

Good Governance Problems and Recent Financial Crises in Some EU Countries

Dragan Gamberger and Tomislav Smuc

Abstract

The starting point for the research has been the list of 147 banking crises within the period 1976–2011 prepared by the International Monetary Fund. The countries with crises have been analysed with respect to publicly available World Bank indicators in the periods of three years before the crises. The machine learning methodology for subgroup discovery has been used for the analysis. It enabled identification of five subsets of crises. Two of them have been identified as especially useful for the characterization of EU countries with banking crises in the year 2008. Fast growing credit activity is characteristic for the first subgroup while socioeconomic problems recognized by non-increasing quality of public health are decisive for the second subgroup. Comparative analysis of EU countries included into these subgroups demonstrated statistically significant differences with respect to World Bank good governance indicator values for the period before the crisis. Control of corruption, rule of law, and government effectiveness are the indicators which are statistically different for these sets of countries. The significance of the result is in the segmentation of the corpus of countries with banking crises and the recognition of connections between banking crises, socioeconomic problems, and governance effectiveness in some EU countries.

The authors present the methodology and the results of their paper in the talk now available as a video:

http://videlectures.net/ktsymposium2013_gamberger_modeling/

JEL C63 E01 H11 J00

Keywords Banking crises; World Bank indicators; subgroup discovery; good governance; corruption

Authors

Dragan Gamberger, ✉ Rudjer Boskovic Institute, Croatia, dragan.gamberger@irb.hr
Tomislav Smuc, Rudjer Boskovic Institute, Croatia

Citation Dragan Gamberger and Tomislav Smuc (2014). Good Governance Problems and Recent Financial Crises in Some EU Countries. *Economics: The Open-Access, Open-Assessment E-Journal*, Vol. 7, 2013-41. <http://dx.doi.org/10.5018/economics-ejournal.ja.2013-41>

1 Introduction

In systemic banking crises multiple banks in a country fail simultaneously and the effects on the economy of the country may be significant. Recent crises (2007–2009) stimulated a large amount of research directed towards analyses of linkages between financial institutions and interbank exposures (Caldarelli et al., 2013; Haldane and May, 2011) and their impact on crises development and systemic risk (Pokutta et al., 2011). This trend has been a consequence of the general opinion that the recent crises are primarily borne of the global financial system. However, in most cases of systemic crises in history, one cannot ignore the mutual dependencies between the real economy and financial sector for the crises' development (Nicolo and Lucchetta, 2011).

The aim of this work is the data driven analysis of financial and socioeconomic risk factors and their combinations that present an environment in which systemic banking crises are more likely to develop. The source of the data is the database with country level indicators publicly available from the World Bank. For the analysis a machine learning methodology aimed at identification of relevant subgroups of cases has been used. The methodology is appropriate for descriptive analysis of available data because it generates rule based models that can be easily interpreted by human experts (Gamberger and Lavrac, 2002). The methodology has been already successfully applied in a few medical domains and in the domain of political stability (Lambach and Gamberger, 2008). The result of the application of this methodology is the identification of five subgroups of banking crises that are relatively homogenous with respect to the values of the World Bank indicators for the period of three years before the crises. Segmentation of the domain of banking crises and the conclusion that banking crises are not a monolithic phenomenon are the most relevant results of the work.

The consequences are potentially manifold, especially for the future research and development of models. We have used the results for the analysis of crises in EU countries in the year 2008. It has been detected that some countries in which crises have been especially strong and existed for a long period of time are characterized by socioeconomic problems recognized by non-increasing or decreasing quality of public health. By comparative analysis it has been detected that these countries had statistically significant decrease in some of the good governance

indicators before the onset of the crisis. The result is surprisingly in accordance with the model constructed by Francis (2003) connecting governance indicators and financial fragility. Our result confirms the correctness of the model on a completely independent set of crises and with a completely different methodology. Francis started by developing a simple theoretical model that connects financial fragility and quality of governance and then demonstrated the correctness of the model by the available data. In contrast to that, we have constructed as homogeneous groups of banking crises as possible by the available data with no a priori assumptions about their characteristics. By evaluating the properties of countries included into different groups we have detected differences with respect to good governance indicators. Additionally, our result demonstrates that Francis's model is not universal, i.e. that there are banking crises that cannot be attributed to the problems of governance.

The organization of the rest of the work is as follows. In the next section we present the preparation of the data set used for the analysis while in Section 3 we give a short description of the used methodology, followed by a presentation of induced subgroups in Section 4 and analysis of the relevance and the meaning of the obtained results in Section 5. The central part of the work is in Section 6 which includes evaluation of crises in detected subsets of EU countries with respect to good governance indicators. The discussion and conclusions are in Section 7.

2 Data

The research presented in this paper is based on the list of banking crises reported by Valencia and Laeven (2012). In total 147 crises in the period 1976–2011 have been described in this document and they have been used as the positive cases for our analysis. 29 out 147 crises are in the period 2008–2011. Examples of crisis cases are China in the year 1998 and the USA in the years 1988 and 2007. As a control group we have used 287 cases. This control group of non-crises or negative cases has been chosen from the same countries that have experienced banking crises, but in such manner that negative cases are in 10 year increments separated from positive cases for the same country. Examples of non-crises are Finland in the years 1971, 1981, and 2001 and the UK in the years 1977, 1987, and 1997.

The reason is that Finland had a banking crisis in the year 1991 and the UK in the year 2007. A period of 10 years without a crisis is assumed as long enough to demonstrate that a country is a good representative of a non-crisis case. The total number of non-crisis cases represents the maximal number of cases that could be generated with the given constraints.

The crisis and non-crisis cases have been described by 105 indicators available from the World Bank dataset. The indicators are publicly available from the World Bank Data website (<http://data.worldbank.org/indicator>). At first we included 5 indicators suggested as potentially relevant by the mentioned International Monetary Fund document (Valencia and Laeven, 2012). They are: current account balance as percentage of GDP, central government debt as percentage of GDP, domestic credit to private sector as percentage of GDP, foreign direct investments as percentage of GDP, and bank capital to assets ratio. Besides them we have included 100 other indicators from various data fields of the World Bank database. Included indicators are from economic policy, health, agriculture, and gender sets of data. From each field we tried to select a few, most representative indicators. In order to be able to compare countries of largely different indicator values (such as GDP) on relatively equal terms, instead of using absolute value indicators we used only indicators of fractional or relative type (examples are: percentage of rural population, life expectancy at birth, percentage of unemployment with tertiary education, and research and development expenditure as percentage of GDP). Good governance indicators have not been included because their values are available only for the period after the year 1996.

An important data preprocessing step was the transformation of basic indicators into values from the temporal window preceding a positive/negative case year. We have used as a representative temporal window the period of 3 years before the event. Besides 3 basic values, 6 new indicators are introduced for each window: mean value for the window, slope, minimum value, maximum value, and relative years related to minimum and maximum value of an indicator in the time window before the event year. The result is a dataset consisting of 147 positive examples and 287 negative examples so that each of these examples is described by 945 (105 times 9) numerical attributes. Some of these attributes have unknown values. The dataset is prepared in a form that may be used by diverse machine learning systems. In our previous work we have analysed crises to the year 2007 (Gamberger

et al., 2012). In this work we have used as negative cases countries that have never experienced banking crises and the countries that have experienced crises but so that negative cases are at least 10 years from periods of crisis. Due to the fact that in the period 2008–2011 crises have been experienced in many developed countries and that this period is the most interesting for the analysis, it has been decided that negative cases can be only from countries that have experienced crises. The reason for such a bias in constructing negative cases is to avoid detection of differences between countries experiencing crises and those that never had a banking crisis, especially to avoid detection of differences between developed economies in which crises are relative often and undeveloped economies in which banking crises are rare. The aim of the work has been to identify differences in indicator values that precede the outburst of crises. The underlying problem of the analysis based on this dataset is its skew between positive and negative cases with respect to the temporal dimension: many positive cases are from the most recent period, while most of the negative cases are from the period before the year 2000. This bias could possibly lead to induction of subgroup descriptions that reflect time-related development of countries rather than the development of crises. We have tested this possibility by repeating the complete subgroup induction process with the same examples but classified so that the 147 most recent crises and non-crises examples (after year 1996) have been positive examples and all other examples (before year 1996) have been set as negative. The resulting subgroups have been significantly different from those obtained for the crises/non-crises classification of cases. Based on this result it may be concluded that in spite of time related skew of data we are able to induce crises related results.

3 Methodology

Subgroup Discovery (SD) was introduced as a data mining methodology by Klösgen (1996) and Wrobel (1997). SD techniques aim to discover distinct but potentially overlapping subsets in data that are statistically unique or interesting and at the same time as large as possible. The goal of subgroup discovery is the induction of human interpretable descriptions of subgroups. The input is a set of cases consisting of a group of positive cases P (here, countries experiencing banking crises

in a specific year) and a control group of negative cases N (countries in a period of no banking crisis). The subgroup discovery algorithm constructs rules that are true for positive cases and false for negative cases. It is not necessary that rules are true for all positive cases and false for all negative cases, but the intention is to find short rules that are true for large subsets of positive cases and at the same time false for large subsets of negative cases. Subgroup sizes are not defined in advance but the algorithm tends to make them as large as possible. A rule with ideal covering properties is true for all positive cases and not true for all negative ones. Positive cases covered by a rule are also called true positives and their number is denoted by TP , while negative cases covered by the rule are called false positives (FP). All remaining negative cases not covered by the rule are called true negatives (TN). An ideal rule has $TP = |P|$ and $TN = |N|$, and because of $|N| = TN + FP$ the ideal rule has $FP = 0$.

The first step in the rule construction process is the construction of all possible features representing elementary rule building blocks (Fürnkranz et al., 2012). For numerical attributes the features have the form *Attribute* > *value* or *Attribute* < *value*. Examples of features for the attributes in the crises/non-crises dataset are percentage of rural population > 40.8 or slope of quasi-liquid liabilities < 0.11. For each input attribute there can be many different features and the process of their construction is well defined. Practically for each pair of one positive and one negative case it is possible to construct one feature for every attribute. For example, if we have a positive case with percentage of unemployment = 10 and a negative case with percentage of unemployment = 15 then a feature percentage of unemployment < 12.5 may be constructed. This feature will successfully discriminate between these two cases because it is true for the positive case and false for the negative one.

The central part of the rule construction process is the search algorithm for selecting combinations of features with optimal covering properties on the given set of cases. Features can be connected only by logical conjunction. This means that a combination of features is true for a case only if all features are true for the case and that a combination of features is false for a case if any of the features is false for it. In the subgroup discovery approach, the following rule quality measure Q is used as the optimization goal in the heuristic search of rules: $Q = TP / (FP + g)$ where g is an appropriately selected generalization parameter. High quality rules will

have a large Q value and they will cover many positive cases (large TP) and a low number of negative cases (small FP). The number of tolerated FP cases relative to the number of TP cases covered by the rule is determined by parameter g . Most relevant rules are typically generated with intermediate values of the parameter but the final decision which model will be selected as the most appropriate depends on human expert evaluation of included conditions, unexpectedness of the result, or on the possible practical relevance of the rule. For the experiments within the crises/non-crisis domain g values were varied in the range between .2 and 5 and the results reported in the next section are obtained with a g value of 0.5.

The subgroup discovery methodology based on ILLM (Inductive Learning by Logic Minimization) system has been implemented at the Rudjer Boskovic Institute, Zagreb, Croatia. At <http://dms1.irb.hr> there is a publically available Data Mining Server that may be used for subgroup discovery tasks on user-submitted data. The server presents a very simple and user-friendly interface to the data analysis process, but is limited to 1000 cases and 1000 descriptors to prevent server overload. At http://dms1.irb.hr/do-illm/examples/list_of_examples.php one can find and download the complete dataset described in this work. The results reported in the next section can be repeated by uploading the dataset to the page, http://dms1.irb.hr/do-illm/bin/levelA/execute_levelA.php, by selecting generalization parameter equal 0.5, selecting model complexity equal 4, and finally by pressing the Start Induction button. Computation time is about 15 minutes.

4 Induced Subgroups of Crises

The result of the descriptive induction process for the prepared dataset is the detection of 5 subgroups of banking crises. The subgroups are defined by the list of included positive cases. The properties of the subgroups are described by a list of necessary conditions and a list of supporting conditions. The necessary conditions are features that are used in the body of the rule and must be satisfied in order that a positive case is included into the subgroup. The supporting conditions are features that are typically true for positive cases included in the subgroup. They are determined by repeating the same subgroup discovery methodology on the dataset which includes only positive examples from this subgroup and all negative cases.

In this case the complexity of constructed rules is limited to one feature with the intention to derive a list of independent features characteristic for the subgroup. The generalization parameter used, value g is identical (equal 0.5) as for the basic subgroup induction.

According to the preliminary expert evaluation of detected necessary and supporting conditions each subgroup received a name.

Subgroup 1: *Fast growing credit activity in economies with aging population*

List of 16 included crises: Sweden in year 1991, USA and UK in year 2007, Belgium, Denmark, France, Greece, Ireland, Island, Italy, Luxemburg, Netherlands, Portugal, Slovenia, Spain and Sweden in year 2008.

The necessary conditions:

- slope of domestic credits to private sector as % of GDP in the period of three years before crisis > 5.8% per year
- life expectancy for females three years before the crisis > 80.2 years.

The supporting conditions:

- under-five mortality rate in the year before crisis < 8.0 (per 1000 live births)
- population ages 14 and less in the year before the crisis < 21.6% of total population
- population ages 65 and above in the year three years before the crisis > 11.0% of total population.
- market capitalization of companies maximal value in the three year period > 51.1% of GDP.

Subgroup 2: *High credit activity in economies with high social security*

List of 16 included crises: Japan in year 1997, UK in year 2007, Austria, Belgium, Germany, Denmark, France, Greece, Luxemburg, Netherlands, Portugal, Spain, Slovenia, Sweden, Switzerland in year 2008.

The necessary conditions:

- maximal under-five mortality rate in the period of three years before crisis < 6.3 (per 1000 live births)
- population ages 65 and above mean value in the three years period before the crisis > 14.2% of total population.

The supporting conditions:

- life expectancy for females two years before the crisis > 80.4 years
- population ages 14 and less in the year three years before the crisis < 19.0% of total population
- domestic credits to private sector in the year before the crisis > 78.5% of GDP
- money and quasi money in the year before the crisis > 64.7% of GDP
- market capitalization of companies in the year before the crisis > 48.9% of GDP.
- expenditures for public health minimal value in the three year period > 6.9% of GDP
- male agriculture employees maximal value in the three year period < 11.6% of male employment.

Subgroup 3: *Increasing credit activity in developing economies*

List of 16 included crises: Chile in year 1976, Israel and Spain in year 1977, Argentina in year 1980, Uruguay and Chile in year 1981, Jordan in year 1989, Lebanon and Brazil in year 1990, Sweden in year 1991, Kuwait in year 1992, Brazil and Mexico in year 1994, Argentina in year 1995, Colombia in year 1998, Uruguay in year 2002.

The necessary conditions:

- domestic credits to private sector as % of GDP in the year before the crisis > maximal value in two previous years.

- population aged 15–64 in the year before the crisis < 64.3% of total population
- rural population in the year three years before the crisis < 33.7% of total population

The supporting conditions:

- under-five mortality rate in the year before the crisis < 60.3 (per 1000 live births)
- life expectancy for females in the year before the crisis > 68.8 years
- annual population growth maximal value in the three years period before the crises > 0.4
- road sector energy consumption in the year two years before the crisis > 13.5% of total energy consumption.

Subgroup 4: *Socioeconomic problems recognized by decreasing life expectancy*

List of 17 included crises: Congo in the years 1991, 1992, and 1994, Latvia in the years 1995 and 2008, Sierra Leone in year 1990, Kenya in year 1992, Uganda and Burundi in year 1994, Belarus, Central African Republic, Cameroon, Kyrgyzstan, Lithuania, Swaziland, Zambia, Zimbabwe in year 1995.

The necessary condition:

- slope of life expectancy for females in the period of three years before crisis < -0.3 years per year

The supporting conditions:

- slope of under-five mortality rate in the period of three years before the crisis > -2.5 (per 1000 live births per year)
- annual population growth three years before the crises > population growth one or two years before the crisis

- mean value added by agriculture in the three years period before the crisis > 10.8% of GDP.

Subgroup 5: *Socioeconomic problems recognized by non-increasing quality of public health*

List of 25 included crises: Sierra Leone in year 1990, Finland, Liberia, Nigeria, Norway, and Sweden in year 1991, Kenya and Poland in year 1992, Burundi in year 1994, Belarus, Central African Republic, Latvia, Lithuania, Swaziland, and Zimbabwe in year 1995, Bulgaria in year 1996, Ukraine in year 1998, Uruguay in year 2002, Belgium, Greece, Hungary, Island, Italy, Portugal, and Spain in year 2008.

The necessary conditions:

- non-increasing life expectancy for females in the period of three years before crisis
- slope of under-five mortality rate in the period of three years before crisis > -0.5 (per 1000 live births per year)

Supporting condition:

- annual money and quasi money growth in the year two years before the crisis > 5.2% of GDP.

5 Analysis of Subgroups

Subgroups 1 and 2 are relevant because they include many banking crises in developed countries in the years 2007 and 2008. There has been a strong "avalanche" effect causing crises in many countries at the same time, but the detected conditions demonstrate that there have been common patterns in many countries characterized by strongly increasing credit activity in economies with an ageing population (Subgroup 1) or high credit activity in economies with high social security (Subgroup 2). The subgroups are very similar and there are 10 countries that are in both of them (like Italy, Netherlands, and Spain). The crisis in the USA in the year 2007

is perhaps the most representative example for the first group while Germany and Switzerland are representative examples for the second one. Although Subgroups 1 and 2 are mostly characteristic for the crises in years 2007 and 2008, it can be noticed that Subgroup 1 includes also the crisis in Sweden in the year 1991 while Subgroup 2 includes the crisis in Japan in the year 1997. This demonstrates the generality of the result and to some extent that the induced model is robust to the temporal imbalance of positive and negative cases.

Subgroup 3 consists of a very different set of countries. Most of them are developing countries characterized by a relatively low percentage of active population (below 64.3%) in which an increase in credit activity has happened. Representative examples are Israel in the year 1977, Argentina in the year 1980, and Uruguay in the year 2002. It is interesting to notice that this subgroup includes also the crisis in Spain in the year 1977 and the crisis in Sweden in the year 1991.

Similarity among Subgroups 1, 2, and 3 is that they all include the same driving force (increased credit activity) in the societies that have a relatively low percentage of active population. In Subgroups 1 and 2 this is mainly due to an aging population while in Subgroup 3 this is due to the relative small percentage of active population that in many developing countries is a consequence of a high percentage of the very young. Our interpretation of this result is that the high availability of credit supports the national economy but also that expansive credit activity, which is over some limit, may result in a banking crisis. At least one part of the credit absorption limit of a country depends on the number of its active population because the profit necessary for the repayment of credit can be formed only by the active work. In the case of the World Bank indicators we do not have an indicator of the credit absorption limit of a country and so it happened that a low percentage of active population and an aging population have been used as its proxy.

It can be noticed that the necessary conditions for Subgroup 2 actually describe all developed economies. This is a consequence of the fact that in the year 2008 we had banking crises in many developed economies and in the used data set there is a small total number of examples of developed economies without crises. Subgroups 1 and 2 significantly overlap and a necessary condition for Subgroup 1 is high slope of domestic credit to the private sector while one of the supporting conditions for Subgroup 2 is the high absolute value of this credit. The appropriate interpretation

is that high credit activity in developed economies is dangerous and not the fact that an economy is developed.

For Subgroups 4 and 5 a common characteristic is a decreasing quality of life. Together they form a larger group of banking crises that we define as socioeconomically driven crises. The interpretation is that a decreasing quality of life is not a cause of banking crises but actually a sign of problems in the country connected with the worsening of the macro-economic situation which may create an environment for the development of banking crises. There are many possible causes of systemic problems including ethnic or civil wars, significant changes within the economic system, and deep political crises. A good example is that all countries in transition had both significant socioeconomic problems and banking crises. With respect to the underlying problems however, individual countries may differ significantly.

Subgroup 4 includes developing countries in which in some cases banking crises are related with turbulent conditions. For example, Congo in the period 1991–1994 is faced with the suspension of military and financial assistance for the Mobutu regime, Kenya in year 1992 experienced significant violence in certain parts of the country before presidential elections, and finally in Belarus, Kyrgyzstan, Lithuania, and Latvia in the year 1995, these post-communist countries are trying to implement novel economical models. In all these cases, a common result has been an instable socioeconomic system prone to banking crises. Besides stagnating life expectancy Subgroup 5 includes the necessary condition of stagnating mortality of children under the age of 5 years. Both conditions to some extent reflect the quality of the health care in the country and its decrease or stagnation seems to be able to signal various socioeconomic problems in that society. It is relevant to notice that besides crises in countries like Nigeria and Ukraine the model is valid for the Finnish, Norwegian, and Swedish crises in the year 1991, as well as for a few EU countries for crises in the year 2008.

Based on the presented analysis it may be concluded that we have two types of subgroups. In the first are Subgroups 1–3 that have in common increased values for the indicator of domestic credit to private sector in the country either as their necessary or sufficient conditions. One of these subgroups has also a supporting condition representing money and quasi-money (M2) as percentage of GDP. Because of these indicators, the first group of subgroups may be recognized as a

"financially driven" type of banking crises. Subgroups 4 and 5 have stagnating or decreasing life expectancy of females as a common necessary condition. Additionally, both subgroups have increasing or stagnating mortality of children as either necessary or supporting conditions. They may be recognized as the "socioeconomic problems" type of banking crises.

6 Evaluation of EU Countries Included into Subgroup 5

Necessary conditions of Subgroup 5 are health related World Bank indicators that identify socioeconomic problems of a country. When accompanied by the supporting condition of significant money growth they present an environment in which banking crises may outburst. The subgroup includes crises in undeveloped and developing countries like the crises in Sierra Leone in the year 1990, in Kenya in the year 1992, in Burundi in the year 1994, and in Bulgaria in the year 1996. It is interesting to notice that all three countries of the Nordic banking crisis in the year 1991 (Finland, Norway and Sweden) are also included in this subgroup. Even more intriguing is that crises in the year 2008 in six EU countries (Belgium, Hungary, Greece, Italy, Portugal, and Spain) are included as well. With the exception of Hungary, five of these crises are also in Subgroup 1. This fact may be interpreted as a sign that although the crises in these countries have been triggered by high credit activity that there exist also socioeconomic reasons for the crises.

An independent set of World Bank indicators has been used in order to test the hypothesis that EU countries included in Subgroup 5 are different from other EU countries with banking crises in the same year. This is the set of Worldwide Governance Indicators (<http://info.worldbank.org/governance/wgi/index.asp>) which consists of six aggregated indicators representing voice and accountability, political stability and absence of violence, government effectiveness, regulatory quality, rule of law, and control of corruption (Kaufmann et al., 2010).

A good characteristic of the indicators is that they are available as absolute values and as percentile ranks (p-rank, rank over the complete set of 215 economies presented in range 0–100 with 0 as the lowest value). The latter is appropriate for comparative analysis of the performance of a single country or a group of countries. We have used it to compute differences in ranking between the years

2007 and 2004, i.e. for the period before the crises in the year 2008. At first the differences have been computed for all six indicators for Belgium, Greece, Hungary, Italy, Portugal, and Spain and the results are presented in the upper part of Table 1. After that we have selected 5 other EU countries that all experienced banking crises in the year 2008 but which have not been included into Subgroup 5. They are Austria, Denmark, France, Germany, and Netherlands. The results for these countries are presented in the middle part of Table 1. Finally, the differences between these two groups of countries have been evaluated by T-test and the levels of statistical significance are presented in the last row. Columns representing governance indicators are ordered by decreasing significance. The result means that the most significant difference with level of 99.9% between these two groups of countries is in respect of control of corruption, followed by rule of law (level 97%) and government effectiveness (level 96%). The differences in the remaining three indicator rankings are not significant but the differences in the total sum (last column) are significant, with a level of 99%.

With respect to this result, it is interesting to look also at the most recent World Bank data available for the year 2011. Table 2 presents lists of EU countries

Table 1: Differences in p-ranks for years 2007 and year 2004 for six governance indicators for two groups of EU countries and the statistical significance of the difference between the two groups.

	Control of corruption	Rule of law	Government effectiveness	Voice and accountability	Political stability	Regulatory quality	Total
Belgium	-4.32	0.00	-1.91	1.92	0.00	3.03	-1.28
Greece	-6.66	-5.74	-4.75	-8.65	0.48	0.22	-25.10
Hungary	-2.80	-1.44	-0.87	-4.33	-2.88	2.10	-10.21
Italy	-6.16	-6.22	-12.02	-1.44	6.25	-2.24	-21.82
Portugal	-4.78	-5.26	-6.23	-1.92	-5.77	-2.28	-26.25
Spain	-7.71	-0.96	-8.20	-4.81	-12.98	-0.85	-35.51
Austria	-0.96	3.35	3.43	2.88	11.06	3.47	23.23
Denmark	0.49	1.44	-0.49	-2.40	3.85	1.97	4.85
France	1.99	-1.91	-3.36	0.00	5.77	1.11	3.59
Germany	-0.94	0.00	3.45	0.00	14.90	3.00	20.41
Netherlands	1.48	0.00	-3.39	1.92	-7.69	-0.47	-8.14
Level of statistical significance	99.9%	97%	96%	Non-sig.	Non-sig.	Non-sig.	99%

with the most significant decrease in relevant governance indicators for the period 2008–2011. It may be noticed that Greece and Italy are still on the top of these lists. A possible interpretation is that the situation in these countries is not improving and the fact has been practically confirmed by their ongoing financial problems in the years 2012 and 2013. It is perhaps even more relevant that Cyprus and Slovenia, for which it is known that they started to have financial crises in years 2012 and 2013, are both highly positioned in these lists.

It must be noted that the results do not demonstrate the predictive quality of good governance data with respect to the banking crises, but only that financial fragility, at least for developed EU countries, is related with good governance. Even more relevant is the growing body of literature which questions the validity of good governance indicators (Arndt and Oman, 2006; Weber Abramo, 2008). The main reason is that the indicators are completely based on perception. For the presented results it means that the deteriorating socioeconomic situation in a country may drive these negative dynamics regardless of the real quality of governance. Additionally, there is some evidence that six different indicators are not able to distinguish among aspects of the quality of governance but that they appear to measure a broad concept of government effectiveness that is potentially best described by the averaged value of the six indicators (Langbein and Knack, 2010). This means that the negative trends of a concrete indicator are not directly actionable for policy-making and that the total score is the best indication of governance related problems.

Table 2: Lists of EU countries with most significant decrease of difference in p-ranks for years 2011 and year 2008 for six governance indicators.

Total for 6 governance indicators		Total for 3 most relevant indicators		Control of corruption indicator	
Greece	-39.49	Greece	-19.57	Italy	-5.76
Malta	-29.58	Malta	-11.35	Cyprus	-5.33
Slovenia	-26.84	Austria	-8.24	Greece	-5.24
Portugal	-25.13	Hungary	-7.59	Austria	-5.09
Ireland	-19.08	Cyprus	-6.62	Malta	-3.84

7 Discussion and Conclusions

The presented models are obtained by descriptive induction based on the subgroup discovery methodology. The results are potentially relevant for human interpretation and for the better understanding of the connections existing between publicly available World Bank indicators and the occurrences of banking crises.

The results confirm that excessive credit activity and the high availability of money and quasi-money present a high risk for the outburst of banking crises. The novelty is that, besides these financial factors, all induced models include as necessary conditions also demographic and/or public health indicators. In Subgroup 1 the life expectancy of females should be interpreted as a sign of a reduced percentage of an active population. This result actually means that the high credit activity is dangerous especially in developed economies with an ageing population. Because most of the crises described by Subgroups 1 and 2 happened simultaneously in the year 2008, inter-country dependences are obviously very strong for this type of crisis. The available data do not include this information and consequently the induced conditions cannot include inter-country dependences but these relationships must be taken into account in the expert evaluation.

Subgroup 5 is much more general because it includes undeveloped, developing, and developed countries in a relative large time frame. The result states that banking crises may be expected in countries with some socioeconomic problems. Although socioeconomic problems may have various origins and causes, it seems that they may be, at least to some extent, identified by indicators like stagnating or decreasing life expectancy and stagnating or increasing mortality of children. The supporting condition for Subgroup 5 is high money and quasi-money growth. The result is useful to understand that a banking crisis in many cases is a normal consequence of problems that are not financially related, and also that appropriate (restrictive) monetary policies may avert a potential banking crisis.

The most relevant result is that socioeconomic problems detected by Subgroup 5 for some EU countries are strongly connected with changes in values of governance indicators for these countries before the outburst of crises. From available data it is not possible to conclude causality relations between banking crises, socioeconomic problems, and governance indicators. It is not clear if socioeconomic problems are the result of problems in governance or vice versa, and how they both

are connected with banking crises, but results demonstrate that selected socioeconomic and governance indicators collected and prepared by the World Bank may be used as warning signs for country-level problems.

The results presented in Table 1 demonstrate that banking crises in some EU countries in the year 2008 and the financial crises that followed have much more complex backgrounds than pure financial causes. The result is in accordance with the model developed by Francis (2003) that connects governance indicators and financial fragility. The correctness of the Francis's model is confirmed on examples and by the methodology that is different from those used for the development of the model. Additionally, we demonstrate that trends (differences) of governance indicators are more relevant than their absolute values.

From the results it is also clear that Francis's model is not able to describe all cases of banking crises. According to our results it is valid only for a relatively small but significant part of crises. It means that appropriate modeling and understanding of banking crises is possible only after successful detection and grouping of similar patterns of events. The results presented in this work are perhaps the first step in this direction.

The importance of the work is in the fact that it demonstrates that future banking and financial crises prevention should also focus on governance effectiveness, more strict law implementation, and measures against corruption.

Acknowledgements: This work reported in this paper was supported by EU FP7 FOC-II project (255987) and Croatian Ministry of Science, Education, and Sport project "Machine Learning Algorithms and their Applications" (098-0982560-2563).

References

- Arndt, C., and Oman, C. (2006). *Uses and Abuses of Governance Indicators*. OECD Development Centre. URL http://www.oecd.org/document/25/0,2340,en_2649_33935_37081881_1_1_1_1,00.html.
- Caldarelli, G., Pammolli, F., Gabbi, G., Riccaboni, M., Battiston, S., and Delpini, D. (2013). Evolution of Controllability in Interbank Networks. *Scientific Reports*, 3(1626). DOI doi:10.1038/srep01626. URL <http://www.nature.com/srep/2013/130409/srep01626/full/srep01626.html>.
- Francis, M. (2003). Governance and Financial Fragility: Evidence from a Cross-Section of Countries. Working Papers 03-34, Bank of Canada. URL <http://ideas.repec.org/p/bca/bocawp/03-34.html>.
- Fürnkranz, J., Gamberger, D., and Lavrač, N. (2012). *Foundations of Rule Learning*. Springer-Verlag. DOI 10.1007/978-3-540-75197-7. URL <http://www.springer.com/978-3-540-75196-0>.
- Gamberger, D., and Lavrac, N. (2002). Expert-Guided Subgroup Discovery: Methodology and Application. *Journal of Artificial Intelligence Research*, 17: 501–527. URL <http://www.jair.org/papers/paper1089.html>.
- Gamberger, D., Lucanin, D., and Smuc, T. (2012). Descriptive Modeling of Systemic Banking Crises. In J.-G. Ganascia, P. Lenca, and J.-M. Petit (Eds.), *Discovery Science*, volume 7569 of *Lecture Notes in Computer Science*, pages 67–80. Springer. URL <http://link.springer.com/book/10.1007/978-3-642-33492-4/page/1>.
- Haldane, A. G., and May, R. M. (2011). Systemic Risk in Banking Ecosystems. *Nature*, 469(7330): 351–355. DOI 10.1038/nature09659. URL <http://www.nature.com/doifinder/10.1038/nature09659>.
- Kaufmann, D., Kraay, A., and Mastruzzi, M. (2010). The Worldwide Governance Indicators : Methodology and Analytical Issues. Policy Research Working Paper Series 5430, The World Bank. URL <http://ideas.repec.org/p/wbk/wbrwps/5430.html>.

- Klösigen, W. (1996). Explora: A Multipattern and Multistrategy Discovery Assistant. In *Advances in Knowledge Discovery and Data Mining*, pages 249–271. Publisher AAAI. URL <http://dl.acm.org/citation.cfm?id=257965&CFID=225928170&CFTOKEN=81206005>.
- Lambach, D., and Gamberger, D. (2008). Temporal Analysis of Political Instability through Descriptive Subgroup Discovery. *Conflict Management and Peace Science*, 25: 19–32. DOI 10.1080/07388940701860359.
- Langbein, L., and Knack, S. (2010). The Worldwide Governance Indicators: Six, One, or None? *The Journal of Development Studies*, 46(2): 350–370. URL <http://EconPapers.repec.org/RePEc:taf:jdevst:v:46:y:2010:i:2:p:350-370>.
- Nicolo, G. D., and Lucchetta, M. (2011). Systemic Risks and the Macroeconomy. Working Paper 16998, National Bureau of Economic Research. URL <http://www.nber.org/papers/w16998>.
- Pokutta, S., Schmaltz, C., and Stiller, S. (2011). Measuring Systemic Risk and Contagion in Financial Networks. *Social Science Research Network Working Paper Series*. URL <http://ssrn.com/abstract=1773089>.
- Valencia, F., and Laeven, L. (2012). Systemic Banking Crises Database: An Update. IMF Working Papers 12/163, International Monetary Fund. URL <http://ideas.repec.org/p/imf/imfwpa/12-163.html>.
- Weber Abramo, C. (2008). How Much Do Perceptions of Corruption Really Tell Us? *Economics: The Open-Access, Open-Assessment E-Journal*, 2(2008-3). DOI 10.5018/economics-ejournal.ja.2008-3. URL <http://dx.doi.org/10.5018/economics-ejournal.ja.2008-3>.
- Wrobel, S. (1997). An Algorithm for Multi-Relational Discovery of Subgroups. In *First European Symposium on Principles of Data Mining and Knowledge Discovery*, pages 78–87. Springer. URL http://link.springer.com/chapter/10.1007%2F3-540-63223-9_108.

Please note:

You are most sincerely encouraged to participate in the open assessment of this article. You can do so by either recommending the article or by posting your comments.

Please go to:

<http://dx.doi.org/10.5018/economics-ejournal.ja.2013-41>

The Editor