Assurances et gestion des risques
Insurance and Risk Management

MACROECONOMIC STRESS-TESTING OF MORTGAGE DEFAULT RATE USING A VECTOR ERROR CORRECTION MODEL AND ENTROPY POOLING

David Ardia, Anas Guerrouaz and Jeanne Rey

Volume 83, Number 3-4, 2016

URI: https://id.erudit.org/iderudit/1091505ar
DOI: https://doi.org/10.7202/1091505ar

See table of contents

Publisher(s)

Faculté des sciences de l'administration, Université Laval

ISSN

1705-7299 (print)
2371-4913 (digital)

Explore this journal

Cite this article

Ardia, D., Guerrouaz, A. & Rey, J. (2016). MACROECONOMIC STRESS-TESTING OF MORTGAGE DEFAULT RATE USING A VECTOR ERROR CORRECTION MODEL AND ENTROPY POOLING. *Assurances et gestion des risques / Insurance and Risk Management*, *83*(3-4), 115–133. https://doi.org/10.7202/1091505ar

Article abstract

We propose a methodology to perform macroeconomic stress-testing on the probability of default of a given borrowers' population (*i.e.*, aggregate probability of default) through simulation from a vector error correction model and entropy pooling (Meucci, 2008).

Assurances et gestion des risques, Vol. 83 (3-4)
Insurance and risk management, Vol. 83 (3-4)

MACROECONOMIC STRESS-TESTING OF MORTGAGE DEFAULT RATE USING A VECTOR ERROR CORRECTION MODEL AND ENTROPY POOLING[1]

David Ardia, Anas Guerrouaz and Jeanne Rey[2]

ABSTRACT

We propose a methodology to perform macroeconomic stress-testing on the probability of default of a given borrowers' population (*i.e.*, aggregate probability of default) through simulation from a vector error correction model and entropy pooling (Meucci, 2008).

RÉSUMÉ

Nous proposons une méthodologie pour effectuer des tests de stress macroéconomique en lien avec la probabilité de défaut d'une population d'emprunteurs (i.e. sur une probabilité de défaut agrégée) par l'entremise d'une simulation de modèle vectoriel à correction d'erreurs et de concentration d'entropie.

Keywords: Mortgage default probability entropy pooling macroeconomic variables stress-testing VECM OFSI

1. Introduction

Stress-testing is an important aspect of risk management, which aims at measuring the implications of adverse market/economic scenarios on a given portfolio. These scenarios are subjective views on possible market/economic realizations and focus on negative outcomes (*e.g.*, a market crash, a recession or an unemployment crisis). More specifically, stress-testing helps: i) providing forward-looking assessments of risk, ii) overcoming limitations of models and historical data, and iii) providing an indication of the appropriate level of capital necessary to endure adverse economic conditions (Basel Committee on Banking and Supervision, 2009).

There are several approaches to stress-testing. In the *forward stress-test* methodology, the goal is to study the impact of a given adverse macroeconomic scenario on the evolution of the default probability. On the contrary, the *inverse stress-test* tries to determine the set of macroeconomic scenarios associated with an increase in the risk of the portfolio; see Grunke (2011) for illustrations. Moreover, forward stress-tests can be broken down into two categories: i) *sensitivity* analysis and ii) *scenario* analysis (Lopez, 2005). Sensitivity analysis studies the impact of the variation of a single model parameter or factor, all others being equal (Plank, 2013). Scenario analysis allows the variation of multiple parameters or factors (Mager and Schmeider, 2007). The selection of adverse factor scenarios can be carried out using historical data or judgmental forecasts, *i.e.*, expert judgments without direct reference to historical data (Rebonato, 2010). In the case of historical data, the scenarios can either be retrieved from the historical distribution of the factors (*i.e.*, non-parametric approach) or from factor scenarios obtained from a parametric model fitted to historical data (Brueur et al., 2012; Assouan, 2012).

In this paper, we develop a methodology to conduct stress-tests on the default probability of a population of mortgage borrowers. A mortgage loan is distributed by a retail lender and is used by purchasers of real property to raise funds to buy real estate or by existing property owners to raise funds for any purpose while using the property being mortgaged as collateral. This question is highly relevant in the current context of high household indebtedness in Canada. In fact the Bank of Canada considers the high level of household debt, largely explained by low interest rates and rising residential prices, as the foremost vulnerability in the Canadian financial system [3]. According to them, income growth is not keeping pace with rising house prices, leading to higher household indebtedness. The household sector is therefore less resilient, and during times of stress, defaults may increase, creating losses for lenders. This risk may materialize in the event of a severe recession and/or a significant rise in unemployment which would yield a significant number of households unable to service their debt, leading them to default. The risk of such an event is rated as *elevated* by the Bank of Canada (see Bank of Canada, 2015, page 23): "The probability of the risk materializing remains low, but the potential impact on the economy and the financial system if the risk were to materialize would be severe."

Our approach falls within the class of forward stress-tests based on scenarios. We first link our variable of interest (*i.e.*, the aggregated default probability[3]) with macroeconomic explanatory variables through a vector error correction model (VECM) and a joint distribution for the innovations, which is used to simulate future levels of the macroeconomic variables. The list of potential explanatory macroeconomic variables on default rates has been established by several studies (see, e.g., Zandi, 1998, for a review). In line with the industry (see, e.g., Fong and Wong, 2008; Assouan, 2012) and regulation (see, e.g., OSFI, 2009) practice[4], the macroeconomic variables have two roles: i) explaining the variable of interest through the multi-factor model, and ii) describing the macroeconomic context considered for the stress-test. Moreover, the auto-induced effect of the multivariate approach integrates the feedback effects between all variables as required by the regulators (see OSFI, 2009, page 6). Even if we only focus on one dimension, our approach leads to a shift of the probabilities attributed to the simulated scenarios. Hence, a statement on a function of a single macroeconomic variable has potentially an impact on other macroeconomic variables due to the dependence between the variables.

The variables' distribution obtained through simulations constitutes the *reference model* in our approach. This reference model (also referred to as the *prior* in Meucci (2008)) results from the attribution of equal probabilities to all of the simulated scenarios. Hence, we do not integrate any information on the future distribution of the variables other than the underlying data generating process. In order to perform stress-tests on the default probability, we optimally distort the prior to reflect possibly adverse macroeconomic scenarios such as an unemployment crisis. These adverse macroeconomic scenarios can be encapsulated into a number of *views*, which can be any statement about functions of our macroeconomic variables that directly contradict the prior (*e.g.*, stating that the one-year ahead expected value of unemployment rate will be greater than 10% when its expected value in the prior is 8%). The *posterior* distribution is the *closest* to the prior while satisfying our stress-test views. We use a *relative entropy* measure in order to quantify the similarity between the two distributions. Our approach falls within the class of scenario stress-testing (as opposed to sensitivity testing) even when the view concerns a function of a single macroeconomic variable as we proceed to modify the probabilities attributed to the simulated scenarios.

The rest of this paper is organized as follows. Section 2 presents the data. Section 3 discusses the model, some results and the simulation procedure. Section 4 describes the stress-test methodology and provides two illustrations. Section 5 concludes.

2. Data

2.1 Data Description and Variables Selection

This technical note is realized in collaboration with the National Bank of Canada in an effort to offer a benchmark model to stress-test the default probability of a portfolio of mortgages distributed in the province of Québec; hence our emphasis on data linked to Québec. However, our introductory remarks remain pertinent as Québec is not spared by the current upward trend in household indebtedness (see Gauthier, 2015, Table 3). Moreover, due to confidentiality reasons, we cannot disclose all macroeconomic variables used in the study. We therefore present hereafter a model based on a subset of relevant and publicly available macroeconomic variables. The methodology can be extended in a straightforward way to a larger set of variables.

We use the proportion of mortgages in arrears compiled by the Canadian Bankers Association (CBA) as a *proxy* for the mortgage default rate, which is not directly observable. In this context, a mortgage is in arrears if payments have been missed for at least three months. This definition of the arrears is fully aligned with the defaulting retail client defined in the Basel II risk framework (see Basel Committee on Banking and Supervision, 2004, Paragraph 452). The CBA combines mortgage arrears data from ten major Canadian commercial banks including the National Bank of Canada. In addition to the mortgage default rate, we select five macroeconomic variables compiled by CANSIM, namely:

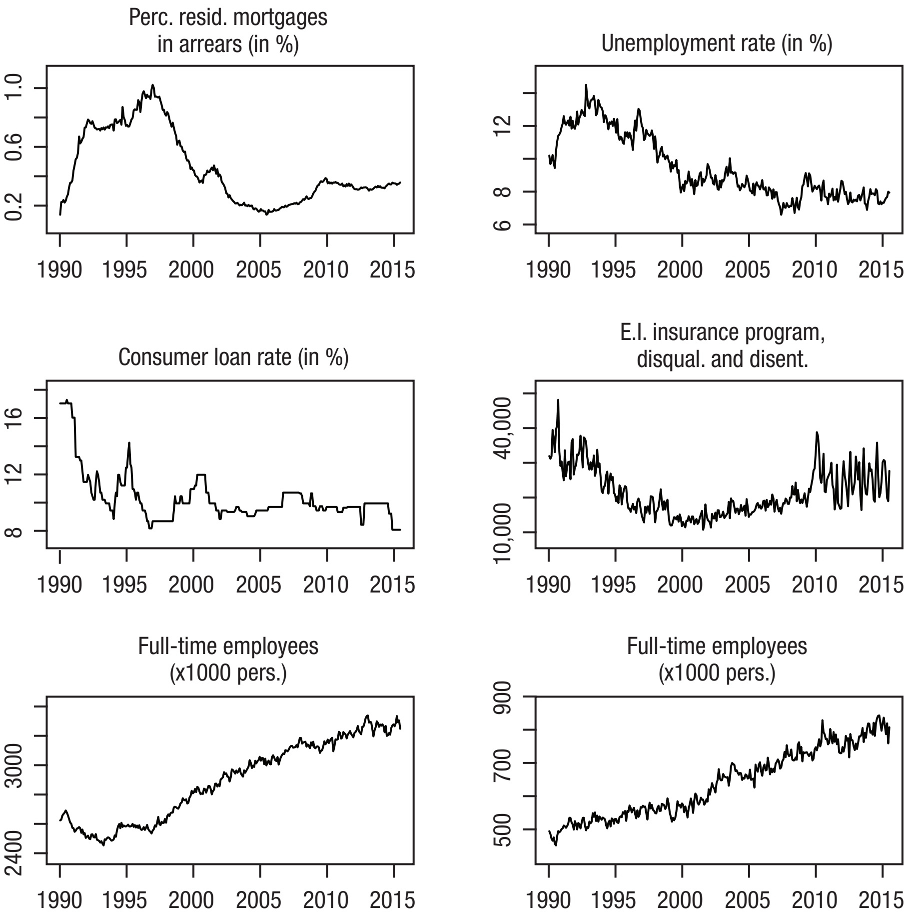
1. The unemployment rate in Québec (in %);
2. The consumer loan rate (in %);
3. The total number of disqualifications and disentitlements from the E.I. insurance program in Québec;
4. The number of full-time employees in Québec;
5. The number of part-time employees in Québec.

We choose to work with monthly data in order to have enough observations for the estimation of the VECM but also to capture short-term dynamic shifts. Overall we have 306 monthly observations ranging from February 1990 to July 2015. The description of the variables is presented in Table 1. Figure 1 displays the time series.

TABLE 1 *Description of variables*

VARIABLE NAME	DESCRIPTION
Proportion of residential mortgages in arrears in Québec	Includes data from BMO, CIBC, HSBC Bank Canada, National Bank of Canada, RBC Royal Bank, Scotiabank, and TD Canada Trust, Canadian Western Bank (as of Apr. 2004), Manulife Bank (as of Apr. 2004) and Laurentian Bank (as of Oct. 2010). A mortgage is considered in arrear after three or more months of missing payments. Compiled by the CBA.
Unemployment rate in Québec	Cansim table 282-0001: Labour force survey estimates. The unemployment rate is the number of unemployed persons expressed as a percentage of the labour force.
Consumer loan rate	Cansim table 176-0043: Chartered banks – nation-wide consumer loan rate as compiled by the Bank of Canada (BoC).
Total number of disqualifications and disentitlements from the employment insurance (E.I.) program in Québec	Cansim table 276-0003: Employment Insurance Program (E.I.), number of disqualifications and disentitlements by province and reason.
Number of full-time employees in Québec (thousands of persons)	Cansim table 282-0001: Labour force survey estimates. Full-time employment consists of persons who usually work 30 hours or more per week at their main or only job.
Number of part-time employees in Québec (thousands of persons)	Cansim table 282-0001: Labour force survey estimates. Full-time employment consists of persons who usually work less than 30 hours per week at their main or only job.

FIGURE 1 *Evolution of the variables over time*



2.2 Data Analyses

First, we test our variables for potential seasonality. To this end, we calculate the seasonal indices using the ratio to moving average (RMA) method (Aczel, 2006) and perform a graphical check to identify the series which display a seasonal pattern. Removing potential seasonality allows us to design a more parsimonious model as we do not have to include seasonal dummy variables. Our analysis indicates a pronounced seasonal pattern for all time series considered except the consumer loan rate. Excluding the latter, we adjust them accordingly by dividing the series' values by the corresponding RMA values.

Second, we determine if the series are stationary. To this end, we perform an augmented Dickey-Fuller test (Said and Dickey, 1984) to determine the order of integration of the series. Our results indicate that the six series are integrated of order one (*i.e.*, $I(1)$).

The dependent variable of interest (*i.e.*, the proportion of mortgages in arrears) is part of the set of $I(1)$ variables so we are interested in testing the six series for potential cointegration. The cointegration property is interesting in the context of a model designed for simulations as it accrues the short-term dynamic with a long-term equilibrium relation between the variables in level. We proceed to test the six series for potential cointegration through the Johansen (1991) procedure, which allows for the existence of multiple cointegration relationships.

The null hypothesis of the Johansen (1991) test with the maximum eigenvalue that there is r cointegration relationships against $r + 1$ ones. We proceed sequentially for $r = 0, 1, 2, \ldots, 5$ and take the first non-rejection of the null as an estimate of r. At the 5% significance level our estimate is $r = 3$.[5]

3 Model

3.1 Model Estimation

Given the Johansen (1991) test results, it is reasonable to estimate a VECM linking the six $I(1)$ variables. The choice of the lag order is guided by the study of the partial auto-correlograms and the cross-correlograms of the differentiated variables. The analysis suggests a three-period lag order in the VECM representation (*i.e.*, four lags in the equivalent VAR representation, using the $I(1)$ variables), as most correlograms fade to zero by the fifth lag. We choose not to include a deterministic trend in the VECM, as it renders the modeling more complex and is hard to justify economically.

In order to understand the VECM dynamics and assess its relevance from an economic point of view, we display in Figure 2 and Figure 3 the VECM impulse response functions. Impulse responses depict the reaction of the VECM variables over time when the innovation of one of them is shocked (*e.g.*, incremented by a value equal to its historical standard deviation) at time zero, all other things being equal. In each plot of Figure 2 and Figure 3, the solid bold curve represents the non-cumulative variation of the mortgage default rate when another macroeconomic variable is shocked in a proportion represented by its respective curve. The dynamic of the mortgage default rate is fully coherent with the direction of the impulses. More precisely, it appears that:

1. A positive shock on the unemployment rate increases the mortgage default rate, which is coherent with a decrease in disposable income; it also triggers a decrease in the number of full-time employees and a small increase in the number of part-time employees, which is consistent with a rise in precarious employment;

2. A positive shock on the consumer loan rate increases the mortgage default rate after four months as the consumers' disposable income decreases and their interest charge increases;

3. An increase of the number of withdrawals from the insurance program slightly decreases the mortgage default rate as this increase may be explained by an increase of the population wealth (and thus a higher propensity to leave the Employment Insurance program) which ultimately improves mortgage payments;

4. An increase in the number of full-time employees means more disposable income and decreased mortgage default rates as observed on the fifth impulse response graph;

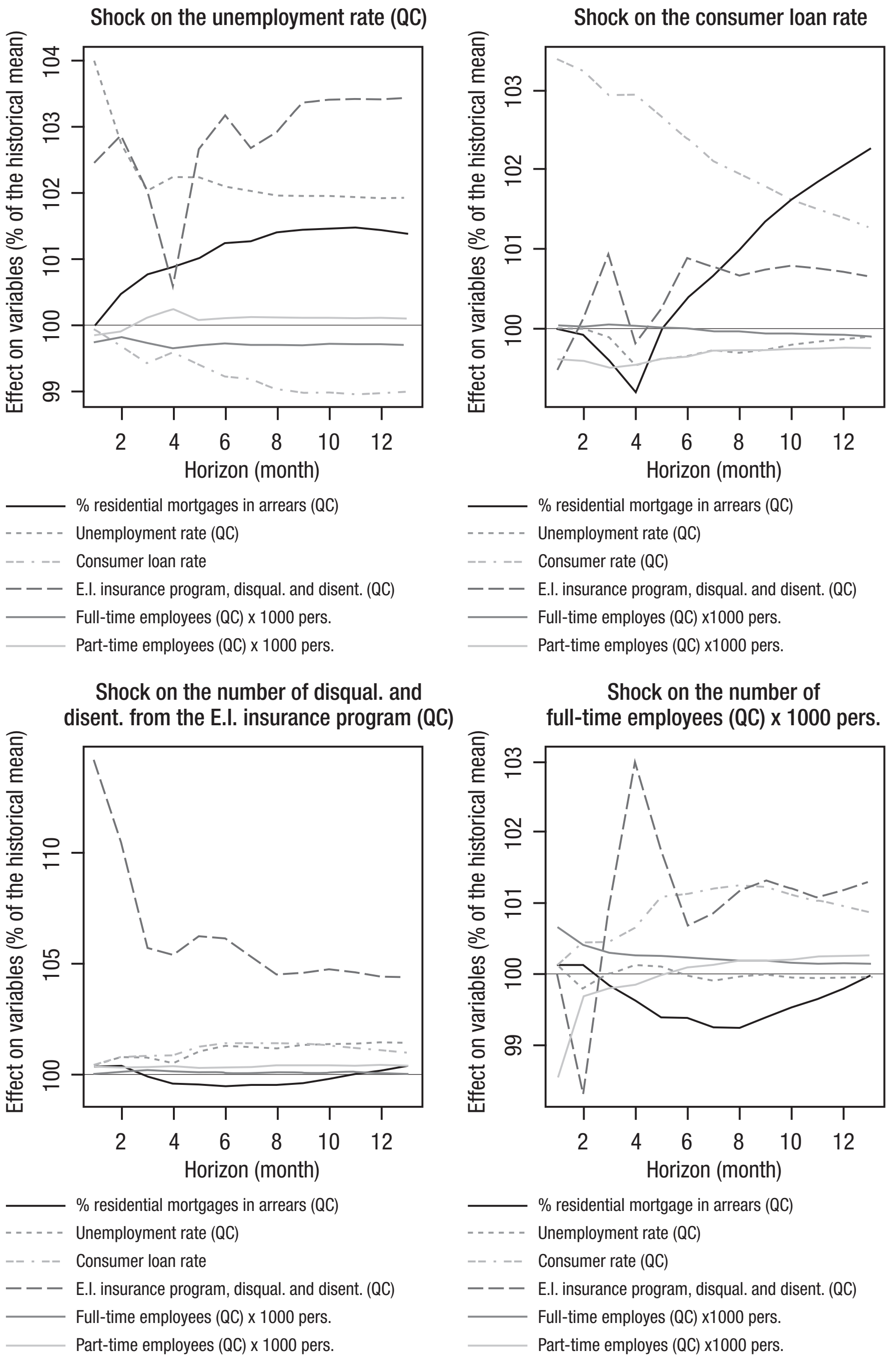
5. An increase in the number of part-time employees tends to lower the mortgage default rate during the first months, but this increase in part-time work increases the mortgage default rate in the long-run, as it implies a more precarious job market.

For simplicity, we work from now on with the vector autoregression (VAR) model representation as the $I(1)$ variables are readily available. Indeed, using the VECM representation for simulations would require differentiating the variables but also recalculate them in levels at each simulation step to obtain the error correction term.[6]

In order to obtain simulations from our VAR model, we make an assumption on the joint distribution for the innovations. We rely on a marginal-copula approach.

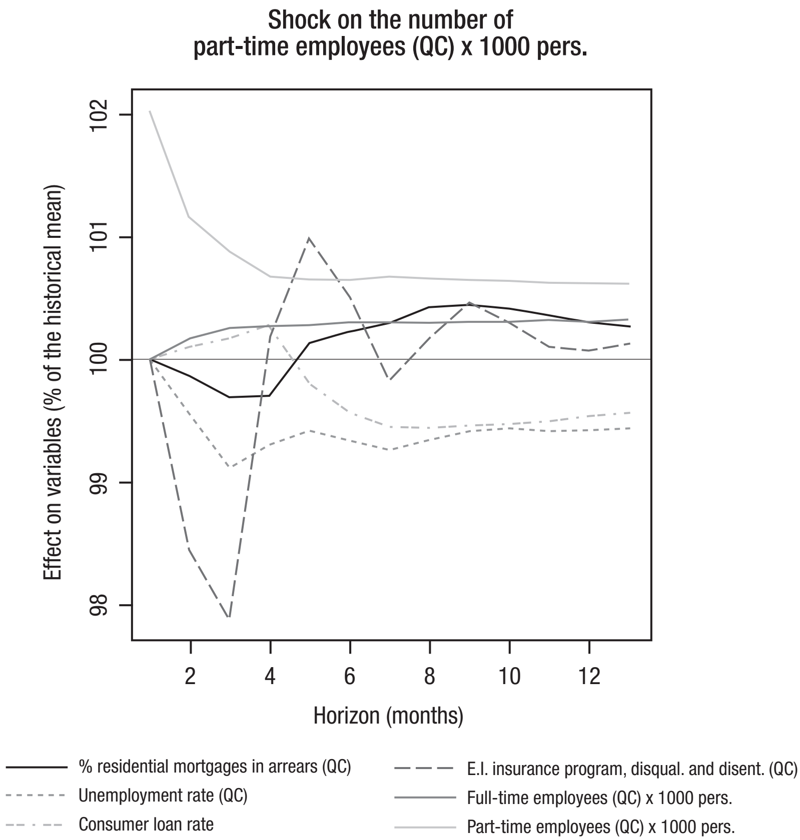
First, we start by modeling the innovations' marginal distributions. We use the flexible skewed Student-t distribution for the marginal distribution. It is fitted by maximum likelihood on the residuals of the VAR model. The test by Fan (2014) cannot reject the chosen specification. To calibrate the copula, we transform the VAR model residuals into their fitted cumulative distribution function. By applying the methodology of Genest et al. (2009), we select a Gaussian copula which is also fitted by maximum likelihood.[7]

FIGURE 2 *Impulse response functions of the vector error correction model*



Impulse responses of each variable are expressed in percentage of the historical mean of the variable. The solid bold line represents the non-cumulative variation of the mortgage default rate when another macroeconomic variable is shocked in a proportion represented by its respective curve.

FIGURE 3 *Impulse response functions of the vector error correction model*



Impulse responses of each variable are expressed in percentage of the historical mean of the variable. The solid bold line represents the non-cumulative variation of the mortgage default rate when another macroeconomic variable is shocked in a proportion represented by its respective curve.

3.2 Simulation

So far, we have modeled the relationship between the variables with a VAR model. The model is estimated by maximum likelihood to historical data. Then, the innovations for the VAR model are modeled from the residuals of the fitted VAR model. First, each marginal is modeled using a skewed Student-t distribution and then a Gaussian copula is used to merge the marginals. This full parametric approach allows us to simulate new scenarios for the variables.

We simulate the future values of our six variables using a rolling simulation. More precisely, for each simulation-path we:

1. Forecast the VAR's values for the next period using the four last observations (as the VAR's lag order is four);
2. Simulate the six innovations with the estimated copula and marginal distributions, then increment the forecast with the innovations;
2. Append the values used to forecast the VAR with the newly simulated values (while the oldest values are discarded);
4. Repeat steps 1 to 3 to simulate the next period.

Following this procedure, we generate $J \equiv 500{,}000$ simulations. At the end of the simulation procedure, we re-introduce seasonality using the inverse step of the method described in Section 2. We assume that the monthly RMA indices remain constant and multiply each simulated value by its corresponding index. Then we aggregate the simulations into future distributions of the variables by attributing equal probabilities to all of the simulated scenarios, hence obtaining our prior or reference model.

4. Stress-testing and Entropy Pooling

4.1. Methodology

Once the prior is obtained, we can distort it using views on any functions of the macroeconomic variables (*i.e.*, the same variables used in the VECM). These views directly contradict the prior and will be the starting point to determine the stressed joint distribution of our variables (*i.e.*, the posterior distribution).

The stressed distribution is obtained by modifying the probabilities attributed to each simulation, *i.e.*, $\{p_j\}_{j=1}^{J}$ where $p_j \equiv 1/J$, such that the views are satisfied. The new set of probabilities defined as $\{p_j^*\}_{j=1}^{J}$ must be as *close* as possible to the set of prior probabilities. To find this optimal set, we rely on the relative entropy measure to quantify the *distance* between the two distributions of probabilities. We follow Meucci (2008, 2010) and use the Kullback-Leibler divergence (Kullback and Leibler, 1951) as the measure of discrepancy. Note that Meucci (2010) relies on past historical scenarios while we use scenarios generated from our prior which is forward-looking.

More formally, given a set of J scenarios $\{y_{1,j},\ldots,y_{6,j}\}_{j=1}^{J}$ on the 6-dimensional variable $(Y_1,\ldots,Y_6)'$ and associated prior probabilities $\{p_j\}_{j=1}^{J}$, the relative entropy measure ε between the prior probabilities and a set of candidate probabilities $\{q_j\}_{j=1}^{J}$ is given by:

$$\varepsilon\left(\{q_j\}_{j=1}^{J};\{p_j\}_{j=1}^{J}\right) \equiv \sum_{j=1}^{J} q_j \ln\left(\frac{q_j}{p_j}\right).$$

If the variable on which we have a view is Y_i and our view is that its expected value is $E(Y_i) \gtreqqless Y_i^*$, with Y_i^* a value set by the user, the view under the simulation framework writes:

$$\sum_{j=1}^{J} y_{i,j} q_j \gtreqqless Y_i^* .$$

In this case, the optimal set of posterior probabilities is found by solving:

$$\{p_j^*\}_{j=1}^{J} \equiv \underset{\{q_j\}_{j=1}^{J}}{argmin}\ \varepsilon\left(\{q_j\}_{j=1}^{J};\{p_j\}_{j=1}^{J}\right)$$
$$\text{subject to } \sum_{j=1}^{J} y_{i,j} q_j \gtreqqless Y_i^* . \qquad (1)$$

Meucci (2008) shows how the optimization problem (1) can be solved efficiently and provides illustrations of (in-)equality views on other quantities of interest such as standard deviation or quantiles.

4.2 Illustrations

We provide two illustrations of the stress-testing methodology by applying stress-test views on the unemployment rate and how it impacts the distribution of the mortgage default rate.

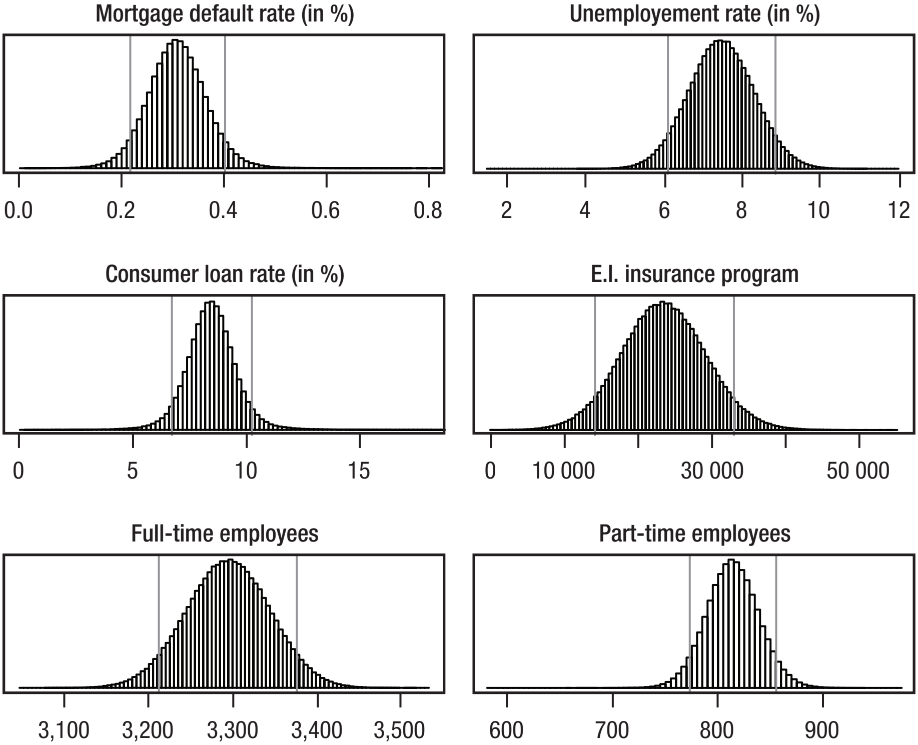
4.2.1 Simple Case

First, we simulate our six variables over the next 12 months (*i.e.*, one year) and stress-test the distribution of the mortgage default rate using a view according to which the one-year ahead unemployment rate is equal or greater than 10%.

Running the simulations and considering the last value of the simulations yields the marginal distributions of the macroeconomic variables at the desired 12-month horizon. In Figure 4, we display the marginal distributions of the variables obtained from our simulation framework. With the simulations from the prior, the expected value of the one-year ahead unemployment rate is estimated at 7.48%. In Figure 5, we display the bivariate distributions, which highlight the dependencies between the variables.

Simulations are generated from the prior and by construction have each a $1/J$ probability of occurrence. We now apply the stress-test scenario (view) on the unemployment rate. We want to stress the mortgage

■ FIGURE 4 *Marginal distributions of the variables in the prior model of Section 4.2.1*

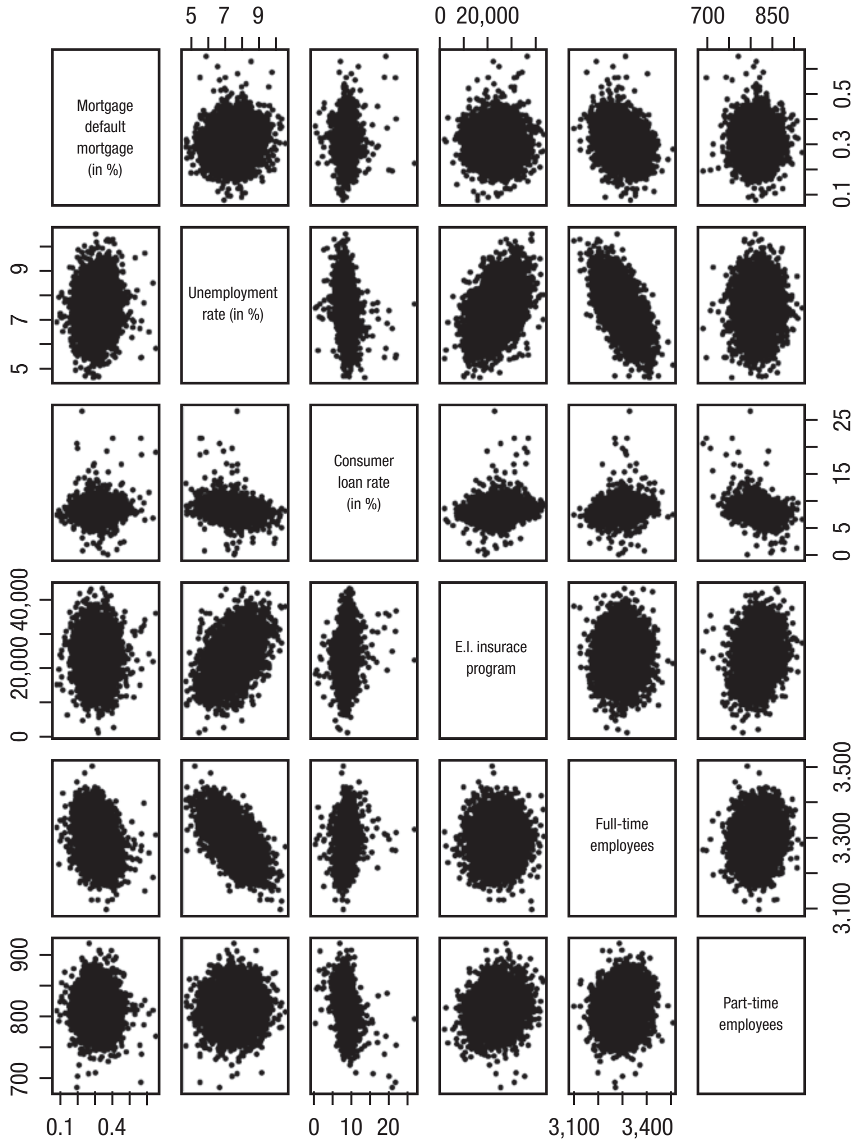


We display the histogram of simulated values over a 12-month horizon of the six variables under the prior model (i.e., with no view). The vertical blue lines report the 90% confidence bands.

default rate's distribution using a statement on the future unemployment rate that directly contradicts the prior. Our (subjective and illustrative) view is that the one-year ahead expected value of the unemployment rate is equal or greater than 10%.

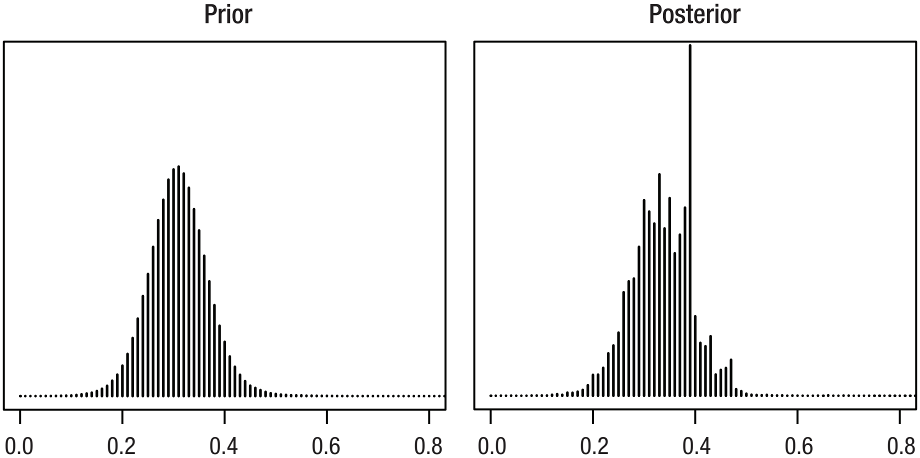
Using the optimization program (1), we compute the set of posterior probabilities. They are then used to aggregate the weighted scenarios to compute the posterior distribution of our variable of interest and use it to extract any relevant indicators. The prior and posterior (*i.e.*, stress-tested) distributions of the mortgage default rate are displayed in Figure 6. We notice how the entropy pooling step distorts the distribution. Summary statistics of the distributions are presented in Panel A of Table 2. We note higher posterior quantiles compared with those of the prior. This shows how the unemployment crisis is captured by the stress-test view as this adverse scenario increases the mortgage default rate. The implementation of such scenario with a value frontier definition would not be as satisfying using a pure VECM. Indeed, its implementation would consist in truncating the stressed variable distribution on the frontier threshold. The resulting distribution would be hard to explain economically. On the contrary, the entropy pooling approach produces a distribution which is *coherent* with the prior.

FIGURE 5 *Bivariate scatter plots of the variables in the prior model of Section 4.2.1*



We display the bivariate scatter plot for each pair of the six variables simulated over a 12-month horizon under the prior model (*i.e.*, with no view).

FIGURE 6 *Mortgage default rate distribution under the prior and posterior model of Section 4.2.1*



We display the histogram of simulated values for the mortgage default rate under the prior (left plot) and the posterior (right plot) models. The posterior is obtained with a stress-test view stating that the expected value of the unemployment rate at the 12-month horizon must be equal or greater than 10%.

TABLE 2 *Prior and posterior results for the two illustrations*

			Quantiles				
	Mean	Std	5th	25th	50th	75th	95th
Panel A							
Prior	0.31%	0.06%	0.22%	0.27%	0.31%	0.35%	0.40%
Posterior	0.36%	0.06%	0.26%	0.33%	0.38%	0.39%	0.43%
Panel B							
Prior	0.30%	0.07%	0.19%	0.25%	0.30%	0.34%	0.41%
Posterior	0.39%	0.08%	0.26%	0.35%	0.38%	0.41%	0.49%

Summary statistics for the prior and posterior distribution of the mortgage default rate. Panel A refers to Section 4.2.1 and Panel B to Section 4.2.2. The prior results are obtained from the prior model simulations. The posterior or stress-test results are obtained from the posterior model simulations. In each case, we report the mean, standard deviation and quantiles of the distribution.

4.2.2 Complex Case: OSFI's Macroeconomic Scenario

The Office of the Superintendent of Financial Institutions (OSFI) issues an annual stress-test scenario for the key Canadian macroeconomic variables, which we use now as our second illustration. The OSFI 2015 scenario provides quarterly values for the main macroeconomic variables from 2015-Q1 to 2019-Q3. We implement simultaneous views regarding the unemployment rate in 2016-Q3 and 2016-Q4 but the approach can be extended to other quarters in a straightforward manner. We use monthly data so we take the OSFI's quarterly scenarios as equal to the quarterly expected value obtained from the simulated unemployment rates.

Our estimation sample ends in July 2015. Hence we simulate the next seventeen months in order to obtain unemployment rates' quarterly averages for the last two quarters of 2016. Our view is that these quarterly averages are equal to the OSFI's values (*i.e.*, 11.2% and 11.3%, respectively). In Panel B of Table 2 we report the summary statistics for the prior and the posterior distributions of the mortgage default rate in 2016-Q4. The monthly analysis of the results show the progressive shift of the mortgage default rate distribution reflecting the time dependency of VECM structure. The significant increase of the whole mortgage default rate distribution fully reflects the aggressiveness of the unemployment OSFI scenario. The quantiles on the mortgage default rate offer margins of conservatism accounting for the modeling errors based on a quantified risk.

Entropy pooling is useful in the implementation of such scenario with path-dependent conditions compared with a pure VECM. In a pure VECM, the quarterly OSFI scenario could be integrated in two manners. First, the VECM may be calibrated on quarterly data, but, as mentioned in the introduction, this would significantly reduce the number of available observations and some of the monthly dependencies may disappear under a larger time scale. Second, keeping the monthly VECM estimation, one could consider flat deterministic monthly scenarios for the stressed variable and make the shocked variable (*e.g.*, the unemployment rate) exogenous. In such a case, the volatility of the mortgage default rate would be underestimated as the volatility of the exogenous variable itself is, compared with the smoother condition on a quarterly average. The use of entropy pooling is thus fundamental in order to fully reflect the stressed scenario and the dependencies between the variables.

5 Conclusion

This paper proposed a methodology for stress-testing the aggregated rate of default of retail mortgages in the Québec market, easily extendable to the Canadian or international markets. The approach is based on a parametric model linking historical mortgage default rates with relevant macroeconomic variables in a first step, and generating simulations with distorted probabilities to meet stress-test views "à-la-Meucci" in a second step.

Several extensions of the proposed method are possible. From an econometric viewpoint, alternative model specifications can be tested and compared (*e.g.*, skewed Laplace instead of skewed Student-*t*, or Student-*t* copula instead of Gaussian copula for modeling the innovations of the VECM model). Also, automatic methods for selecting explanatory variables could be used (*e.g.*, stochastic search methods for variable selection). From a stress-testing viewpoint, a straightforward extension would consist in assigning a confidence level to the stress-test scenario. Several stress-tests with various confidence levels could also be defined, and a stress-test distribution encompassing views from different sources with their own confidence level can easily be constructed by simulation, as presented in Meucci (2008). This extension is interesting as it directly answers recommendation by OSFI's guidelines for stress-testing, which states that (see OSFI, 2009, Section D): "Stress-testing programs should take account of views from across the organization and should cover a range of perspectives and techniques".

References

[1] A.D. Aczel. *Complete Business Statistics*. McGraw Hill Higher Education, 2006.

[2] S. Assouan. Stress testing a retail loan portfolio: An error correction model approach. *Journal of Risk Model Validation*, 6(1):3–25, 2012.

[3] Bank of Canada. *Financial System Review: Assessment of Vulnerabilities and Risks*. Bank of Canada, December 2015.

[4] Basel Committee on Banking Supervision. *Principles for sound stress testing practices and supervision*. Basel Committee on Banking Supervision, March 2009.

[5] T. Breuer, Martin Jandackaa, Javier Menciab, and Martin Summerc. A systematic approach to multi-period stress testing of portfolio credit risk. *Journal of Banking and Finance*, 36(2):332–340, 2012.

[6] Y. Fan. Testing the goodness-of-fit of a parametric density function by kernel method. *Econometric Theory*, 10:316–356, 2014.

[7] C. Genest, B. Rémillard, and D. Beaudoin. Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and Economics*, 44(1):199–214, 2009.

[8] P. Grunke. Reverse stress tests with bottom up approaches. *Journal of Risk Model Validation*, 5(1):71–90, 2011.

[9] S. Johansen. Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica*, 59(6):1551–1580, 1991.

[10] S. Kullback and R.A. Leibler. On information and sufficiency. *Annals of Mathematical Statistics*, 22(1):79–86, 1951.

[11] J.A. Lopez. Stress test: Useful complements to financial risk models. Technical Report 14, Federal Reserve of San Francisco, 2005.

[12] F. Mager and C. Schmieder. Stress testing of real credit portfolios. Technical Report 17, Deutsche Bundesbank, 2007.

[13] A. Meucci. Fully flexible views: Theory and practice. *Risk*, 21(10):97–102, 2008.

[14] A. Meucci. Historical scenarios with fully flexible probabilities. *GARP Risk Professional*, pages 47–51, 2010.

[15] K. Plank. Stress tests of credit risks. Working paper, 2013.

[16] R. Rebonato. *Coherent Stress Testing: A Bayesian Approach to the Analysis of Financial Stress*. Wiley, 2010.

[17] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2015.

[18] S.E. Said and D.A. Dickey. Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*, 71(3):599–607, 1984.

NOTES

1. The authors are grateful to Marie-Claude Beaulieu, Oussama Chakroun, Attilio Meucci, Matthieu Stigler and Anick Yaha. This research was funded by IFSID Montréal (see technical notes http://ifsid.ca/en/publications/technical-notes/). All analyses have been performed in the R statistical language (R Core Team, 2015) and are available from the authors upon request.

2. Email addresses: david.ardia@unine.ch (David Ardia), Anas.Guerrouaz@gmail.com (Anas Guerrouaz) and Jeanne.Rey@bnc.ca (Jeanne Rey).

David Ardia (corresponding author): Institute of Financial Analysis, University of Neuchâtel, Neuchâtel, Switzerland and Department of Finance, Insurance and Real Estate, Laval University, Québec City, Canada

Anas Guerrouaz: Department of Finance, Insurance and Real Estate, Laval University, Québec City, Canada Jeanne Rey: Banque Nationale, Montréal, Canada Correspondence: University of Neuchâtel, Rue A.-L. Breguet 2, CH-2000 Neuchâtel, Switzerland. Phone: +41 32 718 1365.

3. With the loss given default (LGD) and the exposure at default (EAD), the default probability is one of the risk measures banks need to assess in order to estimate the capital amount required from regulatory and business purposes.

4. Most of stress-testing models for retail portfolios are regression-based. The VECM approach offers the most complete representation of the serial and cross-sectional dynamics of a multivariate system.

5. Full analyses of seasonality, stationarity and cointegration are available from the authors upon request.

6. Both VECM and VAR coefficients are available from the authors upon request.

7. All estimation results are available from the authors upon request.