 US crisis could be predicted then the RR analysis would have been silent -- of course, in 2003 any such comparable four-year trend for the United States was completely missing.

*In this paper I show, in contrast, that an early warning system that classifies historical observations into “tranquil” and “pre-crisis” categories could actually predict the recent US case consistently starting in the year of 2003.*

Further, a comparison of any crisis with the “average” of historical crises could be elusive, especially, and again, for the practical purposes of prediction. As indicated by the literature, an actual forecasting exercise will warrant careful examination and utilization of the data in the panel format, which might reveal significant differences across countries. Such differences, however, might be cloaked under the “average” construct, particularly when pronouncedly similar trends of an indicator across a few crisis experiences override the dissimilarities across other crisis experiences. But, obviously these dissimilarities remain in the data in panel format. Hence, any model that must utilize the panel format for an actual forecasting exercise is likely to perform poorly unless such dissimilarities are removed.

The empirical model developed in the next section will show that these are reasons why not all indicators discussed by these authors are “statistically” significant for explaining historical crises and predicting the recent US case.
V. Predicting the US Subprime Crisis

A. The Problem

While the central issue addressed in this paper is of prediction of the US subprime crisis, the problem itself needs to be reset in more specific terms. To reiterate what was asserted in the preceding section, the question I ask is that how many years in advance could the crisis be predicted based on data on a common set of indicators for advanced economies that experienced systemic financial turmoil in the post-war period. In other words, could we foretell the crisis incidence if we were to be on a particular date prior to the year, 2007? The answer, naturally, will depend on how well in advance the post-war historical crises in the advanced economies could be predicted in a within-sample exercise. Accordingly, with a careful examination of the data and some initial statistical estimation it is concluded that for most advanced countries the crises are predictable at best four years in advance.

Therefore, the rest of the paper is devoted to the question as to whether the US subprime crisis could be predicted consistently starting from 2003. Notably, this is same as asking if going back to the year of 2002 or earlier we could predict that the United States was not going to experience any crisis in 2006 or earlier.

B. Data

For estimation and forecasting purposes, all indicators as in RR (see section IV) are included. However, public debt is replaced by its share in GDP. Intuitively, it is really this share that
should provide any relevant information about a future crisis incidence. So, it is expected to bear a more significant relationship with crisis probability. In light of the discussion in Section II, I consider two additional indicators as given below.

   i) \textit{Growth Difference between Average Productivity and Real Wage Rate}

   ii) \textit{Inflation}

The first additional indicator is constructed for finding a relationship between growth of earnings inequality and probability of crisis incidence. The second additional indicator is for accounting for the effect of costs of living.

As for the inclusion of countries in the sample, I consider, as in RR, all advanced economies that experienced systemic crisis in the postwar period, except Greece and Iceland which have a large number of missing observations.

Thus, the panel database is formed with yearly observations on all seven common indicators (five as in RR and two additional ones mentioned in this section) for all advanced economies (as in RR, except Greece and Iceland) that experienced systemic bank-centered crisis in the post-war period. As for the number of observations, only eight prior to the actual occurrence of any crisis are included, where, according to the problem defined in the preceding subsection, the first four observations are labeled as “tranquil” and the last four observations are labeled as “pre-crisis”.

\textsuperscript{6} The indicator is constructed as such since consistent data on Gini ratio or other measures of earnings inequality for all countries under consideration are not available.

\textsuperscript{7} Unavailability of consistent data restrained consideration of longer time series for all indicators and countries. In the literature on crisis prediction generally data on each indicator and each country are intended to start from a common month/year and end in a common month/year so that the panel database will contain the same number of observations on each country. But all studies discussed in Section III had to deal with the problem of a very large
Further, each of the seven indicators is transformed into a normalized variable so that it assumes a maximum value of *one* and a minimum value of *zero*. Specifically, let $X$ denote an indicator and $T$ denote the transformed variable. Further, let $c$ denote a crisis episode and $i$ denote an observation before $c$ occurred. Then, we transform an indicator value, $X_{ci}$, as follows

$$T_{ci} = (X_{ci} - MIN_c)/(MAX_c - MIN_c), \quad i = 1 \text{ to } 8; \quad c = 1 \text{ to } 18^8,$$

where $MIN_c$ and $MAX_c$ are the minimum and maximum of the indicator values related to the crisis, $c$.

C. *Methodology and Criteria*

I adopt the logit/probit based approach due to its advantages (as discussed in Section III) over the indicators approach. But, specifically, based on some initial experiments, a multivariate probit model is employed for all estimation and forecasting purposes.

Accordingly, in the panel data the binary dependent variable is assigned the value *zero* when the observation is “tranquil” and *one* when the observation is “pre-crisis”. After the parameters of the model are estimated, the predicted value of the dependent variable gives the number of missing observations. Hence, the number of observations for each country in the panel database has never been the same across countries. In the current study I address this problem by allowing only a limited but equal number of observations for all countries under consideration. Moreover, the distribution of observations are made symmetric (w.r.t “tranquil” and “pre-crisis” categories), which then sets the prior probability for each category to 50% and provides intuitive justification for a threshold probability of 50%, crossing which an alarm for a crisis will be issued. Understandably, this does not impose any limit to addressing the fundamental question of the paper: Could we predict the US subprime crisis four years in advance? We can even ask a similar question: if we were to be, say, in 1996, could we predict that there was not going to be any financial crisis in the United States within a four year period?

---

8 This is because there are eighteen crisis incidences (including the recent US case) under consideration and for each crisis incidence eight observations are included in the panel database.
probability of a crisis episode within a period of four years. For example, in a within sample exercise, if the predicted value of the dependent variable is 0.2 for Japan in 1986, then the predicted probability of a financial crisis in Japan any time between 1987 to 1989 is also 0.2.

In order to conclude if the predicted probability must imply a crisis incidence within a period of four years, it is checked if the probability crosses a pre-specified threshold value. If the predicted probability crosses the threshold value then we must issue an “alarm” for a crisis within a period of four years. In this paper the results are presented with both 50% and 75% as the threshold value.

It follows that an alarm can be true or false. If an alarm was issued and there is actually a crisis incidence within a period of four years then it is true; the alarm is false if a crisis does not ensue within a period of four years. It is seen readily that the predictive power of a model should be judged not only with the percentage of “pre-crisis” observations correctly called, but also with the percentage of true alarms generated. Indeed, a model could correctly call a high percentage of the crisis episodes, but if it also issues a high percentage of wrong signals then the warning system is elusive – it fails to distinguish between a “tranquil” period and a period that must signal a crisis. This, moreover, will mislead policy makers who must take measures if a crisis is known to be imminent. Hence, following are the two most important criteria for goodness of fit with which one must judge the success of the model as an early warning system.

i) Probability of an alarm conditional on a crisis within four years (percentage of pre-crisis years correctly called)

ii) Probability of a crisis within four years following an alarm (true alarms as percentage of total alarms)
In addition, following the literature, I also consider three calibration scores, as given below.

i) Quadratic probability score
ii) Log probability score
iii) Global squared bias

In specifying the model, in addition to the within-sample performance in terms of the goodness of fit measures and calibration scores listed above, I consider a few additional conventional criteria for goodness of fit: significance of the coefficients, Akaike information criterion, and Schwartz information criterion.

D. Model Specifications and Forecasting

One objective of the paper was outlined as performing a test of whether the indicators considered by RR are significant for explaining the historical crises and predicting the 2007-2008 US subprime crisis. Another objective is to test if income disparity and cost of living could add explanatory and predictive power. Accordingly, the model specifications and forecasts are performed with the following steps.

Step 1:

i) Specify a multivariate probit model with panel data that includes only the five variables considered by RR.

ii) In order to increase the predictive power of the model, choose the specification which has all significant coefficients and the best within sample performance in terms of goodness of fit measures and calibration scores.

iii) Predict the US subprime crisis out of sample with the final specification.
Step 2:

Specify a multivariate probit model with panel data that includes also growth difference between average productivity and real wage rate and inflation, in addition to the indicators considered by RR. Then follow the same procedure as in Step 1.

RR indicates that the US subprime crisis resembles more to the big five crises than all other crises. A recent study (Berg, Candelon, and Urbain, 2008) questions the assumption that crises are homogeneously caused by identical factors and suggests a preliminary step toward finding an optimum country cluster. Hence, a careful scrutiny of the data is performed and it is found that when all seven indicators are included, the similarities are more pronounced across a different subgroup of crises: Finland (1991), Italy (1990), Japan (1992), New Zealand (1987), the United Kingdom (1995), and the United States (1984, 2007-2008). It is therefore thought to be likely that data on this subgroup of countries will predict the recent US experience with a greater accuracy. Hence, the above steps are repeated with panel data on the following groups of countries in turn.

i) All countries as in RR (except Greece and Ireland)


E. Results

The results are reported in Table 1 through Table 3. Table 1 presents the coefficient estimates from four probit regressions. Table 2 presents the within sample calibration scores and goodness of fit measures, and Table 3 presents the same measures for the out-of-sample prediction of the US subprime crisis.

It is seen that when all crisis experiences are included (Specification 1A) then all RR indicators are significant, except share of public debt in GDP. This seems to be on account of the fact that for some individual countries (e.g. Australia, Britain, Spain, and Sweden) this indicator actually had a declining trend – not only over the four year period of reference, but also over an extended period. Thus, differences in the behavior of this indicator across crisis incidences -- which otherwise are cloaked under the “average” construct -- make it insignificant in the panel probit estimation. This problem with the RR analysis was discussed in Section IV, and is again noted here in the specific context.

Nevertheless, significance of all other indicators in Specification 1A corroborates the conclusions by RR. In particular, current account having a negative sign supports the idea that a persistent deficit with a drastic deterioration over a recent period might trigger a crisis. Similarly, significance and the negative sign of real estate price confirm that a crisis episode generally ensues after real estate prices soar significantly (RR; Reinhart & Rogoff, 2008b; Shiller, 2008). It is also consistent with the idea that such bubbles are caused by irrational exuberance but eventually run out of suckers leading to a crisis (Krugman, 2008; Shiller, 2008).
Table 1
Multivariate probit model\(^1\): coefficient estimates with three groups of countries that experienced financial crisis after WWII

<table>
<thead>
<tr>
<th>Variable</th>
<th>All postwar systemic crises</th>
<th>Biggest five crises</th>
<th>A selected</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Specification 1A (P-values)</td>
<td>Specification 1B (P-values)</td>
<td>Specification 2 (P-values)</td>
</tr>
<tr>
<td>REAL GDP PER CAPITA</td>
<td>2.54 (0.00)</td>
<td>2.86 (0.00)</td>
<td></td>
</tr>
<tr>
<td>PERCENT OF CURRENT ACCOUNT IN GDP</td>
<td>-1.27 (0.00)</td>
<td>-1.02 (0.00)</td>
<td></td>
</tr>
<tr>
<td>REAL STOCK PRICE</td>
<td>0.78 (0.08)</td>
<td>0.93 (0.04)</td>
<td></td>
</tr>
<tr>
<td>REAL ESTATE REAL PRICE</td>
<td>-1.51 (0.00)</td>
<td>-1.35 (0.00)</td>
<td>8.17 (0.00)</td>
</tr>
<tr>
<td>SHARE OF PUBLIC DEBT IN GDP</td>
<td></td>
<td></td>
<td>-2.11 (0.00)</td>
</tr>
<tr>
<td>GROWTH DIFFERENCE:</td>
<td></td>
<td>-0.97 (0.01)</td>
<td></td>
</tr>
<tr>
<td>REAL GDP PER EMPLOYED &amp; REAL WAGE RATE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>INFLATION RATE</td>
<td></td>
<td></td>
<td>-1.84 (0.00)</td>
</tr>
<tr>
<td>(AIC)</td>
<td>0.98</td>
<td>0.95</td>
<td>0.56</td>
</tr>
<tr>
<td>(SIC)</td>
<td>1.07</td>
<td>1.05</td>
<td>0.68</td>
</tr>
</tbody>
</table>

\(^1\) The model specification for each group of countries is selected with AIC, SIC, significance of coefficients, and most importantly with the performance results from within sample predictions reported on Table 2

Table 2: Within sample performance results with a multivariate probit model: three groups of countries that experienced financial crisis after WWII (US 2007-2008 excluded)

<table>
<thead>
<tr>
<th>Variable</th>
<th>All postwar systemic crises</th>
<th>Biggest five crises</th>
<th>A selected group(^1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Specification 1A</td>
<td>Specification 1B</td>
<td>Specification 2</td>
</tr>
<tr>
<td>Calibration scores</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quadratic probability score(^2)</td>
<td>0.30</td>
<td>0.28</td>
<td>0.14</td>
</tr>
<tr>
<td>Log probability score(^3)</td>
<td>0.15</td>
<td>-0.12</td>
<td>0.02</td>
</tr>
<tr>
<td>Global squared bias(^4)</td>
<td>0.01</td>
<td>0.006</td>
<td>0.00007</td>
</tr>
<tr>
<td>Goodness-of-fit (threshold probability = 50%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Probability of an alarm conditional on a crisis within 4 years (percent of pre-crisis years correctly called)</td>
<td>86.8</td>
<td>85.3</td>
<td>80.0</td>
</tr>
<tr>
<td>Probability of a crisis within 4 years following an alarm (true alarms as percent of total alarms)</td>
<td>79.7</td>
<td>79.5</td>
<td>94.1</td>
</tr>
<tr>
<td>Goodness-of-fit (threshold probability = 75%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Probability of an alarm conditional on a crisis within 4 years (percent of pre-crisis years correctly called)</td>
<td>69.1</td>
<td>64.7</td>
<td>80.0</td>
</tr>
<tr>
<td>Probability of a crisis within 4 years following an alarm (true alarms as percent of total alarms)</td>
<td>73.9</td>
<td>74.6</td>
<td>94.1</td>
</tr>
</tbody>
</table>


\(^2\) \(QPS = \frac{1}{T} \sum_{t=1}^{T} 2(P_t - R_t)^2\); where \(P_t\) & \(R_t\) are predicted and actual probabilities respectively, and \(T\) is the number of observations predicted

\(^3\) \(LPS = \frac{1}{T} \sum_{t=1}^{T} [(1 - R_t) \ln(1 - P_t) + R_t \ln(P_t)]\); where \(T\), \(P_t\) & \(R_t\) are as in \(QPS\)

\(^4\) \(GSB = 2(P_m - R_m)^2\), where \(P_m = \frac{1}{T} \sum_{t=1}^{T} P_t\), \(R_m = \frac{1}{T} \sum_{t=1}^{T} R_t\), and \(T\), \(P_t\) & \(R_t\) are as in \(QPS\) & \(LPS\)
Table 3
Predicting the U.S., 2007-2008 financial crisis out of sample with a multivariate probit model: performance results with data on three groups of countries that experienced crisis after WWII

<table>
<thead>
<tr>
<th>Variable</th>
<th>All postwar systemic crises</th>
<th>Biggest five crises</th>
<th>A selected group1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Specification 1A</td>
<td>Specification 1B</td>
<td>Specification 2</td>
</tr>
<tr>
<td>Calibration scores</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quadratic probability score²</td>
<td>0.25</td>
<td>0.28</td>
<td>0.13</td>
</tr>
<tr>
<td>Log probability score³</td>
<td>-0.11</td>
<td>-0.02</td>
<td>-0.14</td>
</tr>
<tr>
<td>Global squared bias⁴</td>
<td>0.01</td>
<td>0.0008</td>
<td>0.02</td>
</tr>
<tr>
<td>Goodness-of-fit (threshold probability = 50%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Probability of an alarm conditional on a crisis within 4 years (percent of pre-crisis years correctly called)</td>
<td>100.0</td>
<td>75.0</td>
<td>100.0</td>
</tr>
<tr>
<td>Probability of a crisis within 4 years following an alarm (true alarms as percent of total alarms)</td>
<td>80.0</td>
<td>100.0</td>
<td>80.0</td>
</tr>
<tr>
<td>Goodness-of-fit (threshold probability = 75%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Probability of an alarm conditional on a crisis within 4 years (percent of pre-crisis years correctly called)</td>
<td>50.0</td>
<td>50.0</td>
<td>75.0</td>
</tr>
<tr>
<td>Probability of a crisis within 4 years following an alarm (true alarms as percent of total alarms)</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>


² QPS = 1/T \( \sum_{t=1}^{T} 2(P_t - R_t)^2 \); where \( P_t \) & \( R_t \) are predicted and actual probabilities respectively, and \( T \) is the number of observations predicted

³ LPS = 1/T \( \sum_{t=1}^{T} [(1 - R_t) \ln(1 - P_t) + R_t \ln(P_t)] \); where \( T, P_t \) & \( R_t \) are as in QPS

⁴ GSB = 2(\( P_m - R_m \))^2, where \( P_m = 1/T \sum_{t=1}^{T} P_t \), \( R_m = 1/T \sum_{t=1}^{T} R_t \), and \( T, P_t \) & \( R_t \) are as in QPS & LPS
Specification 1A has promising goodness of fit measures. With a 50% threshold probability, it predicts 87% of the crises and generates only 20% false alarms within sample. Out of sample, it correctly calls all pre-crisis observations (2003 – 2006), but at the same time produces one false alarm in 2000. This means that starting in 2003 the specification could predict the US subprime crisis in 2007-2008, although it would have misled policy makers thereby suggesting that there was also going to be a crisis sometime between 2001 and 2004. When a threshold probability of 75% is applied, not only the within sample performance falls considerably, but also it predicts the subprime crisis starting only from the year of 2005. These, however, are very high rates of success which of course confirm the general findings by RR.

Specification 1B retains all variables of Specification 1A, but includes also growth of earnings inequality, which appears to be significant. But the negative sign of this indicator is counter intuitive. The implication is that the onset of a crisis is explained even as inequality rose in a relatively tranquil period and declined over a few years on the eve of the crises. Inflation as another possible predictive variable was tested but was found to be insignificant. It is noted that signs and significance of all original variables that appear in Specification 1A remain unaltered in Specification 1B.

However, Specification 1B does not represent an improvement over Specification 1A in terms of goodness of fit measures. Within sample, it predicts almost the same number of crises and produces almost equal number of false alarms as does Specification 1A. In the out-of-sample exercise, it produces no false alarms for the US crisis, and its calibration scores are somewhat better than in Specification 1A. But even with a threshold probability of 50% it would have failed
to foretell the US subprime crisis in the year of 2003, although would have predicted it consistently starting in the year of 2004. With a threshold probability of 75% the specification could predict the subprime crisis starting in 2005. These results imply that when all postwar advanced economy crises are considered, inclusion of growth of income inequality does not add any predictive power to the model.

*Specification 2* is based on panel data that include only the big 5 postwar crises. Here, only real estate price, public debt, and inflation rate are significant, but all have apparently counter intuitive signs. The positive sign of real estate price is explained by the fact that all these countries had experienced a real estate boom before the onset of crisis, but the housing prices started declining only around the time of crisis. It is noted that contrary to what RR concludes, current account does not appear to contribute to making a distinction between a tranquil period and a period that could signal the big five crises.

In terms of goodness of fit measures and calibration scores, *Specification 2* represents some improvement over Specifications 1A & 1B, both within sample and out-of-sample. And, like *Specification 1A*, with a threshold probability of 50% it predicts the 2007-2008 US crisis starting in the year of 2003, but also produces a false alarm.

*Specification 3*, which is based on panel data on a different subgroup of countries (see the last subsection), predicts the U.S. subprime crisis consistently starting in the year of 2003 and, at the same time, does not generate a single false alarm. This essential result does not change if the threshold probability is raised from 50% to 75%. Noticeably, this is quite unlike with the last three specifications. Further, it has the best out of sample calibration scores from all four
specifications. The three critical RR indicators, current account, real estate price, and public debt in this specification are significant with the expected signs. In most of these countries the real estate bubble had burst few years before the onset of crisis, and this justifies the negative sign of the related indicator. The negative sign of current account is explained by the fact that a few years before the crisis occurred most countries had started experiencing sharp deteriorations in current account status. Public debt in most of these countries increased either over the entire eight year period under consideration or only over three to four years before the crisis. This explains the positive sign of the indicator.

It is noted with interest that growth of inequality is also significant in Specification 3 with a positive sign. This is accounted for by the increases in inequality growth in all these countries over four to five years before the crisis episodes. The fact that Specification 3 also predicts the US subprime crisis with 100% accuracy with no false alarms is consonant with the conclusion by Warren and Tyagi (2005) and Krugman (2007) that large increases in mortgage debts (and hence the recent crisis) in the United States cannot be explained without accounting for the protracted increases in income disparities and costs of living for the low and middle income earners. It is also consistent with the observation by Shiller (2008) that the subprime boom was more pronounced in the low-price tier than in the high-price tier and subsequently when the housing bubble burst the low price tier suffered the biggest drop.

The results from all four probit regressions presented in this section are consonant with the probable causes of the US subprime crisis discussed in Section II. The crisis can indeed be explained with similarities across systemic crises witnessed in history. But the crisis incidences
are not as strikingly similar as might appear, and unless we also account for the dissimilarities across such incidences we are unable to achieve full accuracy in predicting a future unhappy episode.

VI. Conclusions

Referring to the opening line in Tolstoy’s classic, Anna Karenina, Reinhart and Rogoff (2008a) make the concluding remark that “there can be similarities across unhappy families, too”. Thus, they imply, crisis incidences across the world could share certain similarities. It seems that common patterns across incidences of crisis are an undeniable fact. But, nevertheless, many “unhappy families” across the world under the dire current conditions will unmistakably assume similarities in terms of loss of employment, income, wealth, health insurance, and life-time savings.

The idea of this paper stemmed from these rupturing effects the US subprime crisis currently has in the form of a global meltdown. The motivation for the entire analysis was to enquire if these unhappy events could be averted by predicting the crisis earlier than what some experts did (most notably, Shiller, 2005). The findings of the paper suggest that the economics profession could accomplish again by producing systematic warnings for the policy makers starting in as early as 2003 – a year before the President of the United States was re-elected. It is noted that this conclusion withstands the general pessimism in the literature about systematic prediction of financial crisis. Such systematic alarms, however, could be generated by looking at both similarities and dissimilarities across crisis experiences. A sustained growth of income
inequalities, in particular, is a characteristic that was not seen as a precursor of many crises – which, however, could well contribute -- and as strongly as in the recent case in the United States.

Or perhaps, as Shiller (2008) maintains, the crisis could not be averted. Investment in real estate assets with a belief in ever-rising home prices was accepted as the state of the art, like what happens every time a new technology or a new esoteric market instrument raises false hope and over-optimism. Like every time, it was again delusions of grandeur -- a big enough distance from the reality so that even the economists, policy makers, and the market experts would falter.

But perhaps we learn better every time, the way we learned how to tame inflation. Prudent government policies might avert future meltdowns. Perhaps many “unhappy families” will become woeful tales of the past.

REFERENCES


