ON FACTORS EXPLAINING THE 2008 FINANCIAL CRISIS

Eduardo Acosta-González
Department of Quantitative Methods, University of Las Palmas de Gran Canaria,
Campus de Tafira, 35017 Las Palmas de Gran Canaria, Spain.
eacosta@dmc.ulpgc.es

Fernando Fernández-Rodríguez (*)
Department of Quantitative Methods, University of Las Palmas de Gran Canaria,
Campus de Tafira, 35017 Las Palmas de Gran Canaria, Spain.
Telefone: +0034928451802. Fax: +0034928451829
ffernandez@dmc.ulpgc.es

Simón Sosvilla-RiveroFernando
Department of Quantitative Economics, Complutense University of Madrid,
Campus de Somosaguas, 28223 Madrid, Spain.
sosvilla@ccee.ucm.es

(*) Corresponding Author

Classification JEL codes: E32, E44, E51

Keywords: Financial Crises, Genetic Algorithms

Abstract

Starting with a large number of potential causes and indicators of the 2008 crisis, we select the best econometric model for explaining the severity of the crisis using a statistical methodology only guided by data, which is carried out employing a Genetic Algorithm. Our results show that such a model just contains one factor: "percentage of bank claims in the private sector over deposits in the year 2006", being very robust to alternative measures of the severity of the crisis. When lower penalty is considered for the introduction of factors in the model, the severity of the crisis is robustly explained by just two financial factors: "percentage of bank claims in the year 2006", and "percentage in GDP of Domestic Credit Private Sector in the year 2006". In addition, the presence in the model of the geographical and non financial factor "Central/Eastern European or Central Asian" would suggest that these countries have received special impact during the crisis.

---

1. Introduction

There is an increasing literature concerning potential factors determining the relative performance during the 2008 global financial crisis. In order to study the feasibility of an early warning system to predict the timing of the crises, Rose and Spiegel (2009) have studied the possibility of modelling the causes of the 2008 financial crisis and why its severity differs across countries. Their methodology treats the severity of the 2008 crisis as a latent variable, observed only imperfectly, and which is linked to four observable indicators of the crisis: the real GDP growth over 2008, the declines in the perception of a country's creditworthiness, equity market collapses and exchange rate depreciations. In order to combine appropriately these four factors, a common component is extracted using conventional factor analysis. Additionally, to carry out a sensitivity analysis, three additional variants of this common component are produced, replacing the four observable indicators of the crisis with others obtained with similar, but alternative, methodologies. So, four different variables measuring the severity of the crisis, which are strongly positively correlated, are employed.

As explanatory factors of the 2008 crisis, Rose and Spiegel (2009) have considered over sixty potential factors frequently studied in literature. It covers such categories as: financial system policies and conditions, asset price appreciation in real estate and equity markets, international imbalances and foreign reserve adequacy, macroeconomic policies, and institutional and geographic features.

Starting with the previous paper by Rose and Spiegel (2009) and using their database, the main contribution of our paper consists of selecting the best econometric model for explaining the severity of the 2008 financial crisis from the potential factors frequently considered in the literature. We employ a purely statistical methodology which automatically selects the factors in the econometric model in a process only guided by data. Our methodology properly handles multicolineality problems in the estimations that could arise due to the redundancy of the information provided by the factors. Besides, it offers a parsimonious model of the severity of the crisis, containing few factors with as much information as possible that capture the essential characteristics of the data, and can help in identifying the main reasons explaining the crisis.

2. Our methodology

In order to avoid data mining problems for constructing the best econometric model explaining the crisis, we follow the general to specific approach characteristic of the London School of Economics based on the theory of reduction [Hendry, 1995, ch. 9]. It means that given an endogenous variable Y, which represents the severity of the 2008 crisis, and a set $X_1,....,X_N$ of sixty-plus potential factors explaining it, the problem that we face is to find the best submodel of the form:

$$Y = \beta_0 + \beta_1 X_{i_1} + ... + \beta_K X_{i_K} + \varepsilon \text{, where } \{i_1, i_2 ...., i_K\} \subseteq \{1, 2, ..., N\} \text{, } K < N \square 60$$

In building this multiple regression model, due to the multicolineality produced by the redundancy of the information provided by the complete set of factors, a crucial problem is the selection of regressors that should be included. If a lower amount of regressors are selected in the model, the estimate of the parameters will not be consistent and if a higher amount is selected, its variance will increase.

The problem is that the number of submodels where the criteria developed by econometric theory for selecting the best model is 2^N. So, for $N = 60$, the number of possible models is higher than 10^{18}. In order to resolve this intractable problem, we follow the heuristic strategy proposed by Hoover and Perez (1999) by searching all the feasible reduction paths when simplifying the general model. In particular, we use the methodology developed in Acosta-González and Fernández-Rodríguez (2007). This is a procedure for automatic selection of factors in the model, only guided by the data, which is carried out using a Genetic Algorithm where the lost function is the Schwartz Information Criterion (SIC, henceforth). This methodology avoids the tendency to over-identify models detected in several popular heuristic methods for selecting models, like stepwise (Lovell, 1983).

Therefore, starting with a large number of potential causes and indicators (over sixty possible explanatory factors) of the severity of the crisis considered in the literature, our methodology selects the best econometric model, in the sense of the SIC.

The SIC advises choosing the econometric model which minimizes the expression

$$SIC(m) = \log \hat{\sigma}^2 + c\frac{k}{N}\log(N) \qquad (1)$$

where $\hat{\sigma}^2 = \frac{e'e}{N}$ is the variance of the residuals e, N is the sample size, and k is the number of factors in the model.

The correcting factor c avoids the possibility of over-parameterized models and solves the trade-off between the in sample goodness of fit and the out of sample forecasting ability. The higher the value of c, the higher is the penalty for the introduction of more factors in the model.

Finally, although original data base in Rose and Spiegel´s estimates had initially 107 countries, because of missing data, their results were actually derived from only 85 countries. This procedure leads to a sample selection bias. In this paper we employed a technique of multiple imputation developed by King *et al.* (2001) which permits the approximation of missing data and allow us to make efficient use of all the information contained in the whole sample.

3. Empirical results

Our data series is the same as employed by Rose and Spiegel (2009) and were extracted from the World Bank's *World Development Indicators*, the International Monetary Fund's *International Financial Statistics*, and other key data sets. The entire (STATA 10.0) data set is available at http://faculty.haas.berkeley.edu/arose/MIMICData.zip.

In Table 1 and table 2 we present the results of applying our genetic algorithm in order to search for the best set of factors explaining the four versions of the severity of the crisis considered by Rose and Spiegel (2009). Depending on the value of parameter c in SIC, the number of factors explaining the crisis is different. The case $c = 2$ in equation (1), whose results are in Table 1, corresponds to a high reluctance to introduce

many explicative factors. In this case, our final model contains a constant and the statistically significant financial factor “percentage of bank claims in the private sector over deposits in the year 2006”. This factor is a ratio where the numerator corresponds to claims in the private sector which include gross credit from the financial system (not including Central Bank) to individuals, enterprises, nonfinancial public entities not including under net domestic credit, and financial institutions not included elsewhere. Its denominator corresponds to transferable deposits included in Broad Money and other deposits included in Broad Money. This variable is a typical financial factor that was signalling the 2008 financial crisis since 2006. As can be seen, this model is very robust as shown in the sensitivity analysis since its explanatory role is maintained for all our measures of the severity of the crises. Only when the second version of the severity of the crisis is considered, the best econometric model adds a second factor: “civil liberties in the year 2006”.

Our finding could be consistent with Eichengreen and Mitchener (2003) who conclude that financial crises throughout modern history can be viewed as “credit booms gone wrong”. Moreover, our results suggest that valuable information about macroeconomic and financial stability would be missed if central banks chose to ignore the behaviour of credit aggregates and confine themselves to following inflation targeting rules. Indeed, policymakers are now taking a harder look at how to regulate credit and the procylicality of the financial system (e.g., Turner 2009).

[TABLE 1]

In the case $c = 1$ in equation (1), where SIC permits more explanatory factors, the best econometric model depends on the endogenous variable representing the severity of the crisis. The results are presented in Table 2.

[TABLE 2]

It is important to observe that there is a predominance of financial factors describing the severity of the crisis, in the four cases. Macroeconomic policies and institutions are factors that appear occasionally. The rest of the factors never appear. So, for example, in the first model there are six financial factors, one macroeconomic policies factor and one geographical factor. Besides, there are two financial factors which were selected in each of the four models associated with the four severity measures: "percentage of bank claims in the private sector over deposits in the year 2006", variable that also was selected when $c = 2$, and "percentage in GDP of Domestic Credit Private Sector in the year 2006". A non financial factor given by the variable Central/Eastern European or Central Asian was also selected in three out of four endogenous variables describing the severity of the 2008 crisis. The appearance of this geographical and non financial factor (which is statistically significant in all of the cases) means that these European central Eastern countries and Central Asian Countries have received special impact during the crisis. Its negative sign reveals that this impact has been especially negative, and the magnitude of its coefficient measures, in any sense, the severity of the crisis.

4. Concluding remarks

This paper has empirically tested the financial origins of the 2008 crisis. Our main result is that the intensity of the crisis seems to be strongly linked to financial factors which in 2006 had been warning about the possibility of a financial crisis. This would suggest that the current crisis could have been anticipated by avoiding the strong increase in credit, revaluating the role of credit quantities and factors affecting the supply of credit in the conduct of monetary policy.

References

Acosta-González, E., Fernández-Rodríguez, F., 2007, Model selection via genetic algorithms illustrated with cross-country growth data. Empirical Economics. 33, 313-337.

Eichengreen, B., Mitchener, K. J., 2003. The Great Depression as a Credit Boom Gone Wrong. Working Paper No. 137, Bank for International Settlements, September.

Hendry, D. F., 1995. Dynamic Econometrics, Oxford: Oxford University Press.

Hoover, K. D., Perez, 1999, S.J., 1999. Data mining reconsidered: encompassing and the general- to-specific approach to specification search. Econometrics Journal. 2, 167-191.

King, G, J., Honaker, A., Joseph, Scheve, K., 2001. Analyzing incomplete political science data: an alternative algorithm for multiple imputation. American Political Science Review. 95, 49-69.

Lovell, M.C., 1983. Data mining, Review of Economics and Statistics. 65, 1-12.

Rose, A. K. Spiegel, M. M., 2009. Cross-Country Causes and Consequences of the 2008 Crisis: Early Warning. Federal Reserve Bank of San Francisco. Working Paper 2009-17. Available at http://www.frbsf.org/publications/economics/papers/2009/wp09-17bk.pdf

Turner, A., 2009. The Turner Review: A Regulatory Response to the Global Banking Crisis. London: Financial Services Authority. Available at http://www.fsa.gov.uk/pubs/other/turner_review.pdf

TABLE 1

Severity of the crisis 1 (R2=0.2437)

Factor name	coeficient	t-student	p-value
Constant	0.77060	3.76750	0.00010
Bank Claims, % Deposits 2006	-0.76380	-4.62860	0.00000

Severity of the crisis 2 (R2=0.3595)

Factor name	coeficient	t-student	p-value
Constant	0.22840	1.05550	0.14680
Bank Claims, % Deposits 2006	-0.60240	-4.44190	0.00000
Civil Liberties, 2006	0.15070	3.93110	0.00010

Severity of the crisis 3 (R2=0.2268)

Factor name	coeficient	t-student	p-value
Constant	0.67480	3.36560	0.00050
Bank Claims, % Deposits 2006	-0.66430	-4.23130	0.00000

Severity of the crisis 4 (R2=0.1995)

Factor name	coeficient	t-student	p-value
Constant	0.88290	4.26380	0.00000
Bank Claims, % Deposits 2006	-0.85190	-4.86820	0.00000

TABLE 2

Severity of the crisis 1 (R2=0.4646)

Var_name	coeficiente	t-student	p-value
Constant	0.4245	0.7813	0.2182
Official Supervisory Power, 2003	0.0368	1.8709	0.0321
Domestic Credit Private Sector, %GDP 2006	-0.0064	-3.3033	0.0007
Private Sector Credit Access, 2006	0.1303	1.6968	0.0463
Bank Liquid Reserves, %Assets 2006	-0.0068	-1.5687	0.0599
Bank Capital, %Assets 2006	-0.0413	-1.0293	0.1528
Bank Claims, %Deposits 2006	-0.6718	-3.283	0.0007
CPI Inflation, 2006	-0.0389	-1.4273	0.0782
Central/Eastern European or Central Asian	-0.5003	-2.5317	0.0064

Severity of the crisis 2 (R2=0.3930)

Var_name	coeficiente	t-student	p-value
Constant	0.3914	1.7198	0.0442
Domestic Credit Private Sector, %GDP 2006	-0.0030	-2.1998	0.0150
Bank Claims, %Deposits 2006	-0.4693	-3.2717	0.0007
Civil Liberties, 2006	0.1231	3.1611	0.0010

Severity of the crisis 3 (R2=0.4033)

Var_name	coeficiente	t-student	p-value
Constant	3.4799	2.7759	0.0033
Domestic Credit Private Sector, %GDP 2006	-0.0056	-3.7641	0.0001
Bank Claims, %Deposits 2006	-0.4683	-2.4470	0.0080
Control of Corruption	0.2920	2.3781	0.0096
Central/Eastern European or Central Asian	-0.3678	-1.9292	0.0282
Log(2006 Real GDP per capita)	-0.2762	-2.1542	0.0167

Severity of the crisis 4 (R2=0.3677)

Var_name	coeficiente	t-student	p-value
Constant	0.8078	2.8442	0.0027
Declaring Insolvency Power, 2003	0.2334	1.6306	0.0530
Domestic Credit Private Sector, %GDP 2006	-0.0070	-3.2547	0.0008
Bank Claims, %Deposits 2006	-0.6467	-3.4435	0.0004
Regulatory Quality	0.3318	2.0794	0.0200
Central/Eastern European or Central Asian	-0.5500	-2.4697	0.0076