

Observation Driven Mixed-Measurement Dynamic Factor Models with an Application to Credit Risk

Drew Creal^a, Bernd Schwaab^{b*} Siem Jan Koopman^{c,e}, André Lucas^{d,e}

(^a) *Booth School of Business, University of Chicago*

(^b) *European Central Bank*

(^c) *Department of Econometrics, VU University Amsterdam*

(^d) *Department of Finance, VU University Amsterdam & Duisenberg School of Finance*

(^e) *Tinbergen Institute, Amsterdam*

July 18, 2012

Abstract

We propose an observation driven dynamic factor model for mixed-measurement and mixed-frequency panel data. In this framework time series observations may come from a range of families of parametric distributions, may be observed at different frequencies, may have missing observations, and may exhibit common dynamics and cross-sectional dependence due to shared exposures to dynamic latent factors. A convenient feature of our model is that the likelihood function is known in closed form and can be computed in a straightforward way. This enables parameter estimation using standard maximum likelihood methods. We adopt the new mixed-measurement framework in an empirical study for signal extraction and forecasting of macro, credit, and loss given default risk conditions for U.S. Moody's-rated firms from January 1982 until March 2010.

Keywords: panel data; loss given default; default risk; dynamic beta density; dynamic ordered probit; dynamic factor model.

JEL classification codes: C32, G32.

*We thank Emanuel Moench (our discussant at the 2011 AFA meetings in Denver), and Michel van der Wel for earlier comments. We are grateful to seminar participants at CReATES at Aarhus, the Forecasting and Empirical Methods group at the NBER Summer Institute, Tinbergen Institute, the University of Chicago Booth School of Business, University of British Columbia, University of Wisconsin Madison, University of StGallen, and VU University Amsterdam. Schwaab thanks the C. Willems Stichting for financial support while visiting the University of Chicago, Booth School of Business. Lucas thanks the Dutch National Science Foundation (NWO) for financial support. Corresponding author: Drew Creal, University of Chicago, Booth School of Business, 5807 S. Woodlawn Ave., Chicago, IL 60637; Tel 773.834.5249, Email: dcreal@chicagobooth.edu. The views expressed in this paper are those of the authors and do not necessarily reflect the views of the European Central Bank or the European System of Central Banks.

1 Introduction

Consider an unbalanced panel time series y_{it} for $i = 1, \dots, N$ and $t = 1, \dots, T$, where each variable can come from a different distribution. Such heterogeneous panel data sets may occur in many areas of economics and finance. The need for a joint modeling framework for variables from different distributions with common features has become more apparent in recent years with the increasing availability of data resources. For example, the construction of an accurate and reliable business cycle indicator requires many different measurements of economic activity. Some of these may have a Gaussian distribution (typical macroeconomic time series), whereas others are fat-tailed (such as stock returns), integer-valued or binary (such as the well-known NBER recession dates), or categorical (such as low/moderate/high consumer confidence levels) variables. All of these variables reflect a common exposure to the business cycle, but at the same time each variable requires its own appropriate distribution. We propose an observation driven dynamic modeling framework for the simultaneous analysis of mixed-measurement time series that are subject to common features.

An additional challenge in multiple time series analysis is that the observation frequencies can be different for each time series. Some series are observed every year while other series are observed every quarter or month. A simultaneous analysis of time series with different observation frequencies is a challenging task. Different methodologies have been developed for this purpose. For example, Mariano and Murasawa (2003) adopt a state space approach for the construction of a coincident business cycle index using quarterly and monthly data while Ghysels, Santa Clara, and Valkanov (2006) adopt a mixed-data sampling analysis for predicting the volatility of financial time series using intra-daily returns of different frequencies. Our mixed-measurement modeling framework incorporates a mixed-data sampling approach by explicitly formulating a high-frequency time series process and allowing for missing observations in the analysis.

Our main motivation to develop a mixed-measurement, mixed-frequency dynamic modeling framework is for the estimation, analysis and forecasting of credit risk. Credit risk analysis has become highly relevant in the aftermath of the 2008 financial crisis. Financial institutions and regulators are specifically trying to assess what is the common variation in firm defaults in order to correctly assess risk. In our empirical analysis we focus on the systematic variation in cross-

sections of macroeconomic data, credit rating transitions, and bond loss rates upon default (also known as loss given default). Our data set exhibits the complications as discussed above. While the number of credit rating transitions between rating categories is modeled as a discrete, ordered random variable, the macroeconomic variables are modeled as continuous variables, and the percentage amounts lost on the principal in case of default are modeled as continuous and bounded between zero and one. Some of the macro series are observed quarterly while others are observed monthly. Furthermore, the loss given defaults are only observed if there are defaults, such that we have many missing observations in these series by construction. Finally, all series exhibit some common dynamic features related to the business cycle. Loss rates and defaults both tend to be high during an economic downturn, indicating important systematic covariation across different types of data. In our modeling framework, the commonalities are captured by latent dynamic factors. The total data set forms an unbalanced panel with 19,540 rating transition events for 7,505 companies, 1,342 cases of (irregularly spaced) defaults with associated losses given default, and six selected macroeconomic series of mixed quarterly and monthly frequency.

After the parameters in the model have been estimated, we can use the model to forecast credit risk conditions in the economy and to construct predictive loss densities for portfolios of corporate bonds at different forecasting horizons. The model can therefore be used to stress test current credit portfolios and determine adequate capital buffers using the high percentiles of the simulated portfolio loss distributions. Our modeling framework provides a relatively simple observation driven alternative to the parameter driven frailty models of McNeil and Wenden (2007), Koopman, Lucas, and Monteiro (2008), and Duffie, Eckner, Horel, and Saita (2009). In addition, our proposed model allows for the identification of three components of credit risk simultaneously: macro risk, rating migration/default risk, and loss given default risk. Earlier models have concentrated on defaults only, defaults and ratings, or defaults and macro risk.

Our proposed modeling framework is entirely *observation driven*. This is a distinguishing feature of our approach. It allows parameters to vary over time as functions of lagged dependent variables and exogenous variables. The time-varying parameters are stochastic, but perfectly predictable given the past information. The alternative class of *parameter driven* models, by contrast, does not share this property of perfect predictability; see Cox (1981) for a more detailed discussion of the two classes of models. The main advantage of an observation driven

approach is that the likelihood is known in closed form. This leads to simple procedures for likelihood evaluation and in particular avoids the need for simulation based methods to evaluate the likelihood function. Observation driven time series models have become popular in the applied statistics and econometrics literature. Typical examples of these models include the generalized autoregressive conditional heteroskedasticity (GARCH) model of Engle (1982) and Bollerslev (1986), the autoregressive conditional duration (ACD) model of Engle and Russell (1998), and the dynamic conditional correlation (DCC) model of Engle (2002). In the same spirit, we develop a panel data model for mixed-frequency observations from different families of parametric distributions which are linked by a small set of latent dynamic factors. The likelihood function is available in closed form and can be maximized in a straightforward way.

A number of well-known methods for the modeling of large time series panels and based on latent dynamic factors have been explored in the literature, such as (i) principal components analysis in an approximate dynamic factor model framework, see, for example, Connor and Korajczyk (1988, 1993), Stock and Watson (2002), Bai (2003), Bai and Ng (2002, 2007); (ii) frequency-domain estimation, see, for example, Sargent and Sims (1977), Geweke (1977), Forni, Hallin, Lippi, and Reichlin (2000, 2005); and (iii) signal extraction using state space time series analysis, see for example, Engle and Watson (1981), Watson and Engle (1983), Doz, Giannone, and Reichlin (2006) and Jungbacker and Koopman (2008). When compared to the approaches of (i) and (ii), our current framework provides an integrated parametric framework for obtaining in-sample estimates and out-of-sample forecasts for the latent factors and other variables in the model. When compared to the state space methods of (iii), the likelihood function in our modeling framework is known in closed form, even when the model is fully or partially nonlinear and/or when it includes non-Gaussian densities. In our framework we provide basic and simple procedures for likelihood evaluation and parameter estimation without compromising the flexibility of model formulations to construct effective forecasting distributions.

In Section 2, we introduce observation driven mixed-measurement dynamic factor models. We present our empirical study using the new framework for the joint modeling of macroeconomic dynamics, credit rating and default dynamics, and losses given default dynamics in Section 3. In Section 4, we use the new model to estimate and forecast time-varying credit risk and loss given default risk factors jointly with macroeconomic variables. Section 5 concludes. An online Appendix is available with additional estimation and model specification results.

2 Mixed-measurement dynamic factor models

2.1 Model specification

Consider the $N \times 1$ vector of variables y_t . At time t , N_t elements are observed and $N - N_t$ elements are treated as missing, with $1 \leq N_t \leq N$. Since different time series can be observed at different frequencies and each time series can be observed within different time intervals, missing observations are a common feature in our analysis. The measurement density for the i th element of y_t is given by

$$y_{it} \sim p_i(y_{it}|f_t, \mathcal{F}_{t-1}; \psi), \quad \text{for } i = 1, \dots, N, \quad t = 1, \dots, T, \quad (1)$$

where f_t is a vector of unobserved factors or time-varying parameters, $\mathcal{F}_t = \{y_1, \dots, y_t\}$ is the set of past and concurrent observations at time t , and ψ is a vector of static unknown parameters. In our mixed-measurement framework, the densities $p_i(y_{it}|f_t, \mathcal{F}_{t-1}; \psi)$, for $i = 1, \dots, N$, can originate from different families of distributions. All distributions, however, depend upon the same $M \times 1$ vector of common unobserved factors f_t . We assume a factor model structure in which the y_{it} 's at time t are cross-sectionally independent conditional on f_t and on the information set \mathcal{F}_{t-1} . We then have

$$\log p(y_t|f_t, \mathcal{F}_{t-1}; \psi) = \sum_{i=1}^N \delta_{it} \log p_i(y_{it}|f_t, \mathcal{F}_{t-1}; \psi), \quad (2)$$

where δ_{it} takes the value one when y_{it} is observed and zero when it is missing. The density in (2) may also depend on a vector of exogenous covariates. We omit this extension here to keep the notation simple. We further emphasize that the notation above slightly deviates from Creal, Koopman, and Lucas (2012), where the conditioning information includes f_t, \dots, f_1 . Given that our model is observation driven, extending the conditional information in (2) by f_t, \dots, f_1 is not needed since it is subsumed by the conditioning set \mathcal{F}_{t-1} . Furthermore, f_t is a known, deterministic function of \mathcal{F}_{t-1} and hence the conditioning on f_t is also redundant. We leave f_t in our notation to associate the process of time-varying parameter f_t with (2).

The dynamic factor f_t is modeled as an autoregressive moving average process given by

$$f_{t+1} = \omega + \sum_{i=1}^p A_i s_{t-i+1} + \sum_{j=1}^q B_j f_{t-j+1}, \quad t = 1, \dots, T, \quad (3)$$

where s_1, \dots, s_T is a martingale difference sequence with mean zero, ω is an $M \times 1$ vector of constants and the coefficients A_i and B_j are $M \times M$ parameter matrices for $i = 1, \dots, p$ and $j = 1, \dots, q$. The coefficients can be specified and restricted so that the process f_t is covariance stationary. The unknown static parameters in (1) together with the unknown elements in ω , A_1, \dots, A_p and B_1, \dots, B_q are collected in the static parameter vector ψ . The initial value f_1 is taken as fixed at the unconditional mean of the stationary process f_t .

We follow Creal, Koopman, and Lucas (2012) by setting the innovation s_t in (3) equal to the score of the log-density $p(y_t|f_t, \mathcal{F}_{t-1}; \psi)$, for $t = 1, \dots, T$. In particular, s_t is defined as

$$s_t = S_t \nabla_t, \quad \text{where} \quad \nabla_t = \frac{\partial \log p(y_t|f_t, \mathcal{F}_{t-1}; \psi)}{\partial f_t}, \quad (4)$$

and where S_t is an appropriately chosen scaling matrix. The scaled score s_t in (4) is a function of past observations, factors, and unknown parameters. It follows immediately from the properties of the score that the sequence s_1, \dots, s_t is a martingale difference. The dynamic factors f_t are therefore driven by a sequence of natural innovations.

For particular choices of the measurement density $p(y_t|f_t, \mathcal{F}_{t-1}; \psi)$ and the scaling matrix S_t , Creal, Koopman, and Lucas (2012) show that the modeling framework (2)–(3) reduces to popular models such as the GARCH model of Engle (1982) and Bollerslev (1986), the ACD model of Engle and Russell (1998), the multiplicative error model of Engle and Gallo (2006), as well as other models. For our mixed-measurement model in which we allow for missing observations and for different observation frequencies, we construct the scaling matrix from the eigendecomposition of the Fisher information matrix as given by

$$\mathcal{I}_t = E_{t-1}[\nabla_t \nabla_t'] = E[\nabla_t \nabla_t' | \mathcal{F}_{t-1}].$$

The eigendecomposition of the matrix \mathcal{I}_t is denoted as

$$\mathcal{I}_t = U_t \Sigma_t U_t',$$

with the columns of the $M \times r$ matrix U_t equal to the eigenvectors of \mathcal{I}_t corresponding to its nonzero eigenvalues, and the $r \times r$ diagonal matrix Σ_t containing the nonzero eigenvalues of \mathcal{I}_t . We have implicitly defined r as the rank of \mathcal{I}_t . The scaling matrix is then given by

$$S_t = U_t \Sigma_t^{-1/2} U_t', \quad (5)$$

which can be regarded as the generalized square root inverse matrix of \mathcal{I}_t . Given that S_t is based on the Fisher information matrix, the gradient ∇_t is corrected for the local curvature of the measurement density $p(y_t|f_t, \mathcal{F}_{t-1}; \psi)$ at time t . It also ensures that the martingale difference series s_t has a finite, idempotent covariance matrix. For example, when the information matrix is nonsingular, the covariance matrix of s_t equals the identity matrix for all times t .

In the mixed-measurement setting with measurement densities specified by (1) and (2), the score vector at time t takes the simple additive form

$$\nabla_t = \sum_{i=1}^N \delta_{it} \nabla_{i,t} = \sum_{i=1}^N \delta_{it} \frac{\partial \log p_i(y_{it}|f_t, \mathcal{F}_{t-1}; \psi)}{\partial f_t}, \quad (6)$$

where δ_{it} is defined below (2). Similarly, the conditional information matrix is also additive,

$$\mathcal{I}_t = E_{t-1}[\nabla_t \nabla_t'] = \sum_{i=1}^N \delta_{it} E_{i,t-1}[\nabla_{i,t} \nabla_{i,t}']. \quad (7)$$

It is therefore straightforward to compute the scaling matrix in (5).

We stress the difference between our current approach and the state space approach to dynamic factor analysis; see, for example, Engle and Watson (1981) and Watson and Engle (1983). Our approach here is based on an observation driven time series model as defined in Cox (1981). This means that the value of f_t is known conditional on \mathcal{F}_{t-1} , because f_t is a deterministic function of past data. As a result, the likelihood function is known analytically and parameter estimation is straightforward; see also Section 2.3. The factors can also easily be estimated using the model updating equation (3) by noting that the estimate of f_t depends on past values of y_t only. By contrast, in a state space dynamic factor analysis, the factors f_t are not deterministic when we condition on \mathcal{F}_{t-1} . Instead, they are still subject to their own source of error. The factors are therefore inherently unobserved. Computation of the log-likelihood

function requires integrating over the path space of f_t taking account of the dynamic properties of f_t . Also, filtered estimates of f_t in that case depend on current and past values of y_t , while smoothed estimates of f_t depend on current, past, and future values of the data. Hence our current observation driven approach is relatively simple to implement, while it can still account for most of the flexibility of a standard parameter driven dynamic factor model, including its ability to measure mixed measurements under a common factor structure.

2.2 Measurement of the factors

The estimation of the factors at time t given past observations $\mathcal{F}_{t-1} = \{y_1, \dots, y_{t-1}\}$, for a given value of ψ , is carried out as a filtering process. At time t , we assume that \mathcal{F}_{t-1} and the paths f_1, \dots, f_t and s_1, \dots, s_{t-1} are given. When observation y_t becomes available, we compute s_t as defined in (4) with scaling matrix (5). Subsequently, we compute f_{t+1} using the recursive equation (3). At time $t+j$, once observation y_{t+j} is available, we can compute s_{t+j} and f_{t+1+j} in the same way for $j = 1, 2, \dots$. In practice, the filtering process at $t = 1$ starts with f_1 being set to some fixed value. The initial value f_1 can also be treated as a part of ψ or as a random variable; see, for example, Diebold and Schuermann (2008).

Missing values in data sets are intrinsically handled simply through the specifications of ∇_t and \mathcal{I}_t in (6) and (7), respectively. The variables ∇_t and \mathcal{I}_t enable the computation of s_t . When all entries in y_t are missing, it follows that $s_t = 0$ such that f_{t+1} slowly reverts to its long-term average.

When the time series panel is unbalanced, missing values appear naturally at the beginning and/or at the end of the sample period. They also appear when time series are observed at different frequencies. The overall time index t refers to a time period associated with the highest available frequency in the panel. Time series observed at lower frequencies contain missing values at time points for which no new observations are available. For example, a panel with monthly and quarterly time series adopts a monthly time index. A quarterly time series is then arranged by having two monthly missing values after each (quarterly) observation. The precise arrangement depends on whether the variable represents a stock (measured at a point in time) or a flow (measured as a quantity over time, typically an average).

2.3 Maximum likelihood estimation

Observation driven time series models are attractive because the log-likelihood is known in closed form. For a given set of observations y_1, \dots, y_T , the vector of unknown parameters ψ can be estimated by maximizing the log-likelihood function with respect to ψ , that is

$$\hat{\psi} = \arg \max_{\psi} \sum_{t=1}^T \log p(y_t | f_t, \mathcal{F}_{t-1}; \psi), \quad (8)$$

where $p(y_t | f_t, \mathcal{F}_{t-1}; \psi)$ is defined in (2). The evaluation of $\log p(y_t | f_t, \mathcal{F}_{t-1}; \psi)$ is easily incorporated in the filtering process for f_t as described in Section 2.2.

The maximization in (8) can be carried out using a conveniently chosen quasi-Newton optimization method that is based on score information. The score here is defined as the first derivative of the log-likelihood function in (8) with respect to the constant parameter vector ψ . Analytical expressions for the score function can be developed, but typically lead to a collection of complicated equations. In practice, the maximization of the log-likelihood function is therefore carried out using numerical derivatives.

Identification of the individual parameters in ψ needs to be considered carefully in factor models. A rotation of the factors by some nonsingular matrix may yield an observationally equivalent model. To make sure that all coefficients in ψ are identified, we impose the restriction $\omega = 0$ in (3). We also restrict the set of factor loadings. In particular, we restrict a set of M rows in the factor loading matrix to form a lower triangular matrix with ones on the diagonal. We assume that the matrices A_i and B_j of (3) for $i = 1, \dots, p$ and $j = 1, \dots, q$ are diagonal.

2.4 Forecasting

The forecasting of future observations and factors is straightforward. The forecast f_{T+h} , with $h = 1, 2, \dots, H$, can be obtained by iterating the factor recursion (3) in which the sequence s_{T+1}, \dots, s_{T+H} is treated as a martingale difference. To obtain forecasting expectations of nonlinear functions of the factors, the conditional mean of the predictive distribution needs to be computed by simulation due to Jensen's inequality. Simulating the factors is straightforward given the recursion (3). Simulation is also the appropriate tool if other characteristics of the forecasting distribution are of interest, such as percentiles and quantiles.

Forecasting in our modeling framework has some advantages when compared to the two-step forecasting approach in the approximate dynamic factor modeling framework of Stock and Watson (2002). Forecasting the future observations and factors in our framework does not require the formulation of an auxiliary model. Parameter estimation, signal extraction, and forecasting occurs in a single unified step. In the two-step approach, first, the factors are extracted from a large panel of predictor variables, and, second, the forecasts for the variables of interest are computed via regression with the lagged estimated factors as covariates. Our simultaneous modeling approach retains valid inference results that may be lost in a two-step approach, and it ensures that the extracted factors are related to the variables of interest throughout the estimation and forecasting process.

3 An application to macroeconomic and credit risk

The interest in credit risk analysis has increased considerably since the 2007-2008 financial crises in both the professional and academic finance literature. Credit risk is often discussed in terms of the probability of default (PD) and the loss given default (LGD): PD is the probability that a firm or company goes into default over a specific time period, and LGD is the fraction of the capital that is lost in case the firm enters default. It is argued that both PD and LGD are driven by the same underlying risk factors; see the discussions in Altman, Brady, Resti, and Sironi (2003), Allen and Saunders (2004), and Schuermann (2006). The implication is that LGD is expected to be high when PD is expected to be high as well. As a result, the total credit risk profile of a portfolio increases.

In our empirical study, we apply the general modeling framework of Section 2 to investigate the linkages between macroeconomic and credit risk. We analyze firm-level data on defaults and on changes in credit quality to obtain insight into the dynamic relations between PD, LGD and macroeconomic fluctuations. The model for credit quality is based on a dynamic ordered logit distribution and the model for LGD is based on a dynamic beta distribution. Both of these are new to the literature; see Gupton and Stein (2005) and CreditMetrics (2007) for static versions of our model. The macroeconomic variables are specified as linear Gaussian processes.

3.1 Data

Our available time series panel consists of three groups of variables: macroeconomic, default and ratings, and loss given default (LGD). The macroeconomic group has six time series (five monthly and one quarterly) that we have obtained from the FRED database at the Federal Reserve Bank of St. Louis. General macroeconomic conditions are reflected by three variables: (i) the annual change in log industrial production, $IP_t - IP_{t-12}$, where IP_t is industrial production at the end of month t ; (ii) the annual change in the unemployment rate, $UR_t - UR_{t-12}$, where UR_t is the unemployment rate at the end of month t ; (iii) the annual change in log real gross domestic product (GDP), $RGDP_t - RGDP_{t-12}$ where $RGDP_t$ is real GDP at the start of the first month t of a quarter, and missing otherwise. These three variables are strongly related to the state of the business cycle and intend to capture the extent of economic activity. The change in real GDP thus corresponds to credit events in the same period.

General financial market and credit risk variables are included to account for the market's perception of the probability of default (PD). We include the credit spread, annual change in stock market log-prices (returns), and stock market volatility all at a monthly frequency. The credit spread is measured as the spread $r_t^{Baa} - r_t^{Gov10}$ between the yield r_t^{Baa} on Baa rated bonds and the yield r_t^{Gov10} on 10-year treasury bonds at the end of month t , where the ratings are assigned by Moody's. Credit spread movements capture two components of credit risk: changes in the market's perception of PDs and LGDs; and changes in the price that the market charges for this type of risk. Particularly the first of these two components can be relevant for determining default rate dynamics.

The stock market variables are the monthly observed annual returns $r_t = \log(SP_t / SP_{t-12})$ on the S&P 500 index, where SP_t is the S&P500 index at the end of month t . The volatility is measured by the annualized daily realized volatility computed over the current month, i.e.,

$$(\hat{\sigma}_t^{rv})^2 = \frac{252}{n_t} \sum_{i=1}^{n_t} (R_{t,i} - \bar{R}_t)^2, \quad \bar{R}_t = \frac{1}{n_t} \sum_{i=1}^{n_t} R_{t,i},$$

where n_t is the number of working days in month t , $R_{t,i}$ is the S&P500 return over day i of month t and the value 252 proxies for the number of trading days in a year. Both the stock market return and its volatility can be linked to default risk through the structural model of

Merton (1974) in which firms with higher asset values or lower asset volatilities are less likely to default. In the aggregate, the dynamics of the two can be approximated by equity returns and equity volatilities given that the average debt-equity proportions of the S&P index constituents are relatively stable over time. The sample period January, 1981 to March, 2010 contains 350 monthly observations. All six macroeconomic variables are standardized by subtracting their subsequent sample means and dividing by their sample standard deviations.

As our current empirical study focuses on the joint modeling of a diversity of macroeconomic and credit related variables of very different type in a unified dynamic modeling framework, we restrict ourselves here to the limited but representative set of six macroeconomic variables above. More macro variables can be included in the analysis at the expense of an increase in computation time. Here, however, we seek to enlarge the cross-sectional dimension of our panel data set by mixing the macro variables with credit related time series. This yields a cross-sectional dimension for our panel time series that goes up to 48 dimensions for specific periods.

The default and rating transition variables contain credit ratings assigned by Moody's, which reflect the credit quality of the firm. We re-group the ratings of Moody's into four rating groups: Investment Grade (IG) that contains Moody's rating grades Aaa down to Baa3; double B (BB) that contains Ba1–Ba3; single B (B) that contains B1–B3; triple C (CCC) that contains Caa1–C3. A company that defaults is marked as a transition to the absorbing category D. The vector of possible ratings is $R = (D, CCC, B, BB, IG)'$. To account for all possible transitions, including staying in the current rating group, we have five transition types that a firm can make. Hence we keep track of twenty (four times five) different time series for rating transitions and defaults. In April 1982 and October 1999 Moody's redefined some of their rating categories. This caused a large number of artificial rating transitions for some categories. We handle these events in our model by including two dummy variables for these two months.

The loss given default (LGD) measures the fraction of the total exposure that is lost conditional on a firm defaulting. Our sample contains 1342 defaults from which we obtain 1125 measurements of LGD. The LGD is measured from financial market data using what is known as the market implied LGD. Market implied LGDs are constructed by recording the price of a traded bond just before the default announcement and the market price of the same bond 30 days after the default announcement. The percentage drop in price then defines the loss

fraction or LGD; see McNeil, Frey, and Embrechts (2005) for further details on the different ways to measure LGDs. Missing LGDs in the database in case of default are due to the underlying bonds not being traded in the market or to the unavailability of price information on the bonds in the underlying data sources. In month t , a number of firms $K_t \geq 0$ may default. The dimension K_t of the vector of LGD measurements at time t therefore varies over time from 0 to 22 (maximum number of defaults in one day is 22 in our data set).

3.2 The components of the joint model

The panel time series y_t can be partitioned into three sub-vectors $y_t = (y_t^m, y_t^c, y_t^r)'$ where y_t^m contains the six macroeconomic variables, y_t^c contains the observed proportions of twenty possible credit rating transitions, and y_t^r contains the LGD variables. The business cycle features in the macroeconomic variables are assumed to influence credit ratings and loss given default rates. On top of this, credit ratings and LGDs also share common features in their remaining dynamics. We consider the following observation densities at time t

$$y_t^m \sim N(\mu_t, \Sigma_m | f_t, \mathcal{F}_{t-1}),$$

$$y_{it}^c \sim \text{Ordered Logit } (\pi_{ijt} | f_t, \mathcal{F}_{t-1}),$$

$$y_{kt}^r \sim \text{Beta } (a_{kt}, b_{kt} | f_t, \mathcal{F}_{t-1}),$$

where the mean vector $\mu_t = \mu(f_t)$ is a function of the $M \times 1$ vector of latent factors f_t , the 6×6 variance matrix Σ_m is fixed over time, y_{it}^c is the i th element of y_t^c , the probability $\pi_{ijt} = \pi_{ij}(f_t)$ for the ordered logit density relates to the transition of firm i with rating $R_{it} \in \{\text{CCC, B, BB, IG}\}$ to rating $j \in \{\text{D, CCC, B, BB, IG}\}$ during period t , y_{kt}^r is the k th element of y_t^r , $k = 1, \dots, K_t$, K_t is the number of defaults in month t , and $a_{kt} = a_k(f_t)$ and $b_{kt} = b_k(f_t)$ are the positive shape coefficients for the beta density. The details for the observation densities and the dynamic specification of f_t are discussed below.

3.3 A macro model for y_t^m

Define \tilde{S}_t as the identity matrix I_6 with the rows removed which are associated with entries missing in y_t^m . The log-likelihood contribution of y_t^m at time t is then given by

$$\text{const} - 0.5 \log |\tilde{S}_t \Sigma_m \tilde{S}'_t| - 0.5 \left(\tilde{S}_t (y_t^m - \mu_t) \right)' \left(\tilde{S}_t \Sigma_m \tilde{S}'_t \right)^{-1} \left(\tilde{S}_t (y_t^m - \mu_t) \right), \quad (9)$$

with

$$\mu_t = \mu(f_t) = z^m + Z^m f_t, \quad (10)$$

where z^m is the 6×1 vector of intercepts and Z^m is the $6 \times M$ matrix of factor loadings. As the macroeconomic variables have been standardized, we set $z^m = 0$. The conditional score and information matrix for the Gaussian component are given by

$$\nabla_t^m = \left(\tilde{S}_t Z^m \right)' \left(\tilde{S}_t \Sigma_m \tilde{S}'_t \right)^{-1} \tilde{S}_t (y_t^m - \mu_t), \quad (11)$$

$$\mathcal{I}_t^m = \left(\tilde{S}_t Z^m \right)' \left(\tilde{S}_t \Sigma_m \tilde{S}'_t \right)^{-1} \tilde{S}_t Z^m, \quad (12)$$

from which we can compute s_t in (3) via (6) and (7). It follows that the dynamic updating of f_t is a linear function of the prediction error vector $y_t^m - \mu_t$ and is effectively a generalized least squares computation.

In the online appendix, we provide the score and information matrix for a model with a time-varying mean and an observation density based on the multivariate Student's t distribution. In our empirical study we have found that the results for the Student's t and the Gaussian models do not differ much since the estimated degrees of freedom is relatively high. Nevertheless, the extension to the Student's t model may be useful for applications with higher frequency data where outliers and heavy-tails are of a more prominent concern.

3.4 A rating transition model for y_{it}^c

For the credit rating transitions, we specify a dynamic ordered logit model. Previous research on credit risk has focused on a standard multinomial specification; see the contributions by, for example, Koopman, Lucas, and Monteiro (2008) and Koopman, Kraeussl, Lucas, and Monteiro (2009). The multinomial density does not take into account the fact that ratings are ordered.

By using the ordered logit specification, the ordering of the ratings is taken into account. The model then becomes a dynamic alternative to the static ordered probit model of CreditMetrics, which is one of the industry standards, see Gupton and Stein (2005). However, it is far from evident how to construct an observation driven dynamic ordered logit model as opposed to a dynamic multinomial model. In particular, it is not clear what functions of the data should be chosen to drive the changes in the probabilities for the transitions. Our observation driven modeling framework solves these issues by relying conveniently on the score of the conditional log-likelihood, which in this case becomes the score of the ordered logistic log-likelihood.

We specify the binary probability that the rating of firm i does not exceed rating j at the end of period t by

$$\tilde{\pi}_{ijt} = P[R_{i,t+1} \leq j | \mathcal{F}_{t-1}] = \frac{\exp(\theta_{ijt})}{1 + \exp(\theta_{ijt})}, \quad (13)$$

where R_{it} is the rating of firm i at the start of month t , with $j \in \{\text{D, CCC, B, BB, IG}\}$ and $\tilde{\pi}_{i,\text{IG},t} = 1$. From (13) it follows that the probability of a transition of firm i from rating R_{it} to $R_{i,t+1} = j$ is given by

$$\pi_{ijt} = P[R_{i,t+1} = j | \mathcal{F}_{t-1}] = \tilde{\pi}_{ijt} - \tilde{\pi}_{i,j-1,t}, \quad (14)$$

with $\tilde{\pi}_{i,j-1,t} = 0$ for $j = \text{D}$. The log-likelihood contribution at time t becomes

$$\sum_i \sum_j y_{ijt}^c \log(\pi_{ijt}), \quad (15)$$

where the first summation is over all firms, the second summation is over all five ratings, and the indicator y_{ijt}^c equals unity if firm i has moved from rating R_{it} to rating j during month t , and zero otherwise.

We specify the logit probability θ_{ijt} as a linear function of the time-varying factor f_t ,

$$\theta_{ijt} = \theta_{ijt}(f_t) = z_{ijt}^c - Z_{it}^c f_t, \quad (16)$$

where z_{ijt}^c is a scalar intercept and Z_{it}^c is an $1 \times M$ vector of factor loadings, both of which can vary over time due to dependence on firm specific information such as the firm's initial rating, its industry sector, and time-varying financial ratios. The z_{ijt}^c determine the baseline transition probabilities in the ordered logit specification and need to be estimated.

The conditional score function for the specified ordered logit model is given by

$$\nabla_t^c = - \sum_i \sum_j \left[\dot{\pi}_{ijt} \frac{y_{ijt}^c}{\pi_{ijt}} \right] Z_{it}^{c'}, \quad (17)$$

where

$$\dot{\pi}_{ijt} = \tilde{\pi}_{ijt} (1 - \tilde{\pi}_{ijt}) - \tilde{\pi}_{i,j-1,t} (1 - \tilde{\pi}_{i,j-1,t}). \quad (18)$$

The term y_{ijt}^c / π_{ijt} is intuitive since it compares the actual outcome of y_{ijt}^c with its probability π_{ijt} . In expectation, the ratio y_{ijt}^c / π_{ijt} is one and ∇_t^c is zero since $\sum_j \dot{\pi}_{ijt} = 0$. It is straightforward to show that the corresponding information matrix $E[\nabla_t^c \nabla_t^{c'}]$ is given by

$$\mathcal{I}_t^c = \sum_i n_{it} \left[\sum_j \frac{\dot{\pi}_{ij,t}^2}{\pi_{ij,t}} \right] Z_{it}^{c'} Z_{it}^c, \quad (19)$$

where n_{it} is one if firm i exists at the start of period t , and zero otherwise.

We have included time dummies in the specification (16) for the months of April 1982 and October 1999 in order to handle outliers which are due to the redefinitions of rating categories. As mentioned earlier, these redefinitions have caused substantial incidental rerating activity during these months.

3.5 A loss given default model for y_{kt}^r

The $K_t \times 1$ vector y_t^r of loss given defaults (LGD) has a dimension which in our data set varies over time from $K_t = 0$ to $K_t = 22$, depending on how many firms default and on whether their LGD is recorded. The LGD rates are reported in percentage terms and it is therefore appropriate to model these with a beta distribution. The log-likelihood contribution at time t is then given by

$$\sum_{k=1}^{K_t} (a_{kt} - 1) \log(y_{kt}^r) + (b_{kt} - 1) \log(1 - y_{kt}^r) - \log[B(a_{kt}, b_{kt})], \quad (20)$$

where a_{kt} and b_{kt} are positive scalar coefficients and $B(a_{kt}, b_{kt}) = \Gamma(a_{kt}) \Gamma(b_{kt}) / \Gamma(a_{kt} + b_{kt})$ is the Beta function with $\Gamma(\cdot)$ denoting the Gamma function. The specifications for a_{kt} and b_{kt} are implied by the mean and variance of the beta distribution.

We define the mean of the k th contribution of the beta distribution as $\mu_{kt}^r = \mu_k^r(f_t)$, which is a function of the factor f_t . To let the mean range within the $[0, 1]$ interval, we set

$$\log(\mu_{kt}^r / (1 - \mu_{kt}^r)) = z^r + Z^r f_t, \quad (21)$$

where z^r is a scalar intercept and Z^r is the $1 \times M$ vector of factor loadings. The intercept and loading coefficients are common to all defaults.

Let $\beta_r > 0$ be an unknown scalar and define the variance of the beta distribution as

$$(\sigma_{kt}^r)^2 = \mu_{kt}^r \cdot (1 - \mu_{kt}^r) / (1 + \beta_r). \quad (22)$$

This specification insists that as long as the conditional mean remains within the boundaries of the unit interval, the variance remains positive. From the specifications of the mean μ_{kt} in (21) and the variance $(\sigma_{kt}^r)^2$ in (22), the shape parameters a_{kt} and b_{kt} follow directly,

$$a_{kt} = \beta_r \cdot \mu_{kt}^r, \quad b_{kt} = \beta_r \cdot (1 - \mu_{kt}^r).$$

The conditional score and information matrix in our case are now given by

$$\nabla_t^r = \beta_r \mu_t^r (1 - \mu_t^r) (Z^r)' (1, -1) \sum_{k=1}^{K_t} \left((\log(y_{kt}^r), \log(1 - y_{kt}^r))' - \dot{B}(a_{kt}, b_{kt}) \right), \quad (23)$$

$$\mathcal{I}_t^r = (\beta_r \mu_t^r (1 - \mu_t^r))^2 (Z^r)' (1, -1) \left(\sum_{k=1}^{K_t} \ddot{B}(a_{kt}, b_{kt}) \right) (1, -1)' Z^r, \quad (24)$$

where $\dot{B}(a_{kt}, b_{kt})$ and $\ddot{B}(a_{kt}, b_{kt})$ are the first and second order derivatives of the log Beta function with respect to $(a_{kt}, b_{kt})'$, respectively.

3.6 Further details of the joint model

We have introduced observation driven dynamic logit and dynamic beta model specifications which are, as far as we know, new in the literature. Furthermore, the three different dynamic model specifications (for the normal, the ordered logit and the beta densities) are integrated naturally into a joint observation driven dynamic factor model. The contribution for the log-

likelihood value of all observations at time t is simply obtained by adding the components in (9), (15), and (20). Similarly, the score and information matrix contributions at time t are given by $\nabla_t^m + \nabla_t^c + \nabla_t^r$ and $\mathcal{I}_t^m + \mathcal{I}_t^c + \mathcal{I}_t^r$, respectively.

The dynamic specification of the factors is given by (3), where we set $p = q = 1$. After a preliminary data analysis, we focus for illustrative purposes on models with three or four macro factors, one or two frailty factors, and possibly a separate LGD factor. All these model specifications capture the salient features of our heterogeneous data set.

We impose a recursive block-structure on the factor loading matrix for identification and interpretation purposes. In particular, the transformed signal can be represented as

$$z_t + Z_t f_t = \begin{pmatrix} z^m \\ z_t^c \\ z^r \end{pmatrix} + \begin{bmatrix} Z^m \\ Z_t^c \\ Z^r \end{bmatrix} f_t = \begin{pmatrix} z^m \\ z_t^c \\ z^r \end{pmatrix} + \begin{bmatrix} Z^{mm} & 0 & 0 \\ Z_t^{cm} & Z_t^{cc} & 0 \\ Z^{rm} & Z^{rc} & Z^{rr} \end{bmatrix} \begin{pmatrix} f_t^m \\ f_t^c \\ f_t^r \end{pmatrix}, \quad (25)$$

where the partitioned intercept vectors and block loading matrices have appropriate dimensions, and z_t^c and Z_t^c collect the entries z_{ijt}^c and Z_{it}^c in (16), respectively. For identification purposes, we require that an appropriate selection of rows in the block matrix Z^{mm} is equal to a lower triangular matrix with unit diagonal and dimension $m \times m$, where m is the number of macro factors. The same restrictions apply to loading matrices Z^{cc} and Z^{rr} , but then with dimensions $c \times c$ and $r \times r$, respectively. The three loading matrices Z^{cm} , Z^{rm} and Z^{rc} have no restrictions.

Using specification (25), we allow the macro factors in f_t^m to influence the macro series, the transition probabilities, and the LGD mean and variance. The frailty (or transition) factors f_t^c influence the transition probabilities and the LGD mean and variance. Frailty factors thus capture default and rating transition clustering above and beyond what is implied by shared exposure to common macroeconomic risk factors. Such excess default clustering is also considered in studies by Das, Duffie, Kapadia, and Saita (2007), Duffie, Eckner, Horel, and Saita (2009), Azizpour, Giesecke, and Schwenkler (2010), and Koopman, Lucas, and Schwaab (2011). The LGD factor f_t^r influences the LGD dynamics only and can be used to investigate whether LGD dynamics coincide with macro and/or frailty dynamics. Each model is referred to by its number of composite factors, that is (m, c, r) . For example, a model labeled as $(3, 2, 1)$ denotes a model with three macro factors, two credit risk (frailty) factors, and one LGD factor.

Although some factors do not contemporaneously load into all of the observation densities, information in series with zero loading coefficients can still affect the factors in future periods through the score function that drives the factor recursion (3). For example, the credit (frailty) factors and the LGD factor do not enter the equation for the conditional mean of the macros μ_t in (10), but the macro factors do load into the transformed mean and variance of the LGDs. Therefore, information in the credit rating transitions and LGDs helps to determine the value of the macro factors at time $t + 1$ because they are part of the score vector.

4 Empirical results

The empirical results are presented in four parts. First, we discuss the in-sample estimations results of the model. Second, we present the estimated time-varying factors and a set of diagnostic graphics. Third, we use the model to forecast economic and credit risk scenarios out-of-sample. Fourth, we present the impulse response functions for our nonlinear non-Gaussian model.

4.1 In-sample estimation results

Table 1 contains a list of estimated models with their maximized log-likelihood values and the corresponding Akaike information criterion (AIC) and Schwarz Bayesian information criterion (BIC) values. We have considered for illustrative purposes the models with 3 and 4 macro factors, 0, 1 and 2 frailty factors, and 0 and 1 LGD factors. We observe clear improvements in the log-likelihood values when more factors are added. The likelihood value increases particularly when we include a fourth macro factor or a first frailty factor in the model. The addition of a second frailty factor also provides an increase in the likelihood value, but this increase is more modest. The inclusion of a separate LGD factor appears to have a negligible effect on the likelihood value. It seems sufficient that the LGD variables only depend on the macro and frailty factors.

The model selection criteria point to the (4,2,0) model as the preferred model within the considered range of models. Hence we take the (4,2,0) model as our model of choice in our empirical study. Table 2 presents the estimated parameters and their corresponding standard

errors. To reduce the number of estimated parameters, we have imposed zero restrictions on parameters that have initially been estimated to be insignificant. Standard errors are computed using the inverse Hessian matrix of the maximized log likelihood. Estimation results for the full model, or for models with a different number of macro factors or with a multivariate Student's t observation density for the macros are all reported in our online Appendix that accompanies this paper.

The coefficients in the matrices A and B determine the properties of the dynamic factor process f_t . Their estimates are presented in Table 2 and are clearly significant. The factors appear highly persistent since all the B coefficients are estimated as 0.9 or higher. This implies that rating transition probabilities, including default probabilities, may deviate from their unconditional values as well as from their macro fundamentals for a substantial number of months. From a risk management perspective, it means that capital levels must be set in accordance with an episode (rather than an incidence) of high default rates for any portfolio of credit exposures.

The estimated loading coefficients in Z^m reveal that the first macro factor loads on industrial production growth, real GDP growth, negative changes in the unemployment rate and negative credit spreads. The first macro factor can therefore be interpreted as a mix of business cycle and credit market indicators. Figure 1 reveals that the estimates of the first factor are low during recession periods. In the 1980s the first factor estimates appear to respond to the savings and loans crisis rather than to the state of the business cycle only. The estimates for Z^c corresponding to the first macro factor are only significant for the rating transition probabilities of lower grade companies. It implies that higher default probabilities can be expected for CCC and B rated companies during recession periods.

The second macro factor loads significantly on unemployment rate changes, negative GDP growth and on negative returns on the S&P500; and less significantly on stock market volatility. Also the second factor is related to the business cycle but also to financial market conditions given its reliance on stock markets rather than on credit spreads. The time series of factor estimates show clear peaks at all major recession and crisis periods. However, the recession in the early 1990s is more dominant. The estimated Z^c coefficients for the second macro factor are significant and positive for the sub-investment grade rating groups B and BB. It confirms that the second factor represents financial market conditions to some extent.

The third macro factor estimates depend mainly on credit spreads, negative annual stock returns, and equity market volatility. We may interpret this factor as the perception of economic and risk conditions by financial markets. The corresponding elements of Z^c for the third macro factor are significant for all transition probabilities, but particularly for those of higher grade firms. When markets perceive credit risk to be high, more defaults and downgrades are expected. The element of Z^r for the third macro factor is significant and positive. It implies that a credit risk environment with negative sentiments will lead to higher LGD rates. The fourth macro factor has significant loading coefficients for volatility, real GDP growth and realized returns. This may indicate that it is a proxy for the business cycle only. However, its association with realized volatility implies that uncertainty also drives the factor upwards. Also the fourth factor estimates as displayed in Figure 1 appear to represent more than only the business cycle. The corresponding Z^c coefficient estimates are all negative for higher grade firms which is consistent with the estimated negative loading for LGD rates in Z^r : a higher value for this factor implies lower default and downgrade probabilities and lower LGD rates.

The frailty factors capture the changes in default and downgrade probabilities, and changes in LGD rates that cannot be explained by the macro and finance variables through the first four factors. The first frailty factor is most important as the estimated coefficients for three rating categories in Z^c are substantial and significant. When this frailty factor is high, default and downgrade probabilities increase. The estimated first frailty factor as presented in Figure 1 captures excess default clustering in the early 1990s and in the early 2000s. It also captures the low number of defaults and downgrades in the run-up to the financial crisis, much lower than what we can expect from the macro factors only. This frailty factor is also highly significant for the LGD equation: excessive risk related to credit migrations and LGD clearly move together.

The second frailty factor mainly loads positively on CCC firms and negatively on IG firms. The factor estimates capture two historical features of the corporate bond market. First, in the mid 1980s and in recent years, the number of defaults for investment grade firms is higher than expected, which may explain the significantly negative coefficient for the investment grade loading in Z^c . Second, given the positive loading coefficient for the CCC graded firms, this factor also appears to be affected by the benign default climate before the financial crisis. Since the second factor appears to reflect these historical default periods accurately, its coefficient for the LGD equation is also highly significant. We may therefore conclude that rating migrations,

defaults and LGDs are affected by more than macro factors only.

Other estimated coefficients reported in Table 2 are those for the intercepts and the variances of the disturbances in the macro equations. The estimated intercepts z_{ij}^c for the ordered logit specification reveal that ratings are highly persistent on a monthly basis. For example, by considering only the cut-off points of the ordered logit specification, we find that the probability of remaining in investment grade is about $(1 + \exp(-6.299))^{-1} \approx 99.82\%$, while the probability of a CCC company defaulting over the next month is $(1 + \exp(3.751))^{-1} \approx 2.30\%$.

4.2 Signal extraction and diagnostic checking

The estimated transition and default probabilities presented in Figure 2 are driven by both the macro and frailty factors. Given that we consider transitions at the monthly frequency, all probabilities are close to zero, except the probabilities that ratings remain unchanged. The peaks in the estimated probabilities differ across rating grades. For example, the investment grade class has its highest default probability peaks in the financial and the dotcom crises, and some substantially lower peaks in the mid 1980s and the 1991 recession. By contrast, the CCC class has its highest peak in 1991. However, even though the magnitude of the peaks and troughs differs across rating grades, they are all subject to clusters of high and low probabilities.

To verify whether the model specification is appropriate for our observed data set, we present a selection of diagnostic graphs. In Figure 3 we present the sample correlograms for the six macroeconomic time series and for their one-step ahead prediction errors for the reported (4,2,0) model. We may conclude that most of the dynamic features in the macro series are captured by our model specification. The negatively correlated prediction error at lag 12 points to a seasonal monthly feature that is not captured by the model. Given the disparity of the dynamic features in these six time series, we conclude that the model is appropriate for the economic variables. Further improvements can be made when higher order lags are considered in the dynamic specification for the factors.

In Figure 4 we present diagnostic graphs for the rating transitions. The top graphs with the actual number of downgrades/upgrades on the vertical axis against the one-step ahead predicted number of downgrades/upgrades implied by the model indicate that the model captures movements in credit rating transitions well. The clouds of points cluster around the 45 degree

line. Since the monthly numbers of downgrades tend to be higher, the precision for downgrades also appears higher compared to upgrades. The corresponding standardized prediction errors provide a more balanced picture. Our (4,2,0) model succeeds in the accurate prediction of downgrades and upgrades.

The upper right panel of Figure 5 reveals that LGDs are noisy and irregularly spaced. The extracted factors from our model appear to capture these salient features accurately. LGDs correlate positively with default rate dynamics; they tend to be high and low in the same periods. The effect of time variation in the parameters of the beta distribution is visualized in the left panels of Figure 5. The upper left panel presents the cross-sectional beta distributions for LGDs in the quiet month of June 2006. For the (4,2,0) model, the probability mass is strongly skewed to the left which indicate that LGDs are typically low. The distribution implied by the (4,2,0) model in January 2009 presents the opposite picture: the probability mass is skewed to the right. According to the model, one expects that LGD rates are much higher on average. In comparison with the (4,0,0) model that has four macro factors but no frailty factors, the implied beta distribution for LGDs in June 2006 is substantially different, while the differences are small in January 2009. The differences in the fit of the LGDs by the two models is made apparent in the upper right panel where the signal for the (4,0,0) model appears to miss the periods with exceptionally high and low LGDs. The implication of including frailty factors for credit risk is also evident when we compare the estimated transition probabilities for the two models. For example, in the lower right panel in Figure 5 we present the estimated transition probability from BB to default. The estimated probabilities are substantially different between the two models. The (4,0,0) model misses some peaks and troughs which are captured by the (4,2,0) model.

4.3 Forecasting and credit risk management

The observation driven mixed measurement dynamic factor model can be used for the joint forecasting of macro variables, rating transition and default probabilities, and LGDs. Such forecasts can be used to simulate credit portfolio loss distributions. Lando (2004) argues that the forecasting distributions for portfolio losses are an important ingredient for assessing and managing the risk of credit exposures.

We assume a scenario where at the end of the sample, time T , a financial institution holds a large portfolio of bonds. The risk management department would like to know the distribution of credit losses at a specified future time $T + H$ to set up capital requirements or hedge positions. To simplify the simulation experiment, we concentrate on default losses only. We do not include credit losses due to downgrades in a mark-to-market setting. We also assume a zero discounting rate to cumulate losses that materialize at different points in time. We set the initial portfolio of bonds equal to the current exposures in the last period of our sample. This leads to a portfolio of 1144 firms rated IG, 265 firms rated BB, 615 firms rated B, and 311 firms rated CCC.

Given the portfolio, the model and its parameter estimates, we can generate the dynamic evolution of the macro variables and the rating composition of the portfolio in a simulation study. If one of the firms transits into default, the LGD component with its beta distribution is used to generate the corresponding LGD realization. By simulating the portfolio and macros forward in this way, we obtain a realization of default losses between T and the horizon date $T + H$. This process can be repeated many times. In our simulation study, we use 500,000 simulations for the loss distribution at different horizons.

The forecasts of loss distributions are conditional on the estimated factors at the end of the sample, f_T , rather than on the variables y_T . For policy and risk analyses such as stress tests, it can be useful to condition directly on the variables of interest rather than on the estimated factors. We have two main reasons for conditioning on shocks in the factors rather than on shocks in the variables themselves. First, we model several variables at the same time. An assumption for one variable should therefore be consistent with the dynamic properties of other variables. For example, assuming a large drop in industrial production growth in a stress scenario without simultaneously assuming related effects on other macro variables in the model, such as the unemployment rate, appears contrived. The construction of consistent scenarios for a potentially larger set of macro variables is far from trivial. By concentrating on shocks in factors rather than shocks in variables, we resolve the issue to some extent as all variables in our framework are subject to the same underlying factor structure. Second, in the forecasting analyses we require multi-step scenarios for all macro, rating and LGD variables. If one opts for administering a shock to a specific variable at time T , one therefore also needs to account for the dynamics of such a shock, including its spillover into other variables in the system over time. This is effectively accomplished in our integrated modeling framework.

The model specification and the starting value f_T are important ingredients of the analysis. We expect differences between models with and without frailty factors. We further anticipate differences when f_T has values associated with expansion or with recession periods. Given the high persistence of the factors, these different starting conditions will have different impacts on the loss experience in future periods. Figure 6 presents the initial results of our forecasting study. It presents simulated loss distributions at four different horizons for the (4,2,0) model. At the one month horizon, we clearly visualize the differences in the forecasting distributions. Starting from a recession, the losses in the next month are expected to be higher with more variations. If current economic conditions are good, the losses are expected to be low in the next month with less variations. Even at the 1-month horizon, there is a significant non-overlap of the 99% confidence regions of the two densities.

When the forecast horizon H increases, the two densities gradually start to overlap. This effect is due to the stationary process of the factors in f_t . The high persistence in f_t causes the expected LGD densities to be different for a large number of months so that the cumulated losses remain substantially different. In our simulation study, a current recession turns into an expansion phase, and vice versa, in a period of three years.

The same forecasting study is carried out for the (4,0,0) model and the results are compared with those for the (4,2,0) model in Figure 7. Hence we compare the economic contribution of the frailty factors to capital requirements. When the initial factor f_T is set to zero, the forecasting distributions of models with and without the two frailty factors are similar for the one-month horizon. The differences become clearer for longer forecast horizons. The models with only macro factors produce more concentrated loss distributions. The locations of the two distributions are roughly the same but the one for the (4,2,0) model is much more dispersed. Also, the 99% confidence loss quantiles for the (4,2,0) model are substantially larger than for the (4,0,0) model. We therefore may conclude that the model with only macro factors leads to a severe underestimation of the required capital, particularly at long horizons.

When the initial factors f_T are started with recession values, the differences in the results are even more pronounced, even at shorter horizons. In such a scenario, not only the four macro factors start in a recession, but also the two credit related frailty factors. The anticipated increases in expected loss rates are therefore substantial. The upper quantiles of the loss distribution increase when the frailty factors are added. The increase for the highest quantiles

is close to 100% in most settings, which implies that capital requirements at these horizons need to be doubled. At the longest horizon of 36 months, the distributions start to overlap, similar to the setting with $f_T = 0$, but the differences in dispersion remain. We conclude that models that account for only macro dynamics appear to miss substantial parts of credit loss dynamics and potential credit loss sizes. The absolute numbers for the losses (as a percentage of the notional losses) can be viewed as small. However, typical portfolios hold many investment grade bonds that lead to a high quality portfolio with small losses on average. When we focus on the 99% quantile of the loss distribution at long horizons, including frailty factors in the model can shift the loss quantiles by more than 80%.

4.4 Impulse response analysis

The observation driven mixed measurement model can also be used for an impulse response analysis in which a shock in an unobserved factor can be traced through all variables over time. In particular we are interested in how a shock affects the credit loss distribution over time. The dynamic factor model is non-linear and non-Gaussian and therefore we follow the non-linear impulse response methodology of Koop, Pesaran, and Potter (1996).

The impulse response functions for the macros are presented in Figure 8 and they confirm the estimation results as presented in Table 2. The impulse response function for the first macro factor has the largest impact on the business cycle related variables. A negative shock decreases industrial production growth and real GDP growth and increases the unemployment rate. This impact is persistent as it takes 3 to 4 years before the effect dies out. Similar findings are obtained for the other three macro factors. These effects correspond to the loading estimates for Z^m as reported in Table 2. However, we observe differences in persistence. The impulse response effects for the last two macro factors die out after faster, after approximately 2 years. Given the imposed factor loading structure in our model, the impulse response functions for the frailty factors confirm that they have no direct impact on the macro variables.

The results for the mean and the 90th percentile of the credit loss distributions are presented in Figure 9. Both the means and percentiles reveal similar patterns. The first and third macro factor have the largest and most persistent impact on portfolio credit losses. These are also the factors that directly load on the CCC migrations into default, which is the main determinant of

the large portfolio losses. The effect of a shock can last up to three or four years. The impact of the second macro factor is equally persistent but it is smaller in magnitude. We can verify from Table 2 that this factor mainly affects B and BB downgrade (and default) probabilities but it does not affect the subsequent CCC to default probability. Hence it has a limited impact on portfolio losses. The fourth macro factor has a small impact and it vanishes quickly after one to two years.

The frailty factors have a bigger impact on credit losses. Both frailty factors have an effect similar in size to the business cycle and the financial markets related macro factors. It implies that by omitting the frailty factors from the model, one third to one half of the credit risk is missed. Furthermore, the dynamics of credit losses might be completely misspecified when the frailty factors are not included in the model.

The in-sample estimation results, the out-of-sample forecasting distributions, as well as the impulse response analysis all point in the same direction: the size and dynamics of portfolio credit losses cannot be captured by conditioning only on macroeconomic conditions. The losses appear to have their own additional (frailty) dynamics, which are highly persistent and significant in both statistical and economic terms. Our observation driven mixed measurements modeling framework provides the means for a complete analysis of credit risk. The likelihood function is known in closed form and can be computed in a straightforward way. Hence our framework provides an effective alternative to parameter driven models that require advanced simulation methods for estimation and analysis.

5 Conclusion

We have introduced a new framework for observation driven mixed-measurement dynamic factor models for time series observations from different families of distributions and mixed sampling frequencies. Missing values arise due to unbalanced time series panels and mixed frequencies; they can be accommodated in our framework in a straightforward way. In an empirical application of the mixed-measurement framework we model the systematic variation in US corporate default counts and recovery rates in the period 1982–2010. We estimate and forecast interconnected macro, default and recovery risk conditions. We further demonstrate how to obtain the predictive credit portfolio loss distribution.

Our modeling framework is a useful device for the analysis of a large number of data points in a complex data set. A clear advantage of our framework is that the likelihood remains analytically tractable in closed form and therefore standard likelihood procedures can be used for parameter estimation. The model also lends itself easily to integrated forecasting exercises for joint macro and credit risk developments. In particular, the model can be used in a straightforward way to obtain portfolio loss distributions at multiple horizons. Such distributions can be used as an input for risk analyses. In addition, we can use the simulation framework to conduct impulse response analysis for non-linear model specifications. The impulse response functions can be used directly to study the feed-through mechanism from macro developments to credit losses and thus provide interesting input for both practitioners, regulators, and policy makers.

In our empirical study we have illustrated our new mixed measurement, mixed frequency modeling framework to analyse macro and credit risk. We emphasize however that the framework can be applied more generally. For example, in the context of high-frequency financial data, our approach can be used to model inter-trade durations, discrete tick changes in prices, and general market conditions simultaneously for different assets that are subject to common risk and liquidity factors. Similarly, in the context of a macroeconomic time series panel, we can mix the usual continuous macro variables with indicator data such as the NBER business cycle classifications. Other examples may be found in health and retirement economics, business, marketing, and psychology.

References

- Allen, L. and A. Saunders (2004). Incorporating systematic influences into risk measurements: a survey of the literature. *Journal of Financial Services Research* 26(2), 161–191.
- Altman, E. I., B. Brady, A. Resti, and A. Sironi (2003). The link between default and recovery rates: Theory, empirical evidence, and implications. *Journal of Business* 78, 2203–2227.
- Azizpour, S., K. Giesecke, and G. Schwenkler (2010). Exploring the sources of default clustering. *Stanford University working paper series*.
- Bai, J. (2003). Inferential Theory for Factor Models of Large Dimension. *Econometrica* 71(1), 135–171.
- Bai, J. and S. Ng (2002). Determining the number of factors in approximate factor models. *Econometrica* 70(1), 191–221.

- Bai, J. and S. Ng (2007). Determining the number of primitive shocks in factor models. *Journal of Business and Economic Statistics* 25(1), 52–60.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31(3), 307–327.
- Connor, G. and R. A. Korajczyk (1988). Risk and Return in equilibrium APT: Application of a new test methodology. *Journal of Financial Economics* 21(2), 255–289.
- Connor, G. and R. A. Korajczyk (1993). A test for the number of factors in an approximate factor model. *Journal of Finance* 48(4), 1263–91.
- Cox, D. R. (1981). Statistical analysis of time series: some recent developments. *Scandinavian Journal of Statistics* 8(2), 93–115.
- Creal, D. D., S. J. Koopman, and A. Lucas (2012). Generalized Autoregressive Score Models with Applications. *Journal of Applied Econometrics*.
- CreditMetrics (2007). CreditMetrics (TM) - Technical Document, RiskMetrics Group. www.riskmetrics.com/pdf/dnldtechdoc/CMTD1.pdf.
- Das, S. R., D. Duffie, N. Kapadia, and L. Saita (2007). Common failings: how corporate defaults are correlated. *The Journal of Finance* 62(1), 93–117.
- Diebold, F. X. and T. Schuermann (2008). Exact maximum likelihood estimation of observation-driven econometric models. In R. Mariano, T. Schuermann, and M. J. Weeks (Eds.), *Simulation-based inference in econometrics: methods and applications*. Cambridge: Cambridge University Press. Chapter 8.
- Doz, C., D. Giannone, and L. Reichlin (2006). A Quasi Maximum Likelihood Approach for Large Approximate Dynamic Factor Models. *ECB working paper* 674.
- Duffie, D., A. Eckner, G. Horel, and L. Saita (2009). Frailty correlated default. *Journal of Finance* 64(5), 2089–2123.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* 50(4), 987–1007.
- Engle, R. F. (2002). Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business and Economic Statistics* 20(3), 339–350.
- Engle, R. F. and G. M. Gallo (2006). A multiple indicators model for volatility using intra-daily data. *Journal of Econometrics* 131(1-2), 3–27.
- Engle, R. F. and J. R. Russell (1998). Autoregressive conditional duration: a new model for irregularly spaced transaction data. *Econometrica* 66(5), 1127–1162.
- Engle, R. F. and M. W. Watson (1981). A one-factor multivariate time series model of metropolitan wage rates. *Journal of the American Statistical Association* 76(376), 774–781.
- Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2000). The Generalized Dynamic-Factor Model: Identification and Estimation. *The Review of Economics and Statistics* 82(4), 540–554.

- 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* 100(471), 830–840.
- Geweke, J. (1977). The dynamic factor analysis of economic time series. In D. Aigner and A. Goldberger (Eds.), *Latent Variables in Socio-Economic Models*, pp. 365–383. Amsterdam: North-Holland.
- Ghysels, E., P. Santa Clara, and R. Valkanov (2006). Predicting volatility: getting the most out of return data sampled at different frequencies. *Journal of Econometrics* 131(1-2), 59–95.
- Gupton, G. and R. Stein (2005). LossCalc V2: Dynamic Prediction of LGD. *Moodys KMV Investors Services*.
- Jungbacker, B. and S. J. Koopman (2008). Likelihood-based Analysis of Dynamic Factor Models. *Tinbergen Institute Discussion Paper*.
- Koop, G., M. H. Pesaran, and S. M. Potter (1996). Impulse response analysis in nonlinear multivariate models. *Journal of Econometrics* 74(1), 119–147.
- Koopman, S. J., R. Kraeussl, A. Lucas, and A. Monteiro (2009). Credit cycles and macro fundamentals. *Journal of Empirical Finance* 16(1), 42–54.
- Koopman, S. J., A. Lucas, and A. Monteiro (2008). The multi-state latent factor intensity model for credit rating transitions. *Journal of Econometrics* 142(1), 399–424.
- Koopman, S. J., A. Lucas, and B. Schwaab (2011). Modeling frailty correlated defaults using many macroeconomic covariates. *Journal of Econometrics* 162(2), 312–325.
- Lando, D. (2004). *Credit Risk Modeling: Theory and Applications*. Princeton, NJ: Princeton University Press.
- Mariano, R. S. and Y. Murasawa (2003). A new coincident index of business cycles based on monthly and quarterly series. *Journal of Applied Econometrics* 18(4), 427–443.
- McNeil, A., R. Frey, and P. Embrechts (2005). *Quantitative Risk Management*. Princeton, New Jersey: Princeton University Press.
- McNeil, A. and J. Wenden (2007). Bayesian inference for generalized linear mixed models of portfolio credit risk. *Journal of Empirical Finance* 14(2), 131–149.
- Merton, R. (1974). On the pricing of corporate debt: the risk structure of interest rates. *The Journal of Finance* 19(2), 449–470.
- Sargent, T. J. and C. R. Sims (1977). Business cycle modeling without pretending to have too much a priori economic theory. *Federal Reserve Bank of Minneapolis*.
- Schuermann, T. (2006). What do we know about Loss-given-default? In D. Shimko (Ed.), *Credit Risk Models and Management*. Risk Books, London, UK.
- Stock, J. H. and M. W. Watson (2002). Macroeconomic Forecasting Using Diffusion Indexes. *Journal of Business and Economic Statistics* 20(2), 147–162.
- Watson, M. W. and R. F. Engle (1983). Alternative algorithms for estimation of dynamic MIMIC, factor, and time-varying coefficient regression models. *Journal of Econometrics* 23(3), 385–400.

Table 1: *Likelihoods and Information Criteria*

The table contains the log-likelihood values and information criteria for alternative model specifications. Each model contains a different number of macroeconomic (m), credit or frailty risk (c), and LGD factors (r) which are ordered as (m, c, r) . The maximum log-likelihood value and the minimum AIC and BIC are denoted in bold.

	(3,0,0)	(3,1,0)	(3,2,0)	(3,2,1)
log-Lik	-40056.2	-39817.1	-39780.8	-39780.0
AIC	80242.4	79776.2	79713.6	79716.0
BIC	80991.0	80594.0	80589.0	80615.0

	(4,0,0)	(4,1,0)	(4,2,0)	(4,2,1)
log-Lik	-39828.7	-39596.3	-39561.6	-39560.4
AIC	79805.3	79352.7	79293.2	79294.8
BIC	80658.0	80274.0	80273.0	80297.0

Table 2: Parameter estimates and standard errors for the (4,2,0) model.

This table contains the estimated parameters and their standard errors for our model with (4,2,0) factor structure. The macros are ordered from $i = 1, \dots, 6$ as industrial production growth (IP), unemployment rate change (UR), annual real GDP growth (RGDP), credit spread (Cr.Spr.), annual return on the S&P500 (SP500), and annual realized volatility of the S&P500 returns using the past 252 daily trading days ($\sigma_{S&P}$). Significance at the 10%, 5%, and 1% level is denoted by *, **, and ***, respectively.

	macro ₁	macro ₂	macro ₃	macro ₄	frailty ₁	frailty ₂		
<i>A</i>	0.221*** (0.011)	0.154*** (0.014)	0.300*** (0.016)	0.282*** (0.023)	0.033*** (0.006)	0.036*** (0.010)		
<i>B</i>	0.966*** (0.012)	0.974*** (0.016)	0.924*** (0.022)	0.896*** (0.026)	0.974*** (0.012)	0.981*** (0.012)		
Z^m								
	macro ₁	macro ₂	macro ₃	macro ₄	frailty ₁	frailty ₂	z^m	Σ_m
IP	1.000 —	0.000 —	0.000 —	0.000 —	0.000 —	0.000 —	0.000 —	0.049*** (0.004)
UR	-0.467*** (0.061)	1.000 —	0.000 —	0.000 —	0.000 —	0.000 —	0.000 —	0.049*** (0.004)
RGDP	0.675*** (0.070)	-0.266** (0.120)	0.000 —	0.295*** (0.068)	0.000 —	0.000 —	0.000 —	0.261*** (0.033)
Cr.Spr.	-0.275*** (0.075)	0.000 —	1.000 —	0.000 —	0.000 —	0.000 —	0.000 —	0.058*** (0.004)
$r_{S&P}$	0.000 —	-0.358** (0.157)	-0.293*** (0.076)	1.179*** (0.086)	0.000 —	0.000 —	0.000 —	0.154*** (0.012)
$\sigma_{S&P}$	0.101* (0.096)	0.245* (0.176)	0.563*** (0.084)	1.000 —	0.000 —	0.000 —	0.000 —	0.507*** (0.042)
Z^c								
$R_{it} =$	macro ₁	macro ₂	macro ₃	macro ₄	frailty ₁	frailty ₂		
IG	-0.064* (0.051)	0.000 —	0.217*** (0.053)	-0.110* (0.064)	1.520*** (0.283)	-0.727** (0.270)		
BB	0.000 —	0.204*** (0.057)	0.158*** (0.038)	-0.077* (0.041)	1.000 —	0.000 —		
B	-0.154*** (0.036)	0.130** (0.056)	0.150*** (0.031)	-0.121*** (0.038)	0.914*** (0.137)	0.000 —		
CCC	-0.283*** (0.052)	0.000 —	0.076* (0.048)	0.000 —	1.486*** (0.418)	1.000 —		
z^c								
$R_{it} =$	\leq IG	\leq BB	\leq B	\leq CCC				
IG	6.299*** (0.155)	8.288*** (0.174)	9.626*** (0.224)	10.293*** (0.275)				
BB	-5.197*** (0.086)	4.630*** (0.082)	7.056*** (0.123)	7.453*** (0.140)				
B	-7.636*** (0.132)	-5.455*** (0.080)	4.648*** (0.075)	5.974*** (0.085)				
CCC	-9.071*** (0.403)	-8.154*** (0.295)	-5.341*** (0.200)	3.751*** (0.191)				
	macro ₁	macro ₂	macro ₃	macro ₄	frailty ₁	frailty ₂	z^r	β_r
Z^r	0.000 —	0.000 —	0.274*** (0.045)	-0.077* (0.057)	0.938** (0.315)	0.913*** (0.199)	0.226* (0.151)	2.572*** (0.094)

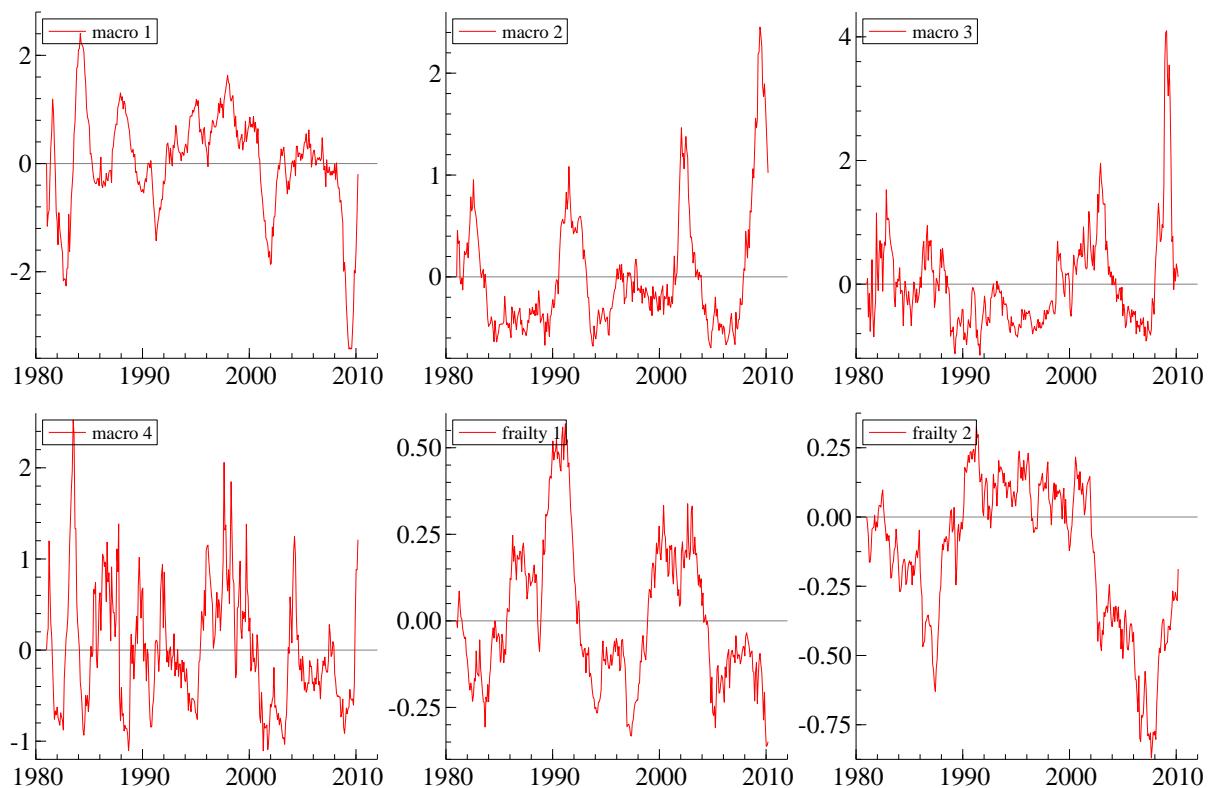


Figure 1: *Estimates of macro and frailty factors in the (4,2,0) model*

This figure contains the estimated factors f_t for a specification with 4 macro and 2 frailty factors.

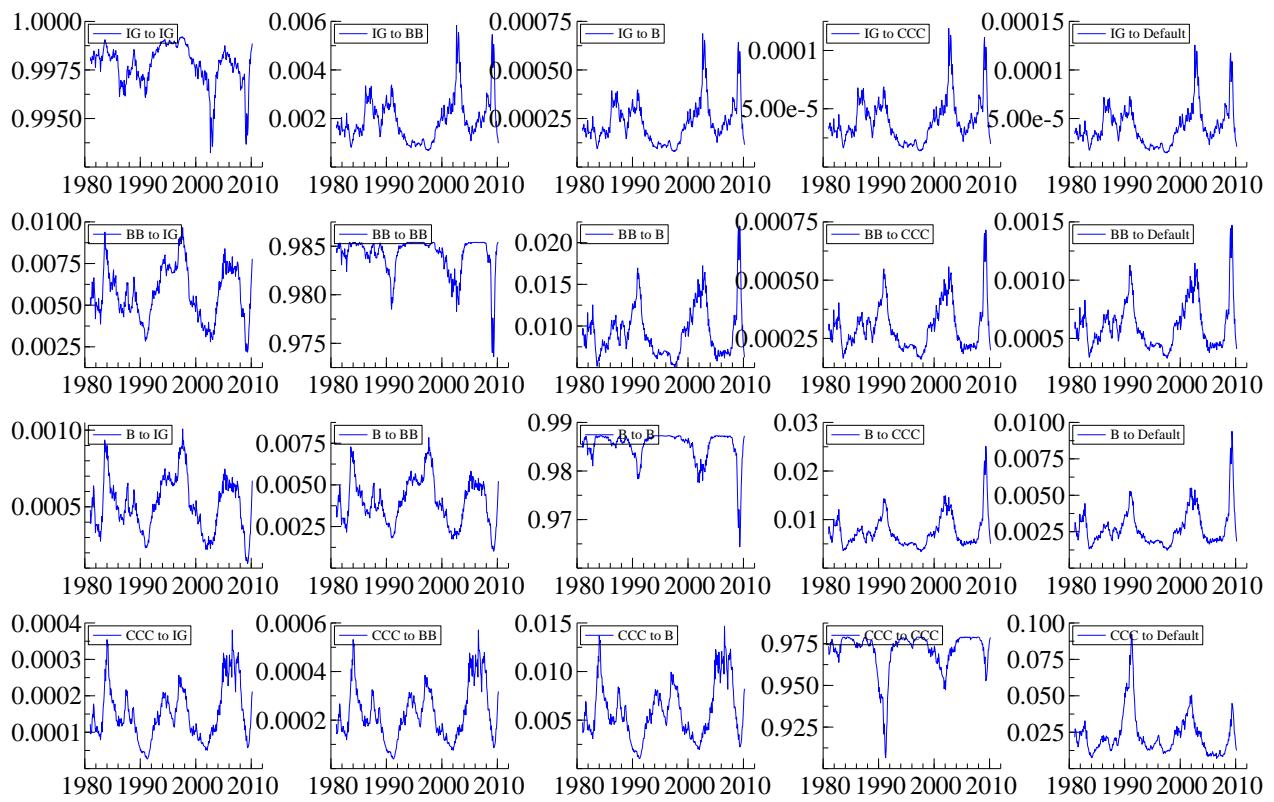


Figure 2: *Time-varying transition probabilities for the $(4,2,0)$ model*

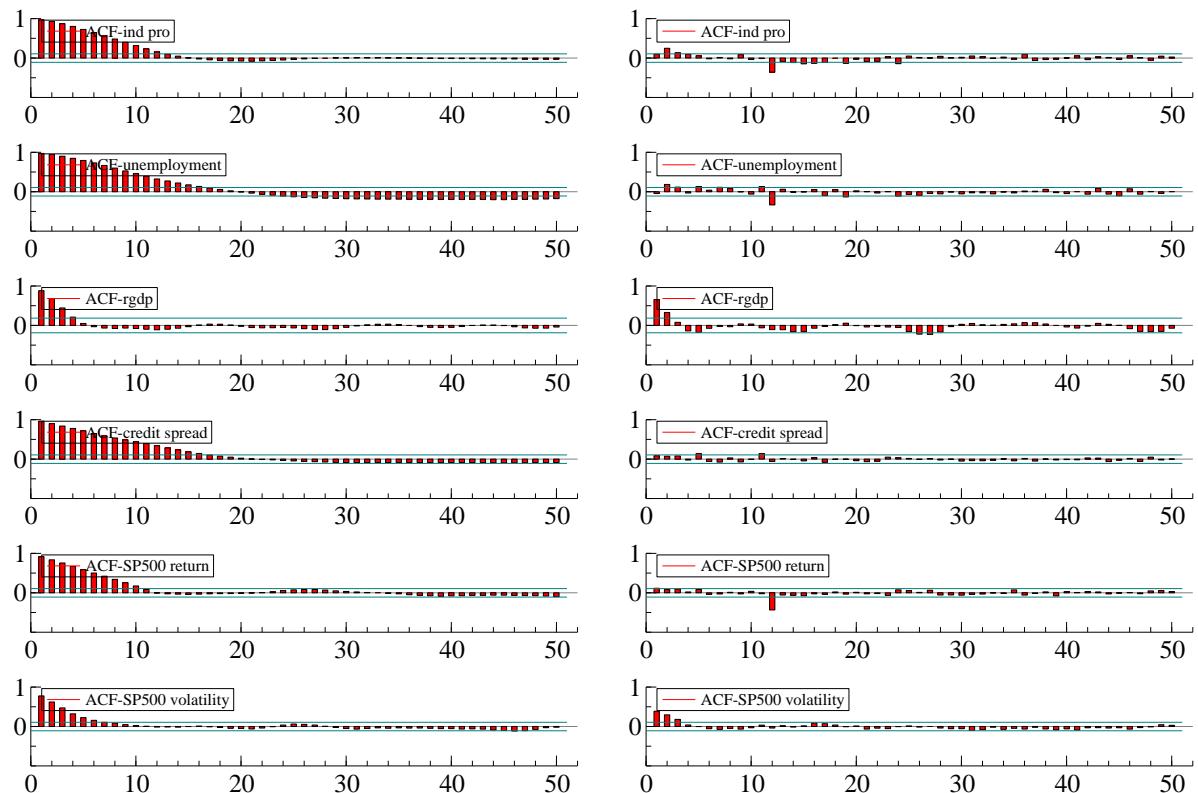


Figure 3: Autocorrelation functions of the data and prediction errors for the (4,2,0) model

This figure presents the autocorrelation functions of the data (left column) and the one-step ahead prediction errors (right column) for the macroeconomic series in the (4,2,0) model.

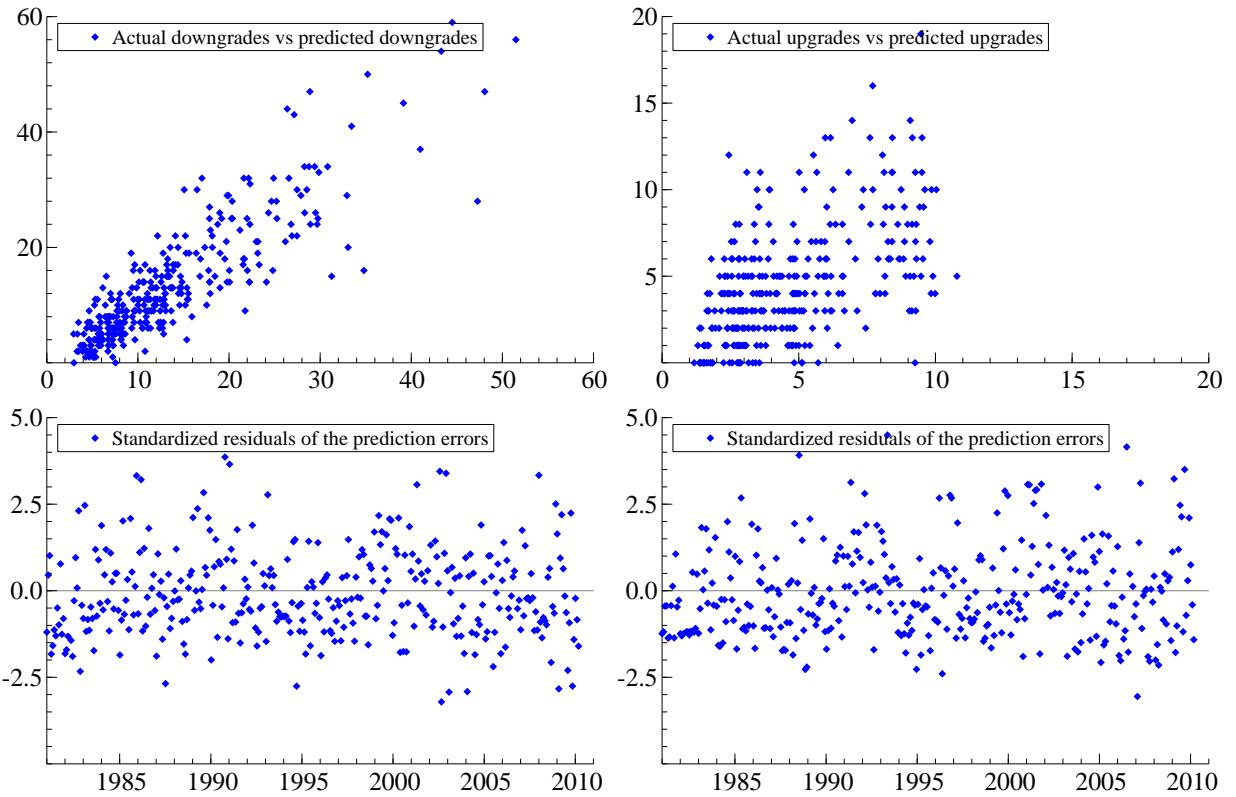


Figure 4: One-step ahead predicted versus realized upgrades and downgrades in the (4,2,0) model

This figure presents diagnostics for the credit-rating transitions produced from the (4,2,0) model. Upper left are the actual downgrades (vertical axis) versus the one-step ahead predicted downgrades (horizontal axis). Upper right are the actual upgrades (vertical axis) versus the one-step ahead predicted upgrades (horizontal axis). Lower left are the standardized prediction residuals for the downgrades. Lower right are the standardized prediction residuals for the upgrades.

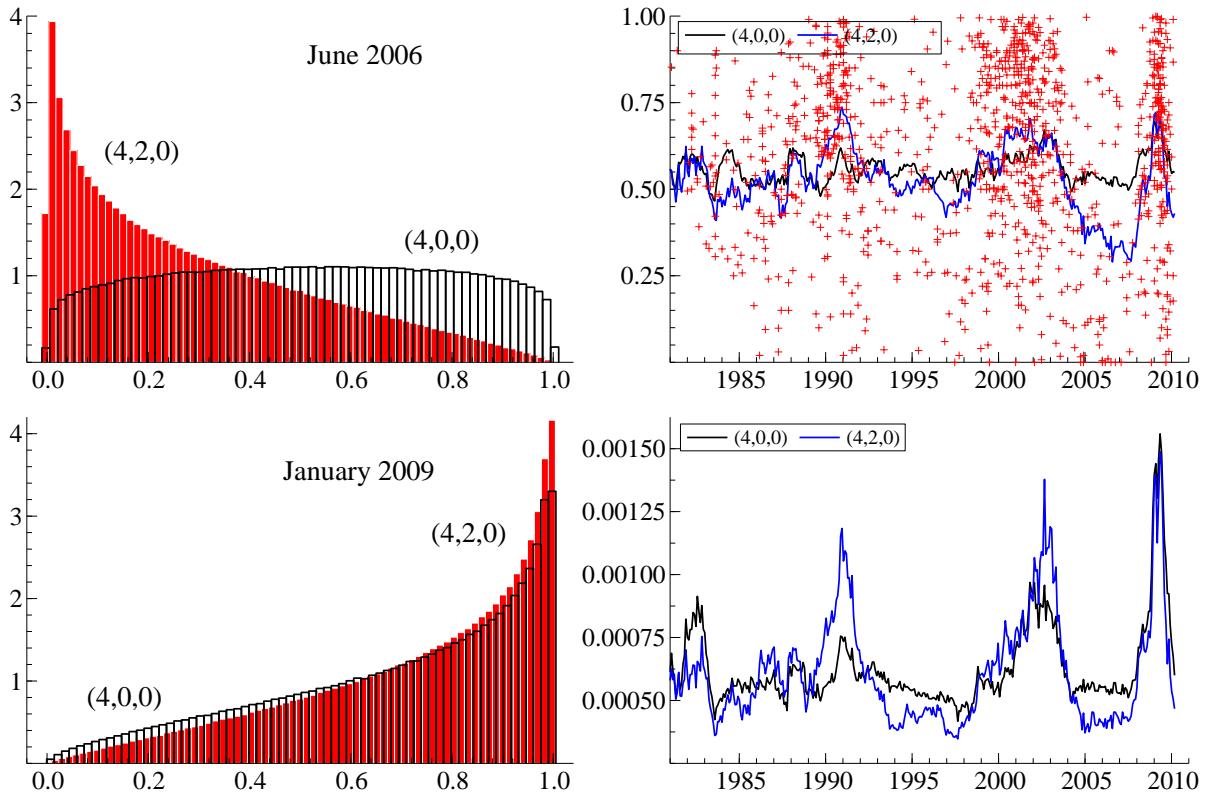


Figure 5: *Loss-Given-Default (LGD) dynamics*

The left panels contain the cross sectional beta distributions applicable in June 2006 and January 2009 for a model with three macro factors $(4,0,0)$ and a model with three macro and two frailty factors $(4,2,0)$. The upper-right panel contains the time series plot of the means of the LGD distributions and its fit to the observed LGD data. The lower-right panel gives the transition probability from BB to Default for the $(4,0,0)$ and $(4,2,0)$ models.

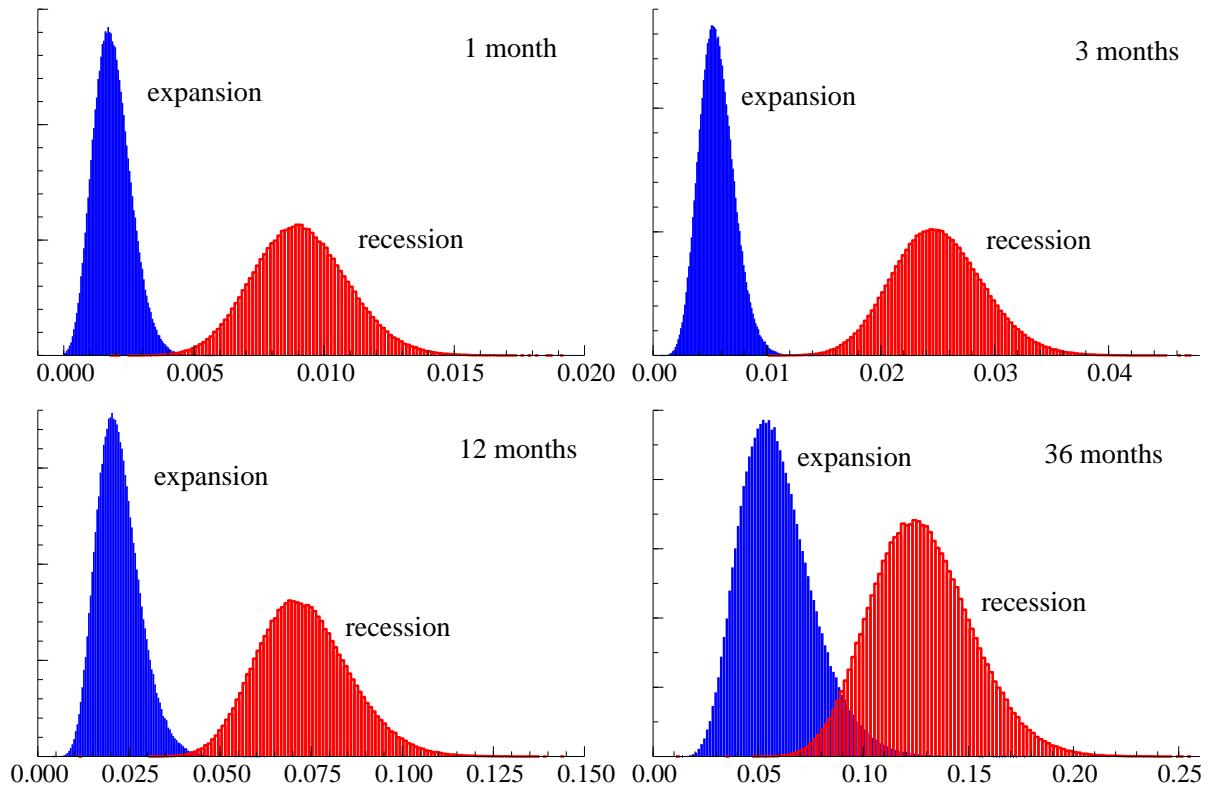


Figure 6: *Comparison of simulated loss distributions for the (4,2,0) model*

For our model with four macro and two frailty factors, the panels present the cumulative losses at different horizons. The difference between the two curves is the starting values for the factors, namely recession and expansion.

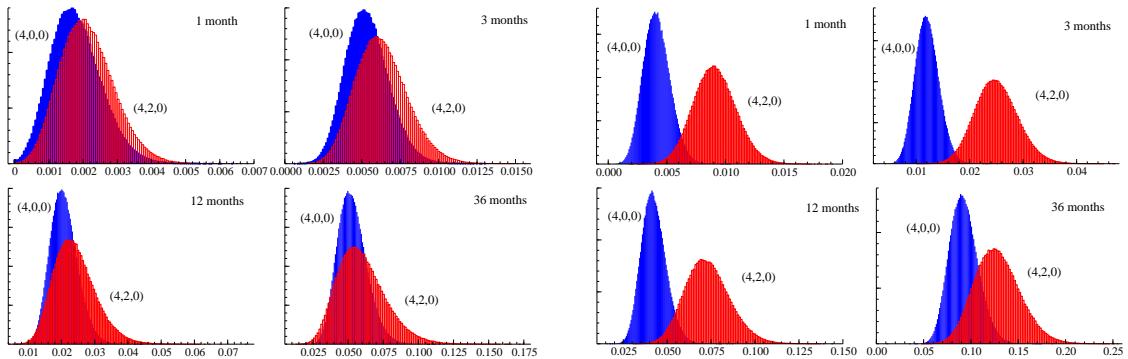


Figure 7: Comparison of simulated loss distributions for the $(4,2,0)$ and $(4,0,0)$ model

For our model with four macro and two frailty factors $(4,2,0)$, respectively only four macro factors $(4,0,0)$, the panels present the cumulative losses at different horizons. The left-hand four panels show the results if the factors f_t are started at zero. The right-hand four panels show the results if the factors are started in a recession period.

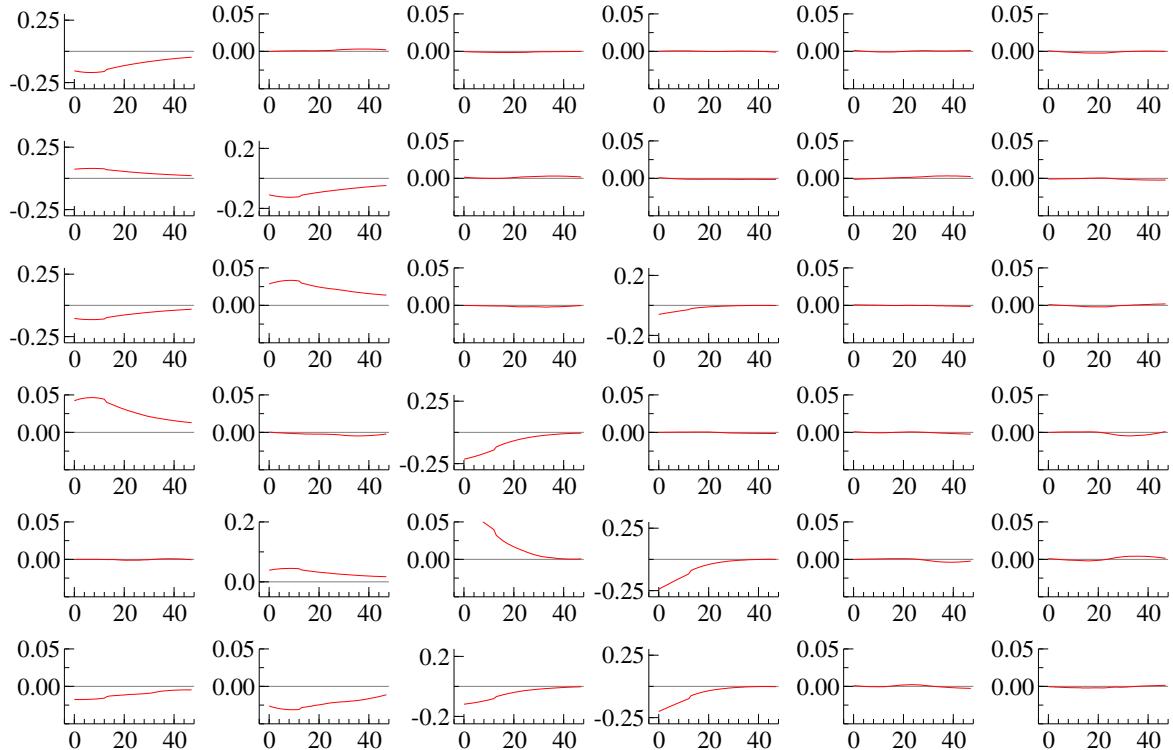


Figure 8: Non-linear impulse response functions for the (4,2,0) model: macros

For our model with three macro and two frailty factors, one of the factors f_t is given a unit size negative shock. All of the remaining stochastics of the model are simulated 48 months forward. The impulse response functions plot the difference between the average of the simulated quantity for a unit size shock to one of the factors and the average of the same quantity where the same factor receives a random model shock. The panels show the results for the 6 macros: rows 1 to 6 for industrial production growth, the change in the unemployment rate, real GDP growth, the credit spread, the S&P500 return, and its volatility, versus columns 1 to 6 for shocking the macro factors 1 to 4 and the 2 frailty factors. Factors are started at their mean values $f_T = 0$.

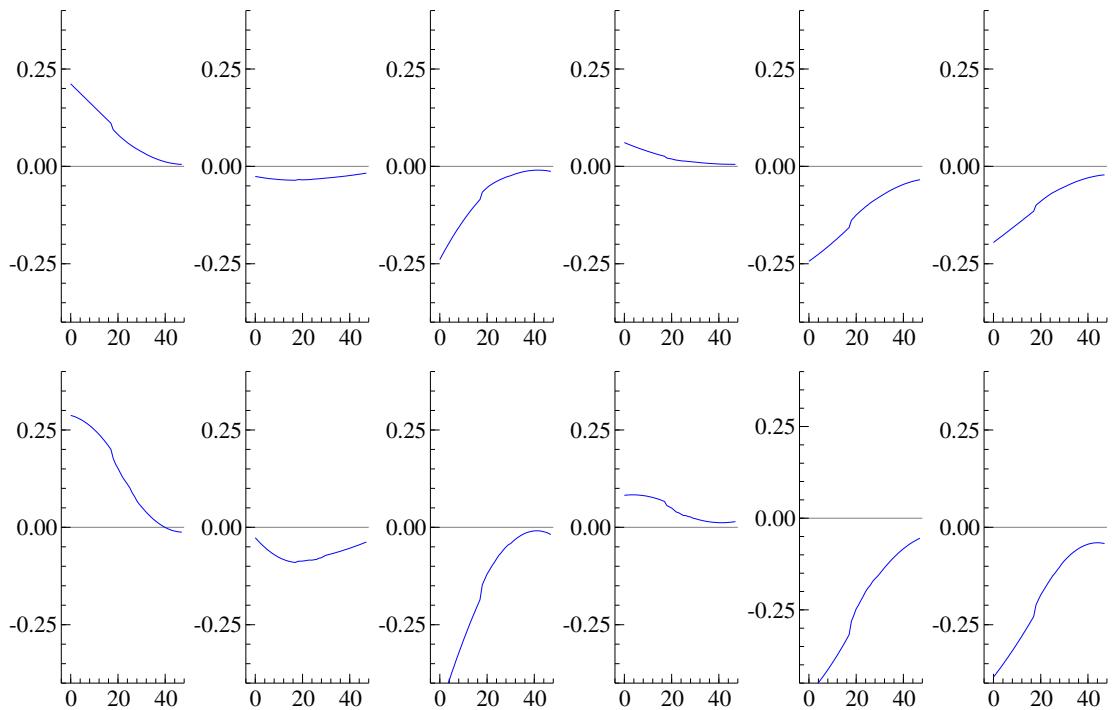


Figure 9: *Non-linear impulse response functions for the (4,2,0) model: portfolio losses*

For our model with three macro and two frailty factors, one of the factors f_t is given a unit size shock. All of the remaining stochastics of the model are simulated 48 months forward. The impulse response functions plot the difference between the average of the simulated quantity for a unit size shock to one of the factors and the average of the same quantity where the same factor receives a random model shock. The panels show the result for the mean portfolio credit loss (top row) and its 90th percentile (bottom row). The portfolio holds 1144 firms rated IG, 265 firms rated BB, 615 firms rated B, and 311 firms rated CCC. Columns 1 to 6 are for a shock to the macro factors 1 to 4 and the 2 frailty factors, respectively. Factors are started at their values fitted at the end of our sample.