The wiiw Balkan Observatory

Working Papers | 100 | November 2012

Ivana Herceg and Danijel Nestić

A New Cluster-Based Financial Vulnerability Indicator: The Analytical Concept and its Application for Stress Testing in a Post-Socialist Economy

The wiiw Balkan Observatory

www.balkan-observatory.net

About

Shortly after the end of the Kosovo war, the last of the Yugoslav dissolution wars, the Balkan Reconstruction Observatory was set up jointly by the Hellenic Observatory, the Centre for the Study of Global Governance, both institutes at the London School of Economics (LSE), and the Vienna Institute for International Economic Studies (wiiw). A brainstorming meeting on Reconstruction and Regional Co-operation in the Balkans was held in Vouliagmeni on 8-10 July 1999, covering the issues of security, democratisation, economic reconstruction and the role of civil society. It was attended by academics and policy makers from all the countries in the region, from a number of EU countries, from the European Commission, the USA and Russia. Based on ideas and discussions generated at this meeting, a policy paper on Balkan Reconstruction and European Integration was the product of a collaborative effort by the two LSE institutes and the wiiw. The paper was presented at a follow-up meeting on Reconstruction and Integration in Southeast Europe in Vienna on 12-13 November 1999, which focused on the economic aspects of the process of reconstruction in the Balkans. It is this policy paper that became the very first Working Paper of the wiiw Balkan Observatory Working Papers series. The Working Papers are published online at www.balkan-observatory.net, the internet portal of the wiiw Balkan Observatory. It is a portal for research and communication in relation to economic developments in Southeast Europe maintained by the wiiw since 1999. Since 2000 it also serves as a forum for the Global Development Network Southeast Europe (GDN-SEE) project, which is based on an initiative by The World Bank with financial support from the Austrian Ministry of Finance and the Oesterreichische Nationalbank. The purpose of the GDN-SEE project is the creation of research networks throughout Southeast Europe in order to enhance the economic research capacity in Southeast Europe, to build new research capacities by mobilising young researchers, to promote knowledge transfer into the region, to facilitate networking between researchers within the region, and to assist in securing knowledge transfer from researchers to policy makers. The wiiw Balkan Observatory Working Papers series is one way to achieve these objectives.

The wiiw Balkan Observatory

Global Development Network
Southeast Europe

This study has been developed in the framework of research networks initiated and monitored by wiiw under the premises of the GDN–SEE partnership.

The Global Development Network, initiated by The World Bank, is a global network of research and policy institutes working together to address the problems of national and regional development. It promotes the generation of local knowledge in developing and transition countries and aims at building research capacities in the different regions.

The Vienna Institute for International Economic Studies is a GDN Partner Institute and acts as a hub for Southeast Europe. The GDN–wiiw partnership aims to support the enhancement of economic research capacity in Southeast Europe, to promote knowledge transfer to SEE, to facilitate networking among researchers within SEE and to assist in securing knowledge transfer from researchers to policy makers.

The GDN–SEE programme is financed by the Global Development Network, the Austrian Ministry of Finance and the Jubiläumsfonds der Oesterreichischen Nationalbank.

For additional information see www.balkan-observatory.net, www.wiiw.ac.at and www.gdnet.org

A NEW CLUSTER-BASED FINANCIAL VULNERABILITY INDICATOR:

THE ANALYTICAL CONCEPT AND ITS APPLICATION FOR STRESS TESTING

IN A POST-SOCIALIST ECONOMY

by

Ivana Herceg* and Danijel Nestić**

*Croatian National Bank
**Institute of Economics, Zagreb

Final paper submitted to the wiiw GDN-SEE-CIS Research Competition 2011
30 November 2012

A NEW CLUSTER-BASED FINANCIAL VULNERABILITY INDICATOR: THE ANALYTICAL CONCEPT AND ITS APPLICATION FOR STRESS TESTING IN A POST-SOCIALIST ECONOMY

ABSTRACT

This paper proposes a new approach to household financial vulnerability analysis by employing cluster analysis techniques in the identification of potentially vulnerable households. The cluster-based vulnerability indicator is combined with a binary dependant variable model and used in stress testing. The proposed methodology is applied to household-level data for a post-socialist economy – that of Croatia – with the specific aim of testing the extent to which the prolonged economic downturn following the Great Recession of 2008-2009 might hurt indebted households. The paper compares the results based on the new approach with those based on traditional stress testing methods. Interest rate shocks had a stronger impact on household vulnerability in the traditional approach, whereas decreases in employment are found to be more disruptive in the cluster-based approach.

Key words: household financial vulnerability, cluster analysis, stress testing, Croatia

1. Introduction

Prior to the outbreak of the 2008 financial crisis, Eastern European countries experienced rapid lending growth driven by fast economic development, financial liberalization and opening, as well as the convergence of their financial systems towards structures found in Western European economies. Household sector debt increased from less than 10% of GDP in most countries in the region in the early 2000s to more than 30% ten years later. Eurostat data shows that the Baltic countries had the most striking expansion of household debt which reached more than 50% of GDP in the cases of Estonia and Latvia. Somewhat slower growth was recorded in Slovenia and Croatia. However, given their fairly high initial level, both countries have remained characterised by relatively high household debt of around 30% and 40% of GDP at the end of 2000s, respectively.

This rising debt relaxed households' financial constraints, allowing them to frontload some of their consumption on expectations of rapidly growing income. However, it also raised concerns about the potential implications on the household sector’s ability to service its debt and resilience to different financial and macroeconomic shocks. Exchange rate risk has been pending due to high proportion of foreign currency indexed loans. ECB (2010) reports that more than 80% of household debt in Estonia and Latvia, and around 70% in Lithuania and Hungary were indexed to foreign currency, primarily to the euro and the Swiss franc. It appears that these concerns for the financial stability of Eastern European economies were not well articulated in the literature in the prosperous times of the early and mid-2000s.

The outbreak of the financial crisis swiftly focused the bulk of literature on household financial vulnerability and the effects of various shocks on household over-indebtedness, financial distress and welfare. Numerous household stress tests and exercises emerged applying various methods and vulnerability indicators. The majority of testing methods have relied on household vulnerability indicators that suffered from a lack of precisely defined limits separating over-indebted and vulnerable households from those which are financially healthy. Weak identification of vulnerabilities has influenced stress testing results and large variations in apportioned risks have emerged.

In an attempt to avoid the drawbacks of the traditional household vulnerability methodology, this paper presents an alternative three-step approach to financial vulnerability analysis. In the first step, a cluster analysis technique should be employed in order to identify potentially vulnerable segments of indebted households. It uses information contained in different vulnerability indicators and combines them into a single multidimensional indicator. The cluster analysis may overcome the arbitrariness which is usually present when those indicators are used separately. The second step models a cluster-based vulnerability indicator in order to extract information on its determinants, which will in turn be used to simulate the impact of adverse stress test scenarios in the third step in order to determine the short- and medium-term effects of the financial crisis on household financial vulnerability. As an illustration of the proposed concept, this paper provides empirical results in the case of a post-socialist economy, namely that of Croatia.

This paper is organized as follows. The next section gives a brief overview of the related literature and elaborates on the traditionally used household financial vulnerability indicators. Section 3 describes the proposed methodological framework for household vulnerability identification and financial resilience testing. In the empirical part of the paper, Section 4 describes the dataset and explains basic vulnerability measures for Croatia, while Section 5 presents the results of household vulnerability identification and stress testing exercises. Section 6 describes our conclusions.

2. Literature review and key concepts

Literature on household financial vulnerability can be separated into two major lines of research: the "macro" and the "micro" approach. Papers that adopt the "macro" approach employ aggregate economic data in dealing with causes of widespread growth in household indebtedness and its consequences (see for example Girouard et al., 2006 and Dynan, 2009). However, observing aggregate household balance sheets and aggregate data on debt service burdens provides a very rough guidance on actual household vulnerabilities due to potentially large differences between groups of households and possible pockets of particularly vulnerable households.

The second line of research adopted the "micro" approach which made it much more intimately intertwined with actual patterns of household indebtedness and vulnerability and gave it much more attention in recent household indebtedness literature. Papers written in this manner predominantly use data on individual households compiled from household surveys. Their aim is to identify the profile and the distribution of household vulnerabilities, in contrast to the "macro" approach where such information remains unknown. Studies that follow this approach to household vulnerability recognized the negative effect that the relaxation of borrowing constraints had on household balance sheets as it mitigated liquidity restrictions and allowed households to increase spending, while reducing their savings. However, a combination of reduced savings buffers and high debt burden have also made households more susceptible to unanticipated shocks that could lead to over-indebtedness and financial distress. In order to identify financially vulnerable households, the most common approach in this line of literature was to model selected household vulnerability indicators and test for the impact of different shocks on household indebtedness levels and their financial resilience (see Betti et al. 2007; Hollo and Papp, 2007; Herrala and Kauko, 2007; Albacete and Fessler, 2010; Georgarakos et al., 2009).

The various definitions of household financial vulnerability come from three main methodological approaches: i) objective, ii) subjective and iii) administrative approach; although variations on these approaches make the actual number of approaches much larger (see Betti et al., 2007; Beck et al., 2010 for more detailed explanations on some of those measures). The first approach, which observes the so-called objective measures of household financial distress, is based on the idea that households are vulnerable in cases when their indebtedness or debt service ratios exceed a certain arbitrarily set threshold.[1] Sometimes indicators of household consumption relative to income are used rather than debt/repayment ratios, with a high consumption ratio being an indicator of possible financial distress (Betti et al. 2007). The derivative of the objective approach is the concept of the so-called financial margin, which lessens some of the problems arising from the use of arbitrary set thresholds. This household vulnerability indicator has gained much popularity in simulating impacts of various shocks on the ranks of vulnerable households, which has made it an industry standard for stress testing exercises (see Hollo and Papp, 2007; Herrala and Kauko, 2007; Albacete and Fessler, 2010). The financial margin refers to the income reserve that remains after debt service and the household-specific poverty line has been subtracted from the household income. Households with a negative financial margin are usually considered to be vulnerable. However, calculation of a household specific financial margin still does not fully resolve the problem of arbitrary threshold setting by researchers, but rather consigns it to the institution setting poverty lines.

The second approach is subjective in the sense that it relies on a subjective evaluation of household balance sheets and the debt servicing burden. Typically, these measures are based on the number of households reporting a degree of hardship in servicing their debts (Herrala and Kauko, 2007). One of the problems with this indicator is that subjective well-

[1] Ratios of debt to income in the range of 450%-600% and debt repayment to income of 30% are commonly used as vulnerability thresholds (European Central Bank, 2007).

being does not necessarily correlate closely with underlying financial distress, but may be influenced by other factors, such as comparisons with the reference group (Georgarakos et al., 2009).

The third approach is the so-called administrative approach where data on actual bankruptcies or debt defaults is used. As most studies use household survey data, a derivation of this approach uses self-reported debt arrears as an indicator of financial distress (Hollo and Papp, 2007). Sometimes the concept of arrears is expanded to include not only arrears incurred towards financial institutions, but also late payments of certain utilities or other bills such as rent. Also, it is possible to vary the thresholds for arrears from one to several months in order to make the criteria more or less stringent or align it with actual banking practices. This definition is the one that is most closely aligned with the concept of bank losses stemming from household financial distress.

Practically all vulnerability measures suffer from a lack of accurately and objectively defined boundaries separating vulnerable households, while recommended threshold values do not consider the differences in life-cycle stages, income levels, risk tolerance and economic conditions between different households (Greninger et al., 1996). Also, indicators of household financial vulnerability typically lag far behind the actual rise of indebtedness and tend to worsen only following an exogenous shock, with high debt levels exacerbating the impact of such shocks (Rinaldi and Sanchis-Arellano, 2006 and Jappelli et al., 2010). Stress-testing exercises based on simulations of adverse macroeconomic scenarios are used to circumvent the backward-looking feature of most household vulnerability indicators. However, even forward looking stress-tests often yield unrealistic and contradictory results, depending on the exact measure of vulnerability used (see Beer and Schürz, 2011; Fuenzalida and Ruiz-Tagle, 2009; Herrala and Kauko, 2007; Daras and Tyrowicz, 2011 and Beck et al., 2010). More importantly, the most commonly used indicators of household vulnerability are often poorly inter-correlated. The choice of a particular indicator critically influences the results and may share only a distant resemblance to the underlying concept it is meant to describe. It appears that results and conclusions of studies on household financial vulnerability critically depend on properties of particular indicators. A possible solution is aggregation of different vulnerability measures in a single multidimensional indicator. That is not a trivial procedure, since typical aggregation approaches (union and intersection) result in a wide range of potentially vulnerable households. Still, "too little attention has been paid to developing practical alternatives to the union, intersection and unidimensional identification approaches" (Alkire and Foster, 2011).

In the quest for an alternative to the standard line of research, this paper proposes a methodological framework which includes three connected steps, with output from one step used as an input in the subsequent step.

3. Methodology

3.1. Identification of financially vulnerable households

The first step of the proposed methodology aims to identify financially vulnerable households by simultaneously examining the whole set of distinct vulnerability indicators and employing a cluster analysis methodology.[2] The major advantage of cluster analysis is that it alleviates many of the problems arising from arbitrary set thresholds to identifying vulnerable households within each particular indicator. The cluster analysis sets the threshold on the basis of information contained in the raw data. The basic idea is to classify households according to different vulnerability measures so that households within one group are similar

[2] The cluster analysis technique was also used for assessing vulnerability at the macro level in order to determine factors and explain different growth-vulnerability patterns in different countries (Ghosh, Sugawara and Zalduendo, 2011).

to one another and different from households in other group(s), without imposing ad hoc thresholds.

In the present context, the latent class cluster analysis (LCCA) can be seen as preferable to traditional clustering techniques.[3] The LCCA classifies objects into two or more latent classes using model-based posterior membership probabilities estimated by maximum likelihood methods. A statistical model is postulated for the population from which the data is obtained by assuming that a mixture of underlying probability distributions generate the data (Magidson and Vermunt, 2002). The basic latent class cluster model is given by:

$$f(y_i \mid \theta) = \sum_{k=1}^{K} \pi_k \prod_{j=1}^{J} f_k(y_{ij} \mid \theta_{jk}) . \tag{1}$$

Here y_{ij} denotes the value of the particular vulnerability indicator j for household i. These observed variables are usually called indicator variables. K is the number of latent classes, i.e. clusters, where $k = 1, \cdots, K$. π_k is the prior probability of belonging to class k and $f(y_{ji} \mid \theta_{jk})$ is a class specific density of vulnerability indicator y_{ij} given model parameters θ_{jk}.

The advantage of the latent class clustering approach compared to more traditional methods is that it enables grouping of objects based on the indicator variables of different scale types. The classical latent class clustering analysis assumes local independence between indicators within each cluster. Local independence assumption essentially implies that the correlation among the observed indicator variables is explained by the latent class variable. This means that the indicator variables are mutually independent so there is no residual

[3] Although various types of clustering algorithms appear in empirical studies, the most popular and commonly used are k-means and hierarchical clustering. These traditional approaches to cluster analysis are quite simple and repose on the construction of some type of distance measurement between objects that are being clustered usually by optimizing certain criteria: minimizing within-cluster variation or maximizing between -cluster variation.

correlation between the indicator variables within the cluster (Clark and Muthen, 2009). However, the local independence assumption can be relaxed by including the direct effect for the pair of indicator variables in the model in order to achieve better grouping of households.[4] Another important advantage of latent class clustering analysis is that indicator variables do not need to be standardized before classification. Also, there are more formal criteria for making decisions about the number of clusters and other model features (Vermunt and Magidson, 2002).

Estimation of the model parameters is not the primary objective of latent class cluster analysis, but classification of households into different groups on the basis of the estimated model. Since this is a probabilistic clustering, households are classified into different classes with specific uncertainty. Therefore, the true class membership of households always remains unknown and every classification implies some degree of classification error. Higher separation between the classes lowers the total classification error of the model. Classifications are based on the posterior probability of belonging to cluster k given a household's observed values of indicator variables y_{ij} (Vermunt and Magidson, 2002):

$$\pi_{k|y_i} = \frac{\pi_k \prod_{j=1}^{J} f_k\left(y_{ij} \mid \theta_{jk}\right)}{\sum_{k=1}^{K} \pi_k \prod_{j=1}^{J} f_k\left(y_{ij} \mid \theta_{jk}\right)} \tag{2}$$

Households can be assigned to different classes based either on the modal rule, when every household is assigned to the class for which estimated posterior probability is the largest, or on the proportional rule, whereby households are assigned to classes with a certain

[4] Dependence of two variables within a cluster implies that there is probably some overlapping information that should be left out when classifying households into different clusters. If such a correlation between indicator variables is ignored, locally dependant indicator variables will have a larger influence on classification compared to other indicator variables. Relaxation of the local independence assumption usually prevents over clustering and enables better classification (Vermunt and Magidson, 2002).

weight, i.e. probability. The modal assignment rule is preferred in this research due to smaller classification error compared to other methods.

3.2. Binary dependant variable model for financial vulnerability

The impact of different determinants on a household's probability of belonging to one cluster relative to the other should be assessed by a binary dependant variable model in the second step of the proposed methodology. A cluster-based vulnerability indicator (z_i), that equals 1 if a household is assigned to the high vulnerability cluster and used as the dependant variable in the logistic regression model of the form:

$$\Pr_i^d\left(z_i \mid x_i, \beta\right) = 1 - \left(\frac{e^{-x_i\beta}}{1+e^{-x_i\beta}}\right) \tag{3}$$

Here x_i denotes a vector of different socio-economic and demographic characteristics of household i. Based on the estimated model parameters and different characteristics, the probability of being vulnerable is assessed for every indebted household in the sample.

3.3. Stress testing simulations

The identified determinants of household vulnerability can be used to assess the possible impacts of different macroeconomic and financial shocks on household financial resilience. This can be done by a more or less standard stress testing exercise that aims to simulate some likely shocks in the particular economy. In the context of small open economies in Eastern Europe facing the 2008-2009 recession and its aftermaths, it is reasonable to focus on labour markets, financial markets and exchange rate shocks.

Empirical tests conducted in this paper include in particular: i) decreases in employment (and the corresponding decline in disposable household income), ii) depreciation of local currency, iii) the rise in lending rates of banks, iv) a combination of employment shock and exchange rate shock and v) a combination of employment shock and interest rate shock. Although all of the first three shocks can occur simultaneously, for the sake of simplifying our presentation simulations include combinations of only two shocks of various intensities. The intensity of shocks is calibrated based on the likely movements of macroeconomic and financial markets.

4. Data and indicators

Illustration of the proposed cluster-based approach is performed on the micro data from the Household Budget Survey (HBS) in Croatia. The HBS has been regularly conducted by the Central Bureau of Statistics on a random sample of private households. Apart from the household-level data on expenditures, income, wealth and housing conditions, the HBS also gives insight into socioeconomic and demographic characteristics of surveyed individuals, allowing for indebtedness analysis based on the characteristics of household members. The data for 2008, 2009 and 2010 is used. The analysis is performed on the subsample of indebted households, regardless of the loan type.

Croatia is a post-socialist country with a bank-centric financial system which is relatively well developed by the standards of the region. Although some measures were imposed by the central bank in an attempt to slow the pre-2008 pace of credit extension (Ljubaj, 2012), it nevertheless progressed strongly. Private sector credit grew by around 17% annually between 2003 and 2008. The Great Recession hit Croatia a bit later than the developed world. September of 2008 is usually taken as the start of the recession due to the sudden drop in industrial production and exports. Banking sectors proved to be quite strong at that time due to previously undertaken monetary policy and macro prudential measures aimed at limiting credit expansion. There was no need for government to be involved in the operation of the financial sector. However, due to a relatively high level of household debt cumulated by the mid-2000s with a structure of bank borrowing that was tilted towards debt with high exposure to exchange rate and interest rate movements, risks originating from the household sector have been rising in the recession that has been ongoing up to 2012.[5]

The HBS data indicates that the characteristics of indebted households remained quite stable between 2008 and 2010 in spite of the recession (Table A1 in the Appendix). Banks were most often granting loans to middle-aged, married males who owned a home in the urban environment and on average had one child. He had a high-school degree, worked in a private or a public company dealing in the tertiary sector of economic activity and had a permanent working contract with full-time working hours.

For the purpose of household clustering regarding financial vulnerability, different measures are considered (Table 1). The share of vulnerable households significantly differed depending on the indicator. In 2010, it went from 3.0%, if assessed as debt in excess of 500% of income, to 29.2%, if based on the households' self-assessment of the financial situation.

Table 1 Financial vulnerability indictors for Croatia

Percentage of the indebted households that are vulnerable according to:	2008	2009	2010
Negative financial margin	15.9	19.9	22.0
Debt in excess of 500% of income	2.5	1.2	3.0
Repayments in excess of 30% of income	13.1	15.1	18.6
Very difficult financial situation	6.2	8.2	7.3
Difficult and very difficult financial situation	26.3	27.6	29.2
Non-performing loan ratio	4.0	5.8	7.8

Sources: Authors' calculations based on HBS and Croatian National Bank for non-performing loan ratio.

5 Croatian GDP fell by 6.9% and 1.4% in 2009 and 2010, respectively. In 2011 there was stagnation, while the first half of 2012 brought another decline of GDP by 1.7%.

Based on their characteristics and distributional features, three vulnerability indicators were chosen for clustering purposes: two continuous objective vulnerability indicators (debt repayment burden and the so-called adjusted debt repayment burden) and one nominal subjective indicator (self-reported financial situation).[6] The debt repayment burden is calculated as the ratio of household loan repayments to disposable income. It is the most often used measure of household over-indebtedness and "the best measure of the ability to handle debt" (Greninger et al, 1996). The adjusted debt repayment burden is essentially derivative of the financial margin indicator since, apart from the amount of loan instalments, it also takes the household-specific poverty line into account. The indicator of a household's subjectively perceived financial situation is a type of Likert item that identifies six degrees of hardship where value 1 is assigned to households whose reported self-financial situation is very difficult and value 6 to households with very good financial position.

5. Empirical results

5.1. Identification of financially vulnerable households

Two clusters of indebted households, a low and a high vulnerability group, were identified based on the LCCA method. Since there is no panel component within the HBS, data from all three years is pooled together for classification purposes.[7] Classification with

[6] The measure of household indebtedness, defined as the ratio of the total amount of household debt to annual disposable income, was not used due to its bias towards households with higher loan amounts that are typical of housing loans. Relatively low coverage of the total loan amount with disposable income does not necessarily indicate that the household has troubles with loan repayments or satisfying a minimal living standard, since a longer period of repayment, lower interest rates due to quality collateral, grace periods and other lending conditions can ease a household's financial burden. Debt arrears of different types were also left out from the analysis due to Croatian HBS data limitations that preclude the distinction between households that are really falling behind on their loan repayment plan and households who have paid less than twelve installments in a year due to the expiration of their loan obligation.

[7] Classification of households based on the pooled sample in the period that covers pre-crisis and crisis years is preferred over estimation year by year due to better assessment of vulnerability in absolute sense. If each year would be observed separately then in 2010, for example, due to worse financial position of all households, some households wouldn't be classified as vulnerable because they would be compared only to the indebted household in that year, relative to whom they are better off. However, if those households are compared to households from 2008, they would be identified as vulnerable.

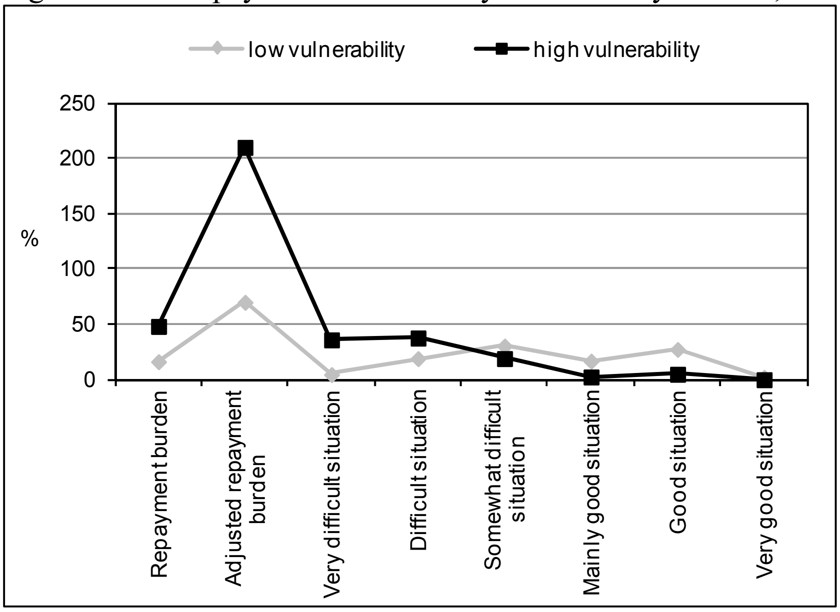
only two clusters was chosen due to the lowest classification error compared with other specifications and its ease of interpretation and appeal with regard to the basic idea and goal of this research, i.e. separation of potentially vulnerable households from those that are not (Table A2 in Appendix). The influences of all three vulnerability indicators on classification are in line with expectations, and discrimination between the clusters is statistically significant. The classification model best explains differences in the level of the adjusted debt repayment burden indicators between households, while differences in the debt repayment burden and subjective opinions regarding the financial situation are accounted for by the model to a much smaller degree.

Households assigned to the high vulnerability group during this period coped with much higher average loan repayment burdens compared to the low vulnerability group (Figure 1). The difference in the financial burden between these two groups was even larger when minimum living costs were combined with the household's loan instalments. Additionally, households from the high vulnerability class had a much higher probability of perceiving their financial position as hard or very hard.

A comparison of household characteristics from the two groups (Table 2) shows that even though the loan instalment of highly vulnerable households is on average lower compared to the debt repayments of households from the low vulnerability group, these households also have lower income available for servicing their loan obligations. Also, representation of households headed by a female is somewhat higher in the high vulnerability group as well as families with more children. In addition, the probability of living in rural areas is much higher for highly vulnerable households. In all three observed years there is a much higher probability that the head of the highly vulnerable household was unemployed and even if she/he was employed, the probability that she/he was working in a public company is considerably smaller.

Separated by period, the share of vulnerable households increased from 5% in 2008 to 9% in 2009 and 11% in 2010, with an even faster growth of the share of total household debt held by vulnerable households (from 7% to almost 15%).

Figure 1 Debt repayment indicators by vulnerability clusters, Croatia, 2008-2010



Note: Debt repayment indicators are average values by clusters. Subjective opinion indicator presents percentage of households in each cluster with a certain perception of their financial difficulties.
Source: Authors' calculations based on HBS.

Table 2 Household profile by vulnerability clusters, Croatia, 2008-2010

	Pooled sample 2008-2010									
	Low vulnerability					High vulnerability				
	Mean	Min.	Max.	Median	Std. Dev.	Mean	Min.	Max.	Median	Std. Dev.
REPAYMENT BURDEN	0.16	0.00	0.58	0.14	0.11	0.48	0.00	3.35	0.36	0.50
ADJUSTED REPAYMENT BURDEN	0.70	0.16	1.60	0.67	0.24	2.10	0.86	16.04	1.61	1.85
SUBJECTIVE OPINION	3.50	1	6	3	1.26	2.02	1	5	2	1.04
LN(DISPOSABLE INCOME)	11.52	9.86	13.25	11.54	0.50	10.48	8.16	12.57	10.55	0.68
LN(LOAN INSTALMENT)	9.42	5.58	12.10	9.57	0.96	9.29	4.80	12.31	9.39	1.29
AGE	52.59	20	93	52	12.65	51.99	21.00	84.00	52.00	12.01
NUMBER OF CHILDREN	0.74	0	5	0	1.00	0.79	0	6	0	1.11
			%					%		
HOUSING LOAN			20.40					20.65		
EDUCATION LOW			21.43					38.87		
EDUCATION HIGH			18.04					4.86		
WORK IN PUBLIC COMP.			25.27					6.88		
ENTERPRENEUR			4.02					4.05		
WORK IN OTHER COMP.			7.27					13.77		
UNEMPLOYED			35.89					56.68		
WOMAN			25.93					31.17		
RURAL AREA			42.60					57.49		

Source: Authors' calculations based on HBS.

5.2. Binary dependant variable model

Following the identification of highly vulnerable households, the next step is modelling of the impact of different household characteristics on the probability of being vulnerable. The binary dependant variable in the logistic regression model equals 1 if the household is assigned to the high vulnerability cluster and 0 otherwise. Different socio-economic and demographic characteristics of household are used as explanatory variables. The estimated coefficients and marginal effects are shown in Table 3.[8]

Table 3 Logistic regression coefficients and marginal effects

	Pooled sample 2008-2010			
Variable	**Coefficient**		**Marginal eff.**	
ln(REAL DISPOSABLE INCOME)	-6.1243	***	-0.038236	***
ln(LOAN INSTALMENT)	1.5794	***	0.009861	***
AGE OF HOUSEHOLD HEAD	0.1742	***	0.001088	**
AGE-SQUARED	-0.0019	***	-0.000012	**
NUMBER OF CHILDREN	0.7229	***	0.004514	***
MINIMUM LIVING COSTS	0.0002	*	0.000001	*
HOUSING LOAN DUMMY	0.5479	**	0.004062	*
EDUCATION (REF. EDUCATION_MIDDLE)				
EDUCATION_LOW	0.1812		0.001189	
EDUCATION_HIGH	0.2965		0.002046	
EMPLOYMENT STATUS (REF. EMPL. IN PRIVATE COMPANY)				
EMPL. IN PUBLIC COMP.	-0.3493		-0.002002	
ENTERPRENEUR	1.2461	**	0.014502	
EMPL IN OTHER COMP.	0.4649		0.003552	
UNEMPLOYED	0.5321	*	0.003586	*
WOMAN	-0.2574		-0.001517	
RURAL AREA DUMMY	0.3041		0.001941	
C	39.6757	***		
Obs with Dep=0	2711			
Obs with Dep=1	247			
Total obs	2958			
McFadden R-squared	0.56			

Notes: * significant at 10% level, ** significant at 5% level and *** significant at 1% level.
Source: Authors' calculations based on HBS.

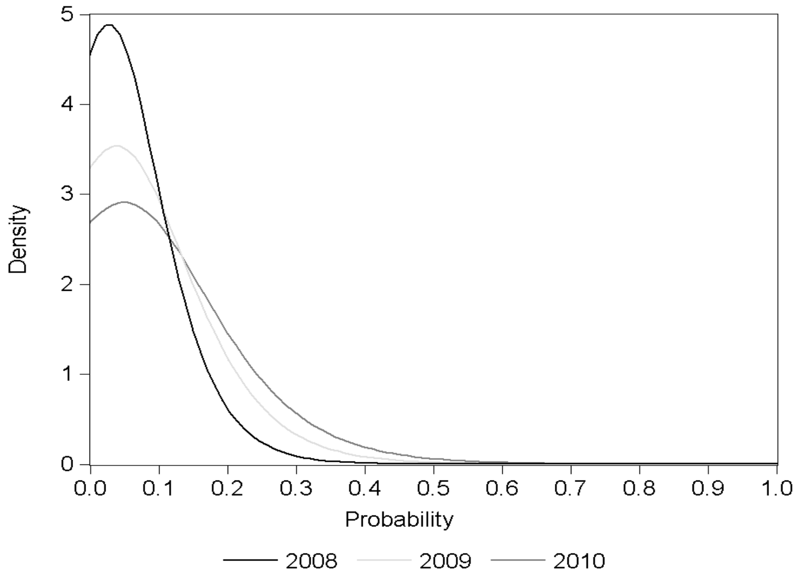
[8] Since the dependant variable in the logistic regression shows the most likely class membership of the household, which necessarily involves some degree of uncertainty and possible misclassification, and not true class membership, estimated logistic coefficients could be downwardly biased. However, due to very good household separation, expressed through a relatively low classification error and high entropy R-squared, and a relatively large sample size, this shouldn't pose a considerable problem in this research, although it needs to be borne in mind when assessing the statistical significance of different households' characteristics on the probability of being vulnerable.

Statistically significant influences on a household's probability of being financially vulnerable are found for a household's real disposable income, the most liquid source of loan repayment, the amount of the loan instalment, the age of the head of the household, the number of children, housing loan repayments and minimum living costs.[9] The signs of the regression coefficients are in line with theoretical expectations. Thus higher loan instalments, higher minimum living costs, more children in the family, a housing loan and an older head increases the probability that the household will be vulnerable. Higher real disposable income has the opposite effect. The statistically significant impact of age and age-squared of household head on the probability of being financially vulnerable confirms the life cycle-permanent income hypothesis according to which the household's consumption and therefore also borrowing function has an inverted U-shaped form, indicating that households increasingly borrow until a certain age is reached when their income peaks and after which their borrowing needs gradually decrease. A household whose head is an entrepreneur or is unemployed has a significantly higher probability of having financial difficulties compared to the reference household whose head is employed in a private company, while households whose head works in a public company where wages are less volatile have lower probability of being vulnerable. For education, gender and location we have not found significant effect on households' vulnerability.

A comparison of the estimated probability distributions for 2008, 2009 and 2010 (Figure 2) shows deteriorated household financial positions after the outbreak of the crisis in 2008. There is a slight increase in the ranks of households with the highest probabilities of being financially vulnerable in 2009 and 2010.

[9] Other binary dependent variable models (probit, LPM) are also estimated, but no significant differences in results are observed.

Figure 2 Regression-based probabilities of being financially vulnerable, Croatia 2008-2010



Source: Authors' calculations based on HBS.

In order to prepare the data set for the stress testing exercise, households are again divided into two vulnerability groups based on the estimated probabilities of being financially vulnerable. After assessing the probability of default based on the estimated logistic regression for every indebted household, a threshold probability that separates higher from lower vulnerability households is determined. Although different threshold values are considered, the probability used as a boundary was calibrated based on the estimated classification model.[10] The number of vulnerable households identified by the logistic model is taken to be the same as in the estimated LCCA model. As a result, the model indicates that 5.2% of the indebted households were financially vulnerable in 2008, 9.0% in 2009 and 10.7% in 2010. At the same time debt held by financially vulnerable households, taken as a percentage of overall household debt, increased from 4.0% in 2008 to 5.3% and 9.4% in 2009 and 2010, respectively.

[10] A possible choice would be a threshold based on probability that simultaneously maximizes the true positive rate (also referred to as sensitivity) and the true negative rate (also referred to as specificity) of the estimated logistic regression model. This is the usual approach when choosing a probability threshold based on a binary dependant variable model. However, it is not the most appropriate approach if the sample is unbalanced and different "costs" are associated with the true detection of objects from different groups (violence of the assumption of an equal cost of misclassifications (Chen et al., 2006)).

5.3. Stress testing simulations

The stress testing exercise is prepared in such a way as to simulate actual or expected developments in the years 2011 and 2012, with details of shocks designed to mirror the situation in the country as closely as possible. It is assumed that there was no change in the total amount of loans granted to the household sector during this two year window. All loans coming due during this period are assumed to have been successfully refinanced in accordance with prevailing market conditions.

In micro-simulations based on survey data, three major shocks are tested - labour market, interest rate and exchange rate shock. The labour market shock involves a random selection of employed individuals who become unemployed, while their labour income is replaced by a certain amount of unemployment benefits based on average salary and work experience according to the eligibility criteria of the current regulations. The impact of the rise in interest rates on the amount of annual loan payments is determined from data on outstanding principal amounts and it is applied to every loan in the data sample. This is because a large majority of total household loans in Croatia are actually granted with interest rates adjustable within one year, as reported by Croatian National Bank (2011). Household loan portfolios are divided into housing loans and other loans to which different interest rates apply. The simulated exchange rate shock takes into account changes in the Croatian kuna/euro and Croatian kuna/Swiss franc exchange rate. Households whose loans are affected by the exchange rate shock, as opposed to households with kuna nominated loans that are not affected by the shock, were randomly chosen. The impact of the shock was assessed based on the results of one thousand repetitions of the simulation, just like the impact of the employment shock. The exchange rate shock is designed to affect 83% of the loans in the sample, out of which 67% are indexed to euro exchange rate movements and 33% are indexed to Swiss franc exchange rates. These proportions are actually recorded in monetary statistics for 2010. The impact of simulated shocks on household vulnerability is approximated by the share of households classified as vulnerable and the share of their debt in terms of total household debt to the banking sector (exposure at default, EAD).

The results of stress tests based on the model for cluster-based indicators are compared with a conventional household stress testing framework where the impact of adverse shocks is estimated on the basis of the household financial margin, i.e. income reserves available to a household after subtracting the minimum living costs and the amount of loan repayments from household disposable income.[11] If disposable income is not sufficient for loan repayment and satisfying minimum living standard, the household is classified as vulnerable.

In the proposed cluster-based approach, simulated shocks should affect a household's probability of having financial difficulties through the value of explanatory variables in the binary dependant variable model, a household's disposable income and the annual amount of loan repayments. A decrease in employment, for example, should reduce a household's disposable income, while interest rate growth and weakening of the kuna exchange rate should increase the amount of loan instalments. Simulations also take into account projected inflation rates and changes in minimum living costs approximated by the year's specific at-risk-of-poverty threshold[12].

The simulated scenario for 2011 combines all three shocks whose intensities were calibrated according to actual movements in macroeconomic and financial variables. For 2012, an array of intensities of these three shocks is simulated, some of which are more likely than others, but all of which are considered plausible. Simulations include employment

[11] Minimum living costs are derived from the at-risk-of-poverty threshold, defined as 60% of median equivalised household income in given year. For households with a certain number of members, the minimum cost is calculated by multiplying the poverty threshold for a one-person household by the equivalent household size. That size is determined by so-called modified OECD equivalence scale where the household head is given the coefficient 1, every other adult aged 14 and over is given the coefficient 0.5, and every child under 14 years of age is given the coefficient 0.3.

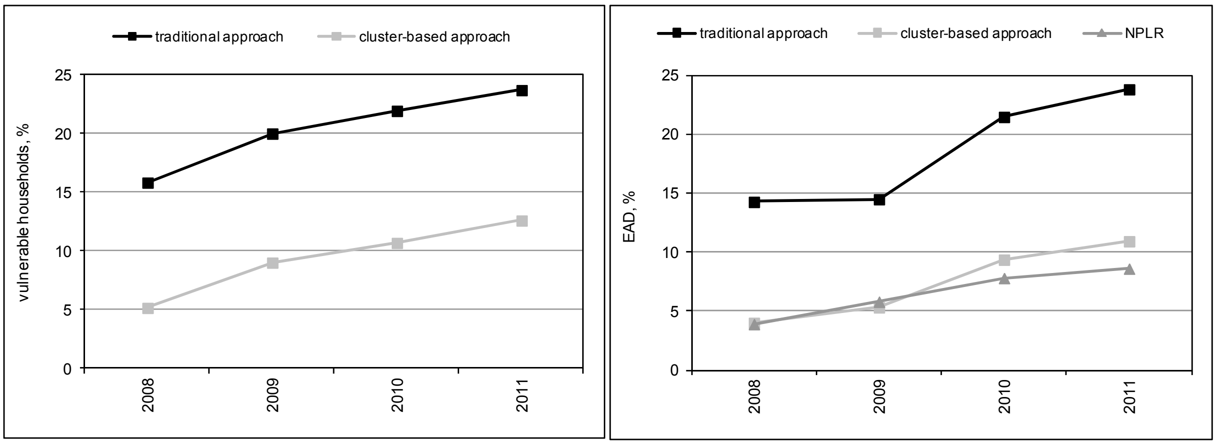
[12] Projection details are available upon request.

decreases ranging from 1% to 5%, a weakening of the kuna exchange rate of between 1%-20% and increases in the lending rates of banks between 1 and 5 percentage points.

5.3.1. Baseline scenario for 2011

Baseline scenario for 2011 combines the impact of all three shocks that already hit the economy, but their impact on household financial vulnerability has not been fully quantified in the official statistics due to data collection and data publication lags. The proposed simulation can help in getting a good sense of vulnerabilities before the official data has been released. In this particular case, simulation includes a drop in employment of 1.2%, the growth of the average net wage by 4.1% in nominal terms, a decrease in interest rates for housing loans by 17 basis points and those for other loans by 11 basis points, and 2.0% depreciation of the Croatian kuna against the euro and 14.2% against the Swiss franc.

Figure 3 Baseline scenario results: proportion of vulnerable households and proportion of debt held by vulnerable households



Sources: Authors' calculations based on HBS and Croatian National Bank for non-performing loans (NPLR).

Simulation results based on the cluster-based vulnerability indicator indicate an increase in the proportion of highly vulnerable households to 12.7% in 2011, whereas their debt amount increases to 10.7% of the total household sector debt. For comparison, simulations based on the traditional approach indicate much higher levels of household financial weakness (Figure 3). Both approaches show that household financial vulnerability continued to rise in 2011 but at a somewhat slower pace than in 2010.[13] The results from both the traditional and the new methodology are in line with dynamics in banks' exposure to vulnerable households measured by the non-performing loan ratio (NPLR), for which 2011 statistics are available due to a much shorter data collection lag. However, the cluster-based approach follows the absolute levels of bad loans that are classified as such in bank accounts much more closely.

5.3.2. Simulations of shocks for 2012

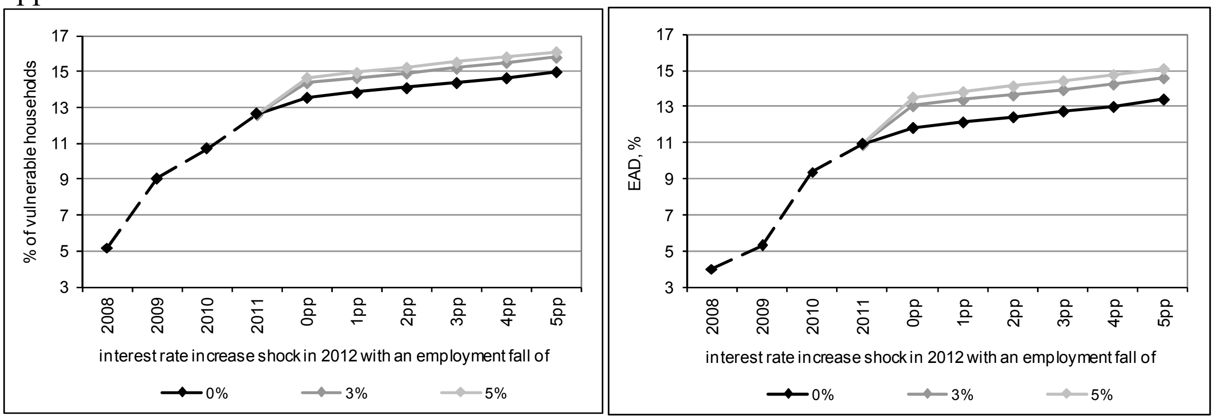
The baseline scenario for 2011 provides the starting point to test the impact of the 2012 shocks. For the sake of simplicity, here are presented only the results of simulations of two shocks at a time, one macroeconomic and one financial. The effect is non-linear and crucially depends on the exact combination and magnitude of the tested shocks. Simulated combinations of shocks in the cluster-based approach show that for any tested decrease in employment, growth in the interest rate by one percentage point has the same effect on household vulnerability as the weakening of the kuna exchange rate by approximately 4% (Figures 4 and 5). In the "worst case scenarios" these shocks result in increase in percentage of vulnerable households from 13% in 2011 to 16%-17% in 2012, while their debt increase from 11% to approximately 15% of total household debt.

In the traditional approach, simulated combination of shocks implies a weaker effect from decreased employment and the stronger influence of the interest rate and exchange rate

[13] A closer look at the contribution of the particular shocks on the household vulnerability increase in 2011 shows that the weakening of the kuna exchange rate, especially against the Swiss franc, has had the largest effect on both testing approaches, while worsening labor market conditions were of secondary importance. Interest rates fell in 2011 and thus had no negative contribution to household vulnerability.

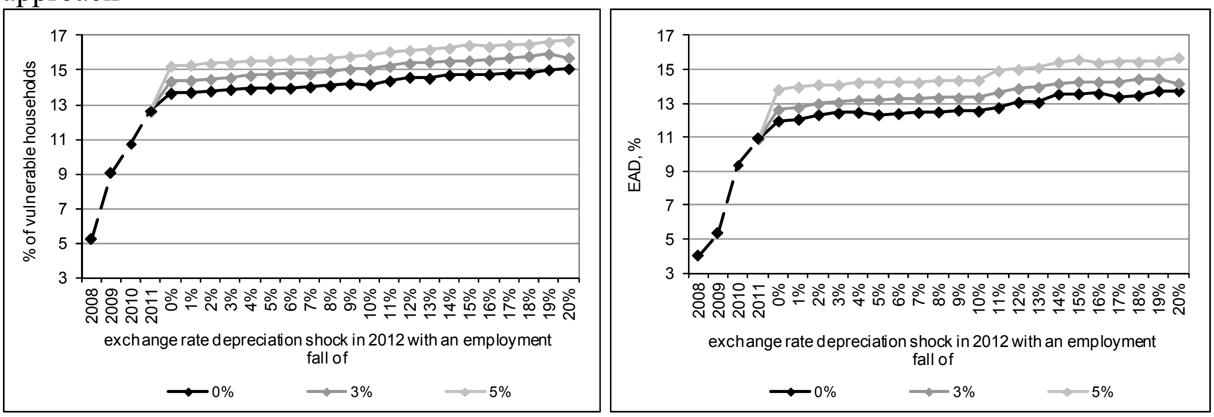
changes than in the cluster-based approach (Figures 6 and 7). The effect of an interest rate shock for a given level of employment is twice as strong as that in the cluster-based approach.[14] A 5% drop in employment together with a simultaneous rise in the amount of loan repayments due to an interest rate increase of 5 percentage points lead to an increase in the percentage of vulnerable households to approximately 27% in 2012, with their debt reaching almost 29% of total household sector debt. A similar result is found in the combination of a 5% drop in employment and depreciation of the kuna exchange rate by 20%.

Figure 4 Effects of combined employment and interest rate shocks in the cluster-based approach



Sources: Authors' calculations based on HBS

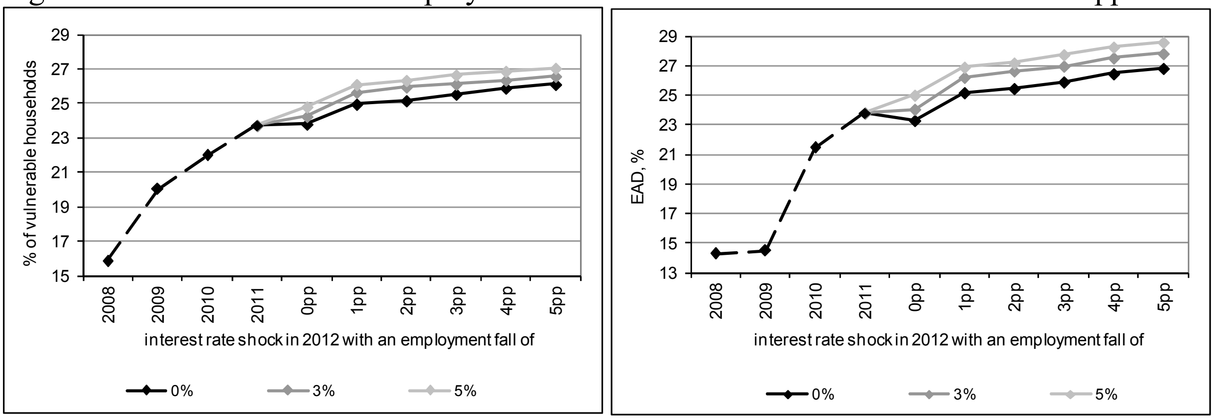
Figure 5 Effects of combined employment and exchange rate shocks in the cluster-based approach



Sources: Authors' calculations based on HBS

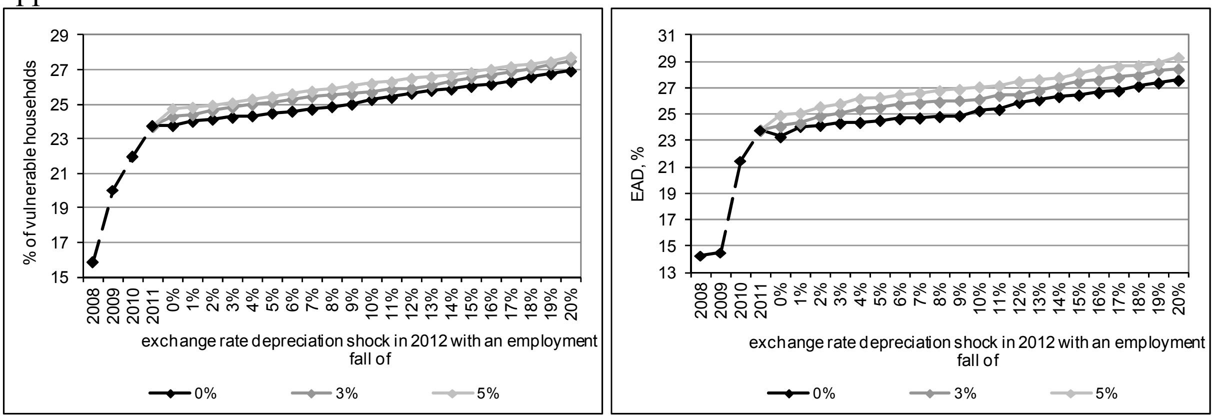
[14] The results of isolated shocks (one shock at a time) confirm that the effect of the employment shocks on the proportion of vulnerable households is stronger in the cluster-based approach than in the traditional approach. The effects of the simulated interest rate and exchange rate shocks are found much stronger in the traditional approach for both the proportion of vulnerable households and the share of their debt.

Figure 6 Effects of combined employment and interest rate shocks in the traditional approach



Sources: authors' calculations based on HBS

Figure 7 Effects of combined employment and exchange rate shocks in the traditional approach



Sources: authors' calculations based on HBS

Both approaches suggest that household vulnerability in Croatia has grown since the financial crisis broke out. The simulations based on likely macroeconomic assumptions indicate that the trend of this increase might have been slowing down in 2011 and 2012.

6. Conclusion

This paper presents an alternative approach to household vulnerability measurement and stress testing. This approach combines information contained in various vulnerability indicators and takes advantage of interactions between them by applying the latent class clustering technique. The resulting cluster based vulnerability indicator is used for assessing the effect of different socio-economic and demographic characteristics of households on their probability of being financially vulnerable. Such a multidimensional approach to vulnerability identification and measurement implemented in stress testing is expected to contribute to better assessments of household financial vulnerability by reducing reliance on arbitrariness.

Application of the proposed cluster-based approach on the actual HBS data for Croatia showed that the outbreak of the financial crisis in late 2008 and its spillover to the real economy in 2009 and 2010 has significantly increased household financial vulnerability. Further deterioration of macroeconomic conditions in 2010, particularly in the labour market, apparently continues to impair household financial resilience. Simulations based on a realistic set of assumptions for Croatia for 2011 and 2012 suggest that household vulnerability continues to grow, although at a somewhat slower pace, primarily due to the combined effects of the weakening of the kuna exchange rate and decreased employment. The expected employment and income drop in 2012, combined with permanent high exposure to interest and exchange rate risks, is likely to further increase the share of vulnerable households.

Comparison of two household stress testing approaches showed effects of differing relative importance for household financial vulnerability. Whereas rising interest rates had a larger impact on vulnerability in the traditional approach, the employment shock proved to be more disruptive for household financial resilience following the proposed cluster-based approach.

This newly presented methodological approach to household stress testing could be further extended and improved in several ways: by including additional explanatory variables in the logistic regression model, estimating the semi-parametric model of household vulnerability, considering other approaches to determining a probability threshold that separates vulnerable households from those that are financially healthy, improving the estimation of the minimum living costs in a specific year and testing for the impact of some other shocks, such as expected cuts in wages in the public sector and other austerity measures.

The proposed methodological framework should help to provide a more complete picture of household financial fragility and enable researchers to estimate fragility depending on various economic scenarios. It may provide a toolkit for policy makers to test the impact of different economic and monetary measures on household vulnerability and consequently financial stability, while also reduce the methodological drawbacks of the traditional household vulnerability measures by taking into account household life cycle stages, income levels, risk tolerance and financial literacy.

References

Albacete, N., and P. Fessler. 2010. "Stress Testing Austrian Households." Financial Stability Report, Oesterreichische Nationalbank (Austrian Central Bank), no. 19: 72-91.

Alkire, S., and J. Foster. 2011. "Understandings and Misunderstandings of Multidimensional Poverty Measurement." *Journal of Economic Inequality* 9, no. 2, 289-314.

Beck, T., K. Kibuuka, and E. Tiongson. 2010. "Mortgage finance in central and eastern Europe opportunity or burden?" Policy Research Working Paper Series 5202, The World Bank.

Beer, C., and M. Schürz. 2011. "Characteristics of household debt in Austria-Does household debt pose a threat to financial stability." *Monetary Policy & the Economy*, no 2: 58-79.

Betti, G., N. Dourmashkin, Rossi, M., and Y.P. Yin. 2007. "Consumer over-indebtedness in the EU: measurement and characteristics" *Journal of Economic Studies* 34, no.2: pp. 136-156.

Chen, J. J., C. A. Tsai, H. Moon, H. Ahn, J.J. Young, and C.H. Chen. 2006. "Decision threshold adjustment in class prediction." SAR and QSAR in Environmental Research 17(3)

Clark, S. L., and B.O. Muthen. 2009. "Relating latent class analysis results to variables not included in the analysis." available from http://www.statmodel.com/download/relatinglca.pdf.

Croatian National Bank. 2011. Financial Stability Report, no. 7, CNB Publications.

Daras, T., and J. Tyrowicz. 2011. "Breeding ones' own subprime crisis: The labour market effects on financial system stability." *Economic Systems* 35, no. 2: 278-299

Dynan, K. E. 2009. "Changing Household Financial Opportunities and Economic Security." *Journal of Economic Perspectives* 23, no. 4: 49–68.

European Central Bank. 2007. EU Banking Sector Stability, November 2007.

European Central Bank. 2010. Financial Stability Review, June 2010.

Fuenzalida, M., and J. Ruiz-Tagle. 2009. "Households' Financial Vulnerability", Working Papers Central Bank of Chile 540, Central Bank of Chile.

Georgarakos, D., A. Lojschova, and M.E. Ward-Warmedinger. 2009. "Mortgage Indebtedness and Household Financial Distress." IZA Discussion Paper no. 4631.

Ghosh, S., N. Sugawara, and J. Zalduendo. 2011. "Finding a Balance between Growth and Vulnerability Trade-Offs: Lessons from Emerging Europe and the CIS." World Bank Policy Research Working Paper Series 5592, The World Bank, Washington D.C.

Girouard, N., M. Kennedy, and C. André. 2006. "Has the rise in debt made households more vulnerable?" OECD Economics Department Working Papers, no 535, OECD, Paris, France.

Greninger, S. A., V. L. Hampton, K.A. Kitt, and J.A. Achacoso. 1996. "Ratios and Benchmarks for Measuring the Financial Well-Being of Families and Individuals", *Financial Services Review* 5, no. l: 57-70

Herrala, R., and K. Kauko. 2007. "Household loan loss risk in Finland - estimations and simulations with micro data." Bank of Finland Research Discussion Paper, no 5, Bank of Finland, Helsinki, Finland.

Holló, D., and M. Papp. 2007. "Assessing household credit risk: evidence from a household survey." MNB Occasional Papers 2008/70, Magyar Nemzeti Bank (The Central Bank of Hungary), Budapest, Hungary.

Jappelli, T., M. Pagano, and M. di Maggio. 2010. "Households' Indebtedness and Financial Fragility." CSEF Working Papers 208, Centre for Studies in Economics and Finance (CSEF), University of Naples, Italy.

Ljubaj, I. 2012. "Estimating the Impact of Monetary Policy on Household and Corporate Loans: a FAVEC Approach", Croatian National Bank Working Paper W-34, Croatian National Bank, Zagreb, Croatia.

Magidson, J., and J.K. Vermunt. 2002. "Latent class models for clustering: A comparison with K-means." *Canadian Journal of Marketing Research*, 20: 36-43.

Rinaldi, L., and A. Sanchis-Arellano. 2006. "Household debt sustainability - What explains household non-performing loans? An empirical analysis." Working Paper Series 570, European Central Bank.

Vermunt, J.K., and J. Magidson. 2002. "Latent class cluster analysis." in: J. Hagenaars and A. McCutcheon (eds.), Applied latent class analysis, 89-106. Cambridge, UK: Cambridge University Press.

Appendix

Table A1 Descriptive statistics

Variable	2008	2009	2010
		Mean (Std. dev.)	
Disposable income (in HRK)	104,640.06	111,142.98	106,588.73
	(56,662.26)	(57,749.14)	(58,659.16)
Loan (in HRK)	114,802.90	115,130.28	134,093.35
	(153,778.62)	(133,419.14)	(145,417.71)
Loan repayment (in HRK)	16,250.27	18,162.28	20,126.46
	(15,704.19)	(14,728.09)	(19,641.55)
No. of employed members	1.58	1.63	1.53
	(1.05)	(1.02)	(1.02)
No. of children	0.74	0.78	0.72
	(0.98)	(1.04)	(1.00)
No. of loans	1.35	1.32	1.23
	(0.67)	(0.66)	(0.50)
		% of total	
New housing loan	4.39	1.16	0.91
Existing housing loan	31.70	17.27	12.53
Characteristics of household head			
<30 years old	2.49	3.24	2.93
30-39 y.o.	10.57	14.60	12.35
40-49 y.o.	27.82	27.11	24.52
50-59 y.o.	29.71	29.08	31.75
>60 y.o.	29.41	25.96	28.45
Male	74.18	75.67	71.55
Female	25.82	24.33	28.45
Homeowner	89.23	87.85	90.21
Tenant	10.77	12.15	9.79
Single	4.09	4.28	5.95
Widow	13.36	12.73	14.82
Married	76.77	78.13	73.47
Separated	5.78	4.86	5.76
Education_low	20.84	22.60	25.00
Education_middle	60.92	61.41	58.52
Education_high	18.25	15.99	16.48
Employee_public sector	24.33	22.25	24.36
Employee_private company	25.72	30.82	24.63
Entrepreneur	4.59	3.82	3.66
Other_employment	8.37	6.14	8.61
Other_doesn't work	36.99	36.96	38.74
Sector of economic activity_primary	9.17	6.13	15.67
Sector of economic activity_secundary	20.84	23.61	29.10
sector of economic activity_tertiary	69.99	70.25	55.22
Working contract_permanent	97.11	96.41	90.66
Working contract_determinante	1.40	2.31	9.34
Working contract_others	1.50	1.27	
Working time_full-time	91.33	93.52	86.42
Working time_part-time	3.89	2.31	6.27
Working time_longer than full-time	4.79	4.17	7.31
Rural area of residence	41.48	42.41	47.16

Source: Authors' calculations based on the HBS.

Table A2 Estimated latent class classification model

Pooled sample 2008-2010			
	Parameter	p-value	R2
Debt repayment burden	0.144	0.000	0.18
Adjusted debt repayment burden	0.632	0.000	0.28
Financial position		0.000	0.02
very difficult	1.603		
difficult	0.951		
somewhat difficult	0.400		
mainly good	-0.266		
good	-0.256		
very good	-2.431		
	Statistics		
Number of clusters	2		
Number of cases	2958		
Number of indicators	3		
Log-likelihood (LL)	-3383.5152		
Classification errors	0.0268		
Entropy R-squared	0.7594		
Standard R-squared	0.7633		

Source: Authors' calculations based on the HBS.