STRESS TESTING BANKS' CREDIT RISK
USING MIXTURE VECTOR AUTOREGRESSIVE MODELS[1]

Prepared by

Tom Pak-wing Fong[2, 3]
Research Department,
Hong Kong Monetary Authority

and

Chun-shan Wong[4, 5]
Department of Finance,
The Chinese University of Hong Kong

ABSTRACT

This paper estimates macroeconomic credit risk of banks' loan portfolio based on a class of mixture vector autoregressive models. Such class of models can differentiate distributions of default rates and macroeconomic conditions for different market situations and can capture their dynamics evolving over time, including the feedback effect from an increase in fragility back to the macroeconomy. These extensions can facilitate the evaluation of credit risks of loan portfolio based on different credit loss distributions.

JEL classification: C15; C32; C53; E37; G21;

Keywords: Stress test; Hong Kong Banking; Credit risk; Mixture autoregressive models; Macroeconomic shocks; Value-at-risk.

EXECUTIVE SUMMARY

- *The underlying distributions of default rates and macroeconomic variables are usually assumed to be unimodal, which does not discriminate between normal and abnormal rises and falls. The resulting distribution may not be "exceptional but plausible" macroeconomic shocks. Although banks are required to estimate the amount of required capital at a very high level of confidence so as to avoid any unexpected losses due to the invalid assumption on normality, the probability of such a loss can be under-estimated if the normality assumption is too strong for the default rate.*

- *This paper considered a class of mixture vector autoregressive (MVAR) models. It extends the classical VAR model which assumed the probability distribution of default rate and macroeconomic variables to be a mixture of normal distributions in order to estimate macroeconomic credit risk. Such an extension can allow multi-modal distributions for different market conditions in estimation so as to provide a more precise estimate of potential loss and simulate exceptional but plausible scenarios when the unimodal assumption is restrictive. Apart from that, the proposed model can capture the dynamics among the default rates and macroeconomic variables, including the feedback effect from an increase in fragility back to the macroeconomy.*

- *Based on the new model, macro stress testing is then performed to assess the vulnerability and risk exposures of banks' overall loan portfolios. By using the framework, a Monte Carlo simulation is applied to estimate the distribution of possible credit losses conditional on the market situation. The results show that the credit risk of the banks' loan portfolio simulated based on a unimodal assumption can be potentially under-estimated. Based on the estimates of the MVAR model, the mean credit losses are found to range from 0.24% to 2.51%, which double the credit losses estimated based on a unimodal assumption.*

I.   INTRODUCTION

Under a stress-testing framework of risk exposures of banks' loan portfolio, the probability distribution of default rates conditional on an adverse macroeconomic shock is usually under-estimated.[6] This is because the underlying distributions of default rates and macroeconomic variables are assumed to be unimodal. Such an assumption does not discriminate between normal and abnormal situations (including abnormal rises and falls). The resulting distribution may therefore reflect the vulnerability of a financial system due to normal market shocks, but not "exceptional but plausible" macroeconomic shocks, since the number of the normal market observations are many more than that of the stressful market situation.[7]

Conventional studies have considered tail events of historical episodes to devise scenarios in order to obtain a more precise estimation of the banks' credit risk when the market is stressed.[8, 9] Bank regulators require banks to estimate the amount of required capital at a very high level of confidence so as to avoid any unexpected losses due to the invalid assumption on normality.[10] In response to this, some banks have taken a level of confidence at 99.99% for the computation of value-at-risk (VaR), representing a probability of 0.01% for banks to experience such severe credit losses under stressful market conditions. This probability seems low. However, it can still be under-estimated if the normality assumption is too strong for the default rate.

Consider Hong Kong's overall loan default rates of retail banks as an example. Chart 1 plots the histogram for the changes of logit-transformed default rates during the period from 1997 Q1 to 2007 Q3.[11] Note that the lower the change in the transformed probability, the higher the probability of default. The chart shows a bimodal distribution with a small minority of observations on the left hand side (ranging from -0.35 to -0.20), representing a severe default situation in the Asian financial crisis. This suggests that models based on the unimodal assumption cannot fully capture the information for the stressful period such that the final credit losses may be underestimated.

---

[6] A comprehensive summary for the framework of stress testing can be found in Sorge (2004).

[7] Some studies have added dummy variables to econometric models to filter such "crisis" effect, however, the stress-testing exercise becomes a test for having a shock under normal market condition. That seems not exactly the purpose of "stress" test banks' loan portfolio.

[8] See details in Froyland and Larsen (2002), Hoggarth and Whitley (2003), Mawdsley et al. (2004) and Bunn et al. (2005).

[9] Some studies have taken into account the probabilistic elements and explicitly considered the correlation among macroeconomic variables and default rates. See details in Wilson (1997a, 1997b), Boss (2002), Virolainen (2004), and Gasha et al. (2004).

[10] Under the Basel II requirement, a higher level of confidence (such as 99.99%) is necessary to compute the required capital amount. See Hugh et al. (2005) for more interpretations.

[11] Detailed definition of the logit transformation can be found in the next section.

Chart 1. The histogram of the logit transformed default rates (from 1997 Q1 to 2007 Q3)
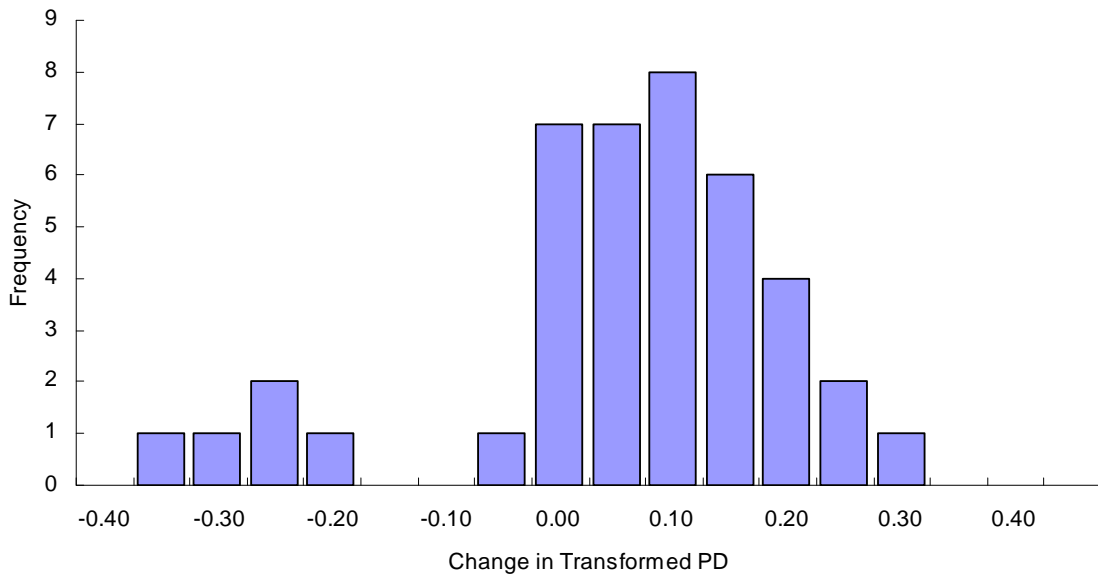
In this paper, the probability distribution of default rates is assumed to be a mixture of normal distributions. A class of mixture vector autoregressive (MVAR) model, proposed by Fong et al. (2007), is considered to estimate macroeconomic credit risk.[12] It is a generalization of vector autoregressive (VAR) models considered in Boss (2002), Virolainen (2004) and Wong et al. (2006).[13] Such an extension can allow multi-modal distributions for different market conditions in estimation so as to (i) provide a more precise estimate of potential loss and (ii) simulate exceptional but plausible scenarios when the unimodal assumption is restrictive. Apart from that, the proposed model can capture the dynamics among the default rates and macroeconomic variables, including the feedback effect from an increase in fragility back to the macroeconomy. Using a credit risk model of this type, adverse macroeconomic scenarios can be generated by using the estimated probability distribution for the stressed market condition. VaR is computed to evaluate how the stressed macroeconomic environment may affect the default probability of banks' loan portfolios.

The paper is organised as follows. Section II discusses the estimation of the MVAR model. Section III studies the application of the MVAR model on the overall default rates of bank loans in Hong Kong. Based on the estimated coefficients, stress testing can be done by simulations of different market situations. The procedure is discussed in Section IV. Section V concludes.

---

[12] Details of the development of the MVAR models can also be found in Martin (1992), Le et al. (1996), Wong and Li (2000), Wong and Li (2001), Berchtold (2003), Lanne and Saikkonen (2003), and Lanne (2004).

[13] Note that the VAR models were used to model the dynamics among macroeconomic variables only. The MVAR model considered in this paper models the dynamics among macroeconomic variables and joint-sector default rates.

II.   ESTIMATION AND INTERPRETATION OF THE MODEL

Similar to Wilson (1997a, 1997b), Boss (2002), Virolainen (2004) and Wong et al. (2006), the framework for stress testing banks' credit exposure to macroeconomic shocks comprises: (i) an empirical model with a system of equations describing credit risk and macroeconomic dynamics, and (ii) a Monte Carlo simulation for generating distribution of possible default rates and credit losses. This section focuses on the estimation of the empirical model.

2.1   The MVAR model

Assume that there are J economic sectors to which banks lend. Let $d_{j,t}$ be the average default rate in sector j observed in period t, where $j = 1,\ldots, J$. A logit-transformation is applied to transform $d_{j,t}$ to $y_{j,t}$. It is defined as:

$$y_{j,t} = \ln\left(\frac{1-d_{j,t}}{d_{j,t}}\right) \tag{1}$$

The transformed variable $y_{j,t}$ is not bound between zero and one instead it can take any value between plus and minus infinity. It can be seen that a higher $y_{j,t}$ is associated with a better credit-risk status. In other words, $d_{j,t}$ and $y_{j,t}$ are negatively related. Suppose that M macroeconomic variables are studied in the analysis and that they are denoted by $x_{m,t}$, where $m = 1,\ldots, M$.

Let $\mathbf{z}_t = (y_{1,t} ,\ldots, y_{J,t}, x_{1,t} ,\ldots, x_{M,t})'$ be a N-dimensional vector, where $N = J+M$. Their relationships are modelled by the MVAR model, denoted by MVAR(N, p_1, p_2), which is defined as

$$z_t \mid \Im_{t-1} \sim \begin{cases} \Phi(\Omega_1^{-1/2}(Z_t - \Theta_{10} - \Theta_{11}Z_{t-1} - \ldots - \Theta_{1p_1}Z_{t-p_1})) \\ \Phi(\Omega_2^{-1/2}(Z_t - \Theta_{20} - \Theta_{21}Z_{t-1} - \ldots - \Theta_{2p_2}Z_{t-p_2})) \end{cases} \text{with probability} \begin{matrix} \alpha_1 \\ \alpha_2 \end{matrix} \tag{2}$$

where $\Im_{t-1}$ indicates the information given up to time t-1, Φ (.) is the multivariate cumulative distribution function of the Gaussian distribution with mean zero and variance-covariance matrix equal to the identity matrix.[14] For the k-th component ($k = 1$ or 2), p_k is specified as the AR order, α_k is the probability and $\alpha_1 + \alpha_2 = 1$, Θ_{k0} is an N-dimensional

---

[14]   The model specified in (2) is a special case of the MVAR(N,2; p_1, p_2) stated in Fong et al. (2006). Note that a mixture of more than two components for a short time series is not common and is not easy to provide a straight forward interpretation. In view of this, we only consider a mixture of two components in this paper.

vector, Θ_{k1} ,..., Θ_{kp_k} are $N \times N$ coefficient matrices, and Ω_k is the $N \times N$ variance covariance matrix. For identifiability, it is assumed that $\alpha_1 \geq \alpha_2 \geq 0$.[15]

The resulting model in equation (2) can be also viewed as a mixture of two Gaussian VAR models with probabilities α_1 and α_2 respectively, which can be represented by

$$Z_t = \begin{cases} \Theta_{10} + \Theta_{11}Z_{t-1} + ... + \Theta_{1p_1}Z_{t-p_1} + \varepsilon_{1t} \\ \Theta_{20} + \Theta_{21}Z_{t-1} + ... + \Theta_{2p_1}Z_{t-p_2} + \varepsilon_{2t} \end{cases} \text{ with probability } \begin{array}{c} \alpha_1 \\ \alpha_2 \end{array} \qquad (3)$$

where $\varepsilon_{1t} \sim MVN(0,\Omega_1)$ and $\varepsilon_{2t} \sim MVN(0,\Omega_2)$, given the information up to the period t -1 $(= \Im_{t-1})$.[16] To be parsimonious, the MVAR $(N,1,1)$ model is considered. Specifically,

$$Z_t = \begin{cases} \Theta_{10} + \Theta_{11}Z_{t-1} + \varepsilon_{1t} \\ \Theta_{20} + \Theta_{21}Z_{t-1} + \varepsilon_{2t} \end{cases} \text{ with probability } \begin{array}{c} \alpha_1 \\ \alpha_2 \end{array} .$$

Given the probabilities for the two components, a bimodal distribution can be seen by simulation. Consider a 100 random numbers drawn from the system in equation (2), a number of $100 \times \alpha_1$ observations will be drawn from the first component which has a mean $\Theta_{10} + \Theta_{11}Z_{t-1} + ... + \Theta_{1p_1}Z_{t-p_1}$ and a variance covariance matrix Ω_1. The remaining observations, a total of $100 \times \alpha_2$, will be drawn from the second component which has a mean $\Theta_{20} + \Theta_{21}Z_{t-1} + ... + \Theta_{2p_2}Z_{t-p_2}$ and a variance covariance matrix Ω_2. When these two means are significantly different from each other, the resulting distribution, combining these two samples, will be bimodal.

As in the literature, this model links the default behaviours in different economic sectors to the macroeconomic conditions in a vector form so that macroeconomic variables determine default rates and the default signals also feedback to the macroeconomy. Our specification is therefore more general than Wilson (1997a, 1997b), Virolainen (2004) and Wong et al. (2006) where y_t is assumed to depend on $\mathbf{x}_t$ and $\mathbf{x}_{t-1}$ without any feedback from y_{t-1} to x_t.[17]

---

[15] Intuitively these avoid the problem of non-identifiability due to the interchange of component labels. See Titterington et al. (1985) and McLachlan et al. (1988) for details.

[16] In fact, a random variable drawn from a simple AR(p) model can be said to follow a one-component mixture Gaussian distribution conditional on the past information.

[17] Wong et al. (2006) has a similar specification as in this paper, but they did not consider the mixture properties.

2.2   Interpretation of the MVAR model

The main contribution of the above generalisation is to allow the unimodal distribution to split into two in mixture. Each distribution in the mixture may reflect different market conditions. Under the assumption that the first component has a higher probability of occurrence, the VAR model in the first component will reflect the dynamic relationship between the macroeconomy and default rates in the range of the transformed probability which is more commonly observed (see Chart 2). When the time series is separately fitted by the MVAR and VAR models, the coefficients of the first component of the MVAR model are expected to be roughly similar to those of the VAR model. Separately, the possibility of sharp changes in transformed probability (i.e. sharp changes in default rates), as shown in Chart 2, can be modelled by the second component. When the macroeconomic conditions worsen drastically in a rather short time-span, a thicker lower tail (i.e. more defaults) of the loss distribution would be obtained, representing a rise in the probability of default. On the other hand, when the market improves from a stressful situation to a more normal situation, a thicker upper tail (i.e. less defaults) of the loss distribution would be obtained.

Chart 2. Mixture of two components at a particular time point
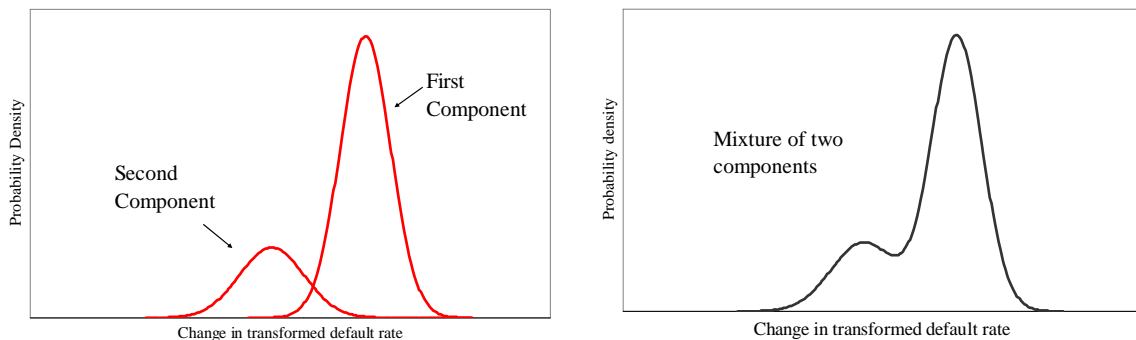
In addition, the MVAR model is a time series model that allows the distribution switching between unimodal and bimodal over time. As the means of the two components depend on past values of the time series, the shape of the predictive distributions for the next period may change from symmetric unimodal to skewed bimodal according to the market situation in the current period. With this property, a shift towards the left hand side of the distribution would allow the MVAR model to capture the increasing possibility of having a severe deterioration in default rates and macroeconomic environments in the next period. [18] On the other hand, a shift towards the right hand side would capture an increasing possibility of having an improvement in the next period.

---

[18]   If the number of components is more than two, the distribution can change to multimodal.

III.   DATA AND EMPIRICAL RESULTS

3.1   Default rate and macroeconomic variables

To illustrate our method, as discussed in Session I, we consider the default rates of Hong Kong's retail banks.[19] The change of the transformed probabilities is denoted by Δy. The equation system on default probability and macroeconomic dynamics is estimated by using retail banks' data covering the period from 1997 Q1 to 2007 Q3. In particular, the default rate is specified to depend on the following macroeconomic variables:

(i)   real GDP growth of Hong Kong (denoted by g_t^{HK}) - GDP governs the ability of agents in the economy to service their debt. For loans used to finance economic activities in the domestic market, the GDP of Hong Kong should be an important factor influencing the ability to repay;

(ii)   real interest rates in Hong Kong (r_t) - the interest rates directly affect the burden of the debt. We use the three-month HIBOR to represent nominal interest rates;[20]

(iii)   real property prices in Hong Kong ($prop_t$) - We consider property prices relevant because real estate is the major item of collateral. If the collateral value declines, the incentive to continue servicing the debt will weaken. The property price index compiled by the Rating and Valuation Department is used to calculate the variations in property prices in Hong Kong.[21]

The macroeconomic variables, real interest rates and real property prices, are $I(1)$, as suggested by the results of an augmented Dickey-Fuller test,[22] so we consider the change of real interest rates and the percentage change of real property prices in the regression. When the data and some selected macroeconomic variables are fitted by a

---

[19]   The default rate is measured as a ratio of the amount of loans which have been overdue for more than three months to the total amount of loans. The time series can be download at http://www.info.gov.hk/hkma/eng/statistics/msb/attach/T0306.xls. The data series on default rate is transformed by the logit formula to produce the y_t series. Results obtained from an augmented Dickey-Fuller test suggest that y_t is an I(1) process. Thus, we opt to model its first difference Δy_t

[20]   Real interest rates are calculated as $[(1+r_t)/(1+\pi_{t+1})]-1$, where r_t and π_{t+1} are the nominal interest rate in period t and the inflation rate in period $t+1$ respectively. This means that we assume agents are rational so that the difference between actual observed inflation and the expected (as of the previous period) rate is uncorrelated with current and past information. We use the seasonally adjusted CPI to calculate the inflation rate.

[21]   The real rate of change of property prices is calculated as $[(1+\triangle prop_t)/(1+\pi_t)]-1$, where $\triangle prop_t$ is the rate of change of nominal property prices in period t.

[22]   The p-values of augmented Dickey Fuller tests for the real interest rates and real property prices are found to be 0.35 and 0.69 respectively, suggesting to accept the null hypothesis of having unit root in these two time series.

VAR model that is based on the unimodal assumption the residuals are found to distribute bimodally (see Chart 3). Therefore, a MVAR model with two components is suggested to allow a bimodal distribution in the analysis.

Chart 3. Residuals of the transformed default probabilities using a VAR model
(from 1997 Q1 to 2007 Q3)
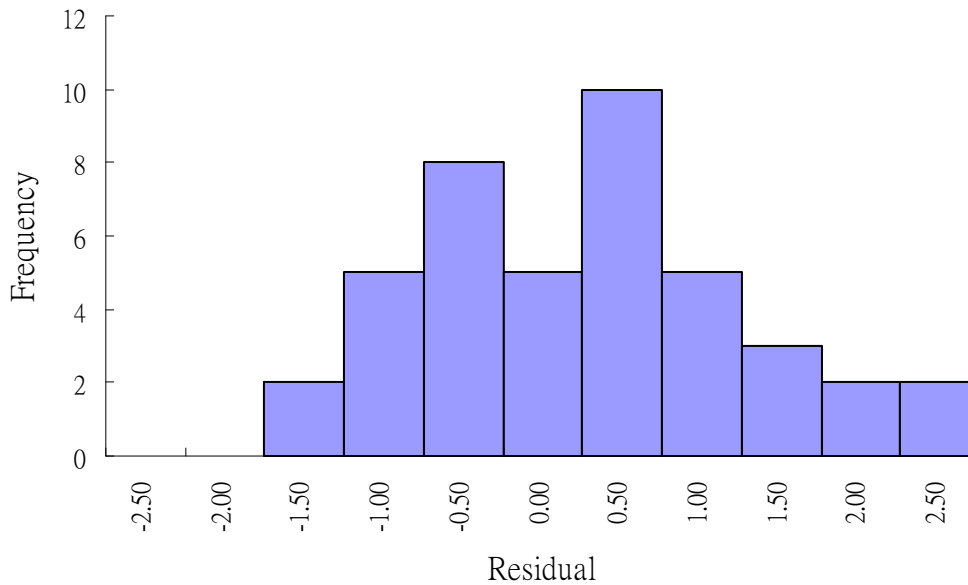
3.2 Estimation results and interpretations

Equation (3) is estimated by the expectation-maximization (EM) algorithm. [23] The results of coefficient estimates are summarized in Table 1. [24] The diagnostic test statistics reported in Table 2 show that the MVAR model is adequately fitted because the standardized residuals are neither serial correlated nor heteroskedastic (see Chart 4). In addition, there is no significant residual cross-correlation for the variable Δy_t and other variables. Moreover, the histogram for the residuals of the MVAR model plots a better bell shape than that of the VAR model, suggesting that the MVAR model is able to capture the extreme market values.

---

[23] An expectation-maximization (EM) algorithm finds maximum likelihood estimates of parameters in probabilistic models which depend on unobserved latent variables. EM alternates between performing an expectation (E) step, which computes an expectation of the likelihood by including the latent variables as if they were observed, and a maximization (M) step, which computes the maximum likelihood estimates of the parameters by maximizing the expected likelihood found on the E step. The parameters found on the M step are then used to begin another E step, and the process is repeated.

[24] The computation of standard errors can be done by Louis' method. Details can be found in Louis (1982), Lutkepohl (1991), and Chin and Wong (2008).

Based on the estimated MVAR model, the predictive distributions, $F(Z_t | \mathfrak{I}_{t-1})$, for the period during the Asian financial crisis (i.e. from 1997Q4 to 1998Q3) and after the crisis (from 1998Q4 to 1999Q3) are shown in Chart 5. The shapes of the predictive distributions are found to be changing over time, from unimodal to bimodal distributions during the crisis period, and finally back to unimodal distributions after the crisis period. For comparison, a VAR model is estimated. It can be seen that its unimodal distributions are usually less dispersed compared to the MVAR model, suggesting that the potential loss and the probability of occurring such loss using a unimodal distribution may be under estimated.

Table 1. Estimation results for the MVAR model (sample period: 1997 Q1 to 2007 Q3)

First Component : $\alpha_1 = 0.7403$ (0.0681)

VAR Part	Dependent Variable			
	Δy_t	g_t^{HK}	Δr_t	$\Delta prop_t$
Intercept	0.0076	-0.0183	-0.1982	**-2.2947****
	(0.0165)	(0.2496)	(0.2347)	(0.7254)
g_{t-1}^{HK}	**0.0320****	**0.6864****	0.0803	-0.2931
	(0.0116)	(0.1745)	(0.1643)	(0.5080)
Δr_{t-1}	**-0.0368****	-0.2983	-0.1100	**-2.2077****
	(0.0140)	(0.2110)	(0.1984)	(0.6135)
$\Delta prop_{t-1}$	**0.0063****	0.0413	0.0249	**0.7042****
	(0.0025)	(0.0373)	(0.0349)	(0.1084)
Δy_{t-1}	**0.1693***	-0.9743	**-2.3712***	**14.3326****
	(0.0973)	(1.4692)	(1.3830)	(4.2702)

Variance-Covariance Matrix	Matrix Variable			
	Δy_t	g_t^{HK}	Δr_t	$\Delta prop_t$
Δy_t	**0.0046****			
	(0.0012)			
g_t^{HK}	0.0082	**1.0478****		
	(0.0126)	(0.2676)		
Δr_t	**-0.0338****	**-0.3188***	**0.9302****	
	(0.0132)	(0.1861)	(0.2361)	
$\Delta prop_t$	0.0603	**1.9580****	-0.7399	**8.8515****
	(0.0380)	(0.6579)	(0.5314)	(2.2688)

Notes:
1. Standard errors estimated via Louis' method are in parentheses.
2. * and ** indicate the ratios of estimate over its standard error are larger than 1.645 and 1.96 in magnitude respectively.

Data Sources: CEIC, Census & Statistics Department of Hong Kong.

Table 1 (cont'). Estimation results for the MVAR model (sample period: 1997 Q1 to 2007 Q3)

Second Component : $\alpha_2 = 0.2597$ (0.0681)

VAR Part	Dependent Variable			
	Δy_t	g_t^{HK}	Δr_t	$\Delta prop_t$
Intercept	0.0403	**2.4770****	-0.0017	1.6250
	(0.0349)	(0.5117)	(0.2229)	(1.0682)
g_{t-1}^{HK}	-0.0200	**-0.4182****	-0.1005	**1.0390****
	(0.0131)	(0.1921)	(0.0837)	(0.4005)
Δr_{t-1}	0.0292	-0.2193	**-0.4094****	0.1250
	(0.0177)	(0.2608)	(0.1136)	(0.5437)
$\Delta prop_{t-1}$	**0.0136***	0.0216	**0.1048****	**0.3948***
	(0.0073)	(0.1065)	(0.0463)	(0.2219)
Δy_{t-1}	0.4351	**12.4094****	**4.9088****	-2.9275
	(0.3373)	(4.9511)	(2.1513)	(10.3006)

Variance-Covariance Matrix	Matrix Variable			
	Δy_t	g_t^{HK}	Δr_t	$\Delta prop_t$
Δy_t	**0.0076****			
	(0.0032)			
g_t^{HK}	0.0356	**1.6351****		
	(0.0354)	(0.7008)		
Δr_t	-0.0242	**-0.5224***	**0.3104****	
	(0.0164)	(0.2679)	(0.1332)	
$\Delta prop_t$	0.0801	1.2226	**-1.3137****	**7.1065****
	(0.0745)	(1.0983)	(0.6015)	(3.05)

Notes:
1. Standard errors estimated via Louis' method are in parentheses.
2. * and ** indicate the ratios of estimate over its standard error are larger than 1.645 and 1.96 in magnitude respectively.

Data Sources: CEIC, Census & Statistics Department of Hong Kong.

Table 2. Summary statistics of the standardized residuals for the MVAR model

Variable	Q^1	Q^2	Q^3	Q^4	Mean	SD	Min	Max
Δy_t	11.33	8.13	-	-	-0.00	1.01	-2.29	1.81
g_t^{HK}	1.77	3.43	8.90	4.85	-0.00	1.01	-2.16	2.22
Δr_t	5.02	5.70	0.96	1.93	0.00	1.01	-2.41	3.20
$\Delta prop_t$	3.73	1.85	7.14	6.30	-0.00	1.01	-1.86	2.28

Note: The Q^1 and Q^2 are the Portmanteau test statistics of the standardized residuals and the squared standardized residuals respectively, while Q^3 and Q^4 are the Portmanteau test statistics of the positive and negative cross-correlation between the standardized residuals of Δy_t and the selected variable. They are testing whether their estimated correlations are jointly equal to zero up to lag 4. At 0.05 and 0.01 levels of significance, the critical values of the test are about 9.48 and 13.3 respectively.

**Chart 4. Residuals of the transformed default probabilities using a MVAR model
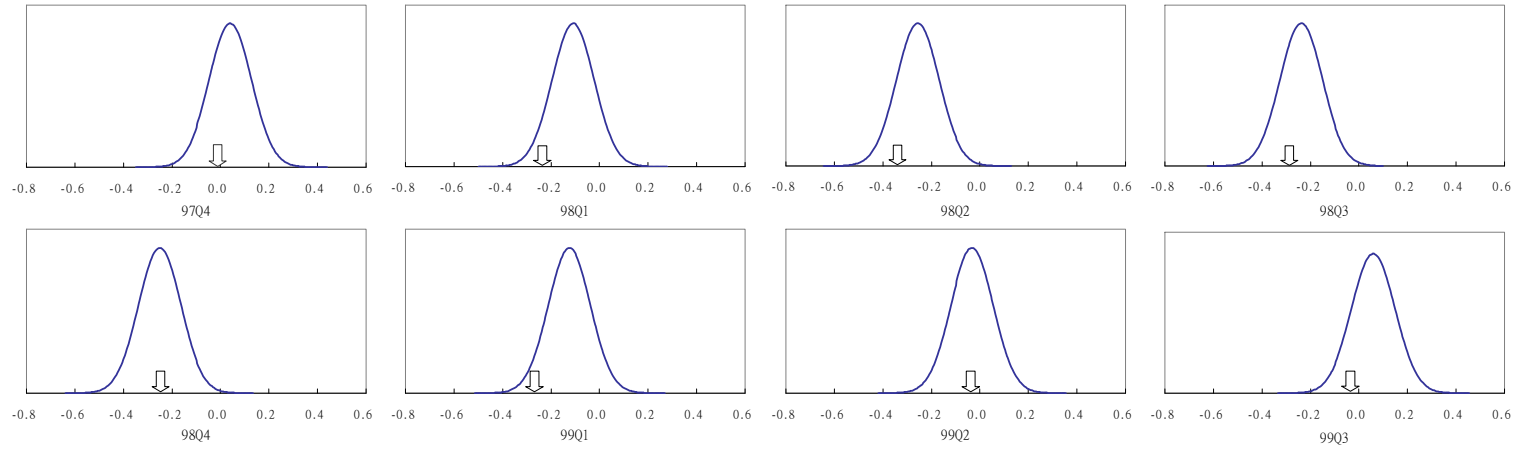(from 1997 Q1 to 2007 Q3)**

Chart 5. Comparison of the predictive distributions of Δy_t between the MVAR and VAR models (from 1997Q4 to 1998Q3)

<u>MVAR models</u>
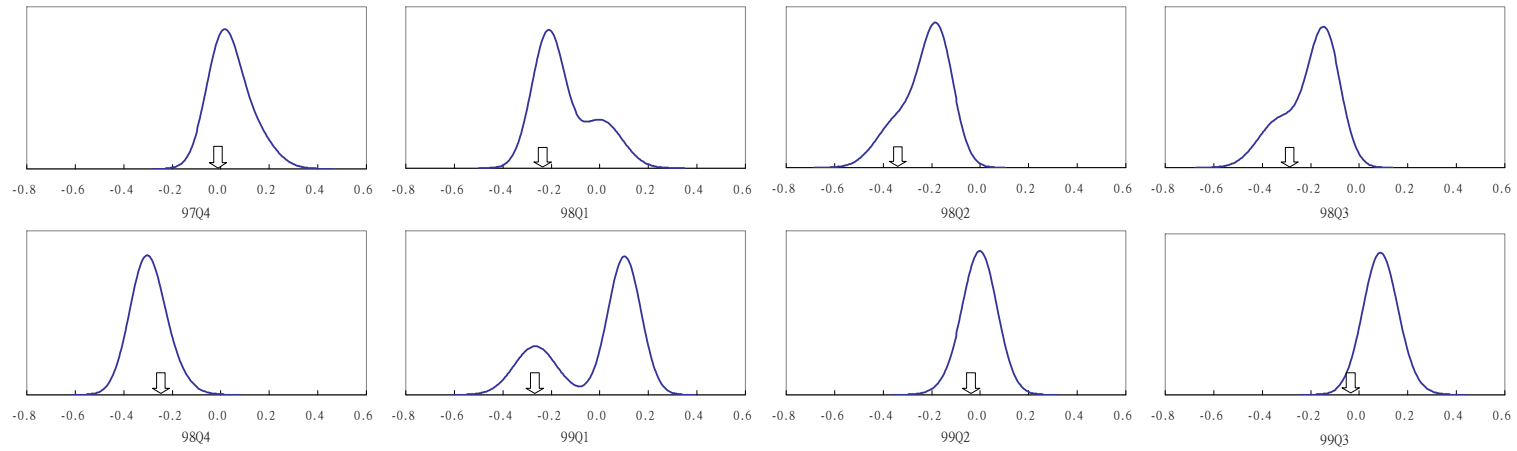

<u>VAR models</u>

Note: The arrows indicate the actual value of Δy_t

After the estimation process, each of the periods can be classified as either a normal or an abnormal market situation. Chart 6 plots the time series of $\triangle y$ and the selected macroeconomic variables. The impulse responses of variables given one-SD shock on GDP growth are depicted in Chart 7. Observations are highlighted when they are estimated to be under an abnormal condition. The results of the estimations are summarized as follows:

(i) Eleven out of 43 time points are estimated to be under the second component of the MVAR models. They are all found to be during periods of either severe economic contractions or significant rebounds from stressful conditions. The estimated probability of observing an abnormal market condition is 26%, while that of observing a normal market condition is 74%. Conditional on an abnormal market, a probability of 45% may be found for market condition being stressful and 55% being booming.

(ii) Under a regime of steady adjustment (based on the significant coefficients in the first mixture component of the MVAR model), the dynamics between the default variable and the macroeconomic variables are:

➢ The variable $\triangle y$ is positively correlated with real GDP growth and real property prices, and negatively correlated with real interest rates. In other words, lower real GDP growth, lower real property prices and/or higher real interest rates will worsen the default problem in the overall loan portfolio, and vice versa. In addition, the default status of overall loan portfolio in the previous period has a significant impact on the current default status. More defaults in the previous quarter will likely worsen the default problem in the current quarter, other things being equal.

➢ The loan default situation in the previous period affects interest rates and property prices in the current period through the feedback effect. An improved default condition in the previous quarter provides feedback to increase property prices and to decrease interest rates, resulting in a more favourable market condition in the current quarter.

(iii) Under a regime of volatile adjustment (based on the significant coefficients in the second mixture component of the MVAR model), the relationships become:

➢ The default probability is mainly driven by real property prices, suggesting that a negative change in property prices will directly worsen the default problem of the overall loan portfolio. In addition, the estimated variance of the default probability in the second component is larger than that in the first component, reflecting that default rate changes estimated by the second component are more volatile than those estimated by the first component.

➤ The loan default situation in the previous period affects GDP growth and interest rates in the current period through the feedback effect. An improved default condition in the previous quarter will increase both GDP growth and interest rates. However, property prices are estimated to be no longer affected by the feedback effect. They become affected mainly by GDP growth in the current period and its own lag.
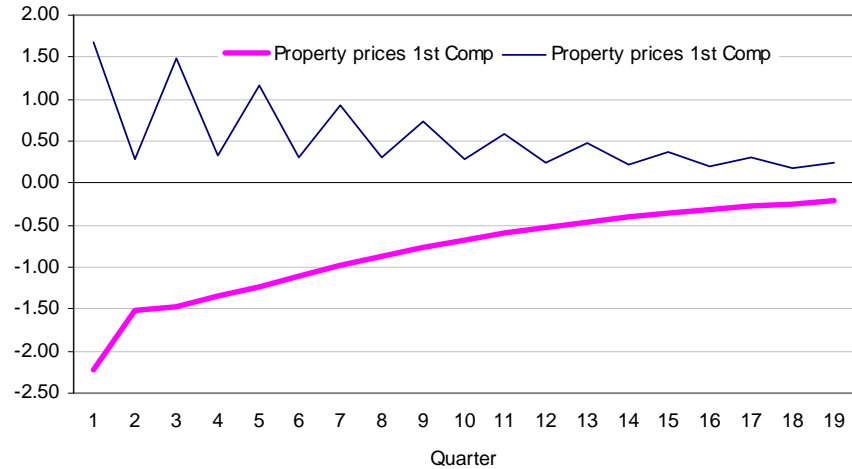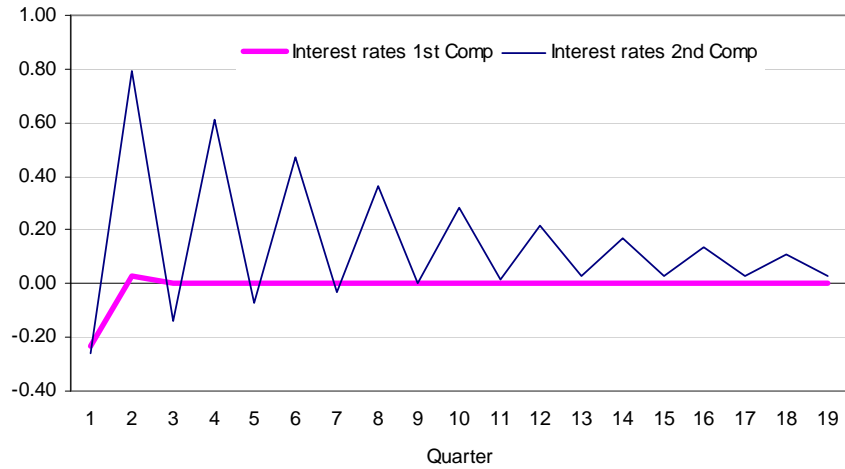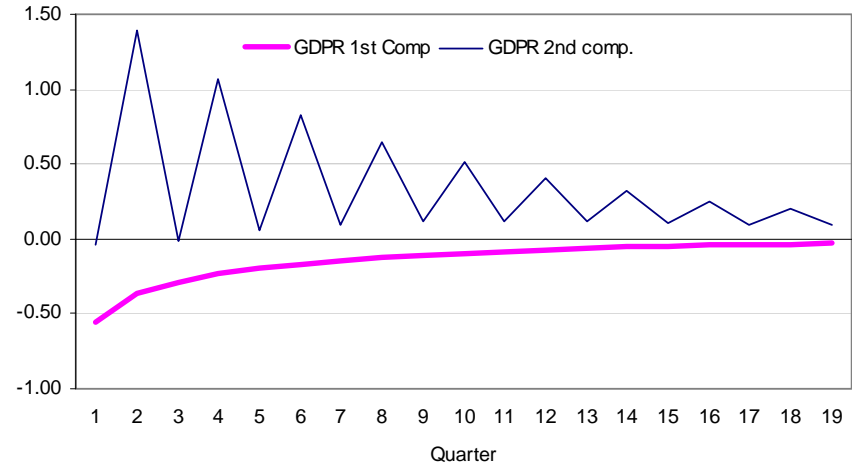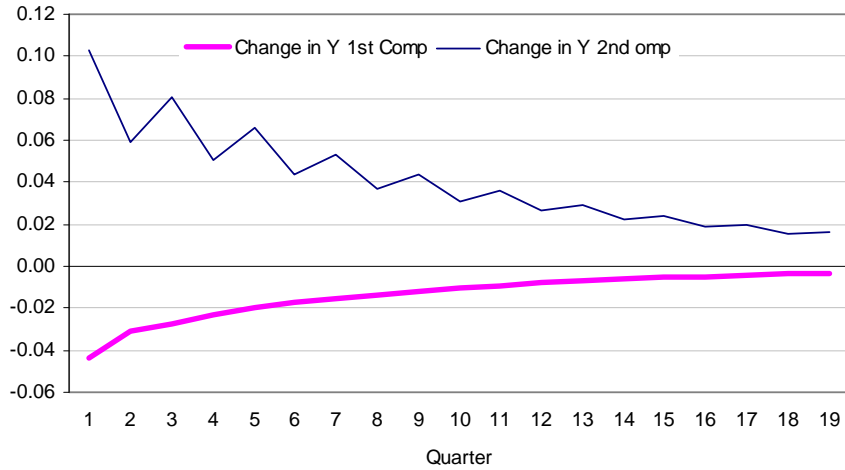
Chart 6. Time series of ΔY and the selected macroeconomic variables
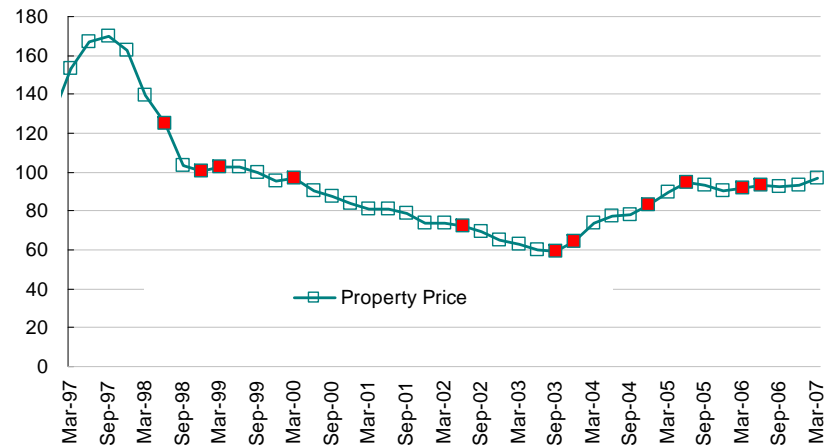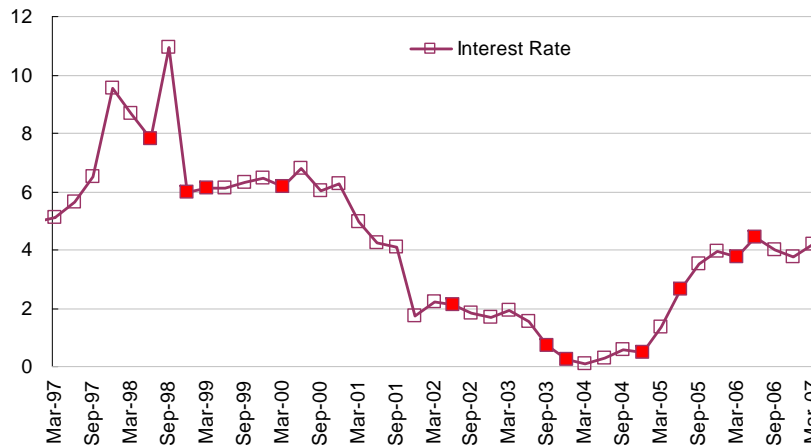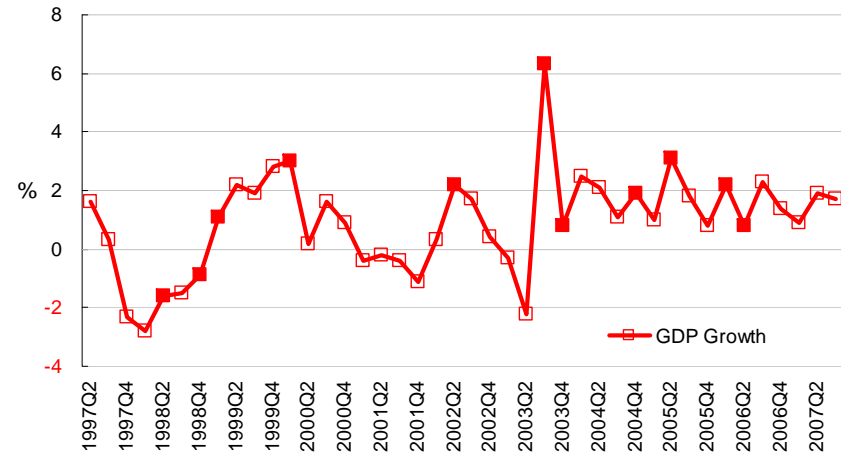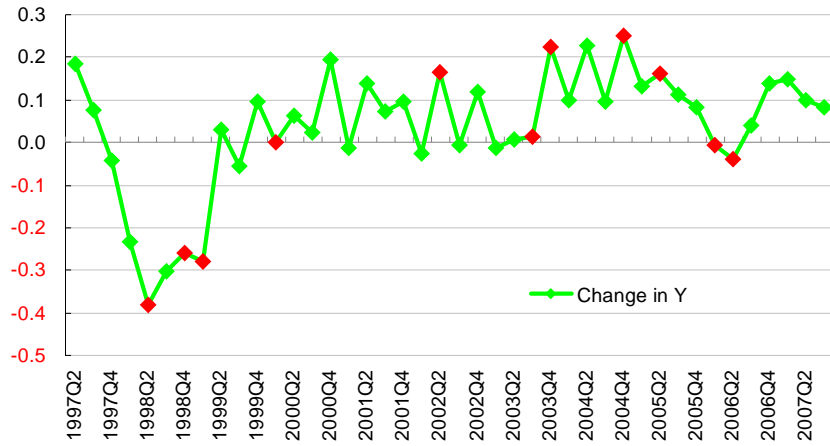
Note: Under the EM-estimation, observations in red colour are used to estimate for the second mixture component, while other observations are used to estimate the first mixture component.

Chart 7. Impulse responses of variables given a one-SD shock on GDPR

IV.   SIMULATION OF FUTURE CREDIT LOSSES AND STRESS-TESTING

This session focuses on the simulation of scenarios using the results of estimation in the previous session. Similar to the conventional approach, estimated frequency distribution of the horizon-end default rates for each sector is obtained from simulating a large number of future default rates.[25] For each period, the estimates of both the first and second components of the MVAR model are used in simulation. The selection between the two components is subject to a probability of mixture (i.e. α_1 and α_2). When the first component is chosen, the market will follow a normal trend. When the second component is chosen, the market will behave abnormally, either going up or down significantly. With this characteristic, the model can simulate the behaviours of the variables under a volatile market with a sharp rise of the credit loss.

Similar to Wong et al. (2006), four scenarios are considered. A baseline scenario is to introduce no artificial shock. Another three scenarios are to study stressed markets with three separate shocks on the three selected macroeconomic conditions. [26] Specifically, they are:

(i)   reductions in Hong Kong's real GDP by 2.3%, 2.8%, 1.6% and 1.5% respectively in each of the four consecutive quarters starting from 2007 Q4;

(ii)   a rise of real interest rates by 300 basis points in the first quarter, followed by no change in the second and third quarters and another rise of 300 basis points in the fourth quarter; and

(iii)   reductions in real property prices by 4.4%, 14.5%, 10.8% and 16.9% respectively in each of the four consecutive quarters starting from 2007 Q4.

These are quarter-to-quarter changes and are supposed to change separately from 2007 Q4 to 2008 Q3. Their magnitudes are in general similar to those during the Asian financial crisis. No further artificial shock is introduced for the subsequent quarters. For each of the baseline scenario and stressed scenarios, we simulate 100,000 future paths and use the simulated 100,000 default rates in 2009 Q3 to construct a frequency distribution of credit loss percentages.

To take into account the loss given default (*LGD*), we assume the *LGD* will vary with property price indices (*PI*) as properties are by far the most important collateral

---

[25]   A random vector of multivariate normal distribution can be obtained by first computing the Cholesky decomposition $\mathbf{C}$ of the variance-covariance matrix $\boldsymbol{\Sigma}$, where $\mathbf{C}$ is defined by $\boldsymbol{\Sigma} = \mathbf{CC'}$. Pre-multiplying a random vector $\mathbf{z}$ whose entries are independently drawn from the standardised normal distribution $N(0,1)$ by $\mathbf{C'}$ gives $\mathbf{r}$.

[26]   Shocks on each variable can pass to others through the use of the Cholesky decomposition matrix $\mathbf{C}$ of the variance-covariance matrix $\boldsymbol{\Sigma}$, where $\mathbf{C'}$ is a lower triangular matrix. When the first element of the vector $\mathbf{z}$ is shocked, other elements of the resulting vector $\mathbf{r} = \mathbf{C'z}$ will also change with the first element in the simulation, resembling a pass-through of shock to other elements in the vector $\mathbf{r}$.

for lending. The *LGD* in 2007 Q3 is assumed to be 0.5 and the *LGD* in 2009 Q3 to be approximately:

$$LGD_{2009Q3} = 0.5 - 0.5 \times \frac{PI_{2009Q3} - PI_{2007Q3}}{PI_{2007Q3}}.^{27}$$

Means and the VaR statistics are presented in Table 3. Chart 8 plots the distributions of credit losses. For the period from 2007 Q4 to 2009 Q3, the credit losses based on the MVAR model are found to be generally larger than the results based on the model of Wong et al. (2006) which suggested an unimodal distribution for the time series.[28] The averages of banks' maximum credit losses are found to increase from a range of 0.22% to 1.08% to a range of 0.24% to 2.51% -- with the maximum loss almost doubled. Under the extreme case for the VaR at the confidence level of 99.99%, the previous estimated losses for the stressed scenario ranged from 6.17% to 7.13%.[29] However, the probability of occurrence of such a loss can be higher based on the MVAR model. For instance, the probability of occurring even a higher loss of around 9% of the portfolio under similar interest rate shock is estimated to be much higher at 5%.

V.   CONCLUDING REMARKS

This paper studied a macro stress testing framework for loan portfolios of banks in Hong Kong. The MVAR model is considered which can provide a mixture of two separate VAR models for different market conditions. Our analysis suggests that the mixture property of default rates cannot be ignored in modelling the relationship between the default rates of bank loans and key macroeconomic factors, including Hong Kong's real GDP, real interest rates, and real property prices.

Macro stress testing is then performed to assess the vulnerability and risk exposures of banks' overall loan portfolios. By using the framework, a Monte Carlo simulation is applied to estimate the distribution of possible credit losses conditional on the market situation. Our model allows the shocks originated from several macroeconomic variables. The results show that the credit risk of the banks' loan portfolio simulated based on a unimodal assumption can be potentially under-estimated. Based on the estimates of the MVAR model, the mean credit losses are found to range from 0.24% to 2.51%, which double the credit losses estimated based on a unimodal assumption.

---

[27] If no formal statistics are available for the loss given default (LGD), some studies assign a rough constant ratio based on market information to obtain the estimated credit loss. If no market information is available, a ratio of 0.5 may be assumed for the calculation of loss figures. In this paper, we assume the LGD will vary with property prices as properties are by far the most important collateral for lending. Property prices should therefore have an impact on how much banks can recover from their losses.

[28] The mean and VaR statistics of simulated credit loss distribution for the period from 2007Q2 to 2009Q1 can be found in the Half-yearly Monetary and Financial Stability Report June 2007.

[29] The estimated credit losses using unimodal distribution can be found in the Half-yearly Monetary and Financial Stability Report December 2007.

Table 3. The mean and VaR statistics of simulated credit loss distributions

| Credit loss (%)	Baseline scenario	Stressed scenarios		
		GDP shock[a]	Property price shock[b]	Interest rate shock[c]
Mean	0.24	1.04	1.64	2.51
VaR at 55% CL[d]	0.17	0.94	1.47	1.34
VaR at 60% CL	0.20	1.02	1.59	1.57
VaR at 65% CL	0.22	1.11	1.72	1.85
VaR at 70% CL	0.26	1.20	1.87	2.21
VaR at 75% CL	0.30	1.32	2.05	2.63
VaR at 80% CL	0.36	1.47	2.27	3.25
VaR at 85% CL	0.44	1.65	2.55	4.15
VaR at 90% CL	0.56	1.93	2.96	5.73
VaR at 95% CL	0.78	2.40	3.70	9.41
VaR at 99% CL	1.53	3.57	5.65	22.48
VaR at 99.9% CL	2.72	5.58	9.06	45.19
VaR at 99.99% CL	4.20	7.90	13.46	65.27

Notes: a) Reductions in Hong Kong's real GDP by 2.3%, 2.8%, 1.6% and 1.5% respectively in each of the four consecutive quarters starting from 2007 Q4.
b) Reductions in real property prices by 4.4%, 14.5%, 10.8% and 16.9% respectively in each of the four consecutive quarters starting from 2007 Q4.
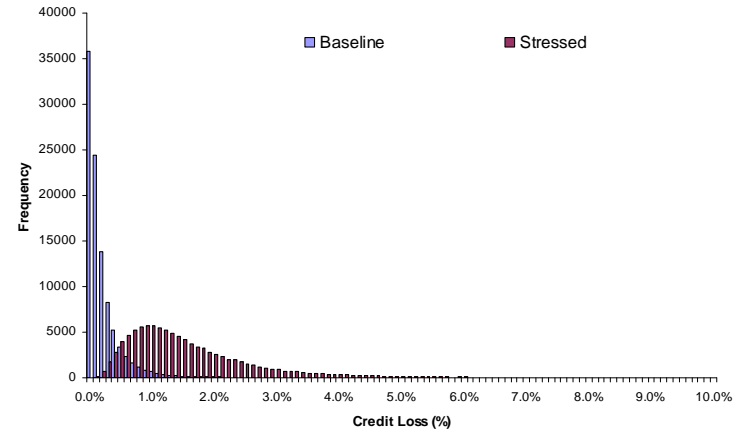c) A rise of real interest rates by 300 bps in the first quarter, followed by no change in the second and third quarters and another rise of 300 bps in the fourth quarter.
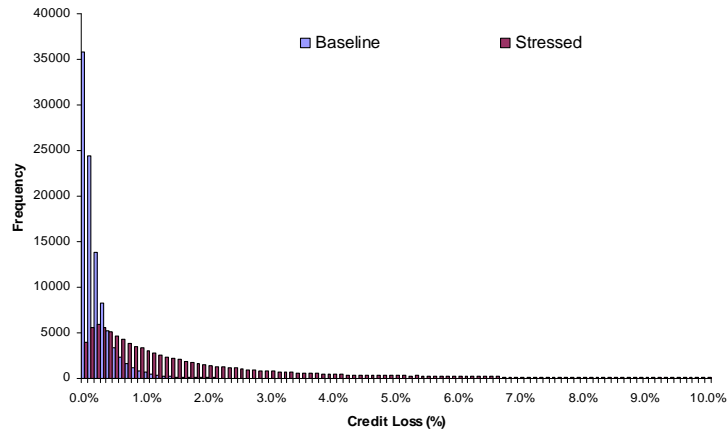d) CL denotes the confidence level

Chart 8. Simulated frequency distributions of credit loss under baseline and stressed scenarios
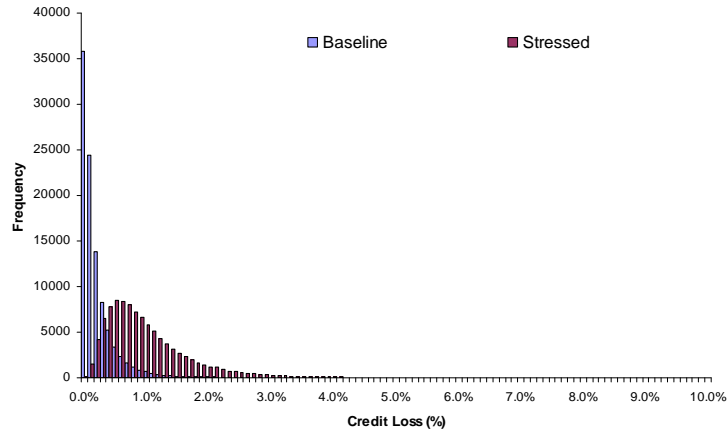
GDP Shock

Property price shock

Interest rate shock

Note: Each distribution is constructed with 100,000 simulated future paths of default rates.

REFERENCES

Berchtold, A. (2003). "Mixture transition distribution (MTD) modeling of heteroskedastic time series". *Computational Statistics and Data Analysis*, 41, 3-4, 399-411.

Boss, M. (2002). "A Macroeconomic Credit Risk Model for Stress Testing the Austrian Credit Portfolio," *Financial Stability Report* 4, Oesterreichische Nationalbank.

Bunn, P. Cunningham, A. and Drehmann, M. (2005). "Stress Testing As a Tool for Assessing Systemic Risks," *Financial Stability Review*, June 2005, Bank of England.

Chin, H.Y. and Wong, C.S. (2008). "Constrained Estimation of Mixture Vector Autoregressive Model", unpublished manuscript.

Fong, P.W., Li, W.K., Yau, C.W., and Wong, C.S. (2007). "On a Mixture Vector Autoregressive Model", *The Canadian Journal of Statistics*, 35,1, pp. 135-150.

Froyland E. and Larsen K. (2002). "How Vulnerable are Financial Institutions to Macroeconomic Changes? An Analysis Based on Stress Testing," *Economic Bulletin.* October 2002, Norges Bank.

Gasha, J. Giancarlo and Morales, R. Armando (2004) "Identifying Threshold Effects in Credit Risk Stress Testing", IMF Working Paper No. WP/04/150

Hugh Thomas and Zhiqiang Wang (2005) "Interpreting the Internal Ratings-Based Capital Requirements in Basel II", *Journal of Banking Regulation*, 6, 3, 274-289.

Hoggarth, G. and Whitley, J. (2003). "Assessing the Strength of UK Banks through Macroeconomic Stress Tests," *Financial Stability Review*, June 2003, Bank of England.

Lanne, M. (2004). "Nonlinear dynamics of interest rate and inflation". Bank of Finland Discussion Papers.

Lanne, M. and Saikkonen, P. (2003). "Modeling the U.S. short-term interest rate by mixture autoregressive processes". *Journal of Financial Econometrics*, 1, 96-125.

Le, N. D., Martin, R. D. and Raftery, A. E. (1996). "Modeling flat stretches, bursts, and outliers in time series using mixture transition distribution models". *Journal of the American Statistical Association*, 91, 1504-1514.

Louis, T. A. (1982). "Finding the observed information matrix when using the EM algorithm". *Journal of the Royal Statistical Society. Series B*, 44 (2), 226-233.

Lütkepohl, H. (1991). *Introduction to multiple time series analysis*. Berlin: Springer.

Martin, V.L. (1992). "Threshold time series models as multimodal distribution jump processes". *Journal of Time Series Analysis*, 13, 79-94.

Mawdsley, A., McGuire, M. and O'Donnell, N. (2004). "The Stress Testing of Irish Credit Institutions," *Financial Stability Report*, 2004, Central Bank and Financial Services Authority of Ireland.

Sorge, M. (2004). "Stress-testing Financial Systems: An Overview of Current Methodologies," BIS Working Papers, no. 165.

Virolainen, K. (2004). "Macro Stress-testing with a Macroeconomic Credit Risk Model for Finland," Bank of Finland Discussion Paper, no. 18/2004.

Wilson, T. C. (1997a). "Portfolio Credit Risk (I)," *Risk*, vol. 10, issue 9, pp. 111-17.

Wilson, T. C. (1997b). "Portfolio Credit Risk (II)," *Risk*, vol. 10, issue 10, pp. 56-61.

Wong, J., Choi, K.F., and Fong, T. (2006). "A Framework for Stress testing Banks' Credit Risk," HKMA Research Memorandum,15.

Wong, C. S. and Li, W.K. (2000). "On a mixture autoregressive model". *Journal of the Royal Statistical Society*, Series B, 62, 95-115.

Wong, C.S. and Li, W.K. (2001). "On a mixture autoregressive conditional heteroskedastic model". *Journal of the American Statistical Association*, 96, 982-995.