CAHIER D'ÉTUDES
WORKING PAPER

N° 75

AN EARLY-WARNING AND DYNAMIC FORECASTING FRAMEWORK OF DEFAULT PROBABILITIES FOR THE MACROPRUDENTIAL POLICY INDICATORS ARSENAL

Xisong JIN

Francisco NADAL DE SIMONE

JUNE 2012

BANQUE CENTRALE DU LUXEMBOURG
EUROSYSTÈME

An Early-warning and Dynamic Forecasting Framework of Default Probabilities for the Macroprudential Policy Indicators Arsenal

Xisong Jin
Luxembourg School of Finance

Francisco Nadal De Simone
Banque centrale du Luxembourg

Draft: June 13, 2012

Abstract

The estimation of banks' marginal probabilities of default using structural credit risk models can be enriched incorporating macro-financial variables readily available to economic agents. By combining Delianedis and Geske's model with a Generalized Dynamic Factor Model into a dynamic t-copula as a mechanism for obtaining banks' dependence, this paper develops a framework that generates an early warning indicator and robust out-of-sample forecasts of banks' probabilities of default. The database comprises both a set of Luxembourg banks and the European banking groups to which they belong. The main results of this study are, first, that the common component of the forward probability of banks' defaulting on their long-term debt, conditional on not defaulting on their short-term debt, contains a significant early warning feature of interest for an operational macroprudential framework driven by economic activity, credit and interbank activity. Second, incorporating the common and the idiosyncratic components of macro-financial variables improves the analytical features and the out-of-sample forecasting performance of the framework proposed.

JEL Classification: C30, E44, G1

Keywords: financial stability; macroprudential policy; credit risk; early warning indicators; default probability, Generalized Dynamic Factor Model; dynamic copulas; GARCH.

The views expressed in this paper are the authors' and do not necessarily represent those of the BCL or its staff. We thank participants at the BCL Financial Stability Review seminar in April 2012 for their comments, in particular Paolo Guarda, Abdelaziz Rouabah and Dominique Marchal as well as the FNR for its financial support. All errors and omissions remain our own. Correspondence can be sent to: Xisong Jin, Luxembourg School of Finance, 4 Rue Albert Borschette L-1246 Luxembourg, Tel: (+352) 46 66 44 5626; E-mail: xisong.jin@uni.lu; Francisco Nadal De Simone, Banque centrale du Luxembourg, 2, boulevard royal L-2983 Luxembourg, Tel: (+352) 4774-4518; E-mail: Francisco.NadaldeSimone@bcl.lu.

I. Motivation

A relatively broad characterization of the objective of macroprudential policy is to limit systemic risk so as to minimize the costs of financial instability on the economy (ECB, June 2010). The literature on financial system risk has made a distinction between three different sources of systemic risk (ECB, December 2009): first, the exposure of all financial institutions to common, simultaneous macro-financial shocks; second, the sequential contagion from an idiosyncratic shock affecting a financial institution that spreads to other financial institutions and eventually to the real sector of the economy and; third, financial imbalances that build up over time and may unravel in a disorderly manner. Limiting financial systemic risk requires having indicators that provide a measure, albeit "fuzzy", of financial stability, and a set of instruments to maintain and restore financial stability, when it is perturbed (Borio and Drehmann, 2009). Like the sources of systemic risk, indicators of systemic risk cover the cross-sectional dimension of systemic risk (e.g., Segoviano and Goodhart, 2009) and the time-dimension of systemic risk (e.g., Borio and Lowe, 2002). This paper contributes to several strands of the literature on both dimensions of systemic risk. Its objective is to develop a framework that generates an early warning indicator of overall credit risk in the banking sector that identifies as early as possible the build up of endogenous imbalances; that recognizes exogenous shocks timely; that factors in some manner dependence among financial institutions and; that provides robust out-of-sample forecasts of probabilities of default.[1]

One of the biggest challenges for credit risk models is modelling dependence between default events and between credit quality changes. Dependence modelling is necessary to understand the risk of simultaneous defaults, the ensuing distribution of losses and the effects on financial stability. Failing to account for dependence, therefore, underestimates potential losses (Lando, 2004). This is crucial for meaningful stress testing exercises, for instance, as well as more generally, for the development of measures of systemic risk. To incorporate dependence, there are basically three broad approaches or mixtures of them: (1) to let probabilities of default be affected by common underlying observable variables; (2) to let probabilities of default be affected by underlying latent variables and; (3) to let direct contagion from a default event affect other firms. However, whether by using a mixture of distributions to model dependence or by using copula or network analysis, models require the estimation of default probabilities as a first step. This study uses two of the structural credit risk models studied in Jin and Nadal De Simone (2011a) and Jin *et al* (2011b), Merton (1974) model and Delianedis and Geske (2003) model, to estimate implied neutral probabilities of

[1] The issue of financial institutions' contributions to systemic risk is not addressed in this paper, but in an accompanying study.

default. To model dependence among financial institutions' default probabilities, this paper uses the Generalized Dynamic Factor Model (GDFM) of Forni *et al* (2005), which has been used extensively to exploit the information from a large dataset and also for forecasting (e.g., Kabundi and Nadal De Simone, 2011, De Nicolò and Lucchetta (2012), and D'Agostino and Giannone, forthcoming).[2] However, Forni *et al* (2003) forecasting method is not easily applicable to a large number of underlying assets simultaneously, and does not generate the distributions of forecasts. As a result, this paper introduces a novel approach that combines the GDFM with a dynamic t-copula to improve the GDFM forecasting capacity.

Copula theory provides an easy way to deal with (otherwise) complex multivariate modeling (Jin and Lehnert, 2011). The advantage of the copula approach is its flexibility, because the dependence structure can be separated from the univariate marginal components, and hence the dependence structure between these marginal components can be modeled in the second stage, after the univariate distributions have been calibrated. Therefore, the copula approach provides a robust and consistent method to estimate banks' dependence. Correlation analysis, which usually refers to linear correlation, depends on both the marginal distributions and the copula, and is not a robust measure given that a single observation can have an arbitrarily high influence. The conditional dynamic t-copula is relatively easy to construct and simulate from multivariate distributions built on marginals and dependence structure. In fine, the GARCH-like dynamics in the copula variance and rank correlation offers multi-step-ahead predictions of the estimated GDFM common and idiosyncratic components simultaneously.

The framework of this study, therefore, shares the main core features suggested for an appropriate measure of systemic risk according to Schwaab *et al* (2010): a broad definition of systemic risk such as the ECB's, an international focus, the incorporation of macroeconomic and financial conditions, unobserved factors, and the calculation of probabilities of defaults.

The main results and contributions of this paper to the time-dimension of systemic risk are, first, to show that the common component of the forward probability of banks' defaulting on their long-term debt, conditional on not defaulting on their short-term debt, contains a significant early warning feature of interest for an operational macroprudential

[2] Mechanisms for obtaining default dependence are versions of, and possible mixtures of three issues: (1) PDs are influenced by common observable variables and there must be a way of linking the joint movement of a reduced set of factors and how PDs depend on them; (2) PDs depend on unobserved background variables and credit events result in an update of the latent variables which updates PDs and; (3) direct contagion from a credit event.

framework mostly driven by measures of economic activity, credit growth and interbank activity. This is in the tradition recently surveyed by Frankel and Saravelos (2010). As such, the proposed framework measures the relative riskiness of the system in a non-crisis mode. It becomes a useful macroprudential gauge to help policymakers decising when and what remedial actions to take as systemic risk increases over time. Second, that incorporating the common and the idiosyncratic components of macro-financial variables improves the analytical features of the framework proposed, in agreement with recent work by Koopman *et al* (2010) and Schwaab *et al* (2010). Finally, and a novel contribution, the paper's framework produces robust out-of-sample forecasting of overall banking sector credit risk, especially at the individual bank level.

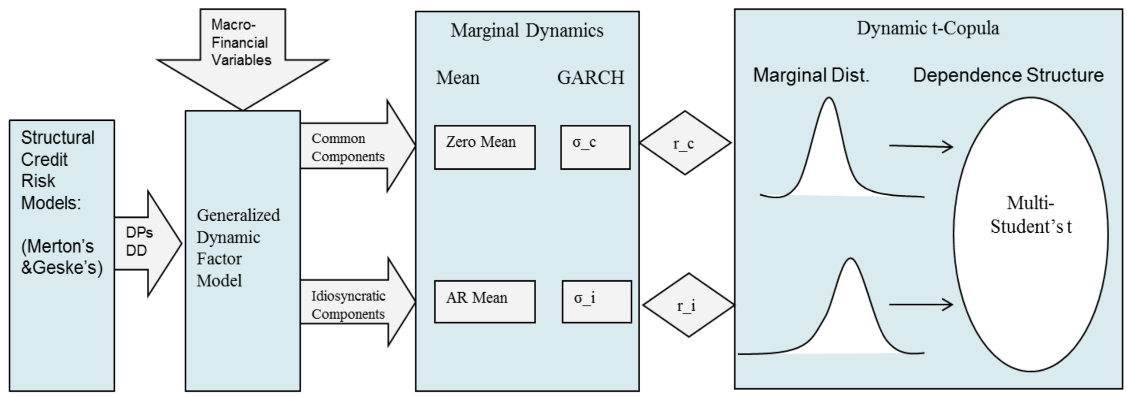
The remainder of the study is organized as follows. Next section introduces a novel integrated modeling framework, and explains how to combine the GDFM with a dynamic t-copula into a dynamic forecasting framework of default probabilities. Section III discusses the data, and section IV examines the empirical results. Section V concludes.

II. An Integrated Modeling Framework

The purpose of this framework is to break down complex information into several smaller, less complex and more manageable sub-tasks that are solvable by using existing tools, and then combining their solutions in order to solve the original problem.

The decomposition approach is frequently used in statistics, operations research and engineering. For example, decomposition of time series is considered to be a practical way to improve forecasting. The usual decomposition into trend, cycle, seasonal and irregular components was motivated mainly by business analysts, who wanted to get a clearer picture of the state of the economy (Fisher, 1995). Ideally, the selected models are expected to be integrated into the same theoretical framework. However, the models are developed to solve specific questions from different strands of literature. For instance, the structural credit risk model is a model for assessing credit risk typically developed from option pricing literature; dynamic factor models have become a standard econometric tool to perform factor analysis on large datasets for measuring comovements in, and forecasting of macroeconomic time series; copulas are a fundamental tool for modeling multivariate distributions, which is extremely useful for active risk management. Despite the difficulties involved in integrating these models, there are already some examples of using the existing tools together one way or another. For example, De Nicolò and Lucchetta (2012) use a dynamic factor model with many predictors combined with quantile regression techniques. Alessi, Barigozzi and Capasso (2007a&b) propose two new methods for volatility forecasting, which combine the GDFM and the GARCH model, and have been proved to outperform the standard univariate GARCH in most cases by exploiting cross-sectional information.

This paper presents an integrated framework which examines credit risk emanating from the macro environment and from banks' interconnectedness. The first contribution of the paper is to generate an early warning framework that in the tradition of Borio and Lowe (2002), associates the buildup of banking sector vulnerabilities with the real economy cycle and credit growth. The text graph illustrates the information flow of the first part of this framework:

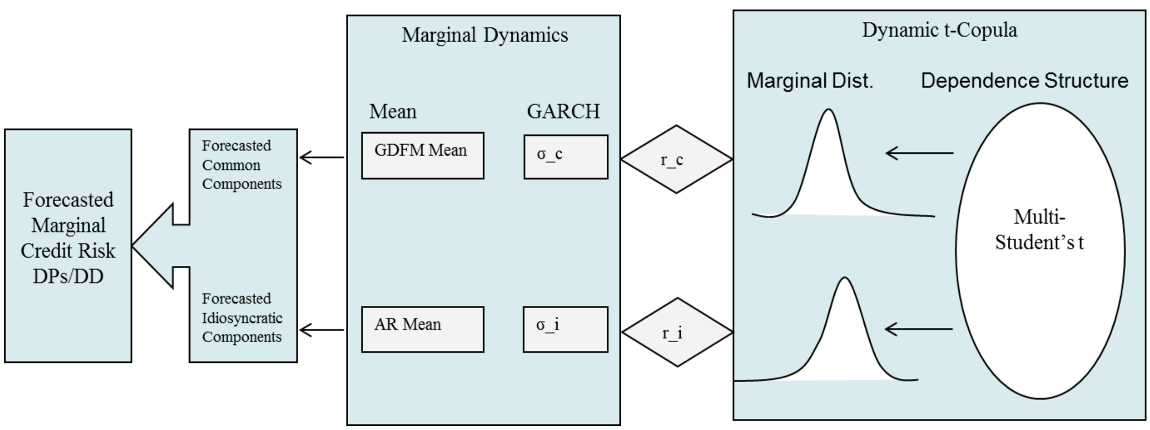


1. The market-based / book value-based Merton and Delianedis and Geske models are used to filter out the tail risks or the probabilities of distress (probabilities of default or PDs, and distance to default or DD);
2. The probabilities of distress together with a large database of macro-financial variables are decomposed into common components and idiosyncratic components by the GDFM;
3. Those components are then broken down into their means and volatilities by the marginal dynamics of AR-GARCH models. For the in-sample estimation, a zero mean is assumed for the common components in order to keep their multi-step-ahead prediction from the GDFM;
4. The standardized residuals from the marginal dynamics, which are $iid(0,1)$ usually with skewness and fat tails, are glued together by a dynamic t-copula with a multivariate GARCH structure;[3]

[3] The converse of Sklar's theorem implies that it is possible to couple together any marginal distributions, of any family, with any copula function and a valid joint density will be defined. The corollary of Sklar's theorem is that it is possible to extract the implied copula and marginal distributions from any joint distribution (Nelsen, 1999). This framework alleviates the statistical inefficiency associated with the fact that PDs (and DD) are generated regressors.

5. By the copula approach, the standardized residuals can be further decomposed into two subsets of information: (i) information of each random variable; i.e., the marginal distribution of each variable; and (ii) information about the dependence structure (nonlinear) among the random variables.

The second main contribution of this paper is to advance a dynamic forecasting framework of credit risk for each bank by simulation from multivariate distributions built on marginal distributions and dependence structure. As shown by the following text graph, the simulation is actually an information loading process through the dynamic structures built in the first step. The forward dependence information is first generated from a multi-student's t copula, and then marginal information is loaded up to get the forward standardized residuals. The forecasted common components and idiosyncratic components are projected by plugging-in marginal dynamics which enables customizing the information of means and volatility clusters. Last, the forecasted marginal credit risk measures are the sum of these two components. Thus, reverse engineering uncovers the tail risk or the probabilities of distress by using not only information from individual banks, but also from a large data set of macro-financial variables.



The remainder of this section reviews the methodological and statistical approaches used to estimate credit risk. First the selected models to estimate default probabilities are briefly described, and then the GDFM to nest macro-financial variables is outlined. Last, the multivariate GARCH techniques are extended into the t-copula to introduce the dynamic forecasting framework.

1. Selected models to estimate default probabilities

In order to develop tools to measure and assess financial stability it is necessary to characterize instability. Approaches to deal with instability include, for instance, modelling financial institutions' default, analysing the financial system using extreme value theory, and allowing for episodes of market illiquidity. The approach taken in this study instead is to apply contingent claim analysis to the analysis and measurement of credit risk, or as it is commonly referred to, structural credit risk modeling. Structural credit risk models attempt to assess the creditworthiness of a firm by modeling the evolution of the firm's asset values as a stochastic process, and by viewing bankruptcy as an endogenous random event linked to the value of the firm's assets.

1.1. The Merton Model

In the Merton model, equity owners are viewed as holding a call option on the firm's value after outstanding liabilities have been paid off. They have the option to default if the firm's asset value falls below the present value of the notional amount—or book value—of outstanding debt ("strike price") owed to bondholders at maturity. In other words, when the market value of the firm's assets is less than the strike price, the value of equity is zero. Similarly, bond holders are viewed as writing a European put option to equity owners, who hold a residual claim on the firm's asset value if the firm does not default. Bond holders receive a put option premium in the form of a credit spread above the risk-free rate in return for holding risky corporate debt (and bearing the potential loss) due to equity owners' limited liability.

According to the Merton model, the market value of a firm's underlying assets follows a geometric Brownian motion (GBM) of the form:

$$dV_A = \mu V_A dt + \sigma_A V_A dW$$

where V_A is the firm's assets value, with an instantaneous drift μ (the expected rate of return of the firm), and an instantaneous asset return volatility σ_A. W is a standard Wiener process. If X is the book value of the debt which is due at time *T*, Black and Scholes' formula provides the market value of equity, V_E:

$$V_E = V_A N(d_1) - Xe^{-r(T-t)} N(d_2)$$

where $d_1 = \dfrac{ln(\frac{V_A}{X}) + (r + \frac{1}{2}\sigma_A^2)(T-t)}{\sigma_A \sqrt{T-t}}$, $d_2 = d_1 - \sigma_A \sqrt{T-t}$, r is the risk-free rate, and $N()$ is the cumulative density function of the standard normal distribution. The PD of the firm is the probability that its assets' value will be less than the book value of its liabilities. The corresponding implied neutral PD is $\pi_N = N(-d_2)$. The "actual" PD is $\pi_A = N(-DD)$, where the distance-to-default, DD, is simply the number of standard deviations that the firm is away from default:

$$DD = \frac{ln(\frac{V_A}{X}) + (\mu - \frac{1}{2}\sigma_A^2)(T-t)}{\sigma_A \sqrt{T-t}}$$

To get to "actual" PDs from neutral PDs, the latter must be adjusted by the market price of risk, which is estimated using a capital-asset pricing model in this study: $u = \rho_{A,M} \frac{\sigma_A}{\sigma_M}(u_M - r) + r$, where σ_M is market asset return volatility, and $\rho_{A,M}$ is the correlation between firm's asset return and market asset return. Alternatively, historical recovery rates can be used to move from risk neutral to "actual" PDs. The derived "actual" PDs, however, could be still much higher than the observed PDs, the so-called "credit spread puzzle" (Huang and Huang, 2003). Moody's KMV, instead, maps distance to default (DD) into historical default probabilities. Chen, Collin-Dufresne and Goldstein (2009) try to adjust the level by calibrating the pricing kernel to equity returns and aggregate consumption. Fortunately, the rankings are more meaningful than the levels, given the objective of this study. A complication of CCA to calculate PDs is that the dynamics of the underlying asset value is not directly observable. To calculate σ_A, Moody's KMV iterative procedure is used.[4] For quoted financial institutions, the KMV approach implies taking daily equity data from the past 12 months to calculate historical assets volatility.[5] Regarding the value of debt, the KMV approach takes all debt due in one year, plus half of the long-term debt. The KMV method is a simple two-step iterative algorithm to solve for assets volatility. The procedure uses an initial guess for volatility to determine the asset value and to de-lever the equity returns. The volatility of the resulting asset returns is used as the input to the next iteration of the procedure which, in turn, determines a new set of asset values and hence a new series of asset returns. The procedure continues in this manner until it converges.

1.2. *The Delianedis and Geske Compound Option-based Risk Model*

Debt maturity influences liquidity risk and PDs. However, Merton model and most credit risk models consider only a single debt maturity. This is an important drawback for a

[4] Duan et al, (2004) show that the KMV estimates are identical to maximum likelihood estimates (MLE).
[5] See next sub-section for the approach followed in the case of Luxembourg banks for which quoted stock prices or options on stock are not available.

central bank or a supervisor interested in assessing and tracking banks' solvency. Geske (1977) and Delianedis and Geske (2003) consider a multi-period debt payment framework to which they apply compound option theory. This enables to account for the influence of the time structure of debt on the estimated PD.

Assume that a bank has long term debt, M_2, which matures at date T_2, and short term debt, M_1, which matures at date T_1. Between T_1 and T_2, the Merton model is valid as the bank's equity equals a call option giving the shareholder the right to buy the bank at the second payment date, T_2, by paying the strike price M_2. If at date T_1, the call option with the bank's value $\overline{V}$ equals at least the face value of the short term debt, M_1:

$$M_1 = \overline{V}N(k_2 + \sigma_A\sqrt{T_2 - T_1}) - M_2 e^{-r_{F1}(T_2 - T_1)} N(k_2)$$

then the bank can roll over its debt. So, the refinancing problem, the right to buy the simple call option of the second period by paying the strike price at the first payment date, is exactly a compound option as follows:

$$V_E = V_A N_2(k_1 + \sigma_A\sqrt{T_1 - t}, k_2 + \sigma_A\sqrt{T_2 - t}; \rho) - M_2 e^{-r_{F2}(T_2 - t)} N_2(k_1, k_2; \rho) - M_1 e^{-r_{F1}(T_1 - t)} N(k_1)$$

where $\rho = \sqrt{\frac{T_2 - t}{T_1 - t}}$, $N_2()$ is a bivariate cumulative normal distribution, and,

$$k_1 = \frac{ln(\frac{V_A}{\overline{V}}) + (r_{F_1} - \frac{1}{2}\sigma_A^2)(T_1 - t)}{\sigma_A\sqrt{T_1 - t}}, \; k_2 = \frac{ln(\frac{V_A}{M_2}) + (r_{F_2} - \frac{1}{2}\sigma_A^2)(T_2 - t)}{\sigma_A\sqrt{T_2 - t}}.$$

The richness of the model allows to calculate the following risk neutral PDs: (1) the total or joint probability of defaulting at either date T_1 or date T_2, i.e., $1 - N_2(k_1, k_2; \rho)$; (2) the short-run probability of only defaulting on the short-term debt at date T_1, i.e., $1 - N(k_1)$ and; (3) the forward probability held today of defaulting on the long-term debt at date T_2, conditional on not defaulting on the short-term debt at date T_1, i.e., $1 - \frac{N_2(k_1, k_2; \rho)}{N(k_1)}$. Similar to the Moody's KMV iterative procedure, the Delianedis and Geske model is estimated by the two-step iterative algorithm. Regarding the maturity of the debt value, this study takes all short term obligations due in one year as a one-year maturity debt, and all long-term debt as a ten-year maturity debt.

1.3. The Book Value-Based Merton and Delianedis and Geske Models

As Luxembourg bank subsidiaries are not publicly quoted, an alternative approach to calculate PDs has to be followed. Hillegeist at al. (2004) demonstrate that the market-based Merton's PD provides significantly more information about the probability of bankruptcy than do the popular accounting-based measures. However, Bharath and Shumway (2008) also examine the accuracy and PDs forecasting performance of the Merton model and find that most of its predictive power comes from its functional form rather than from the estimation method: the firm's asset value, its asset risk, and its leverage. In an application to Brazilian and Mexican banks, Souto *et al* (2009) and Blavy and Souto (2009), respectively, show that the book-based Merton's credit risk measures are highly correlated with market-based Merton's credit risk measures.[6] This suggests that banks' financial statements are a crucial piece of information when forming market expectations about the probability of banks' default. Regarding the estimation of volatility, in empirical work, a dynamic volatility model is often preferred in order to track risks more timely. However, most dynamic volatility models require many more data points than are available for Luxembourg banks. The RiskMetrics (RM) filter/model instead assumes a very tight parametric specification. The book value asset RM variance can be defined as:

$$h^B{}_{t+1} = (1-\zeta)(ln(V_t^B / V_{t-1}^B))^2 + \zeta h^B{}_t$$

where the variance forecast h_{t+1}^B for period *t+1* is constructed at the end of period *t* using the square of the return observed at the end of period *t* as well as the variance on period *t*. Although the smoothing parameter ζ may be calibrated to best fit the specific historical returns, RiskMetrics often simply fixes it at 0.94. To avoid the calibration difficulties for our limited data points, ζ is also assumed to be same for all banks, and estimated by numerically optimizing the composite likelihoods (Varin et al, 2011), here the sum of quasi maximum likelihood functions of the estimation sample over all banks simultaneously:

$$QMLE(\zeta) = -\frac{1}{2}\sum_{i=1}^{N}\ \sum_{t=1}^{T}(ln(h_{t,i}) + (V_{t,i}{}^B / V_{t-1,i}^B)^2 / h_{t,i})$$

[6] See also Gray and Jones, 2006, for an early application of this idea.

where N is number of banks, and there is a time series of T observations for each banks. The recursion is initialized by setting the initial σ_0^B equal to the first year book value asset volatility, and the means of quarterly assets returns in a large sample are assumed to be zeros to avoid the noises brought by the sample means to the RM variance process. The estimated value of ζ is 0.83.

In order to have a more forward-looking measure, the variance forecast $\sigma^B{}_{t+1}$ can be used to calibrate PDs at time t. The book-value risk neutral PDs of the Merton model can be directly estimated by:

$$\pi_B = N(-\frac{ln(V^B / X) + (r - \frac{1}{2}\sigma_B^2)(T-t)}{\sigma_B \sqrt{T-t}}).$$

Similarly the three book-value risk neutral PDs of the Delianedis and Geske model can be estimated by substituting V_B and σ_B into k_1 and k_2 in the Geske model. Given σ_B, the critical book value of total assets $\overline{V}^B$ at T_1 is calculated first. Similarly, this study takes all short term debt due in one year as a one-year maturity debt, and all long-term debt as a ten-year maturity debt.

2. The Generalized Dynamic Factor Model

In recent years, large-dimensional dynamic factor models have become popular in empirical macroeconomics. The GDFM enables the efficient estimation of the common and idiosyncratic components of very large data sets. The GDFM assumes that each time series in a large data set is composed of two sets of unobserved components. First, the common components, which are driven by a small number of shocks that are common to the entire panel—each time series has its own loading associated with the shocks. Second, the idiosyncratic components, which are specific to a particular variable and orthogonal with the past, present, and future values of the common shocks. The common component of PDs is best viewed as the result of the underlying unobserved systemic risk process, and it is thus expected that it will be relatively persistent. The idiosyncratic component instead reflects local aspects of credit risk that while far from negligible, especially in the short term, are transient.

Assume a vector of n series expressed as $x_t^i = \alpha^i(L)u_t + v_t^i$ where $x_t = \left(x_t^1, x_t^2, \ldots, x_t^n\right)'$ is a n-dimensional vector of stochastic stationary process with zero mean and variance

1; $u_t = \left(u_t^1, u_t^2, \ldots, u_t^q\right)'$ is a q-dimensional vector of mutually orthogonal common shocks with zero mean and unit variance, and with $q < n$; $v_t = \left(v_t^1, v_t^2, \ldots, v_t^n,\right)'$ is a n-dimensional vector of idiosyncratic shocks; and $\alpha^i(L)'$is a $(n \times q)$ matrix of rational functions with the lag operator L. The model allows for correlation between v_t^i variables, but the variances of v_t^i bounded as $i \to \infty$. When n is large, the idiosyncratic components, which are poorly correlated, will vanish, and only the common components will be left, and thus they will be identified (see Forni and others, 2000, for a technical proof).

The GDFM model is estimated using the one-sided estimator proposed by Forni *et al* (2005). The procedure comprises two steps: first, estimating the spectral density matrix of the vector stochastic process x_t^i and, second, using the calculated q largest (real) eigenvalues—and their corresponding eigenvectors—of the spectral density matrix to estimate the generalized common components. In this study, the x_t^i $(t \times n)$ vector stochastic stationary process has t = 93 monthly observations and n includes 283 market indexes and macroeconomic variables for Euro area, Belgium, Canada, Denmark, France, Germany, Greece, Japan, Netherland, Italy, Spain, Sweden, Switzerland, United Kingdom, United States, and Luxembourg. Adding the macroeconomic variables to the PDs, there are 496 (354) series for Delianedis and Geske's (Merton's). The number of dynamic factors is q = 3 underlying PDs or DD. Accordingly, there are 496 (354) idiosyncratic shocks. In the $\alpha^i(L)'$ $(n \times q)$ matrix of rational functions with the lag operator L, the number of lags is 2, and total the number of static factors is 9.[7]

Since the common factors are derived on the standardized first difference of PDs or DD, the accumulated common component is constructed from the initial PDs or DD, and the standard deviation (STD) and mean (M) of the first difference of PDs or DD. For example, in the case of PDs, $PDs_t^{AccumulatedCC} = M + \alpha_i(L)u_t STD + PDs_{t-1}^{AccumulatedCC}$, and $PDs_0^{AccumulatedCC} = PDs_1$. Therefore the accumulated common component shows the hypothetical evolving path of credit risk if purely driven by the common factors. The accumulated idiosyncratic component is simply the residual risk between PDs or DD and its accumulated common component. The correlation between the accumulated

[7] See Hallin and Liska (2007) for the *log criterion* to determine the number of dynamic factors, and Alessi, Barigozzi and Capasso (2009), who modify Bai and Ng (2002) criterion for determining the number of static factors in a more robust manner.

common component and the accumulated idiosyncratic component can be statistically significant even the idiosyncratic component is orthogonal with the common factors.

3. A Dynamic Forecasting Framework

Forni *et al* (2005) provide a good framework for multi-step-ahead predictions of the common component. Nevertheless, the idiosyncratic (credit risk) component also plays an important role for financial instability, which cannot be neglected (see Schwaab *et al*, 2010). The idiosyncratic component is in general autocorrelated and therefore can be predicted. Forni *et al* (2003) construct a linear forecasting model with the contemporaneous common component and the lagged idiosyncratic component. However, their forecasting method is not easily applied to a large number of underlying assets simultaneously, and also does not generate the distribution of these forecasts. The input to the GDFM is a vector of stochastic covariance-stationary processes with zero means and finite second-order moments. However, currently there is no structural credit risk model directly combined with the GDFM. Therefore, the standardized first difference of PDs or DD (difference stationary processes) can be regarded as exogenous inputs to the GDFM. The common and idiosyncratic components are assumed to be asymptotically stationary and orthogonal to each other, whereas the idiosyncratic component can be mildly cross-correlated. Similar to the algorithms for combining GDFM and GARCH in Alessi, Barigozzi and Capasso (2007a&b), this study introduces a novel approach to combine the GDFM with a dynamic t-copula. The AR (or zero mean)-GARCH model can be applied to both the common components and the idiosyncratic components for all variables. Afterwards, a dynamic t-copula is used to glue together the standardized residuals or innovations from those marginal components. Formally, the dynamic forecasting model becomes:

$$
\begin{aligned}
&X_{t+1}^{F} = X_{t+1}^{CC\_F} + X_{t+1}^{IC\_F} \\
&X_{t+1}^{CC\_F} = X_{t+1}^{GDF\_F} + \sigma_{t+1}^{CC}\varepsilon_{t+1}^{CC} \\
&X_{t+1}^{IC\_F} = \sum_{i=1}^{p} X_{t+1-i}^{IC} + \sigma_{t+1}^{IC}\varepsilon_{t+1}^{IC} \\
&\sigma_{t+1}^{2} = \alpha_0 + \alpha(\sigma_t\varepsilon_t)^2 + \beta\sigma_t^2 \\
&\varepsilon_{t+1} \sim iid(0,1) \\
&F(\varepsilon_{t+1}^{1}, \varepsilon_{t+1}^{2}, \ldots, \varepsilon_{t+1}^{2n}) = C_T(F_1(\varepsilon_{t+1}^{1}), F_2(\varepsilon_{t+1}^{2}), \ldots, F_3(\varepsilon_{t+1}^{2n}); R_t, v_t),
\end{aligned}
$$

where the forecast X_{t+1}^{F} of the marginal credit risk is the sum of its forecasted common component $X_{t+1}^{CC\_F}$ and idiosyncratic component $X_{t+1}^{IC\_F}$; $X_t^{CC} = \alpha_i(L)u_t$ is the common component, and $X_t^{IC} = v_t^i$ is the idiosyncratic component from the GDFM. Both common and idiosyncratic components are simply assumed to follow a GARCH (1,1) process. The mean of $X_{t+1}^{CC\_F}$ is the prediction of the common component $X_{t+1}^{GDF\_F}$ by the GDFM as in Forni *et al* (2005), whereas the mean of $X_{t+1}^{IC\_F}$ is an autoregressive process of order p, AR (p). The multivariate distribution $F(\varepsilon_{t+1}^1, \varepsilon_{t+1}^2, ..., \varepsilon_{t+1}^{2n})$ for i=1,2,...,2n, which includes standardized residuals from both the common and the idiosyncratic components and has a time-varying t-copula form.

The copula is a fundamental tool for modeling multivariate distributions. It provides a robust method of consistent estimation for dependence, and is much flexible. Correlation, which usually refers to Pearson's linear correlation, depends on both the marginal distributions and the copula, is not a robust measure given that a single observation can have an arbitrarily high influence on it. Instead, using the conditional dynamic copula, it is relatively easy to construct and simulate from multivariate distributions built on marginal distributions and dependence structure. Drawing on Jin and Nadal De Simone (2011a), a PD index of banking sector overall credit risk is constructed aggregating the individual banks' PD estimates weighted by, say, their respective implied asset values.[8] The following sections explain in detail the modelling of marginal dynamics, dynamic t-copulas, and forward simulation procedures.

3.1. Modelling Marginal Dynamics

This study does not specify marginal distributions, but adopts a semi-parametric form for the marginal distributions. Misspecification of marginal distributions can lead to dangerous biases in dependence measure estimation. This is why the semi-parametric approach is quickly becoming the standard in joint multivariate modelling. Time series data, like the common and idiosyncratic components of financial processes, usually reveal time-varying variance and heavy-tailedness. While keeping the multi-step-ahead prediction of the common components from Forni *et al*, 2005, a GARCH (1,1) process is fitted to the common components and an AR(p) - GARCH (1,1) process is fitted to the idiosyncratic components. The proposed marginal dynamics are formally defined as:

[8] Weights other than asset values are used and discussed below.

$$X_t^{CC} = \sigma_t^{CC}\varepsilon_t^{CC}$$
$$X_t^{IC} = \sum_{i=1}^{p} X_{t-i}^{IC} + \sigma_t^{IC}\varepsilon_t^{IC}$$
$$\sigma_t^2 = \alpha_0 + \alpha(\sigma_{t-1}\varepsilon_{t-1})^2 + \beta\sigma_{t-1}^2$$
$$\varepsilon_t \sim iid\,(0,1),$$

where X_t^{CC} is the common component, and X_t^{IC} is the idiosyncratic component from Forni *et al* (2005). The model is estimated directly by Quasi-Maximum Likelihood. The best AR (p) - GARCH (1,1) can be selected by an automatic model selection criteria, such as the Akaike Information Criterion Corrected Version (AICC). Since in the database, book-value data are actually quarterly, an AR (3) process is used to track dynamic changes, which is especially important for macroprudential policy.

Given the standardized i.i.d. residuals ε_t from the estimation of the marginal dynamics, the empirical cumulative distribution function (cdf) of these standardized residuals is estimated with a Gaussian kernel. This smoothes the cdf estimates, eliminating the rugged shape of the sample cdf. Howevr, although non-parametric kernel cdf estimates are well-suited for the interior of the distribution where most of the data are found, they tend to perform poorly when applied to the upper and lower tails. Therefore, to improve the efficiency of the tails of the distribution's estimates, the upper and lower, e.g. 10% thresholds of the residuals, are reserved for each tail. Then, the amount by which those extreme residuals in each tail fall beyond the associated threshold is fitted to a parametric Generalized Pareto distribution (GP) by maximum likelihood. Since in our study there are only 93 monthly observations, 20% thresholds are used to ensure that there are sufficient data points at the tails to conform well to a GP. Extreme Value Theory (EVT) in general, and in particular the GP distribution, provide an asymptotic theory of tail behavior. Under the assumption of a strict white noise process, i.e. an independent, identically distributed process, the theory shifts the focus from modelling the whole distribution to modelling tail behaviour, and hence, even asymmetry may be examined directly by estimating the left and right tails separately. In addition, EVT has the advantage of requiring just a few degrees of freedom. This approach is often referred to as the distribution of exceedances or peaks-over-threshold method (see, for instance, McNeil (1999), McNeil and Frey (2000) or Nystrom and Skoglund (2002a&b)).

3.2. The Dynamic Conditional t-Copula

As stated above, copula theory provides an easy way to deal with (otherwise) complex multivariate modeling. The main advantage of the copula approach is its flexibility. It allows the definition of the joint distribution through the marginal distributions and the dependence between the variables. In addition, copulas are often relatively parsimoniously parameterized, which facilitates calibration. Recently, copula theory has been extended to the conditional case, allowing the use of copulas to model dynamic structures, such as in Dias and Embrechts (2004), Patton (2004, 2006a&b), and Jondeau and Rockinger (2003, 2006). The conditional copula can be a very powerful tool for active risk management as shown by Fantazzini (2009), and Jin and Lehnert (2011).

The t-copula is a good candidate for the high-dimensional problem dealt with in this paper allowing for non-zero dependence in the extreme tails. The copula of the multivariate standardized *t* distribution is the t-copula, and the conditional dynamic t-copula is defined as follows[9]:

$$C(\eta_1,\eta_2,...,\eta_n;R_t,v_t) = T_{R_t,v_t}(t_{v_t}^{-1}(\eta_1),t_{v_t}^{-1}(\eta_2),...,t_{v_t}^{-1}(\eta_n)),$$

where $\eta_n = F_n(\varepsilon_n)$ for i=1,2,...,n, and $\varepsilon_t \sim iid(0,1)$, are the innovations from the marginal dynamics introduced in the previous section. R_t is the rank correlation matrix, and v_t is the degrees of freedom. $t_{v_t}^{-1}(\eta_n)$ denotes the inverse of the t cumulative distribution function. R_t and v_t can be assumed to be constant, or a dynamic process through time.

Engle (2002) proposes a class of models - the Dynamic Conditional Correlation (DCC) class of models - that preserves the ease of estimation of Bollerslev (1990)'s constant correlation model while allowing correlation to change over time. These kinds of dynamic processes can also be extended into t-copulas. The simplest rank correlation dynamics considered empirically is the symmetric scalar model where the entire rank correlation matrix is driven by two parameters:

$$Q_t = (1-\alpha-\beta)\bar{Q} + \alpha_{dcc}(\varepsilon_{t-1}^{*}\varepsilon_{t-1}^{*'}) + \beta_{dcc}Q_{t-1},$$

where $\alpha_{dcc} \geq 0, \beta_{dcc} \geq 0, \alpha_{dcc} + \beta_{dcc} \leq 1$, $\varepsilon_t^{*} = t_{v_t}^{-1}(\eta_n = F_n(\varepsilon_n))$, $Q_t = | q_{ij,t} |$ is the auxiliary matrix driving the rank correlation dynamics, the nuisance parameters $\bar{Q} = E[\varepsilon_t^{*}\varepsilon_t^{*'}|$ with

[9] See Patton (2006b) for the definition of a general conditional copula.

sample analog $\bar{Q} = T^{-1}\sum_{t=1}^{T} E[\varepsilon_t^{*}\varepsilon_t^{*'}|$, so that R_t is a matrix of rank correlations $q_{ij,t}$ with ones on the diagonal, $\rho_{ij,t} = \frac{q_{ij,t}}{\sqrt{q_{ii,t}}\sqrt{q_{jj,t}}}$.

Given that the correlation between the Gaussian rank correlation $\rho_{GR} = Corr(\Phi^{-1}(u)\Phi^{-1}(v)$ and a t-copula rank correlation $\rho_{TR} = Corr(t_v^{-1}(u)t_v^{-1}(v)$ is almost equal to one, R_t can be well approximated by the $R_t^{Gaussian}$ from the dynamic Gaussian Copula. For convenience, this study adopts a two-step algorithm for estimation which means that R_t is estimated from the dynamic Gaussian copula first, and then, with R_t fixed, the degrees of freedom are recovered from the t-copula.

The dynamic multivariate Gaussian copula is defined similarly to the t-copula as follows:

$$C(\eta_1,\eta_2,...,\eta_n;R_t) = \Phi_{R_t^{Gaussian}}(\Phi^{-1}(\eta_1),\Phi^{-1}(\eta_2),...,\Phi^{-1}(\eta_n)),$$

where $\eta_n = F_n(\varepsilon_n)$ for i=1,2,...,n, and $\varepsilon_t \sim iid(0,1)$ are again the innovations from the marginal dynamics introduced in the previous section. $R_t^{Gaussian}$ is the Gaussian rank correlation matrix. The rank correlation dynamics is similarly driven by the two parameters listed above for the t-copula. However, $\varepsilon_t^{*} = \Phi^{-1}(\eta_n = F_n(\varepsilon_n))$.

While the quasi-likelihood function for the dynamic Gaussian copula could be computed, convergence is not guaranteed in high dimensions, and sometimes it fails or is sensitive to the starting values. This incidental parameter problem causes likelihood-based inference to have economically important biases in the estimated dynamic parameters, with specially α displaying a significant downward bias. As a result, Engle, Shephard and Sheppard (2008) suggest an approach to construct a type of composite likelihood, which is then maximized to deliver the preferred estimator:

$$CL(\psi) = \sum_{t=1}^{T}\frac{1}{N}\sum_{i=1}^{N} \log f(Y_{j,t};\psi),$$

where $Y_{j,t}$ is composed of all unique pairs of data, ψ is a set of parameters, N is the number of all pairs, and t=1,2,...,T. The composite likelihood is based on summing up the quasi-likelihood of all subsets. Each subset yields a valid quasi-likelihood, but this quasi-likelihood is only mildly informative about the parameters. By summing up many subsets, it is possible to construct an estimator which has the advantage of not making necessary the inversion of large dimensional covariance matrices. Further, and vitally, the estimator is not affected by the incidental parameter problem discussed above. It can also be very fast, and does not have the biases intrinsic in the usual likelihood estimator when the cross-section is large. This dynamic Gaussian copula can also be estimated by maximizing m-profile subset composite likelihood (MSCL)[10] using contiguous pairs, which is attractive from statistical and computational viewpoints for large dimensional problems, at least compared with the m-profile composite likelihood (MCLE) using all the pairs. Therefore, to avoid the known estimation difficulties of high-dimensional t-copula, m-profile subset composite likelihood (MSCL) are maximized using contiguous pairs, where the degrees of freedom for the t-copula is simply the 50th quantile of all degrees of freedom derived from pairwise t-copulas.

3.3. Forward Simulation

Using conditional dynamic copulas, it is relatively easy to construct and simulate from multivariate distributions built on marginal distributions and dependence structure. The GARCH-like dynamics in both variance and rank correlation offers multi-step-ahead predictions of the common and the idiosyncratic components simultaneously.

The following steps illustrate the one-step-ahead simulation:

1. Draw independently $\varepsilon^{*i1}_{t+1},\ldots,\varepsilon^{*im}_{t+1}$ for each component from the n-dimensional *t* distribution with zero mean, forecast correlation matrix R_{t+1}, and degrees of freedom v_{t+1} to obtain $\mu^{i1}_{t+1},\ldots,\mu^{im}_{t+1}$ by setting $\mu^{ik}_{t+1} = t_{v_{t+1}}(\varepsilon^{*ik}_{t+1})$, where k=1,...,m, the total paths of simulation, i=1,...,n, the number of components;
2. Obtain $\varepsilon^{i1}_{t+1},\ldots,\varepsilon^{im}_{t+1}$ by setting $\varepsilon^{ik}_{t+1} = F_i^{-1}(\mu^{ik}_{t+1})$, where F_i is the empirical marginal dynamics distribution for component i;
3. Obtain $z^{i1}_{t+1},\ldots,z^{im}_{t+1}$ by setting $z^{ik}_{t+1} = \varepsilon^{ik}_{t+1}\sigma^{i}_{t+1}$, where σ^{i}_{t+1} is the forecast standard deviation using a GARCH (1,1) model for component i;

[10] A moment-based profile likelihood, or m-profile likelihood for short, in which the nuisance parameters are not maximum quasi-likelihood estimators but attractive moment estimators.

4. Obtain $X_{t+1}^{i1},\ldots,X_{t+1}^{im}$ by setting $X_{t+1}^{ik} = \lambda_{t+1}^{i} + z_{t+1}^{ik}$, where λ_{t+1}^{i} is the forecast mean using an AR (p) model for the idiosyncratic component i, and the prediction of the common component using Forni *et al* (2005);
5. Finally sum the predicted idiosyncratic and common components at t+1.

In a similar way, several period predictions can be obtained. Both the idiosyncratic and common components are derived on the standardized first difference of the PD index. The simulated cumulative PDs have to be truncated by $Max(DPs^{Simulated},0)$. This forward simulation approach therefore integrates the one-sided forecasting features of the GDFM into the dynamic copula framework.

III. Data

This study is applied to 32 major European banking groups, to their respective 37 subsidiaries active in Luxembourg, and to two 100%-Luxembourg banks; surveillance of banking stability cannot stop at national borders. Market data used for the major European banking groups include government bond yields, stock prices and stock indices, production, employment and GDP data, consumer prices, housing prices, exchange rates, credit data, as well as the number of outstanding shares, and book value data from Bloomberg, DataStream, BIS, Eurostat, and ECB (see Appendix 1 for a detailed list of country and euro area series and sources). The market data start in May 2000 and finish in September 2011.

One difficulty is that short-term borrowing (BS047) and long-term debt (BS051) from Bloomberg have annual, semi-annual, and quarterly frequencies. To make the data consistent, four filtering rules as described in Appendix 2 are used. To get the “actual” PDs from neutral PDs, the expected returns are estimated using a capital-asset pricing model. The implied equity risk premiums data (Damodaran 2011) are downloaded from Damodaran Online at http://pages.stern.nyu.edu/~adamodar/. For consistency with that source, stock market returns are represented by the returns on the S&P 500 index.

All the Luxembourg banks are unlisted, so quarterly book value data from the BCL database going back to 2003Q1 are used.[11] The 37 subsidiaries registered in Luxembourg represent about 63 percent of the total assets of the Luxembourg banking industry. When the two 100% Luxembourg banks are added to the list, the database represents nearly 70 percent of the total assets of the industry. For all the selected

[11] See Jin and Nadal De Simone, 2011a, for a detailed discussion of the estimation of credit risk models using balance sheet data when banks are not publicly listed.

Luxembourg banks, short term debt includes demand and time deposits of up to one-year maturity, short term funding, and repos, while the long term debt includes time deposits of over one-year maturity and other long term funding.

IV. Empirical Results

As stated above, timeliness in reflecting credit risk events is a plus in macroprudential supervision. Banks' marginal PDs estimated from structural credit risk models and aggregated in an index weighted by the share of respective banks' assets in total assets do a relatively good job at tracking changes in credit risk, both across European and Luxembourg banks (Jin and Nadal De Simone, 2011a). Further, the GARCH structural credit risk model, despite its more sophisticated modeling approach, typically underperforms more basic models, but the combined Merton/GARCH-MIDAS model performs best by reflecting important market events earlier than other approaches (Jin *et al* 2011b). Unfortunately, this latter model cannot be used for Luxembourg banks due to the lack of sufficiently long data that would allow the robust modeling of the short- and long-run components of credit risk. For these reasons, this study estimates neutral marginal PDs and DDs from two structural credit risk models, Merton (1974) model and Delianedis and Geske (2003) model, and given its objective of accounting for overall banking sector credit risk, it incorporates dependence among banks' PDs by using the GDFM Model (Forni et al, 2006) with a dataset including macroeconomic and financial variables. In addition to generating an indicator of overall banking sector credit risk that recognizes exogenous shocks timely, this framework identifies the build up of endogenous imbalances and it improves on the GDFM forecasting capacity by combining it with a dynamic t-copula. As stated above, the objective is twofold: first, to test the capacity of the framework to anticipate financial vulnerabilities reflected in, e.g., a persistent increase in PDs; second, to obtain an out-of-sample forecast distribution of overall banking sector credit risk.

The rest of this section discusses first the Kendall correlation of asset-weighted PDs and DDs between European banking groups and their Luxembourg affiliates. It then addresses the early-warning capabilities of the framework at the level of banks' individual PDs and DDs, and at the level of indexes of banks' PDs and DDs with weights suggested in the literature that proposes methods to identify systemic important institutions. Finally, it reports results on the out-of-sample forecasting capabilities of the framework, both for individual PDs and DDs and for total asset-weighted PDs and DDs.

1. Asset-weighted PDs and DDs

As expected, there is a high degree of correlation (Kendall correlation) among European banking groups and Luxembourg banks PDs and DDs (Tables 1a and 1b, respectively). However, these correlations vary over time and also in sign depending on whether the ST or the FW components of PDs are considered and on whether the common or the idiosyncratic components of PDs and DDs are considered.[12]

During the whole sample period, correlation of PDs and DDs between both set of banks are very highly significant for the whole time structure of PDs and for the common components. Interestingly, correlations are negative when the idiosyncratic components are involved, especially those of the banking groups PDs. These results suggest that the parent banks and their affiliates are subject to bank specific factors that may diverge at a given point in time. Finding the causes of this behavior is certainly beyond the scope of this study. Nevertheless, it can be conjectured that this may result from the different business models of Luxemburg affiliates that overwhelmingly are net suppliers of liquidity to parent banks. This working hypothesis seems reasonable when the same analysis is applied to the pre-crisis period, 2004-07, the crisis period, 2008-09, and the post-crisis period, 2010-2011. It seems that it is the FW idiosyncratic components of PD and DDs that are mostly significant and move in the opposite direction between group banks and Luxembourg affiliates during the pre-crisis period.

During the crisis period, as is well known, correlations increase—banks' interdependence increases—as reflected in the rise of correlations between the group banks and Luxembourg banks PDs (also in the increase in the number of significant correlations). DDs idiosyncratic components, instead, move in the same direction, as it is to be expected when banking sector credit risk increases.

Finally, in the post-crisis period, there is again an increase in the importance of idiosyncratic components which move in disparate directions at the parent and at the affiliate banks. This is more the case with respect to the ST PDs than with respect to the FW PDs, however, which is an important difference with the pre-crisis period and possibly a reminder of the persistence of short-term solvency issues across some banking groups in Europe.

[12] Please note that the common and the idiosyncratic components are linearly orthogonal, so that the Pearson linear correlation among them is statistically insignificant. Results are available upon request.

A comparison of results between European and Luxembourg banks suggests that the crisis affected Luxembourg banks relatively less than European banks. Similarly, Luxembourg banks' recovery was less dramatic. Given the liquidity shortages that characterized the crisis, especially in its onset and before policy measures alleviated it, traditional liquidity-providers such as Luxembourg banks were relatively less distressed.

Summarizing, the strongest (and negative) correlation between the common components of banking groups' PDs and their affiliates regards relatively more the FW PDs. This is visible in the pre- and post-crisis periods, and the reason may be the different business model of Luxembourg banks which are net liquidity providers. During the crisis period, however, the comovement of common components increased, in particular with respect to ST PDs. This is in agreement with the observed regularity of rising correlations during stressed periods.

2. In-sample Early-warning Features of Single-bank PDs and Weighted Indexes of PDs

As stated above, a macroprudential policymaker is interested not only in the timeliness feature of measures of credit risk, both at the bank level and at the systemic level, but ideally would like to have on real time, and as early as possible, some indication of the buildup of vulnerabilities in the financial system. The first feature is particularly important in the case of banks that are not public given the lags in the availability of book value data. The second feature is crucial for taking preventive actions to preserve financial stability and reduce the likelihood of systemic crises. To assess the strength of the framework proposed in this study to achieve those objectives, three approaches are followed. To assess the timeliness features of measures of credit risk as well as the contribution of the GDFM to that aim, first, a set of in-sample Granger causality tests is performed between the common component of the estimated PDs/the macrofinancial factors and estimated PDs.[13] Second, the degree of in-sample commovement and leads and lags between the common components and estimated PDs is studied using spectral methods. Finally, to assess the strength of the framework to signal the build up of vulnerabilities over time, the forward PD of Delianedis and Geske is related to the macrofinancial data set to extract the factors responsible for its steady growth in the run up to the crisis.

[13] Jin *et al* (2011b) studied lead-lag relationships across models' PDs predictions, but had no reference to macrofinancial conditions.

2.1. Granger causality tests

Tables 2a and 2b summarize the results of the Granger causality tests applied to each bank's estimated PDs and to indexes of PDs weighted in the different ways suggested in the literature on indicators of banks' systemic importance, respectively (Drehmann and Tarashev, 2011): total assets, interbank lending and interbank borrowing. Individual bank data on interbank lending and borrowing at quarterly frequency are available for Luxembourg banks only.[14] Table 2a reports the ratios according to Granger Causality test at the p-values of 1%, 5% and 10%. The ratio is the percentage of cases when X Granger causes another measure Y, and Y does not Granger causes X at a given confidence level over the available banks for both banking groups and Luxembourg banks. The ratios under common component mean that the common component Granger causes DPs and DPs do not Granger cause the common component; similarly, for DPs.[15] At p-values of 1%, for example, the common component of the estimated DPs Granger causes banking groups Geske All PDs and DDs in 31% and 25% of the cases, respectively. It also Granger causes Luxembourg banks' PDs in 26% and DDs in 36% of the cases. In all cases, the opposite is much less frequent. Importantly, the common component has a clearer anticipatory feature with respect to DDs than PDs either for banking groups or for Luxembourg banks.

Moreover, the framework's best performance is with respect to the FW PDs of Luxembourg banks, i.e., 50%. This feature is likely due to the use of book-value data for estimating Luxembourg banks' PDs, which is less timely than the information contained in share prices available for estimating banking groups' PDs. This in-sample leading information on the common component of PDs is a particularly useful feature of the proposed methodology for Luxembourg banks as they are not quoted.

Table 2b shows the Granger causality tests between the weighted common components and PDs both for the whole sample period and for the period immediately prior to Lehman's collapse. Those indexes of PDs have been constructed using equal weights and weights suggested in the literature on the determination of the systemic nature of banking institutions.[16] During the whole sample period, in general, the common

[14] Indexes were constructed weighing estimated PDs using banks' shares in total assets, in total interbank lending and total interbank borrowing.
[15] Only standardized measures are displayed; non-standardized measures provide the same results.
[16] Drehmann and Tarashev (2011) propose three measures for determining banks' systemic importance. Two measures are top down: the *participation approach* (i.e., expected losses incurred by a given bank' non-bank creditors) and the *contribution approach* (i.e. expected losses from a bank's exposure to exogenous shocks, from its contribution to losses via propagation and from its idiosyncratic exposure to shocks). Another measure is *bottom up*, i.e. the expected losses of the whole banking system conditional on a given bank being in default. The authors show that *size* is a good proxy of all measures, that *interbank lending*

component does not Granger cause the PDs or DDs indexes suggesting, by comparison to Table 2a, that averaging hides important information. During the run up to Lehman's collapse, results are marginally better: When all PDs are attributed equal weight, the common component Granger causes DDs and DDs do not Granger cause the common component for banking groups and Luxembourg banks. Using total assets as weights, for Luxembourg banks, the common component Granger causes PDs for Geske ST and FW and PDs do not Granger cause the common component. Clearly, the use of weighting schemes hides information embedded in the common factors and variable loadings making it more difficult to draw conclusive evidence using Granger causality tests. These weights, although suggested in the literature dealing with the determination of the systemic nature of individual banks, is not useful to construct indices of PDs (or DDs) that can (visually or statistically) provide a timely measure of credit risk.

2.2. Frequency-domain analysis

The bivariate test in the previous section clearly suffers from the averaging across periods which, in addition to the presence of nonlinearities and feedback effects in financial markets, may mask the lead/lag relationships between common components and estimated PDs. To take that into account, this section briefly looks at the comovement between PDs and its common components using spectral methods. In particular, the coherence (squared) and the phase angle are estimated.[17] Figures 1a to 1d display the estimated coherences and phase angles of the common components and Geske ST and FW PDs for banking groups and Luxembourg banks. The complicated interrelations and feedback effects between the common components and measures of PDs evince clearly.

In general, the common components lag estimated ST PDs for banking groups only at periodicities between 1 and 2 years. The common components lead ST PDs in cycles between 2.5 years and 8 years, that is to say, roughly during the *minor* (2 to 4 years) and the *major* (4 to 8 years) business cycles' durations (cycle definitions according to the National Bureau of Economic Research). The common components lead FW PD during cycles of 2 years and cycles of between 3 to 5 years, that is to say, during most of the minor cycle and the first part of the major cycle.

---

proxies well the *participation* and the *contribution approaches* whereas *interbank borrowing* proxies well the *contribution* and the *bottom-up approaches*.

[17] Coherence (squared) is the proportion of the variance of a series which can be explained by the other series, period (or frequency) by period (by frequency). The phase lead is the fraction of a cycle by which one series leads (lags) the other at each period or frequency. The phase lead is significant only at the periods (or frequencies) at which the coherence is significant.

In the case of Luxembourg banks, the common components lag estimated ST and FW PDs around periodicities of 1 year, and between 1.5 and 2.5 years for ST PDs and about 2 years for FW PDs. Otherwise, the common components lead ST and FW PDs at periodicities of about 3 quarters and in the longer run, at periodicities ranging between 3 and (over) 8 years for the ST PDs, and between 4 and (over) 8 years for the FW PDs.

Summarizing, the results support the leading features of information embedded in the common components at relatively high frequency (i.e., roughly 3 quarters) and at relatively lower frequency (i.e., between around 3 years and 8 years). These results are consistent with the visual inspection of Figures 2b and 2d.

2.3 In-sample early warning features of forward probabilities of default

Given the macroprudential policymaker's interest in preventing the buildup of financial vulnerabilities that could unravel disorderly, and the well-known nonlinearities and feedback between PDs or DDs and their common components, it is advisable to look at matters even further. First, some leading features of the common component for Luxembourg banks' PDs can be visualized in the set of figures 2 (a to d) and 3 (a to d) which show aggregated PDs in indexes for banking groups and for Luxembourg banks. Equally-weighed PDs are in Figures 2, and total asset-weighed PDs are in Figures 3. What is of particular interest here is the leading behavior of the common component of estimated FW PDs.. For banking groups, starting in 2005 (Figure 2b), and for Luxembourg banks, since early 2006 (Figure 2d), there is a clear, persistent increase in FW PDs which suggests a buildup of credit risk long-term vulnerabilities—a feature also found by Koopman *et al* (2010).

It is useful to compare the equal weighted index with the index weighted by total assets, i.e., Figure 2b compared to Figure 3b, and Figure 2d compared to Figure 3d. When PDs are aggregated weighted by total assets, the level of PDs and of the common component tend to be somewhat lower suggesting that in the buildup of vulnerabilities, the relatively largest banks had lower PDs and were affected less negatively by markets turmoil than the relatively smaller banks. A policy message is that a macroprudential authority may wish to estimate both versions of the index, or look at median and quantiles distributions of PDs and common components.

This salient feature of the framework proposed in this paper merits further analysis. In order to further analyze the in-sample early warning feature of the common components of forward probabilities of default, the following approach is followed: first, for each bank, the first difference of the accumulated common component of FW PDs is regressed on the first difference of the equal weighted accumulated common component of FW PDs; then, the bank i with positive beta β_i (the coefficients of the PD index) are selected, and beta weights are constructed as follows, $\omega_i = \beta_i / sum(\beta)$. Second, the absolute value of each of the factor loadings l_i for each selected bank are chosen such that $f_i = l_i \omega_i$, and f are summed up across the banks giving F_j for each of the j factors. This enables to construct the weighted scores of all the factors using the expression: $FactorScore_j = F_i / sum(F)$. Third, all input variables from GDFM are categorized into five classes: real variables (GDP in volume and nominal, industrial production, the unemployment rate, the HICP, and agricultural and industrial property prices), funding prices (short- and long-term interest rates, foreign exchange rates, stock market prices, stock price volatility, house prices), funding quantities (total credit, loans to households, mortgages, loans to non-financial firms, and interbank lending and borrowing), confidence (various indices of consumer and business sentiment), and PDs (All, ST, and FW). The Classification Score by factor j is constructed as the share of factor loadings in absolute value of each set of economic classification for all variables, or for selected variables by quantiles (the distribution of the factor loadings of all input variables from the GDFM). Finally, the scores by economic classification (i.e., real, funding prices, funding quantities, confidence, and PDs) are aggregated from each factor as follows: Factor Score * Proportion of Variance Explained by the Factor * Classification Score by Factor.

The same exercise is also applied to the accumulated common component of ST PDs. Tables 3a and 3b display the differences of the classification scores between the FW PDs and the ST PDs for all sets of macro variables by quantiles for the pre-crisis period (2004-2007) and the crisis period (2008 to 2011), respectively. This enables the analysis of the GDFM factors contribution to the buildup of the forward FW PDs.

Consistent with early work by Borio and Lowe (2002), and more recent work by Koopman et al (2010), real economic activity, credit growth and interbank activity explain the buildup of vulnerabilities of large European banking groups in the run up to the crisis as well as credit growth and interbank activity explain it for Luxembourg banks. For example, in the run up to the crisis, real variables explain over 6% more of the FW than of the ST at the 10% and at the 20% quartiles. This results from real variables explaining over 16% of the FW PD at those quantiles and real variables explaining less than 10% of the ST PD at those quantiles (not shown). In general, differences are smaller for Luxembourg banks, a likely outcome of the use of book value as opposed to market data for estimating PDs in the case of Luxembourg. Finally, notice that for banking groups, there is a slight tendency for differences to fall during the crisis period when compared to the pre-crisis period, especially for real variables, but the persistence of vulnerabilities as reflected in the accumulated common components is striking.

As suggested by the analysis, the FW probability of default estimated from Delianedis and Geske model is a useful early warning measure of banks' vulnerabilities, and it should be part and parcel of macroprudential policy tools to monitor financial stability and risk build up over time.

3. Out-of-sample Forecasting

In-sample results say nothing about the out-of-sample framework performance. Therefore, this section addresses the out-of-sample forecasting capabilities of the proposed framework. However, the short number of data points available constrains a full-fledged, standard evaluation of the out-of-sample forecasting capabilities of the framework. Table 4 reports the coverage ratios, root-mean squared errors, as well as the bias, the variance and the covariance decomposition of Theil's inequality coefficient from 2010 to 2011 across all estimated Geske's PDs (Table 4a) and DDs (Table 4b) for banking groups and Luxembourg banks.[18] The coverage ratio is the share of banks whose empirical simulated cdf at each of the estimated PDs or DDs is within the range of the respective quantiles. Under the null hypothesis that this forecasting framework correctly estimates the dynamics of PDs or DDs, the coverage ratio should approximate the range of quantiles, if the number of underlying banks were large enough (which recall is not the case). For example, during the first month of out-of-sample forecasts, 77% of bank PDs forecasted using only the common component are within the 5%-95% quantiles of the forecasted cdf of PDs. The percentage just falls to 70% at month six of the out-of-sample forecasts. When not only the common, but also the idiosyncratic components, are forecasted, 86% percent of the forecasted PDs fall within the 5%-95% quantiles and the percentage increases to about 88% at month six of the out-of-sample forecasts. Decomposing Theil's inequality coefficient into bias, variance, and covariance, it seems that the improvement in forecasting ability by adding the idiosyncratic component results from an improvement in the model's capacity to replicate the degree of variance in PDs (column "Variance Proportion") and from reducing unsystematic error (column "covariance Proportion").

Tables 5a and 5b show the results of the out-of-sample forecast evaluation for banking groups and Luxembourg banks, respectively, using an index of PDs weighted by total

---

[18] The model is re-estimated recursively adding one period at a time and forecasting always 6 months forward.

assets. Results suggest that, in contrast to the out-of-sample forecasts of individual banks' PDs, the use of idiosyncratic components deteriorates the forecast for banking groups PDs, but slightly improves the forecast for Luxembourg banks. For Luxembourg banks, overall, idiosyncratic component forecasts seem to improve the out-of-sample performance of the framework, albeit during the first quarter, but not the second quarter Again, aggregation does hide idiosyncratic features of banks, and it is therefore not surprising that the idiosyncratic component forecast does not help improving the overall out-of-sample forecast. Figures 2a to 3d illustrate the out-of-sample forecasts of the common and of the common and idiosyncratic components and the 10%-90% quantiles for an index of banking groups and Luxembourg banks by equal weighted and total assets weighted. Visual inspection confirms the results of Tables 5a and 5b.

V. Conclusions and macroprudential policy implications

This study develops a framework that recognizes exogenous shocks timely and generates an early warning indicator of overall banking sector credit risk that identifies early build ups of endogenous imbalances. In addition, it provides robust out-of-sample forecasts of probabilities of default. It applies it to a set of European banking groups and their affiliates in Luxembourg given that banking stability cannot stop at national borders.

It uses a two-step approach to proxy banks' default dependency. First, marginal PDs are estimated using Delianedis and Geske compound option model, a structural credit risk model that distinguishes between the probability of default at the end of year one and the probability of default in the long run, conditional on not defaulting the first year. Second, the framework offered by the generalized dynamic factor model applied to a large macrofinancial dataset extracts the common component of banks' marginal PDs, both at the banking group and at the subsidiary levels, showing how a set of common systemic factors affect both of them simultaneously, albeit with different weights. Beyond real economic activity, different credit aggregates as well as the amount of interbank lending and borrowing are important systemic drivers of European banking groups' risk, as suggested by Borio and Lowe (2002), and by Drehmann and Tarashev (2011), respectively. For Luxembourg, credit aggregates and interbank activity are the main drivers. Therefore, the proposed framework contains the salient feature of measuring the relative riskiness of the banking system in a non-crisis mode, a particularly useful characteristic associated with one key aspect of systemic risk. As such, it allows to take remedial actions as risk increases following changes in FW PDs over time consistent with the business cycle and with growth in credit aggregates and wholesale funding. This is a useful tool of the macroprudential toolkit.

In addition, the same framework identifies in a robust manner the idiosyncratic component of banks' PDs for the banking groups and their respective Luxembourg affiliates making it possible to model them structurally, a task, however, which is beyond the scope of this paper. This two-step approach permits to track in advance over a couple-of-year time span a persistent increase in credit risk for the banking system in the tradition of early warning indicators. This rise in credit risk can be interpreted as an increase in the vulnerability of the financial system. As such, the framework of this study, by separating the role of system developments from individual banks' idiosyncratic features, is an important step toward building macro-financial models of systemic risk that contain early-warning features with a realistic characterization of episodes of financial instability. This work contributes to the systemic risk literature incorporating the externalities that financial intermediaries exert on the rest of the financial system and on the economy in general by signaling out the role of common systemic forces affecting all banks and also by showing the buildup of credit risk or widespread imbalances over time, another interpretation of systemic risk. It contributes to the macroprudential literature with a method to monitor systemic risk.

Also important for macroprudential policy is the policymaker's capacity to project or forecast increases in the banking sector credit risk at any given point in time. This study contributes as well to the macroprudential literature by suggesting a framework to forecast credit risk changes. By using a dynamic conditional t-copula, this framework helps forecasting both the common as well as the idiosyncratic components of credit risk. This remediates the well known feature that simply aggregating banks' marginal PDs provides a downward-biased measure of banking systemic risk. Indeed, by incorporating the common and the idiosyncratic components of a broad set of macro-financial variables, the framework improves the analytical features and the out-of-sample forecasting performance of the model.

Useful extensions of this work include the development of an indicator of (systemic) joint probabilities of credit risk changes by including in a dynamic copula the banks that at each point in time have the combination of PDs and size (or share in interbank lending) that is the highest. This would proxy each institution's contribution to systemic risk. Also, explicit modeling of the volatility component of structural credit risk model like it is done with CDOs would allow to better link PDs with macroeconomic and financial variables understanding the risk of simultaneous defaults and facilitating thereby policymakers' decision-making process.

References

Alessi, L., Barigozzi, M., and M., Capasso, 2007a, "Generalized Dynamic Factor Model + GARCH: Exploiting Multivariant Information for Univariate Prediction", LEM Working Paper.

Alessi, L., Barigozzi, M., and M., Capasso, 2007b, " Dynamic Factor GARCH: Multivariate Volatility Forecast for a Large Number of Series", LEM Working Paper.

Alessi, L., Barigozzi, M., and M., Capasso, 2009, "A Robust Criterion for Determining the Number of Factors in Approximate Factor Models", Working Paper.

Arondel, P. and A. Rouabah, 2008, "L'extraction des anticipations des acteurs du marché à partir des prix des options", *Financial Stability Review*, Banque centrale du Luxembourg, pp. 142-149.

Bai, J., Ng, S., 2002, "Determining the Number of Factors in Approximate Factor Models", Econometrica 70(1), 191-221.

Basel Committee on Banking Supervision, 2011, "Globally Systemically Important Banks: Assessment Methodology and the Additional Loss Absorbency Requirement", *Bank for International Settlements*.

Bharath, S.T., and T.Shumway, 2008, "Forecasting Default with the Merton Distance to Default Model", *The Review of Financial Studies*, 21, No 3, pp. 1339-1369.

Billo, M., M. Getmansky, A. W. Lo, and L. Pelizon, 2010, "Econometric Measures of Systemic Risk in the Finance and Insurance Sectors", *NBER working paper 16223.*

Black, F., and M. Scholes, 1973, "The Pricing of Options and Corporate Liabilities", *Journal of Political Economy*, 81, pp.637-654.

Blavy, R., and M. Souto, 2009, "Estimating Default Frequencies and Macrofinancial Linkages in the Mexican Banking Sector", *IMF Working Paper*, WP/09/109.

Borio, C. and P. W. Lowe, 2002, "Asset Prices, Financial and Monetary Stability: Exploring the Nexus", *BIS Working Paper No. 114*, Bank for International Settlements.

Borio, C. and M. Drehmann, 2009, "Towards an Operational Framework for Financial Stability: "Fuzzy" Measurement and its Consequences", Working Paper No. 218, *Bank for International Settlements*.

CGFS, 2010, "The Role of Margin Requirements and Haircuts in Procyclicality", No. 36, *Bank for International Settlements*.

D'Agostino, A. & McQuinn, Kieran & O'Brien, Derry, 2011. "NowCasting Irish GDP", MPRA Paper 32941, University Library of Munich, Germany.

Damodaran, A., 2011, "Equity Risk Premiums (ERP): Determinants, Estimation and Implications – The 2011 Edition", Working Paper, Stern School of Business, NYU.

Delianedis, G., and R. Geske, 2003, "Credit Risk and Risk Neutral Default Probabilities: Information about Rating Migrations and Default", Working Paper, University of California at Los Angeles.

De Nicolò G.D. and M. Lucchetta, 2012, "Systemic Real and Financial Risks: Measurement, Forecasting, and Stress Testing", IMF Working Paper WP/12/58, International Monetary Fund.

Dias, A., and P. Embrechts, 2004, "Dynamic Copula Models for Multivariate High Frequency Data in Finance", Working Paper, ETH Zurich: Department of Mathematics.

Drehmann, M. and N. Tarashev, 2011, "Systemic Importance: some Simple Indicators", *BIS Quarterly Review*, March, Bank for International Settlements.

Duan, J.C., G. Gauthier, J-G Simonato,2004, "On the Equivalence of KMV and

Maximum Likelihood Methods for Structural Credit Risk Models", Working Paper.
Engle, R. 2002, "Dynamic Conditional Correlation: A Simple Class of Multivariate GARCH Models, Journal of Business and Economic Statistics, 20, 339-350.

Engle, R., Shephard, N., and K. Sheppard, 2008, "Fitting Vast Dimensional Time-varying Covariance Models", Technical Report, Stern Business School.
European Central Bank, "The Concept of Systemic Risk", *Financial Stability Review,* pp. 134-142, December 2009.
European Central Bank, "Macro-prudential Policy Objectives and Tools", *Financial Stability Review*, pp. 129-137, June 2010.
European Central Bank, "Analytical Models and Tools for the Identification and Assessment of Systemic Risk", *Financial Stability Review*, pp. 138-146, June 2010.
Fantazzini, D. 2009, "Dynamic Copula Modelling for Value at Risk", Frontiers in Finance and Economics, Forthcoming.
Fischer, B., "Decomposition of Time Series - Comparing Different Methods in Theory and Practice", Eurostat Working Paper, 1995.
Forni M., Hallin M., Lippi M. and Reichlin L., 2003, "Do Financial Variables Help Forecasting Inflation and Real Activity in the EURO Area?", Journal of Monetary Economics 50, pp. 1243-55.
Forni, M., M. Hallin, M. Lippi, and L. Reichlin, 2005, "The Generalized Dynamic Factor Model One-sided Estimation and Forecasting", *Journal of the American Statistical Association* Vol. 100, No. 471, pp. 830-840.
Frankel, J. and G. Saravelos, 2010, "Are Leading Indicators of Financial Srises Useful for Assessing Country Vulnerability? Evidence from the 2008-09 Global Crisis", *NBER Working Paper 16047.*
Galati, G. and R. Moessner, 2011, "Macroprudential Policy – A Literature Review", BIS Working Papers No. 337, *Bank for International Settlements.*
Gray, D. and M. Jones, 2006, "Indonesia: Selected Issues Paper, "Measuring Sovereign and Banking Risk in Indonesia: An Application of the Contingent Claims Approach", IMF Country Report No. 06/318, *International Monetary Fund*.
Geske, R., 1977, "The Valuation of Corporate Liabilities as Compound Options", *Journal of Financial and Quantitative Analysis,* 12, pp. 541-552.
Goodhart, C. and B. Hofmann, 2007, "House Prices and the Macroeconomy", *Oxford University Press*.
Gray, D. F. and S. W. Malone, 2008, *Macrofinancial Risk Analysis*, John Wiley & Sons, Ltd.
Hallin and Liska, 2007, "Determining the Number of Factors in the General Dynamic Factor Model", Journal of the American Statistical Association, 102, 603-617.
Heston, S. and S. Nandi, 2000, "A Closed-form GARCH Option Pricing Model", *Review of Financial Studies* 13, pp. 585-626.
Hillegeist, S. A., E. K. Keating, D. P. Cram, and K. G. Lundstedt, 2004, "Assessing the Probability of Bankruptcy", *Review of Accounting Studies*, 9, No. 1, pp. 5-34.
Huang, J., and M. Huang, 2003, "How Much of the Corporate-treasury Yield Spread is Due to Credit Risk? A New Calibration Approach", Working Paper, *Stanford University*.
Jin, X. and F. Nadal De Simone, 2011a, "Market- and Book-based Models of Probability of Default for Developing Macroprudential Policy Tools", Working Paper No. 65, *Banque centrale du Luxembourg*.
Jin, X., T. Lehnert, and F. Nadal De Simone, 2011b, "Does the GARCH Structural Credit Risk Model Make a Difference?", Research Working Paper Series No. 11-06,

Luxembourg School of Finance, University of Luxembourg.
Jin, X., T. Lehnert, 2011, "Large Portfolio Risk Management and Optimal Portfolio Allocation with Dynamic Copulas", Research Working Paper Series No. 11-10, Luxembourg School of Finance, University of Luxembourg.
Jondeau, E., and M. Rockinger, 2003, "Conditional Volatility, Skewness, and Kurtosis: Existence, Persistence, and Comovements", Journal of Economic Dynamics and Control 27, 1699-1739.
Jondeau, E., and M. Rockinger, 2006, "The Copula-GARCH Model of Conditional Dependencies: An International Stock Market Application", Journal of International Money and Finance 25, 827-853.
Koopman, S. J., A. Lucas and B. Schwaab, 2010, "Macro, Industry and Frailty Effects in Defaults: The 2008 Credit Crisis in Perspective", *Tinbergen Institute Discussion Paper*, TI 2010-004/2.
Lando, D., 2004, *Credit Risk Modeling. Theory and Applications*, Princeton.
Malone, S.W., A. Rodriguez, and E.T. Horst, 2008, "The GARCH Structural Credit Risk Model: Simulation Analysis and Application to the Bank CDS Market During the 2007-2008 Crisis", Working paper.
McNeil, A.J., 1999, "Extreme Value Theory for Risk Managers", Internal Modeling CAD II, *Risk Books*, 93--113.
McNeil, A.J., and R. Frey, 2000, "Estimation of Tail-related Risk Measures for Heteroscedastic Financial Time Series: an Extreme Value Approach, *Journal of Empirical Finance*", 7, 271--300.
Merton, R., 1974, "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates", *Journal of Finance* 29, pp. 449-470.
Nelsen, R., 1999, "An Introduction to Copulas", Lecture Notes in Statistics 139, Springer.
Nystrom, K., and J. Skoglund, 2002a, "Univariate Extreme Value Theory, GARCH and Measures of Risk", Preprint, *Swedbank*.
Nystrom, K., and J. Skoglund, 2002b, "A Framework for Scenario-Based Risk Management", Preprint, *Swedbank*.
Patton, A. 2004, "On the Out-of-Sample Importance of Skewness and Asymmetric Dependence for Asset Allocation, Journal of Financial Econometrics, 2, 130-168.
Patton, J. 2006a, "Estimation of Multivariate Models for Time Series of Possibly Different Lengths", Journal of Applied Econometrics, 21, 147-173.
Patton, J., 2006b, "Modelling Asymmetric Exchange Rate Dependence", *International Economic Review*, 47, 527-556.
Scwaab, B., A. Lucas and S. J. Koopman, 2010, "Systemic Risk Diagnostics", *Duisimberg School of Finance and Tinbergen Institute Discussion Paper, TI 10-104/DSF 2.*
Segoviano, M. and C. Goodhart, 2009, "Banking Stability Measures", *IMF Working Paper*
WP/09/04, International Monetary Fund.
Souto, M., Tabak, B., and F. Vazquez, 2009, "Linking Financial and Macroeconomic Factors to Credit Risk Indicators of Brazilian Banks", *Banco Central do Brasil*, Working Paper No. 189.
Varin, C., Reid, N., and D. Firth, 2011, "An Overview of Composite Likelihood Methods", *Statistica Sinica* 21.

Appendix

A. Filtering Rules for BS_ST_BORROW and BS_FW_BORROW:

Short-term borrowing (BS047) and long-term debt (BS051) from Bloomberg have annual, semi-annual, and quarterly data. The four filtering rules applied to make them consistent are the following:

I. Take any zero as missing data.

II. If annual data exist and are not equal to the semi-annual/quarterly data, then let the semi-annual/quarterly data be equal to the annual data—this gives priority to annual data assuming them to be relatively more reliable.

III. If annual data do not exist for the current fiscal year and both the semi-annual/quarterly data and annual data exist for the previous and next fiscal years, but semi-annual/quarterly data are very different to the corresponding annual data at the previous and next fiscal year, then treat the semi-annual/quarterly as missing data—this is done to avoid unreliable semi-annual /quarterly data.

IV. If annual data do not exist for the current fiscal year and only annual data exist at both previous and next fiscal year, but they are very different to the semi-annual /quarterly data, then treat the semi-annual/quarterly as missing data—this should avoid having unreliable and too choppy semi-annual/quarterly data between previous and next fiscal year.

B. Data Sources for market indexes and macroeconomic variables

Bloomberg:
- Interest Rates Index (3M, 6M, 1Y, 10Y)
- Eurostat Industrial Production Eurozone Industry Ex Construction YoY WDA
- Eurostat Industrial Production Eurozone Industry Ex Construction MoM SA
- European Commission Economic SentiMent Indicator Eurozone
- European Commission Manufacturing Confidence Eurozone Industrial Confidence
- Sentix Economic Indices Euro Aggregate Overall Index on Euro area
- European Commission Consumer Confidence Indicator Eurozone
- European Commission Euro Area Business Climate Indicator

DataStream:
- DS Market - PRICE INDEX
- DS Banks - PRICE INDEX
- EURO STOXX - PRICE INDEX
- EURO STOXX 50 - PRICE INDEX
- VSTOXX VOLATILITY INDEX - PRICE INDEX
- EU BANKS SECTOR CDS INDEX 5Y

The Bank for International Settlements (BIS):
- Property Price Statistics

Eurostat:
- GDP
- HICP
- Unemployment Rates

European Central Bank (ECB):
- Exchange Rates
- Loan to Households
- Loan to Non-Financial Corporations

Table 1: Total Asset Value Weighted PDs and PDs' Component Rank Correlation between Banking Groups and Luxembourg Banks

	Lux Geske Total	*Lux Geske ST*	*Lux Geske FW*	*Lux Accumulated Common Component Total*	*Lux Accumulated Common Component ST*	*Lux Accumulated Common Component FW*	*Lux Accumulated Idiosyncratic Component Total*	*Lux Accumulated Idiosyncratic Component ST*	*Lux Accumulated Idiosyncratic Component FW*
						2004-2011			
Group Geske Total	**0.51**	**0.50**	**0.45**	**0.47**	**0.45**	**0.66**	**-0.17**	**-0.19**	**-0.12**
Group Geske ST	**0.52**	**0.50**	**0.47**	**0.46**	**0.43**	**0.66**	**-0.15**	**-0.18**	*-0.10*
Group Geske FW	**0.35**	**0.36**	**0.25**	**0.49**	**0.51**	**0.51**	**-0.28**	**-0.30**	**-0.24**
Group Accumulated Common Component Total	**0.62**	**0.60**	**0.57**	**0.34**	**0.31**	**0.54**	*-0.03*	*-0.06*	*0.03*
Group Accumulated Common Component ST	**0.54**	**0.53**	**0.49**	**0.44**	**0.41**	**0.62**	**-0.14**	**-0.16**	*-0.05*
Group Accumulated Common Component FW	**0.50**	**0.50**	**0.42**	**0.17**	**0.19**	**0.35**	**0.12**	*0.09*	**0.12**
Group Accumulated Idiosyncratic Component Total	**-0.57**	**-0.57**	**-0.47**	**-0.28**	**-0.30**	**-0.41**	*-0.02*	*0.01*	*-0.05*
Group Accumulated Idiosyncratic Component ST	**-0.46**	**-0.46**	**-0.36**	**-0.42**	**-0.44**	**-0.47**	**0.13**	**0.16**	*0.06*
Group Accumulated Idiosyncratic Component FW	**-0.51**	**-0.49**	**-0.43**	**-0.15**	**-0.15**	**-0.34**	**-0.19**	**-0.18**	**-0.15**
						2004-2007			
Group Geske Total	*-0.04*	*-0.04*	*0.10*	*0.09*	*0.01*	**0.60**	*-0.14*	*-0.08*	**-0.38**
Group Geske ST	*0.00*	*-0.01*	*0.14*	*0.06*	*-0.02*	**0.57**	*-0.09*	*-0.04*	**-0.32**
Group Geske FW	**-0.46**	**-0.41**	**-0.37**	**0.38**	**0.32**	**0.30**	**-0.59**	**-0.55**	**-0.60**
Group Accumulated Common Component Total	**0.38**	**0.35**	**0.50**	**-0.38**	**-0.47**	*0.09*	**0.28**	**0.33**	**0.18**
Group Accumulated Common Component ST	*0.09*	*0.07*	**0.19**	*-0.01*	*-0.10*	**0.40**	*-0.07*	*-0.04*	*-0.13*
Group Accumulated Common Component FW	**0.36**	**0.34**	**0.43**	**-0.46**	**-0.47**	*-0.12*	**0.42**	**0.44**	**0.38**
Group Accumulated Idiosyncratic Component Total	**-0.25**	**-0.27**	**-0.25**	**0.43**	**0.41**	**0.26**	**-0.20**	**-0.21**	**-0.36**
Group Accumulated Idiosyncratic Component ST	**0.17**	*0.14*	**0.19**	*-0.09*	*-0.12*	*0.05*	**0.26**	**0.28**	*0.07*
Group Accumulated Idiosyncratic Component FW	**-0.37**	**-0.35**	**-0.38**	**0.53**	**0.53**	**0.20**	**-0.52**	**-0.54**	**-0.47**
						2008-2009			
Group Geske Total	**0.42**	**0.38**	*0.17*	**0.51**	**0.38**	*0.07*	*-0.20*	**-0.30**	**0.31**
Group Geske ST	**0.44**	**0.36**	*0.20*	**0.53**	**0.36**	*0.09*	*-0.21*	**-0.28**	**0.34**
Group Geske FW	**0.37**	**0.48**	*-0.13*	**0.32**	**0.53**	*-0.23*	*-0.10*	*-0.20*	*0.16*
Group Accumulated Common Component Total	**0.45**	**0.40**	*0.16*	**0.52**	**0.40**	*0.06*	*-0.17*	**-0.28**	**0.30**
Group Accumulated Common Component ST	**0.46**	**0.39**	*0.20*	**0.55**	**0.38**	*0.10*	*-0.19*	**-0.26**	**0.35**
Group Accumulated Common Component FW	**0.41**	**0.53**	*-0.14*	**0.36**	**0.58**	*-0.24*	*-0.07*	*-0.15*	*0.15*
Group Accumulated Idiosyncratic Component Total	**-0.69**	**-0.78**	*0.08*	**-0.45**	**-0.59**	*0.19*	**-0.32**	*-0.20*	*-0.03*
Group Accumulated Idiosyncratic Component ST	**-0.38**	**-0.49**	*0.18*	*-0.20*	**-0.39**	**0.27**	**-0.25**	*-0.17*	*-0.01*
Group Accumulated Idiosyncratic Component FW	**-0.55**	**-0.49**	*-0.17*	**-0.54**	**-0.36**	*-0.03*	*-0.20*	*-0.17*	*-0.22*
						2010-2011			
Group Geske Total	**0.61**	**0.61**	**0.54**	**0.42**	**0.37**	**0.49**	**0.35**	**0.35**	*0.07*
Group Geske ST	**0.60**	**0.60**	**0.53**	**0.42**	**0.37**	**0.45**	**0.35**	**0.35**	*0.10*
Group Geske FW	**0.42**	**0.42**	**0.33**	*0.05*	*0.00*	**0.42**	**0.48**	**0.50**	*-0.15*
Group Accumulated Common Component Total	**0.70**	**0.70**	**0.61**	**0.27**	*0.22*	**0.54**	**0.54**	**0.54**	*0.05*
Group Accumulated Common Component ST	**0.67**	**0.67**	**0.62**	**0.33**	**0.27**	**0.59**	**0.46**	**0.46**	*0.02*
Group Accumulated Common Component FW	**0.41**	**0.41**	**0.32**	*-0.18*	*-0.23*	**0.48**	**0.65**	**0.70**	*-0.17*
Group Accumulated Idiosyncratic Component Total	**-0.43**	**-0.43**	**-0.34**	*0.12*	*0.15*	**-0.50**	**-0.61**	**-0.67**	*0.19*
Group Accumulated Idiosyncratic Component ST	**-0.48**	**-0.48**	**-0.41**	*0.05*	*0.08*	**-0.51**	**-0.57**	**-0.63**	*0.11*
Group Accumulated Idiosyncratic Component FW	**-0.44**	**-0.44**	**-0.29**	*0.19*	*0.24*	**-0.45**	**-0.66**	**-0.71**	*0.14*

The table reports the Kendall correlation matrix of the monthly PDs and their components between banking groups and Luxembourg banks. For Luxembourg banks, monthly PDs are assumed to be same within each quarter. A bold value with underscore indicates significance at the 95% level, whereas a bold value without underscore indicates significance at the 90% level.

Table 1b: Total Asset Value Weighted DD and DD's Component Rank Correlation between Banking Groups and Luxembourg Banks

	Lux Merton DD	*Lux Accumulated Common Component*	*Lux Accumulated Idiosyncratic Component*
		2004-2010	
Group Merton DD	**0.47**	**0.48**	*-0.11*
Group Accumulated Common Component	**0.41**	**0.41**	**-0.12**
Group Accumulated Idiosyncratic Compone	**0.43**	**0.50**	**-0.18**
		2004-2007	
Group Merton DD	**-0.37**	*0.06*	**-0.59**
Group Accumulated Common Component	**-0.54**	*-0.16*	**-0.62**
Group Accumulated Idiosyncratic Compone	**-0.22**	**0.24**	**-0.46**
		2008-2009	
Group Merton DD	*0.24*	**0.54**	*0.05*
Group Accumulated Common Component	*0.18*	**0.53**	*0.00*
Group Accumulated Idiosyncratic Compone	**0.25**	*-0.12*	**0.39**
		2010-2011	
Group Merton DD	**0.30**	*-0.12*	*0.02*
Group Accumulated Common Component	*0.24*	*-0.05*	*-0.06*
Group Accumulated Idiosyncratic Compone	*0.09*	*-0.20*	*0.11*

The table reports the Kendall correlation matrix of the monthly DD and its components between Banking Groups and Luxembourg Banks. For Luxembourg Banks, monthly DD (risk neutral) are assumed to be same within each quarter. A bold value with underscore indicates significance at the 95% level, whereas a bold value without underscore indicates significance at the 90% level.

Table 2a: Granger Causality Test between Common Components and PDs for Each Bank

	At p-value of 1%		At p-value of 5%		At p-value of 10%	
	Common Component	PDs	Common Component	PDs	Common Component	PDs
Group Geske All	0.31	0.13	0.41	0.13	0.44	0.13
Group Geske ST	0.28	0.19	0.28	0.16	0.34	0.16
Group Geske FW	0.28	0.21	0.28	0.21	0.34	0.21
Group DD	0.25	0.00	0.41	0.00	0.56	0.00
Lux Geske All	0.26	0.08	0.28	0.05	0.31	0.05
Lux Geske ST	0.21	0.08	0.28	0.05	0.28	0.05
Lux Geske FW	0.50	0.03	0.42	0.08	0.44	0.06
Lux DD	0.36	0.00	0.41	0.00	0.44	0.00

This table reports the ratios according to Granger Causality test at the p-values of 1%, 5% and 10% respectively. The measures are ranked by calculating the ratio of the times X Granger causes another measure Y and Y does not Granger causes X to the number of the available banks for banking groups and Luxembourg banks. The ratios under Common Component mean that the Common Component Granger causes PDs and PDs do not Granger cause the Common Component; similarly, for PDs. The standardized measure is constructed by (x-mean(x))/std(x).

Table 2b: Granger Causality Test between Common Components and PDs for Weighted Index

	1/30/2004 - 9/30/2011				**1/30/2004 - 6/30/2008**			
	p-Value				p-Value			
	Common Component	PDs	Common Component	PDs	Common Component	PDs	Common Component	PDs
	Equal Weighted		Total Asset Value Weighted		Equal Weighted		Total Asset Value Weighted	
Group Geske All	**0.02**	**0.39**	0.71	0.54	0.00	0.00	0.00	0.00
Group Geske ST	0.90	0.13	0.24	0.11	0.00	0.00	0.00	0.00
Group Geske FW	0.00	0.00	0.01	0.02	0.00	0.00	**0.51**	**0.00**
Group DD	**0.00**	**0.58**	0.30	0.79	**0.00**	**0.42**	0.38	0.40
Lux Geske All	0.79	0.95	0.74	0.91	0.00	0.00	0.92	0.64
Lux Geske ST	0.45	0.49	0.59	0.84	**0.04**	**0.00**	**0.00**	**0.02**
Lux Geske FW	**0.68**	**0.00**	**0.91**	**0.00**	**0.90**	**0.00**	**0.00**	**0.86**
Lux DD	0.34	0.62	0.88	0.45	**0.00**	**0.48**	0.00	0.00
	Total Interbank Lending Weighted		Total Interbank Borrowing Weighted		Total Interbank Lending Weighted		Total Interbank Borrowing Weighted	
Lux Geske All	0.84	0.70	0.84	0.77	0.84	0.68	**0.00**	**0.05**
Lux Geske ST	0.67	0.94	0.82	0.96	0.90	0.63	**0.00**	**0.75**
Lux Geske FW	**0.62**	**0.00**	0.68	0.86	**0.95**	**0.00**	0.00	0.00
Lux DD	0.69	0.61	0.94	0.85	0.00	0.00	0.00	0.00

This table reports the p-value according to Granger Causality tests. The p_value under Common Component refers to the test of the null hypothesis that the Common Component does not Granger cause PDs; similarly, for PDs. The standardized measure is constructed by (x-mean(x))/std(x).

Table 3a: Banking Groups - Factors' Contributions to the Early-Warning Features of FW PDs (relative to ST PDs)

Classification	FW Scores minus ST Scores	Scores Difference at 1% Top Quantile	Scores Differenceat 5% Top Quantile	Scores Differenceat 10% Top Quantile	Scores Difference at 20% Top Quantile	FW Scores minus ST Scores	Scores Difference at 1% Top Quantile	Scores Differencea t 5% Top Quantile	Scores Differencea t 10% Top Quantile	Scores Difference at 20% Top Quantile
	30-Jan-2004 to 31-May-2007					31-Jan-2008 to 30-Sep-20011				
Real	**1.91%**	**4.87%**	**7.12%**	**6.56%**	**6.27%**	**6.33%**	**3.56%**	**5.48%**	**5.12%**	**5.12%**
Funding prices	-2.37%	3.28%	-1.39%	-2.38%	-5.05%	-5.03%	0.89%	-2.26%	-2.93%	-4.94%
Funding quantities	**1.14%**	**2.37%**	**2.67%**	**2.39%**	**2.95%**	**4.11%**	**2.38%**	**2.55%**	**2.22%**	**2.58%**
Confidence	-0.42%	0.00%	0.00%	-0.59%	-0.69%	0.14%	0.00%	0.00%	-0.53%	-0.55%
PDs All	-0.99%	-10.20%	-7.39%	-4.40%	-3.60%	-4.29%	-8.19%	-6.15%	-3.58%	-2.96%
PDs ST	-1.21%	-10.51%	-6.57%	-6.31%	-3.80%	-4.48%	-10.03%	-5.66%	-5.31%	-3.22%
PDs FW	**1.93%**	**10.20%**	**5.57%**	**4.73%**	**3.91%**	**3.22%**	**11.38%**	**6.03%**	**5.02%**	**3.97%**

Table 3b: Luxembourg Banks - Factors' Contributions to the Early-Warning Features of FW PDs (relative to ST PDs)

Classification	FW Scores minus ST Scores	Scores Difference at 1% Top Quantile	Scores Differenceat 5% Top Quantile	Scores Differenceat 10% Top Quantile	Scores Difference at 20% Top Quantile	FW Scores minus ST Scores	Scores Difference at 1% Top Quantile	Scores Differencea t 5% Top Quantile	Scores Differencea t 10% Top Quantile	Scores Difference at 20% Top Quantile
	30-Jan-2004 to 31-May-2007					31-Jan-2008 to 30-Sep-20011				
Real	-0.14%	-0.31%	-0.63%	-0.19%	-0.05%	0.14%	0.04%	0.00%	0.59%	0.88%
Funding prices	-0.65%	-2.87%	-1.60%	-1.23%	-1.25%	-1.46%	-4.70%	-3.12%	-2.78%	-2.96%
Funding quantities	**0.28%**	**2.02%**	**1.81%**	**1.06%**	**0.84%**	**0.37%**	**1.98%**	**1.61%**	**1.08%**	**1.12%**
Confidence	-0.04%	0.00%	0.00%	-0.11%	0.03%	-0.10%	0.00%	0.00%	-0.23%	-0.04%
PDs All	0.11%	1.70%	-0.02%	0.00%	-0.05%	0.05%	1.01%	-0.80%	-0.42%	-0.46%
PDs ST	-0.05%	-2.01%	-0.77%	-0.46%	-0.26%	-0.11%	-3.77%	-1.31%	-1.00%	-0.56%
PDs FW	**0.48%**	**1.47%**	**1.22%**	**0.93%**	**0.75%**	**1.11%**	**5.43%**	**3.62%**	**2.77%**	**2.03%**

Note: Real variables are GDP in volume and nominal, industrial production, the unemployment rate, the HICP, and agricultrual and industrial property prices. Funding prices are short- and long-term interest rates, foreign exchange rates, stock market prices, stock price volatility, house prices. Funding quantities are total credit, loans to households, mortgages, loans to non-financial firms, and interbank lending and borrowing. Confidence includes various indices of consumer and business sentiment.

Table 4a: Geske PDs Forecast Evaluation for Banking Groups and Luxembourg Banks

	Coverage Ratio												
	Q 5%-95%	Q 10%-90%	Q 15%-85%	Q 20%-80%	Q 25%-75%	Q 30%-70%	Q 35%-65%	Q 40%-60%	Q 45%-55%	RMS Error	Bias Proportion	Variance Proportion	Coviance Proportion
							Common Component						
2nd Month	0.724	0.607	0.501	0.412	0.342	0.271	0.200	0.133	0.067	0.038	0.004	0.019	0.977
3th Month	0.709	0.566	0.462	0.376	0.316	0.246	0.187	0.119	0.060	0.044	0.010	0.025	0.965
4th Month	0.705	0.559	0.457	0.383	0.314	0.242	0.181	0.124	0.063	0.051	0.016	0.028	0.956
5th Month	0.707	0.555	0.461	0.379	0.310	0.238	0.180	0.122	0.059	0.057	0.015	0.027	0.958
6th Month	0.704	0.563	0.456	0.381	0.316	0.251	0.186	0.125	0.065	0.063	0.014	0.026	0.960
							Common & Idiosyncratic Component						
1th Month	0.857	0.773	0.689	0.596	0.503	0.404	0.310	0.220	0.115	0.031	0.004	0.014	0.981
2nd Month	0.854	0.747	0.649	0.568	0.475	0.381	0.300	0.207	0.108	0.042	0.005	0.015	0.980
3th Month	0.864	0.748	0.649	0.554	0.467	0.369	0.274	0.184	0.093	0.047	0.010	0.016	0.974
4th Month	0.870	0.751	0.649	0.555	0.469	0.377	0.282	0.188	0.105	0.055	0.015	0.018	0.967
5th Month	0.874	0.753	0.647	0.556	0.475	0.382	0.288	0.190	0.095	0.061	0.012	0.014	0.973
6th Month	0.875	0.759	0.648	0.549	0.463	0.374	0.284	0.191	0.095	0.066	0.011	0.012	0.977

The table reports the coverage ratios, root mean square erros, and the proportions of bias, variance, and covariance respectively from 2010 to 2011 across all Gesk's PDs for both banking groups and luxembourg banks. The coverage ratio is the proportion of banks whose empirical cdf (simulated) at each of the observed PDs are within the range of quantiles.

Table 4b: Merton DD Forecast Evaluation for Banking Groups and Luxembourg Banks

	Coverage Ratio												
	Q 5%-95%	Q 10%-90%	Q 15%-85%	Q 20%-80%	Q 25%-75%	Q 30%-70%	Q 35%-65%	Q 40%-60%	Q 45%-55%	RMS Error	Bias Proportion	Variance Proportion	Covariance Proportion
							Common Component						
2nd Month	0.690	0.596	0.512	0.420	0.332	0.259	0.189	0.121	0.055	0.360	0.003	0.000	0.997
3th Month	0.653	0.538	0.450	0.365	0.283	0.223	0.151	0.097	0.043	0.433	0.010	0.000	0.990
4th Month	0.652	0.517	0.434	0.357	0.293	0.237	0.174	0.122	0.061	0.503	0.018	0.000	0.982
5th Month	0.648	0.527	0.434	0.352	0.282	0.214	0.153	0.106	0.048	0.561	0.012	0.001	0.987
6th Month	0.639	0.504	0.412	0.336	0.268	0.202	0.151	0.107	0.061	0.623	0.009	0.005	0.986
							Common & Idiosyncratic Component						
1th Month	0.896	0.835	0.738	0.660	0.570	0.452	0.347	0.232	0.118	0.252	0.001	0.002	0.997
2nd Month	0.886	0.805	0.712	0.621	0.522	0.422	0.311	0.212	0.097	0.368	0.000	0.001	0.999
3th Month	0.890	0.792	0.694	0.600	0.511	0.406	0.299	0.196	0.101	0.435	0.001	0.001	0.999
4th Month	0.888	0.788	0.691	0.585	0.496	0.408	0.314	0.213	0.118	0.504	0.002	0.001	0.997
5th Month	0.882	0.789	0.696	0.596	0.510	0.415	0.325	0.210	0.100	0.562	0.000	0.004	0.996
6th Month	0.882	0.806	0.712	0.594	0.491	0.382	0.287	0.197	0.098	0.621	0.000	0.009	0.991

The table reports the coverage ratios, root mean square erros, and the proportions of bias, variance, and covariance respectively from 2010 to 2011 across all Merton's DPs for both banking groups and luxembourg banks. The coverage ratio is the proportion of banks whose empirical cdf (simulated) at each of the observed DDs are within the range of quantiles.

Table 5a: Geske Total Asset Weighted PDs Forecast Evaluation for Banking Groups and Luxembourg Banks

	RMS Error	Bias Proportion	Variance Proportion	Coviance Proportion	RMS Error	Bias Proportion	Variance Proportion	Coviance Proportion	RMS Error	Bias Proportion	Variance Proportion	Covariance Proportion
	Groups Geske All				Groups Geske ST				Groups Geske FW			
					Common Component							
2nd Month	0.022	0.025	0.074	0.901	0.020	0.014	0.088	0.899	0.003	0.000	0.008	0.992
3th Month	0.021	0.008	0.217	0.776	0.019	0.003	0.192	0.805	0.003	0.005	0.151	0.844
4th Month	0.023	0.001	0.286	0.713	0.022	0.004	0.270	0.726	0.003	0.032	0.244	0.724
5th Month	0.025	0.000	0.193	0.807	0.023	0.001	0.172	0.827	0.003	0.036	0.244	0.720
6th Month	0.028	0.002	0.033	0.965	0.026	0.002	0.013	0.985	0.004	0.085	0.336	0.579
					Common & Idiosyncratic Component							
1th Month	0.017	0.005	0.325	0.670	0.015	0.002	0.279	0.719	0.002	0.023	0.006	0.971
2nd Month	0.024	0.000	0.144	0.856	0.021	0.005	0.102	0.893	0.004	0.081	0.024	0.895
3th Month	0.022	0.003	0.264	0.733	0.020	0.001	0.157	0.842	0.004	0.177	0.289	0.533
4th Month	0.024	0.067	0.383	0.550	0.023	0.022	0.246	0.732	0.004	0.308	0.387	0.305
5th Month	0.027	0.055	0.261	0.684	0.024	0.016	0.171	0.813	0.005	0.289	0.369	0.341
6th Month	0.031	0.031	0.062	0.907	0.027	0.003	0.017	0.981	0.007	0.307	0.397	0.296
	Lux Geske All				Lux Geske ST				Lux Geske FW			
					Common Component							
1th Month	0.009	0.224	0.043	0.734	0.008	0.229	0.031	0.741	0.001	0.104	0.117	0.779
2nd Month	0.012	0.323	0.028	0.649	0.012	0.359	0.016	0.625	0.002	0.017	0.114	0.869
3th Month	0.014	0.482	0.013	0.505	0.014	0.527	0.006	0.467	0.002	0.029	0.157	0.814
4th Month	0.017	0.464	0.050	0.486	0.017	0.512	0.034	0.454	0.003	0.069	0.211	0.720
5th Month	0.020	0.472	0.081	0.446	0.019	0.528	0.058	0.414	0.003	0.057	0.291	0.651
6th Month	0.021	0.512	0.120	0.369	0.020	0.572	0.087	0.341	0.003	0.055	0.422	0.523
					Common & Idiosyncratic Component							
1th Month	0.008	0.240	0.041	0.720	0.008	0.266	0.019	0.716	0.002	0.031	0.133	0.837
2nd Month	0.010	0.330	0.027	0.643	0.011	0.416	0.006	0.578	0.002	0.004	0.085	0.911
3th Month	0.012	0.490	0.019	0.491	0.013	0.605	0.001	0.394	0.002	0.003	0.188	0.810
4th Month	0.015	0.404	0.072	0.524	0.015	0.527	0.031	0.441	0.003	0.000	0.217	0.783
5th Month	0.017	0.378	0.119	0.503	0.017	0.506	0.063	0.432	0.003	0.009	0.268	0.723
6th Month	0.018	0.388	0.181	0.431	0.018	0.523	0.103	0.374	0.003	0.021	0.384	0.595

The table reports the root mean square errors, and the proportions of bias, variance, and covariance respectively from 2010 to 2011 across Geskr's PDs for both banking groups and Luxembourg banks.

Table 5b: Merton Total Asset Weighted DDs Forecast Evaluation for Banking Groups and Luxembourg Banks

	Group Merton				Lux Merton			
	RMS Error	Bias Proportion	Variance Proportion	Coviance Proportion	RMS Error	Bias Proportion	Variance Proportion	Coviance Proportion
				Common Component				
2nd Month	0.294	0.000	0.013	0.986	0.061	0.177	0.002	0.821
3th Month	0.334	0.002	0.000	0.997	0.074	0.311	0.001	0.688
4th Month	0.377	0.014	0.001	0.985	0.082	0.446	0.012	0.542
5th Month	0.429	0.002	0.001	0.998	0.091	0.540	0.041	0.419
6th Month	0.466	0.000	0.000	1.000	0.102	0.649	0.080	0.271
				Common & Idiosyncratic Component				
1th Month	0.192	0.012	0.016	0.972	0.040	0.001	0.002	0.998
2nd Month	0.299	0.007	0.015	0.979	0.055	0.008	0.000	0.992
3th Month	0.340	0.001	0.000	0.999	0.062	0.027	0.001	0.972
4th Month	0.382	0.001	0.001	0.998	0.062	0.034	0.017	0.950
5th Month	0.442	0.003	0.001	0.996	0.064	0.031	0.071	0.898
6th Month	0.483	0.013	0.000	0.987	0.064	0.056	0.187	0.758

The table reports the root mean square erros, and the proportions of bias, variance, and covariance respectively from 2010 to 2011 across Merton's DDs for both banking groups and Luxembourg banks.

Figure 1a - Coherence and Phase Angle between Common Components and Banking Groups' ST PDs

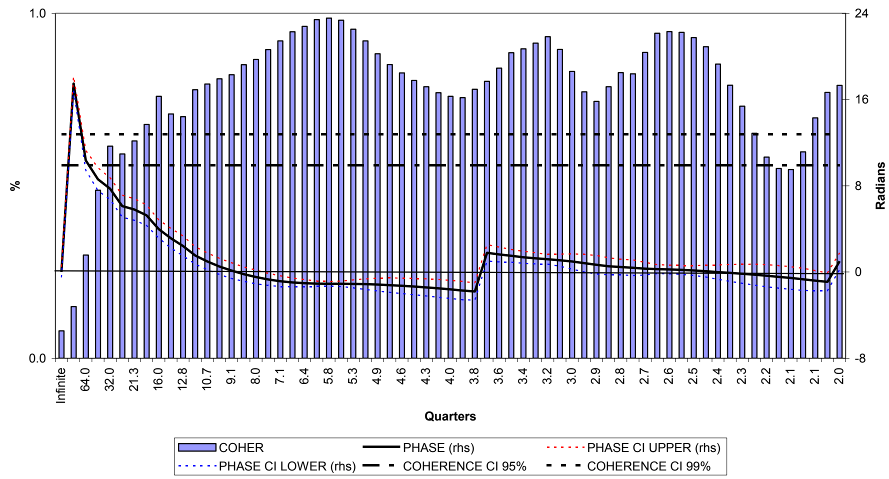


Figure 1b - Coherence and Phase Angle between Common Components and Banking Groups' FW PDs

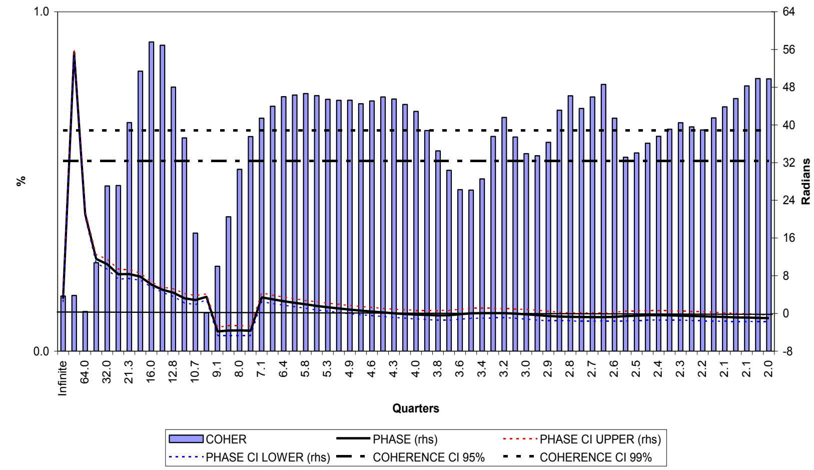


Figure 1c - Coherence and Phase Angle between Common Components and Luxembourg Banks' ST PDs

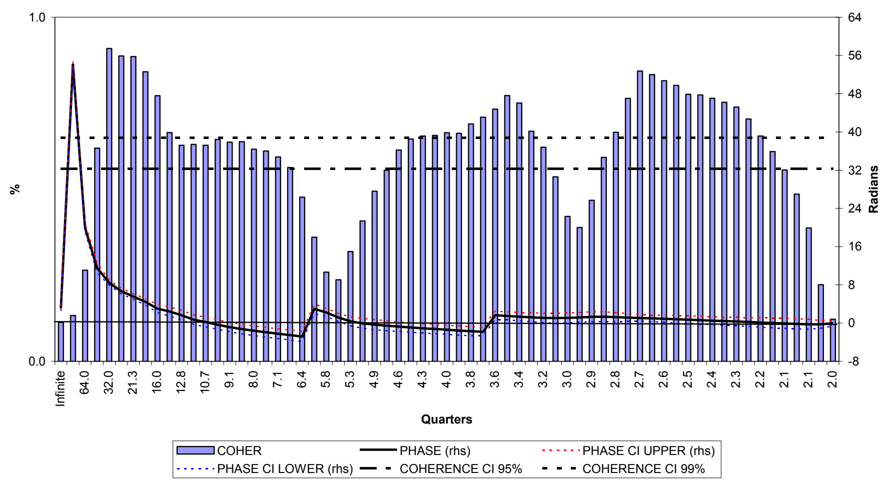


Figure 1d - Coherence and Phase Angle between Common Components and Luxembourg Banks' FW PDs

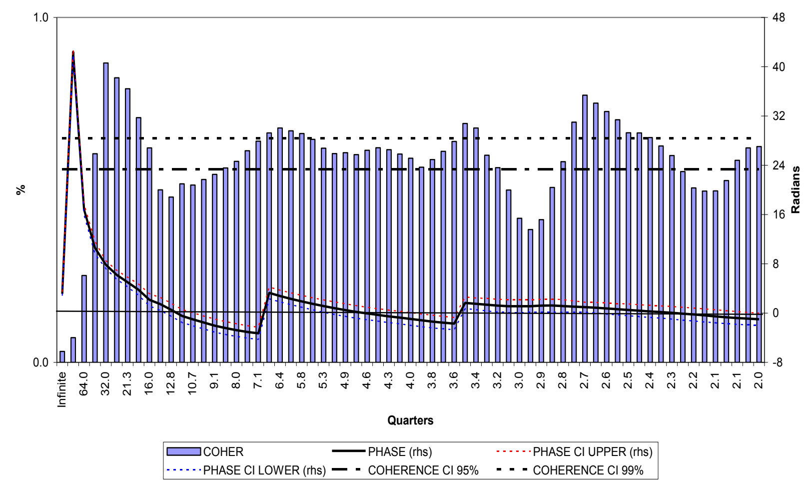


Figure 2a - Equal Weighted DP Index, Accumulated Common Factor and One-period Forecasts for Banking Groups (Geske ST)

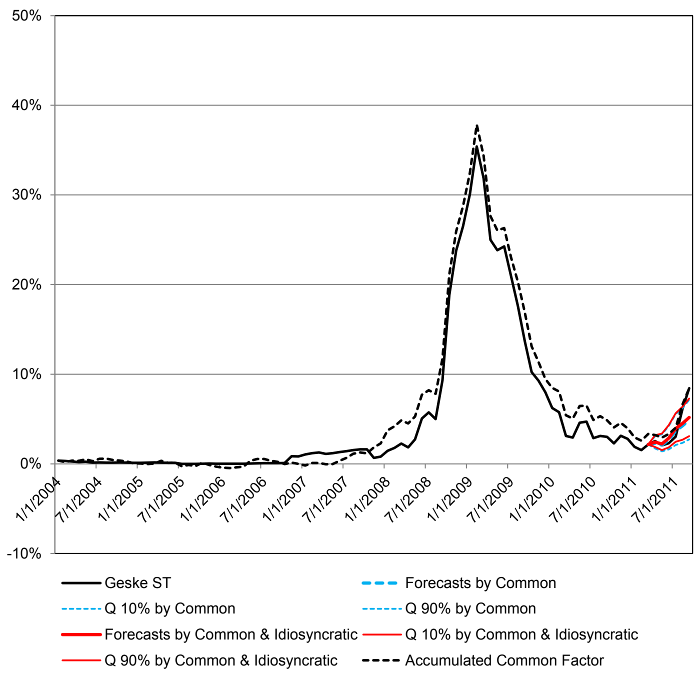


Figure 2b - Equal Weighted DP Index, Accumulated Common Factor and One-period Forecasts for Banking Groups (Geske FW)

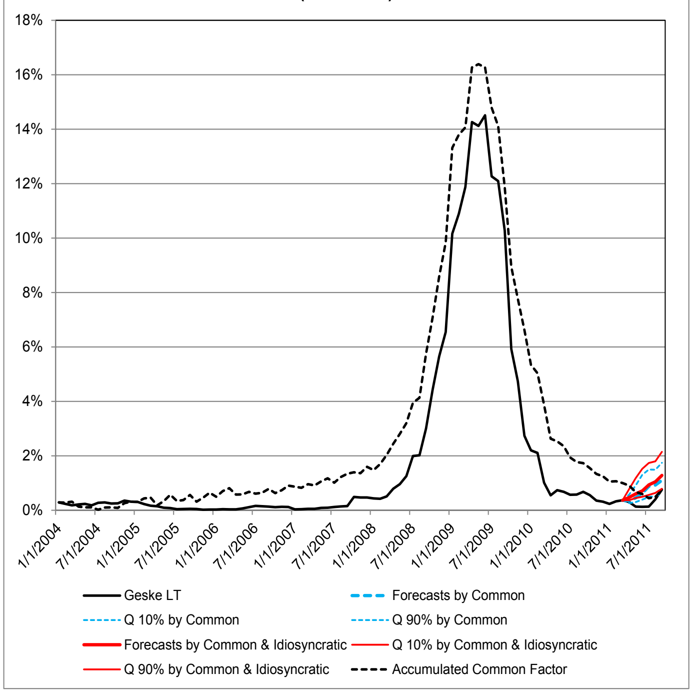


Figure 2c - Equal Weighted DP Index, Accumulated Common Factor and One-period Forecasts for Luxembourg Banks (Geske ST)

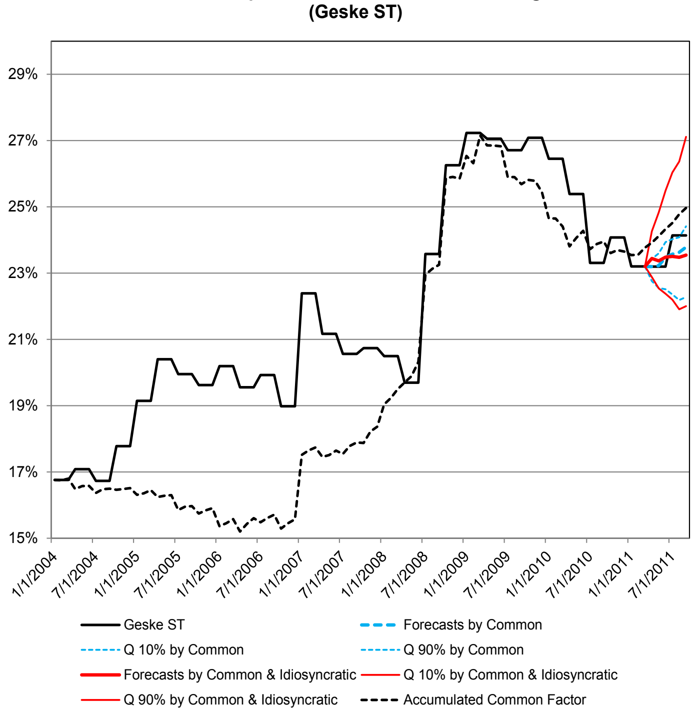


Figure 2d - Equal Weighted DP Index, Accumulated Common Factor and One-period Forecasts for Luxembourg Banks (Geske FW)

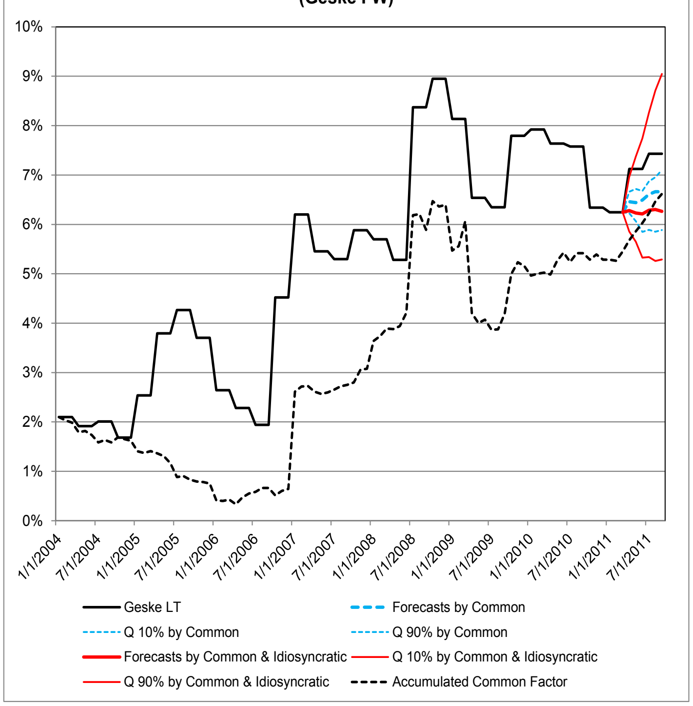


Figure 3a - Asset Weighted DP Index, Accumulated Common Factor and One-period Forecasts for Banking Groups (Geske ST)

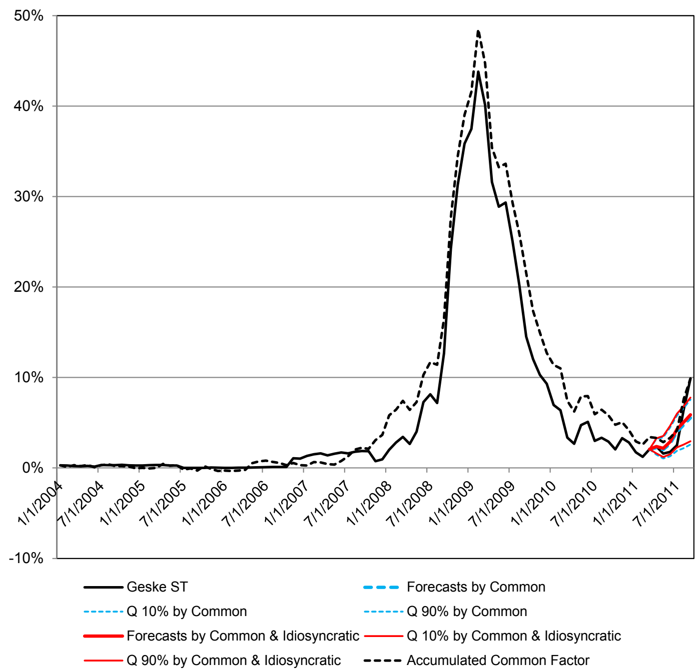


Figure 3b - Asset Weighted DP Index, Accumulated Common Factor and One-period Forecasts for Banking Groups (Geske FW)

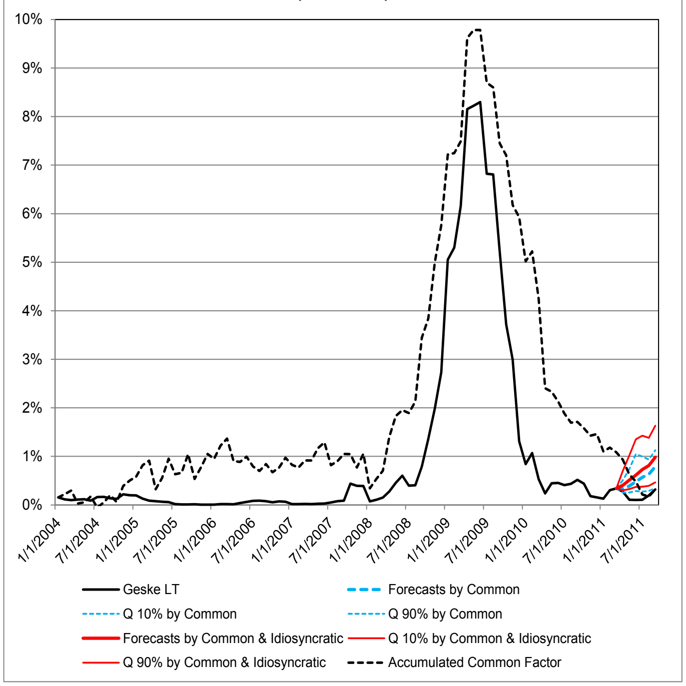


Figure 3c - Asset Weighted DP Index, Accumulated Common Factor and One-period Forecasts for Luxembourg Banks (Geske ST)

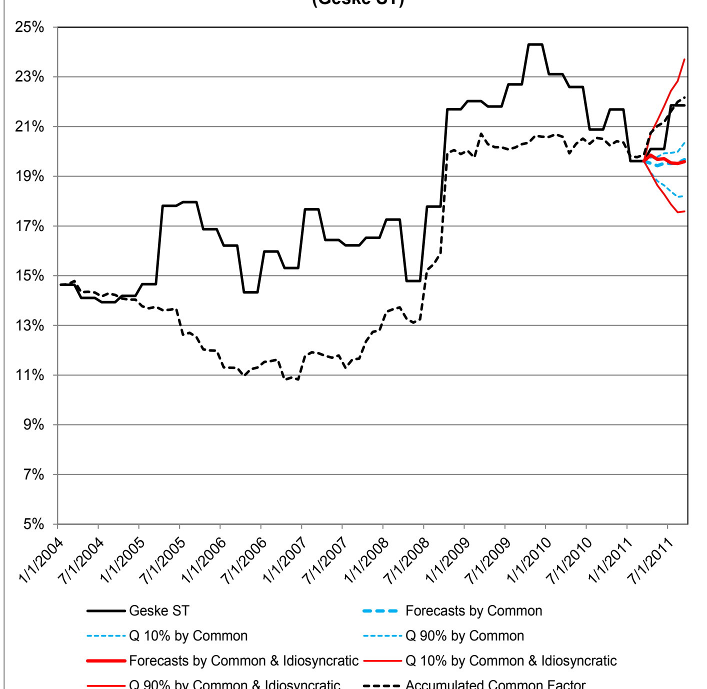
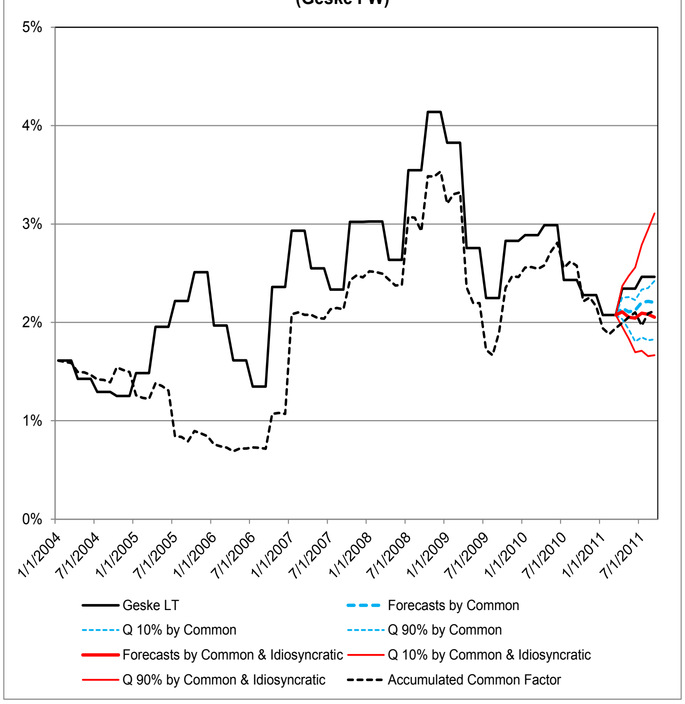


Figure 3d - Asset Weighted DP Index, Accumulated Common Factor and One-period Forecasts for Luxembourg Banks (Geske FW)



BANQUE CENTRALE DU LUXEMBOURG
EUROSYSTÈME

2, boulevard Royal
L-2983 Luxembourg

Tél.: +352 4774-1
Fax: +352 4774 4910

www.bcl.lu • info@bcl.lu