Modeling Frailty-correlated Defaults Using Many Macroeconomic Covariates *

Siem Jan Koopman\textsuperscript{(a,c)} Andre Lucas\textsuperscript{(b,c,d)} Bernd Schwaab\textsuperscript{(b,c)}

\textsuperscript{(a)} Department of Econometrics, VU University Amsterdam
\textsuperscript{(b)} Department of Finance, VU University Amsterdam
\textsuperscript{(c)} Tinbergen Institute
\textsuperscript{(d)} Duisenberg school of finance

December 9, 2008

Abstract

We propose a new econometric model for estimating and forecasting the panel dynamics in disaggregated corporate default data. The model captures changes in systematic default risk using dynamic factors from a large panel of continuous macroeconomic data and a non-Gaussian latent factor model for default counts. In an empirical application we show that a latent frailty factor is required even after taking into account many macroeconomic covariates. In the proposed ‘mixed factor’ framework we capture a large part of default dependence across rating and industry groups, and substantially improve the forecasting accuracy associated with time varying conditional default probabilities, particularly in years of high default stress.

\footnote*{Corresponding author: Siem Jan Koopman, VU University Amsterdam, Department of Econometrics, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands, Tel +31 20 598 6019, Fax +31 20 598 6020, Email s.j.koopman@feweb.vu.nl. We are grateful to Peter Boswijk, Darrel Duffie and Michel van der Wel for their comments on different versions of this paper. We further would like to thank participants in presentations at the BIS workshop Stress Testing of Credit Portfolios in Amsterdam, the ESEM 2008 Meeting in Milan, the Factor Structures for Multivariate Time Series Data conference in Maastricht, the International Conference on Price, Liquidity, and Credit Risks in Konstanz, the 11th Symposium on Finance, Banking, and Insurance in Karlsruhe, the American Economic Association 2009 meeting (ASSA session) in San Francisco, and seminar participants at Tinbergen Institute, Utrecht University, and VU University Amsterdam. All remaining errors are our own.}
Keywords: Non-Gaussian Panel Data; Common Factors; Unobserved Components; Forecasting Conditional Default Probabilities

JEL classification: C33; G21.
1 Introduction

Recent research indicates that observed macroeconomic variables and firm-level information are not sufficient to capture the large degree of default clustering present in observed corporate default data. In an important study, Das, Duffie, Kapadia, and Saita (2007) reject the joint hypothesis of well-specified default intensities in terms of observed macroeconomic variables and firm-specific information and the conditional independence (doubly stochastic default times) assumption. This is bad news for practitioners, since virtually all available credit risk models build on conditional independence. Also, there is substantial evidence for an additional dynamic unobserved risk factor, see also Duffie, Eckner, Guillaume, and Saita (2008). This factor causes default dependence above and beyond what is implied by observed macroeconomic and financial data alone. We refer to this phenomenon as frailty-correlated defaults.

The econometric literature which explicitly allows for unobserved risk factors in the econometric specification consists mostly of recent contributions, including McNeil and Wendin (2007), Das et al. (2007), Duffie et al. (2008), Azizpour and Giesecke (2008), Koopman, Lucas, and Monteiro (2008), and Koopman and Lucas (2008). When default events depend on dynamic latent components, advanced econometric techniques based on simulation methods are required. For example, both McNeil and Wendin (2007) and Duffie et al. (2008) employ Bayesian sampling methods. In contrast, this paper adopts a maximum likelihood framework by using importance sampling techniques derived for multivariate non-Gaussian models in state space form, see Durbin and Koopman (1997, 2001) and further extensions in Koopman and Lucas (2008). The dependence on simulation methods is one reason why to our knowledge all unobserved component models allow for only a small number of observable macro variables alongside the unobserved factor.

In this paper we develop a model for the estimation and forecasting of time-varying default risk conditions in a frailty setting. In modeling the dynamics of conditional default probabilities we take into account the information from a large array of macroeconomic and financial variables by focusing on what they have ‘in common’. A latent frailty component allows us to capture the remaining default-specific dynamics. In effect, the proposed model combines a non-Gaussian panel specification with a dynamic factor model for continuous
time series data as used in for example Stock and Watson (2002a, 2005). We focus on
the nesting of these two strands of literature in high-dimensional multivariate time series
modeling.

This paper makes three contributions to the literature. First, we demonstrate how a
panel data specification for discrete default counts can be combined with an approximate
dynamic factor model for continuous macroeconomic time series data. The resulting model
inherits the best of both worlds. Factor models readily permit the use of information from
large arrays of relevant predictor variables. The non-Gaussian panel model in state space
form allows for unobserved frailty components, captures the cross-sectional heterogeneity of
firms, and easily accommodates missing values. Missing values arise easily in default count
data at a highly disaggregated level.

Second, we find that a frailty component is indispensable for capturing the high degree
of default clustering in ‘bad times’. We therefore strengthen the finding of Das et al. (2007)
in the sense that even a very large amount of macroeconomic and financial information -
while helpful - is not sufficient to explain all observed default clustering. A dynamic latent
component is needed. Given the large amount of macroeconomic information used in this
study it is further unlikely that the latent component can be explained by omitted relevant
macroeconomic data, an explanation offered in Duffie et al. (2008). Instead, the ‘default
cycle’ and the ‘business cycle’ appear to depend on two different processes. Although we use
many macroeconomic variables to capture business cycle variation in default counts, it still
leaves a substantial amount of time variation unaccounted for.

Third, we show that both kinds of factors - common factors from observed macroeconomic
time series, as well as latent factors - are useful for out-of-sample forecasting of default risk.
In a forecasting experiment we find that adding an unobserved component to a set of common
macroeconomic factors improves forecasting accuracy considerably. Feasible improvements
are substantial, particularly in years of high default stress such as 2001. Thus, the model
appears to deliver good forecasts when they are needed most. Improved models for the
time-variation in conditional default probabilities over a large cross-section of firms are most
relevant for financial credit risk management and banking supervision. The model-implied
default probabilities can be used as inputs for the calculation of Value-at-Risk based capital
buffers, for stress testing selected parts of the loan book, and the commonly used actuarial
approach to short-term loan pricing.

This paper proceeds as follows. In Section 2 we introduce the econometric framework of the mixed factor model which combines a non-Gaussian, nonlinear panel time series model with a dynamic factor model for many covariates. Section 3 shows that a close correspondence exists between the proposed econometric model and a multi-factor firm value model for dependent defaults. In Section 4 we discuss the estimation of unknown parameters in the model. Section 5 introduces the two panel data sets for our empirical study, and presents estimation and out-of-sample forecasting results. Section 6 concludes.

2 The econometric framework

In this section we present our reduced form econometric model for dependent defaults. The economic interpretations of this framework are discussed in Section 3. We denote the default counts of cross section \( j \) at time \( t \) as \( y_{jt} \) for \( j = 1, \ldots, J \), and \( t = 1, \ldots, T \). The index \( j \) denotes a combination of firm characteristics, such as industry sector, current rating class, and company age. Defaults are correlated in the cross-section through exposure to the same business cycle, financing conditions, monetary policy, consumer sentiment, etc. These macroeconomic forces are summarized by exogenous factors \( F_t \). A frailty factor \( f_{uc}^t \) (where ‘uc’ refers to unobserved component) captures default clustering above and beyond what is implied by observed macro data. Subject to the conditioning on observed and unobserved risk factors, defaults occur independently in the cross section, see e.g. CreditMetrics (2007) and the textbook exposition of Lando (2003, Chapter 9). The multivariate time series of counts \( y_t = (y_{1t}, \ldots, y_{Jt})' \) is therefore modeled by

\[
y_{jt} | f_{uc}^t, F_t \sim \text{iid Binomial}(k_{jt}, \Pi_{jt}).
\]  (1)

When modeling default counts as a Binomial sequence we interpret \( y_{jt} \) as the total number of default ‘successes’ from \( k_{jt} \) independent Bernoulli-trials with time-varying default probability \( \Pi_{jt} \). In our case, \( k_{jt} \) denotes the number of firms in cell \( j \) that are active at the beginning of period \( t \). We recount \( k_{jt} \) at the beginning of each quarter.

Estimating and forecasting conditional default probabilities \( \Pi_{jt} \) is our central focus. We
model $\Pi_{jt}$ as the logistic transform of an index function $\theta_{jt}$. Therefore, $\theta_{jt}$ may be interpreted as the ‘log-odds’ or logit transform of $\Pi_{jt}$. Probit and other transforms are also possible. Each specification implies a different model formulation and (slightly) different estimation results. We prefer the logit transformation as it leads to convenient expressions. The default probabilities are specified by

$$
\Pi_{jt} = \frac{1}{1 + e^{-\theta_{jt}}},
$$

(2)

$$
\theta_{jt} = \lambda_j + \beta_j f_{uc}^t + \gamma_j' F_t,
$$

(3)

where $\lambda_j$ is a fixed effect for each cross section, and loading coefficients $\beta_j$ and $\gamma_j$ capture risk factor sensitivities which may depend on firm characteristics such as industry sectors and rating classes. The default signals $\theta_{jt}$ do not contain idiosyncratic error terms. Instead, idiosyncratic randomness is captured in (1). The ‘log-odds’ of conditional default probabilities may vary over time due to the variation in the frailty component, $f_{uc}^t$, and variation in the macroeconomic factors, $F_t$.

The frailty factor $f_{uc}^t$ is unobserved and needs to be modeled explicitly. We specify the dynamics of $f_{uc}^t$ by a stationary autoregressive process of order 1, that is

$$
f_{uc}^t = \phi f_{uc}^{t-1} + \sqrt{1 - \phi^2} \eta_t, \quad \eta_t \sim \text{NID}(0,1),
$$

(4)

where $0 < \phi < 1$. The innovations are normalized such that $\mathbb{E}(f_{uc}^t) = 0$, $\text{Var}(f_{uc}^t) = 1$, and $\text{Cov}(f_{uc}^t, f_{uc}^{t-h}) = \phi^h$. The normalization identifies the coefficient $\beta_j$ in (3) as the standard deviation, or volatility, of $f_{uc}^t$ for firms in cross section $j$. Extensions to more unobserved components for firm-specific heterogeneity as well as other dynamic specifications for $f_{uc}^t$ are possible, see Koopman and Lucas (2008).

Modeling the dependence of firm defaults on observed macro variables is an active area of current research, see e.g. Duffie et al. (2007) and references therein. The set of considered macroeconomic variables is usually small and differs across studies. Instead of opting for a specific selection, in our current study we collect a large number of macroeconomic and financial variables denoted $x_{it}$ for $i = 1, \ldots, N$. This time series panel of macroeconomic predictor variables may contain even more regressors than all regressors used in previous
studies. The panel is assumed to adhere to a factor structure as given by

\[ x_{it} = \Lambda_i F_t + e_{it}, \]  

(5)

where \( F_t \) is a vector of principal components, \( \Lambda_i \) is a row vector of loadings, and \( e_{it} \) is an idiosyncratic error term. Equation (5) is the static representation of an approximate dynamic factor model and is described in Stock and Watson (2002a). The static representation (5) can be derived from a dynamic specification such as

\[ x_{it} = v_i(L) f_t + e_{it} \]

by assuming that the lag polynomials \( v_i(L) \) operating on the factors \( f_t \) are of finite (low) order, see Stock and Watson (2002b). The coefficients in \( v_i \) are stacked in \( \Lambda_i \), while the contemporaneous and lagged factors are stacked in \( F_t \). The estimated \( F_t \) represents current and lagged factors in the economy. The methodology of relating a variable of interest to macroeconomic factors has been employed in the forecasting of inflation and production data, see e.g. Massimiliano, Stock, and Watson (2003), or asset returns and volatilities, see e.g. Ludvigson and Ng (2007). These studies have reported favorable results when such factors are used for forecasting.

The main advantage of our current framework is that \( F_t \) can be estimated consistently using the method of principal components. This method is expedient for several reasons. First, dimensionality problems do not occur even for high values of \( N \) and \( T \). This is particularly relevant for our empirical application, where \( T, N > 100 \) in both the macro and default datasets. Second, the method of principal components works under relatively weak assumptions and coincides with the maximum likelihood estimator if idiosyncratic terms are assumed to be Gaussian. Finally, the obtained factors can be used directly for forecasting purposes, see Forni, Hallin, Lippi, and Reichlin (2005). Equations (1) to (5) combine the approximate dynamic factor model with a non-Gaussian panel data model by inserting the elements of \( F_t \) from (5) into the signal equation (3). Statistical model formulation and estimation is discussed in Section 4.

3 The financial framework

In this section we relate the econometric model from Section 2 to a multi-factor CreditMetrics-type model for dependent defaults. By establishing this link we can give an economic inter-
pretation to our econometric model and clarify the economic mechanisms at work. Single-
and multi-factor models for firm default risk are widely used in risk management practice,
see e.g. Lando (2003, Chapter 9) and Tasche (2006).

In a standard static one-factor credit risk model for dependent defaults the values of the
obligors’ assets, \( V_i \), are usually driven by a common, standard normally distributed factor
\( Y \), and an idiosyncratic standard normal noise term \( \epsilon_i \),
\[
V_i = \sqrt{\rho Y} + \sqrt{1 - \rho} \epsilon_i,
\]
for \( i = 1, \ldots, I \), see e.g. CreditMetrics (2007). A dynamic version of the single-factor specifi-
cation lets \( V_i \) vary over time. To introduce multiple factors we generalize the model further
by
\[
V_{it} = \delta_0 f_{it}^{\text{uc}} + \delta_i F_{1,t} + \cdots + \delta_R F_{R,t} + \sqrt{1 - (\delta_{0i})^2 - (\delta_{1i})^2 - \cdots - (\delta_{Ri})^2} \epsilon_{it}
\]
\[
= \delta_i' \tilde{f}_t + \sqrt{1 - \delta_i' \delta_i} \epsilon_{it},
\]
(6)
where \( \tilde{f}_t := (f_{it}^{\text{uc}}, F_{1,t}, \ldots, F_{R,t})' \), \( \delta_i := (\delta_{0i}, \delta_{1i}, \ldots, \delta_{Ri})' \), with \( \delta_i' \delta_i \leq 1 \), and \( \text{Var}[\epsilon_{it}] = 1 \).
The unobserved component \( f_{it}^{\text{uc}} \) is designed to identify credit cycle conditions that are not
captured by the \( R \) macro factors \( F_{1,t}, \ldots, F_{R,t} \). Without loss of generality we assume that
all risk factors have zero mean and unit variance. Furthermore, the risk factors in \( \tilde{f}_t \) are
constructed such that they are uncorrelated. This implies that \( \text{E}[V_{it}] = 0 \) and \( \text{Var}[V_{it}] = 1 \)
regardless of the assumed distribution for the idiosyncratic noise component \( \epsilon_{it} \).

Following Merton’s (1974) firm value-model, we assume that a default occurs as soon as
a firm’s net asset value \( V_{it} \) drops below a specified default barrier, say \( c_i \). This default barrier
may depend on the current rating class, the industry sector, and the time elapsed from the
initial rating assignment. Given these assumptions, a default occurs when
\[
V_{it} < c_i \iff \delta_i' \tilde{f}_t + \sqrt{1 - \delta_i' \delta_i} \epsilon_{it} < c_i
\]
\[
\iff \epsilon_{it} < \left( \frac{c_i - \delta_i' \tilde{f}_t}{\sqrt{1 - \delta_i' \delta_i}} \right).
\]
Denoting information up to time $t$ as $\mathcal{F}_t$, the default probability is given by

$$\Pi_{jt} = \Pr (\epsilon_{it} < \frac{c_j - \delta_j \tilde{f}_t}{\sqrt{1 - \delta_j \delta_j^*}} \bigg| \mathcal{F}_t ).$$

(7)

The default probability in (7) has an intuitive interpretation. Favorable credit cycle conditions, i.e. high values of factors $\tilde{f}_t$, are associated with low default probabilities $\Pi_{jt}$.

Equation (7) can also be related directly to the econometric model specification in (2) and (3). Specifically, if $\epsilon_{it}$ is logistically distributed, we obtain

$$c_i = \lambda_i \sqrt{1 - a_i},$$
$$\delta_{0,i} = -\beta_i \sqrt{1 - a_i},$$
$$\delta_{r,i} = -\gamma_{r,i} \sqrt{1 - a_i},$$

where $a_i = (\beta_i^2 + \gamma_{1,i}^2 + \gamma_{2,i}^2) / (1 + \beta_i^2 + \gamma_{1,i}^2 + \gamma_{2,i}^2)$. Therefore, the parameters in the econometric model of Section 2 have a direct interpretation in a widely used portfolio credit risk model such as CreditMetrics.

### 4 Estimation and state space form

The details of parameter estimation and signal extraction of the factors in model (1) to (5) are discussed in this section. The estimation procedure for the macro factors is discussed in Section 4.1. The state space representation of the econometric model is provided in Section 4.2. We estimate the parameters using computationally efficient (Monte Carlo) maximum likelihood and signal extraction procedures based on importance sampling methods. A brief outline of these procedures is given in Section 4.3. All computations are implemented using the Ox programming language and the associated set of state space routines from SsfPack, see Koopman, Shephard, and Doornik (1999, 2008) and Doornik (2007).

#### 4.1 Estimation of the macro factors

The common factors $F_t$ from the macro data are estimated by minimizing the objective function given by
\[
\min_{\{F_1, \ldots, F_T, \Lambda\}} V(F, \Lambda) = (NT)^{-1} \sum_{t=1}^{T} (X_t - \Lambda F_t)'(X_t - \Lambda F_t), \tag{8}
\]

where the Nx1 matrix \(X_t\) contains observed stationary macroeconomic variables. The time series are demeaned and standardized to unit variance. Concentrating out \(F_t\) and rearranging terms shows that (8) is equivalent to

\[
\max \bar{V}(\Lambda) = \text{tr} \left( \Lambda' \left[ \sum_{t=1}^{T} X_t X_t' \right] \Lambda \right) = T \text{tr} \left( \Lambda' S_{X'X} \Lambda \right) \tag{9}
\]

subject to \(\Lambda' \Lambda = I_r\), where \(S_{X'X} = T^{-1} \sum_t X_t X_t'\) is the sample covariance matrix of the data, see Lawley and Maxwell (1971) and Stock and Watson (2002a). The resulting principal components estimator of \(F_t\) is given by \(\hat{F}_t = X_t' \hat{\Lambda}\), where \(\hat{\Lambda}\) collects the normalized eigenvectors associated with the \(R\) largest eigenvalues of \(S_{X'X}\).

In case the variables in \(X_t\) are not completely observed for \(t = 1, \ldots, T\), we employ the Expectation Maximization (EM) procedure as devised in the Appendix of Stock and Watson (2002a). This iterative procedure takes a simple form under the assumption that \(x_{it} \sim \text{NID}(\Lambda_i F_t, 1)\), where \(\Lambda_i\) denotes the \(i\)th row of \(\Lambda\). In this case \(V(F, \Lambda)\) from (8) is affine to the complete data log-likelihood \(L(F, \Lambda | X^m)\), where \(X^m\) denotes the missing parts of the dataset \(X_1, \ldots, X_T\). Since \(V(F, \Lambda)\) is proportional to \(-L(F, \Lambda | X^m)\), the minimizers of \(V(F, \Lambda)\) are also the maximizers of \(L(F, \Lambda | X^m)\).

The procedure for obtaining the principal components in case of missing data in our setting is as follows. The objective function (8) is given by

\[
\min_{\{F_1, \ldots, F_T, \Lambda\}} V^*(F, \Lambda) = \sum_{i=1}^{N} \sum_{t=1}^{T} I_{it} (x_{it} - \Lambda_i F_t)^2, \tag{10}
\]

where \(I_{it} = 1\) if \(x_{it}\) is observed, and zero otherwise. Equation (10) is minimized iteratively, using the following two step EM algorithm:

1. Obtain initial values for \(\hat{F}_t\) and \(\hat{\Lambda}\). The estimates can be obtained from a smaller set of time series which contains no missing values.
2. For the Expectation-step: Replace missing values by their estimates

\[
\hat{x}_{it}^* = \begin{cases} 
    x_{it} & \text{if } x_{it} \text{ is observed}, \\
    \hat{\Lambda}_i \hat{F}_t & \text{if } x_{it} \text{ is missing.}
\end{cases}
\]

3. Maximization-step: Update \( \hat{F}_t \) and \( \hat{\Lambda} \) by performing the eigenvalue/-vector decomposition on the estimated covariance matrix of the balanced data,

\[
S_{XX}^* = T^{-1}(\hat{X}_t'\hat{X}_t^*),
\]

where \( X_t^* = (x_{1t}^*, \ldots, x_{Nt}^*)' \).

We iterate the EM steps 2 and 3 until convergence has taken place. To formulate a stopping criterion, the objective function \( V(F, \Lambda) \) can be computed as the squared Frobenius matrix norm of the \( TxN \) error matrix \( E = \hat{X}^* - \hat{F}\hat{\Lambda}' \), since \( V(F, \Lambda) = (NT)^{-1}\text{tr}(E'E) \). The iterations stop when the changes in the objective function become negligible, say smaller than \( 10^{-7} \).

4.2 The factor model in state space form

In this subsection we formulate the model (1) to (4) in state space form where \( F_t \) is treated as given. In our implementation, \( F_t \) will be replaced by \( \hat{F}_t \) as obtained from the previous section. The estimation framework can therefore be characterized as a two step procedure. By first estimating principal components to summarize the variation in macroeconomic data we ensure the computational feasibility and the conceptual simplicity of the overall approach. In Section 4.4 we present simulation evidence to illustrate the adequacy of our approach for parameter estimation and for uncovering the factors \( F_t \) and \( f_t^{ue} \) from the data. The joint modeling of (Binomial) default counts and (Gaussian) macro data through sets of common dynamic factors is an interesting direction for future research but will typically be more demanding in computational terms.

The conditionally Binomial log-density function of the model (1) is given by

\[
\log p(y_{jt}| \Pi_{jt}) = y_{jt} \log \left( \frac{\Pi_{jt}}{1 - \Pi_{jt}} \right) + k_{jt} \log(1 - \Pi_{jt}) + \log \left( \frac{k_{jt}}{y_{jt}} \right),
\]

(11)

where \( y_{jt} \) is the number of defaults and \( k_{jt} \) is the number of firms in cross-section \( j \), for \( j = 1, \ldots, J \) and \( t = 1, \ldots, T \). By substituting (2) for the default probability \( \Pi_{jt} \) into (11)
we obtain the log-density in terms of the log-odds ratio $\theta_{jt} = \log(\Pi_{jt}) - \log(1 - \Pi_{jt})$ given by

$$\log p(y_{jt}|\theta_{jt}) = y_{jt}\theta_{jt} + k_{jt}\log(1 + e^{\theta_{jt}}) + \log(k_{jt}).$$

(12)

The log-odds ratio is specified as

$$\theta_{jt} = Z_{jt}\alpha_t,$$
$$Z_{jt} = (e'_j, F_t' \otimes e'_j, \beta_j),$$

where $e_j$ denotes the $j$th column of the identity matrix of dimension $J$, the state vector $\alpha_t = (\lambda_1, \ldots, \lambda_J, \gamma_{1,1}, \ldots, \gamma_{R,J}, f_t^{uc})'$ consists of the fixed effects $\lambda_j$ together with the factor loadings $\gamma_{r,j}$ associated with macro factors for $r = 1, \ldots, R$, and the unobserved component $f_t^{uc}$. The system vector $Z_{jt}$ is time-varying due to the inclusion of $F_t$.

The state updating equation is given in its general form as

$$\alpha_{t+1} = T_t\alpha_t + B_t\xi_t, \quad \xi_t \sim N(0, Q_t),$$

(13)

where system matrices are given by

$$T_t = \text{diag}(I, \phi), \quad B_t = \begin{bmatrix} 0 \\ \sqrt{1 - \phi^2} \end{bmatrix}, \quad Q_t = 1,$$

such that $\xi_t$ is a scalar set equal to $\eta_t$ in (4). The initial elements of the state vector have a diffuse prior distribution with mean zero, except for $f_t^{uc}$ whose prior is given by $N(0,1)$.

Equations (12) and (13) form a non-Gaussian state space model as discussed in Durbin and Koopman (2001) part II, and Koopman and Lucas (2008). We note that equation (12) replaces the more familiar observation equation associated with a linear Gaussian model. In this formulation, most unknown coefficients are part of the state vector $\alpha_t$ and are estimated as part of the filtering and smoothing procedures described in Section 4.3. This increases the computational efficiency of our estimation procedure. The remaining parameters are collected in a coefficient vector $\psi = (\phi, \beta_1, \ldots, \beta_J)'$ and are estimated by the Monte Carlo maximum likelihood methods that we discuss in the next section.
4.3 Parameter estimation and signal extraction

Parameter estimation for a non-Gaussian model in state space form can be carried out by the method of Monte Carlo maximum likelihood. Once we have obtained an estimate of $\psi$, we can compute the conditional mean and variance estimates of the state vector $\alpha_t$. In both cases we make use of importance sampling methods. The details of our implementation are given next.

For notational convenience we suppress the dependence of the density $p(y; \psi)$ on $\psi$. The likelihood function of our model (1) to (4) can be expressed by

$$p(y) = \int p(y, \theta) d\theta = \int p(y|\theta)p(\theta)d\theta = \int p(y|\theta) \frac{p(\theta)}{g(\theta|y)} g(\theta|y)d\theta = E_g \left[ \frac{p(y|\theta)p(\theta)}{g(\theta|y)} \right], \quad (14)$$

where $y = (y_{11}, y_{21}, \ldots, y_{JT})^t$, $\theta = (\theta_{11}, \theta_{21}, \ldots, \theta_{JT})^t$, $p(\cdot, \cdot)$ is a density function, $p(\cdot, \cdot)$ is a joint density, $p(\cdot|\cdot)$ is a conditional density, $g(\theta|y)$ is a Gaussian importance density, and $E_g$ denotes expectations with respect to $g(\theta|y)$. The importance density is chosen as a close approximation to $p(\theta|y)$. Conditional on $\theta$, we can evaluate $p(y|\theta)$ by

$$p(y|\theta) = \prod_{i,j} p(y_{ji}|\theta_{ji}).$$

It follows from (3) and (4) that the marginal density $p(\theta)$ is Gaussian and therefore $p(\theta) = g(\theta)$. From Bayes’ identity $g(\theta|y)g(y) \equiv g(y|\theta)g(\theta)$ we obtain

$$p(y) = E_g \left[ p(y|\theta) \frac{p(\theta)}{g(y|\theta)} \frac{g(y)}{p(y|\theta)} \right] = E_g \left[ g(y) \frac{p(y|\theta)}{g(y|\theta)} \right] = g(y)E_g [w(y, \theta)], \quad (15)$$

where $w(y, \theta) = p(y|\theta)/g(y|\theta)$. A Monte Carlo estimator of $p(y)$ is therefore given by

$$\hat{p}(y) = g(y)\bar{w},$$

with

$$\bar{w} = M^{-1} \sum_{m=1}^{M} w^m = M^{-1} \sum_{m=1}^{M} \frac{p(y|\theta^m)}{g(y|\theta^m)},$$
where \( w^m = w(\theta^m, y) \) is the value of the importance weight associated with the \( m \)-th draw \( \theta^m \) from \( g(\theta|y) \), and \( M \) is the number of such draws. The Gaussian importance density \( g(\theta|y) \) is chosen for convenience and since it is possible to generate a large number of draws \( \theta^m \) from it in a computationally efficient manner using the simulation smoothing algorithms of de Jong and Shephard (1995) and Durbin and Koopman (2002). We estimate the log-likelihood as 
\[
\log \hat{p}(y) = \log \hat{g}(y) + \log \bar{w},
\]
and include a bias correction term discussed in Durbin and Koopman (1997).

The Gaussian importance density is an approximating model which is found by matching the first and second derivative of \( \log p(y|\theta) \) and \( \log g(y|\theta) \) with respect to the signal \( \theta \). This matching takes place around a current guess of the mode of \( \theta \). The next guess of the mode is obtained as the smoothed estimate of \( \theta \) from the current approximating model. Iterations proceed until convergence is achieved, which usually occurs very fast (less than 5 iterations). A new approximating model needs to be constructed for each trial evaluation of \( \log p(y) \) for a different value of parameter vector \( \psi \). Finally, standard errors for the parameters in \( \psi \) are constructed from the numerical second derivatives of the log-likelihood,
\[
\hat{\Sigma} = \left[ -\frac{\partial^2 \log p(y)}{\partial \psi \partial \psi'} \bigg|_{\psi = \hat{\psi}} \right]^{-1}.
\]

For signal extraction, we require the estimation of the conditional mean of \( \theta \) as given by
\[
\bar{\theta} = \mathbb{E}[\theta|y] = \int \theta p(\theta|y) d\theta = \int \theta p(\theta|y) \frac{g(\theta|y)}{g(\theta|y)} d\theta = \mathbb{E}_g \left[ \theta \frac{p(\theta|y)}{g(\theta|y)} \right].
\]
Using Bayes’ identity, the expression obtained in (15), and the fact that \( p(\theta) = g(\theta) \) we obtain
\[
\bar{\theta} = \frac{\mathbb{E}_g [\theta w(\theta, y)]}{\mathbb{E}_g [w(\theta, y)]},
\]
where \( w(\theta, y) \) are the importance sampling weights as defined above. An obvious estimator for \( \bar{\theta} \) is then given by
\[
\hat{\theta} = \hat{E}[\theta|y] = \left[ \sum_{m=1}^{M} w^m \right]^{-1} \sum_{m=1}^{M} \theta^m w^m.
\]
The associated conditional variances are given by

$$\text{Var}[\theta_{it}|y] = \left( \sum_{m=1}^{M} w^m \right)^{-1} \sum_{m=1}^{M} (\hat{\theta}_{it}^m)^2 w^m - (\hat{\theta}_{it})^2,$$

see Durbin and Koopman (2001, Chapter 11) for further details.

### 4.4 Simulation experiments

In this subsection we investigate whether the econometric methods of Sections 4.1 and 4.3 can distinguish the variation in default conditions due to changes in the macroeconomic environment from changes in unobserved frailty risk. The first source is captured by principal components $F_t$, while the second source is estimated via the unobserved factor $f_{uc}^t$. This check is important since estimation by Monte Carlo maximum likelihood should not be biased towards attributing variation to a latent component when it is due to an exogenous covariate. For this purpose we carry out a simulation study that is close to our empirical application in Section 5. The variables are generated by the equations

$$F_t = \Phi_F F_{t-1} + u_{F,t}, \quad u_{F,t} \sim N(0, I - \Phi_F \Phi_F'),$$
$$e_t = \Phi_I e_{t-1} + u_{I,t}, \quad u_{I,t} \sim N(0, I - \Phi_I \Phi_I'),$$
$$X_t = \Lambda F_t + e_t,$$
$$f_{uc}^t = \phi_{uc} f_{uc}^{t-1} + u_{f,t}, \quad u_{f,t} \sim N(0, 1 - \phi_{uc}^2),$$

where the elements of the matrices $\Phi_F$, $\Phi_I$, and $\Lambda$ are generated for each simulated dataset from the uniform distribution $U[. . .]$, that is $\Phi_F(i, j) \sim U[0.6, 0.8]$, $\Phi_I(i, j) \sim U[0.2, 0.4]$, and $\Lambda(i, j) \sim U[0, 2]$, where $A(i, j)$ is the $(i, j)$th element of matrix $A = \Phi_F, \Phi_I, \Lambda$. We refer to Stock and Watson (2002a) for a similar setup. Non-Gaussian default counts are generated by the equations

$$\theta_{jt} = \lambda_j + \beta f_{uc}^t + \gamma^t F_t,$$
$$y_{jt} \sim \text{Binomial} \left( k_{jt}, (1 + \exp [-\theta_{jt}])^{-1} \right),$$
where \( f^{uc}_t \) and \( F_t \) represent their simulated values, and exposure counts \( k_{jt} \) come from the dataset which is explored in the next section. The parameters \( \lambda_j, \beta, \gamma \) are chosen similar to their maximum likelihood values reported in Section 5. For computational convenience we consider \( F_t \) to be a scalar process, \( R = 1 \). Simulation results are based on 1000 simulations. Each simulation uses \( M = 50 \) importance samples during simulated maximum likelihood estimation, and \( M = 500 \) importance samples for signal extraction.

A selection of the graphical output from our Monte Carlo study is presented in Figure 1. We find that the principal components estimate \( \hat{F} \) captures the factor space \( F \) well. The average \( R^2 \) statistic is 0.94. The conditional mean estimate of \( f^{uc} \) is close to the simulated unobserved factor, with an average \( R^2 \) of 0.73. The sampling distributions of \( \phi_{uc} \) and \( \lambda_0 \) appear roughly symmetric and Gaussian, while the distributions of factor sensitivities \( \beta_0 \) and \( \gamma_1 \) appear skewed to the right. This is consistent with their interpretation as factor standard deviations. The distributions of \( \phi_{uc}, \beta_0, \lambda_0 \), and \( \gamma_1 \) are all centered around their true values. We conclude that the model succeeds in discriminating between the different possible sources of default rate variation. The procedure gives correct estimates - on average - for parameters contained in \( \psi \) as well as in the state vector \( \alpha \).

Finally, the standard errors for the estimated factor loadings \( \gamma \) do not take into account that the principal components are estimated with some error in a first step. We would like to investigate whether this impairs inference on these factor loadings. In each simulation we therefore estimate parameters and associated standard errors using true factors \( F_t \) as well as their principal components estimates \( \hat{F}_t \). The bottom panel in Figure 1 plots the empirical distribution functions of t-statistics associated with testing the null hypothesis \( H_0 : \gamma_1 = 0 \) when either \( F_t \) or \( \hat{F}_t \) are used. The t-statistics are very similar in both cases. Other standard errors are similarly unaffected. We conclude that the substitution of \( F_t \) with \( \hat{F}_t \) has negligible effects on inference on model parameters. This is in line with the R-squared results above.
5 Estimation results and forecasting accuracy

This section describes the macroeconomic and default data used for fitting the model. We report the main empirical findings and apply the fitted model in an out-of-sample forecasting exercise.

5.1 The Data: Macro Variables and Default Counts

The mixed dynamic factor model is analyzed using data from two main sources. First, a panel of 120 macroeconomic time series is constructed from the Federal Reserve Economic Database FRED. The aim is to select series which contain information about systematic credit risk conditions. The variables are grouped into five broad categories: (1) bank lending conditions, (2) macroeconomic and business cycle indicators, including labor market conditions and monetary policy indicators, (3) open economy macroeconomic indicators, (4) micro-level business conditions such as wage rates, cost of capital, and cost of resources, and (5) stock market returns and volatilities. Table 1 presents a listing of the series for each category. The macroeconomic panel contains both current information indicators such as real GDP, unemployment rate, new orders, as well as forward looking variables such as stock prices and interest rates.

A second dataset is constructed from the Standard and Poor’s CreditPro 7.0 database which consists of rating transition histories and (possibly) a default date for all S&P-rated firms from 1981 to mid-2005. This data contains the information to determine quarterly values for $y_{jt}$ and $k_{jt}$ in (1). We distinguish 13 industries which we pool into $D = 7$ industry groups: consumer goods, financials, transport and aviation, leisure, utilities, high tech and telecom, and health care sector. We further consider four age cohorts: less than 3, 3 to 6, 6 to 12, and more than 12 years from the time of the initial rating assignment. Age cohorts are included since default probabilities may depend on the age of a company. A proxy for

\footnote{http://research.stlouisfed.org/fred2}
age is the time since the initial rating has been established. Finally, there are four rating groups, an investment grade group $AAA - BBB$, and three speculative grade groups $BB$, $B$, and $CCC$. Pooling over investment grade firms is necessary since defaults are rare in this segment.

![insert Figure 3 around here ]

Time series of disaggregated default fractions can be observed from Figure 3. Default fractions cluster most visibly around the recession years of 1991 and 2001. In addition, default clustering appears more pronounced for firms in higher rating groups than for firms with lower ratings. Systematic default risk may be higher for investment grade firms than for speculative grade firms. This finding is corroborated in Section 5.3.

5.2 The macro factors

We first report the results from applying principal components to the macro panel introduced in Section 5.1. We employ the EM procedure of Section 4.1 to iteratively balance the panel before estimating the factors. Figure 4 presents graphs of the first four principal components from the balanced panel. The first factor exhibits clear peaks around NBER US business cycle troughs. The second factor also has peaks around these periods, but the association with a business cycle is less strong. Factors three and four do not exhibit the pronounced cyclical swings which are present in the first two factors.

![insert Figure 4 around here ]

To determine the number of factors we compute the panel information criteria (IC) as suggested by Bai and Ng (2002). The information criteria suggest two common factors, which together capture about 44% of the total variation in the macro panel. As it is often found in the macro factor literature, we find that the first factor is mainly associated with production and employment data, as well as a selection of business cycle indicators. According to its corresponding eigenvalue the first factor accounts for about 30% of the total variance. The second principal component loads mainly on series associated with firm profit margins, such as interest rates, the cost of intermediate inputs and resources, and prices of final goods. It
accounts for about 14% of the data variance. The third and fourth factor account for 7% and 6% of the total variation, respectively. Based on the IC values, we take \( R = 2 \) for the analysis below.

5.3 Major empirical results

We next discuss the parameter and risk factor estimates for the complete model. Since defaults are rare events we cannot freely and reliably estimate all coefficients \( \lambda_j, \beta_j \) and \( \gamma_{r,j} \) for each cross section \( j \) and macro factor \( r = 1, 2 \). Instead we propose a parsimonious model structure that is sufficiently flexible to address the key issues. The loadings and constants are given by

\[
\begin{align*}
\lambda_j &= \bar{\lambda}_0 + \lambda_{1,d_j} + \lambda_{2,a_j} + \lambda_{3,s_j}, \\
\beta_j &= \bar{\beta}_0 + \beta_{1,d_j} + \beta_{2,a_j} + \beta_{3,s_j}, \\
\gamma_{r,j} &= \bar{\gamma}_{r,0} + \gamma_{r,1,d_j} + \gamma_{r,2,a_j} + \gamma_{r,3,s_j},
\end{align*}
\]

where \( \bar{\lambda}_0, \bar{\beta}_0 \) and \( \bar{\gamma}_{r,0} \) are common coefficients, adjusted by specific coefficients \( \lambda_{i,k}, \beta_{i,k}, \) and \( \gamma_{r,i,k} \), for \( i = 1, 2, 3 \), and indices \( d_j = 1, \ldots, 7 \), \( a_j = 1, \ldots, 4 \), \( s_j = 1, \ldots, 4 \), and \( r = 1, 2 \) refer to a specific industry sector, age cohort, rating group, and macro factor, respectively. This specification empathizes that exposure to systematic risk may differ across industry, age, and rating classes.

In a preliminary analysis of our data we have found that factor loadings \( \beta_j \) and do not vary with age, \( \beta_{2,a_j} = 0 \), while loadings \( \gamma_{r,j} \) vary mostly over rating groups. Due to common parameters \( \bar{\lambda}_0, \bar{\beta}_0 \) and \( \bar{\gamma}_{r,0} \) we need to restrict specific (dummy) coefficients for each class, \( \lambda_{1,7} = \lambda_{2,4} = \lambda_{3,4} = \beta_{1,7} = \beta_{3,4} = 0 \). We also set \( \gamma_{r,1,d_j} = \gamma_{r,2,a_j} = 0 \), such that \( \gamma_{r,j} = \bar{\gamma}_{r,0} + \gamma_{r,3,s_j} \) varies only over rating classes.

In Table 2 we report the parameter estimates for three different specifications of the signal equation (3). Model 1 has \( \beta_j = 0 \) (only macro factors). Model 2 has \( \gamma_{r,j} = 0 \) (only the latent risk factor), and Model 3 has no restrictions in (3).

[insert Table 2 around here]

The estimates of the fixed effects in \( \lambda_j \) are mostly significant and the values are similar
across models. We observe a highly significant monotonic pattern in the coefficients for different rating classes, $\lambda_{3,s}$. This pattern indicates that firms in lower rating classes are more likely to default than firms in higher rating groups. The dummy coefficients for age cohort $\lambda_{2,a}$ show a similar monotonic pattern. The estimates suggest that a firm which has just recently acquired access to the capital market is less likely to default. The age cohorts capture a nonconstant baseline hazard as a result. Finally, there is considerable heterogeneity across industry groups $\lambda_{1,d}$. For example, firms categorized as part of the leisure or utility sector are less likely to default even after taking account of a firm’s rating class.

Our empirical results indicate an important role for the unobserved component or frailty factor. This is still the case after common macro factors are included. The impact of the unobserved component differs across rating and industry groups. For example, financial firms are found to have lower systematic risk than firms from the high tech and telecom sector.

The estimated coefficients $\gamma_{r,j}$ are statistically and economically significant. In all specifications, investment grade firms appear to have the highest systematic risk. Conversely, defaults from the lowest rating class appear to be largely unrelated to the current macroeconomic climate. This stylized fact in default rates is recognized in current banking regulation, see Basel Committee on Banking Supervision (2004).

The large increase in log-likelihood from Model 1 to 3 by more than 65 points is convincing. However, this difference should not be used as a basis for computing a likelihood ratio test statistic since the null hypothesis sets parameters on their boundaries, see Nyblom and Harvey (2000). However, the increase is indicative of a very substantial improvement in model fit. The increase in log-likelihood from Model 2 to 3 by 10 points is statistically significant at a 5% level. Thus, all factors are both statistically and economically significant and help to explain the systematic co-movement in the cross section.

[insert Figure 5 around here]

In Figure 5 we present the smoothed estimates of the default signals $\theta_{jt}$ for investment grade firms. The high default intensities associated with the 1983, 1991, and 2001 recession years are clearly visible from the graph. The dynamics in default rates differ from those in typical business cycle indicators as represented by the estimated macro factors $\hat{F}_t$. The
default pattern with the frailty factor is ‘smoother’ at the top during bad times than the more ‘peaked’ principal components would like to suggest. Default intensities may rise before the economy hits recession, and may remain high after the recession is over. These ‘default specific’ dynamics are captured by the frailty factor.

Figure 6 provides graphical evidence for an improved model fit when a latent factor is included. Each graph compares the observed (aggregate) US quarterly default rate for S&P rated firms with the respective counterpart from a fitted model. We distinguish four models, which either contain (i) no time-varying covariates, (iii) the first two principal components from macro data, or (iv) the first two principal components and a latent factor. As an alternative benchmark, the model in panel (ii) adopts three observed regressors. We select the US unemployment rate detrended by the Hodrick-Prescott filter, changes in filtered unemployment, and Moody’s Baa corporate yield spread over treasuries. Similar covariates have been used by Metz (2008). The model-implied fit to aggregate default rates differs most notably during ‘bad times’, where the latent component seems to be most valuable. This finding is corroborated when we investigate the out-of-sample forecasting accuracy of different model specifications.

5.4 Out of sample forecasting accuracy

We compare the out-of-sample forecasting performance between models by considering a number of competing model specifications. Such a forecast assessment is interesting for two reasons. First, accurate forecasts are valuable in conditional credit risk management, short-term loan pricing, and credit portfolio stress testing. Second, out-of-sample forecasting is possibly the most stringent diagnostic check available to time series econometricians. Good forecasts imply an acceptable model specification. In the analysis presented below we forecast cross sections of conditional default probabilities one year ahead.

Measuring the forecasting accuracy of time-varying default probabilities is not straightforward. Observed default fractions are only a crude measure of the ‘true’ default probabilities. We can illustrate this inaccuracy by considering a cell with, say, 5 firms. Even if the default
probability for this cell is forecast perfectly, it is unlikely to coincide with the observed default fraction of either 0, 1/5, 2/5, etc. The forecast error may therefore be large but it does not indicate a bad forecast. Observed default fractions are a useful measure only for a sufficiently large number of approximately similar firms per cell. For this reason we pool default and exposure counts over age cohorts and focus on two broad rating groups containing firms rated AAA to BBB investment grade, and BB and below known as speculative grade. Also, we focus on predicting an annual quantity instead of quarterly fractions. A mean absolute error (MAE) and root mean squared error statistic (RMSE) are computed by

$$\text{MAE}(t) = \frac{1}{D} \sum_d \left| \hat{\Pi}_{d,t+4}^{an} - \bar{\Pi}_{d,t+4}^{an} \right|,$$

and

$$\text{RMSE}(t) = \left( \frac{1}{D} \sum_d \left[ \hat{\Pi}_{d,t+4}^{an} - \bar{\Pi}_{d,t+4}^{an} \right]^2 \right)^{\frac{1}{2}},$$

where index $d = 1, \ldots, D$ denotes industry groups. The estimated and realized annual probabilities are given by

$$\hat{\Pi}_{d,t+4}^{an} = 1 - \prod_{h=1}^{4} \left( 1 - \hat{\Pi}_{d,t+h|t}^{qu} \right),$$

and

$$\bar{\Pi}_{d,t+4}^{an} = 1 - \prod_{h=1}^{4} \left( 1 - \frac{y_{d,t+h|t}}{k_{d,t+h}} \right),$$

respectively. To obtain the required default signals we first forecast all factors $\hat{F}_t, \hat{f}_{t|t}^{nc}$ jointly using a low order vector autoregression. This approach takes into account that the factors are conditionally correlated. We then predict the conditional default probabilities using equations (2) and (3).

Table 3 reports the forecast error statistics for five competing models. Model M0a does not contain common factors. It thus corresponds to the common practice of estimating default probabilities using long-term historical averages. This yields relatively small forecast errors in cases where the unobserved risk factors are close to their unconditional averages. We use Model M0a as our lower benchmark. As an infeasible best case, Model M4 uses the estimated (smoothed) factors over the complete sample, and holds the model parameters fixed at their end-of-sample values. The upper bound for improvements on average forecast errors over years 1997 to 2005 is about 26% for both rating groups. The reductions in MAE are largest when risk factors are far from their long-term averages. For instance, the reduction in MAE associated with the year 2000 in-sample forecast of the recession year 2001 is 67% for investment grade firms, and about 48% for speculative grade firms, see columns
Model M0b adopts three observed variables instead of common factors to forecast conditional default probabilities. These covariates are described at the end of Section 5.3, and serve as an alternative more realistic benchmark. It mimics the approach of the literature, where a limited set of observed descriptors is used to capture default rate dynamics.

The results for models M1, M2, and M3 in Table 3 correspond to out-of-sample forecasts using the models whose parameter estimates are reported in Table 2. The results show that the common macro factors $F_{1,t}$ and $F_{2,t}$ contribute to the out-of-sample forecasting performance. The observed reduction in MAE is between 1% and 7% on average over the years 1997-2004. However, forecasts improve considerably when an unobserved component is added to the model. The MAE then reduces between 11% and 18%. Reductions in MAE are again highest when risk factors are far from their long term averages. The MAE associated with the year 2000 forecast of 2001 default conditions is reduced by about 37% (IG) and 26% (SG) when compared to Model M0b which contains three observed macroeconomic variables. Such improvements are substantial and have clear practical implications for the computation of capital requirements.

6 Conclusion

We propose and motivate a novel non-Gaussian panel data time series model with regression effects to estimate and forecast the dynamics of disaggregated corporate default count data. The model is the first to combine a non-Gaussian panel data specification with the principal components of a large number of macroeconomic covariates. The model integrates two different types of factors, i.e., common factors from (continuous, Gaussian) macroeconomic and financial time series as well as an unobserved latent component for (discrete, count) default data. In an empirical application, the factors combine to capture a statistically and economically significant share of the time series dynamics in disaggregated default counts. In an out-of-sample forecasting experiment, we achieve substantial reductions between 10%
and 35% in mean absolute forecast error when forecasting conditional default probabilities, particularly in years of high default stress.

We continue to find a large and significant role for a dynamic frailty component even after taking into account many macroeconomic covariates. We strengthen the findings of Das et al. (2007) who point out the need for a latent component in a specification with few macroeconomic predictors. Therefore, the presence of a latent factor may not be due to omitted macroeconomic covariates. The smoothed estimate of the frailty component peaks at about the same time as the troughs of the business cycle. It implies additional default clustering in economically bad times. Practitioners who rely on observed macroeconomic and firm-specific data alone may underestimate their economic capital requirements and crisis default probabilities as a result.
References


Bai, J. and S. Ng (2002). Determining the number of factors in approximate factor models. Econometrica 70(1), 191–221.


### Table 1: Predictor Time Series in the Macro Panel

<table>
<thead>
<tr>
<th>Main category, sub-category</th>
<th>Summary of time series in category</th>
<th>Total no</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bank lending conditions</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Size of overall lending</td>
<td>Total Commercial Loans</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total Real Estate Loans</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total Consumer Credit outstanding</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Debt of Domestic Sector</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Commercial &amp; Industrial Loans</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bank loans and investments</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Household obligations/income</td>
<td></td>
</tr>
<tr>
<td>Extend of problematic</td>
<td>Non-performing Loans Ratio</td>
<td></td>
</tr>
<tr>
<td>banking business</td>
<td>Net Loan Losses</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Return on Bank Equity</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-performing Commercial Loans</td>
<td></td>
</tr>
<tr>
<td><strong>Macro and BC conditions</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General macro indicators</td>
<td>Real GDP</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Industry Production Index</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Private Fixed Investments</td>
<td></td>
</tr>
<tr>
<td></td>
<td>National Income</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Manuf. Sector Output</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Manuf. Sector Productivity</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Government Expenditure</td>
<td></td>
</tr>
<tr>
<td>Labor market conditions</td>
<td>Unemployment rate</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Weekly hours worked</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Employment/Population-Ratio</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Unemployed, more than 15 weeks</td>
<td></td>
</tr>
<tr>
<td>Business Cycle leading/</td>
<td>New Orders: Durable goods</td>
<td></td>
</tr>
<tr>
<td>coinciding indicators</td>
<td>New orders: Capital goods</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Capacity Util. Manufacturing</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Capacity Util. Total Industry</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Light weight vehicle sales</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Housing Starts</td>
<td></td>
</tr>
<tr>
<td></td>
<td>New Building Permits</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Final Sales of Dom. Product</td>
<td></td>
</tr>
<tr>
<td>Monetary policy indicators</td>
<td>M1 Money Stock</td>
<td></td>
</tr>
<tr>
<td></td>
<td>M2 Money Stock</td>
<td></td>
</tr>
<tr>
<td></td>
<td>M3 Money Stock</td>
<td></td>
</tr>
<tr>
<td></td>
<td>UMich Infl. Expectations</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Personal Savings</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gross Saving</td>
<td></td>
</tr>
<tr>
<td>Corporate Profitability</td>
<td>Corp. Profits</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Net Corporate Dividends</td>
<td></td>
</tr>
<tr>
<td>Intern'l competitiveness</td>
<td>Terms of Trade</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trade Weighted USD</td>
<td></td>
</tr>
<tr>
<td></td>
<td>USD/GER Exchange Rate</td>
<td></td>
</tr>
<tr>
<td>Balance of Payments</td>
<td>Current Account Balance</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Balance on Merchandise Trade</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Real Exports Goods, Services</td>
<td></td>
</tr>
<tr>
<td><strong>Macro-level conditions</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Labour cost/wages</td>
<td>Unit Labor Cost: Manufacturing</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total Wages &amp; Salaries</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Wholesale Trade Wages</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Management Salaries</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Technical Services Wages</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Wages &amp; Salaries: Other</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Employee Compensation Index</td>
<td></td>
</tr>
<tr>
<td>Cost of capital</td>
<td>1Month Commerical Paper Rate</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3Month Commerical Paper Rate</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AAA Corporate Bond Yield</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BAA Corporate Bond yield</td>
<td></td>
</tr>
<tr>
<td>Cost of resources</td>
<td>PPI All Commodities</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PPI Interm. Energy Goods</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PPI Finished Goods</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PPI Crude Energy Materials</td>
<td></td>
</tr>
<tr>
<td>Equity market conditions</td>
<td>Equity Indexes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>and respective volatilities</td>
<td></td>
</tr>
<tr>
<td></td>
<td>S&amp;P 500</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nasdaq 100</td>
<td></td>
</tr>
<tr>
<td></td>
<td>S&amp;P Small Cap Index</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dow Jones Industrial Average</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Russell 2000</td>
<td></td>
</tr>
</tbody>
</table>

---

28
Table 2: Estimation results

Note: Coefficients $\lambda$ combine to capture fixed effects. The factor sensitivities $\beta$ pertain to an unobserved component, while sensitivities $\gamma$ pertain to principal components from macro data. Sensitivities may depend on a firm’s industry sector and/or rating class. The groups mnemonics are given by fin: financial, tra: transport and aviation, lei: leisure, utl: utilities, htc: high tech and telecom, hea: health care. The consumer goods industry constitutes the reference group. The results are calculated using 1000 importance samples. Estimation sample is quarterly data from 1981:I to 2005:II.

<table>
<thead>
<tr>
<th>par</th>
<th>Model 1: Only $F_t$</th>
<th>Model 2: Only $F_t^{\text{nc}}$</th>
<th>Model 3: All Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>val</td>
<td>se</td>
<td>t-val</td>
</tr>
<tr>
<td>$\lambda_0$</td>
<td>-1.47</td>
<td>0.13</td>
<td>11.47</td>
</tr>
<tr>
<td>$\lambda_{1, fin}$</td>
<td>-0.47</td>
<td>0.13</td>
<td>3.57</td>
</tr>
<tr>
<td>$\lambda_{1, tra}$</td>
<td>-0.11</td>
<td>0.09</td>
<td>1.26</td>
</tr>
<tr>
<td>$\lambda_{1, lei}$</td>
<td>-0.48</td>
<td>0.11</td>
<td>4.30</td>
</tr>
<tr>
<td>$\lambda_{1, utl}$</td>
<td>-0.37</td>
<td>0.10</td>
<td>3.58</td>
</tr>
<tr>
<td>$\lambda_{1, htc}$</td>
<td>-0.20</td>
<td>0.11</td>
<td>1.84</td>
</tr>
<tr>
<td>$\lambda_{1, hea}$</td>
<td>-0.34</td>
<td>0.13</td>
<td>2.70</td>
</tr>
<tr>
<td>$\lambda_{2, 0-3}$</td>
<td>-0.73</td>
<td>0.12</td>
<td>6.24</td>
</tr>
<tr>
<td>$\lambda_{2, 4-5}$</td>
<td>-0.33</td>
<td>0.12</td>
<td>2.77</td>
</tr>
<tr>
<td>$\lambda_{2, 6-12}$</td>
<td>-0.43</td>
<td>0.12</td>
<td>3.49</td>
</tr>
<tr>
<td>$\lambda_{3, IG}$</td>
<td>-6.22</td>
<td>0.21</td>
<td>29.81</td>
</tr>
<tr>
<td>$\lambda_{3, BB}$</td>
<td>-3.96</td>
<td>0.13</td>
<td>29.52</td>
</tr>
<tr>
<td>$\lambda_{3, B}$</td>
<td>-2.37</td>
<td>0.08</td>
<td>30.50</td>
</tr>
<tr>
<td>$\phi$</td>
<td>0.87</td>
<td>0.08</td>
<td>11.70</td>
</tr>
<tr>
<td>$\beta_0$</td>
<td>0.70</td>
<td>0.18</td>
<td>3.72</td>
</tr>
<tr>
<td>$\beta_{1, fin}$</td>
<td>-0.20</td>
<td>0.17</td>
<td>1.19</td>
</tr>
<tr>
<td>$\beta_{1, tra}$</td>
<td>0.06</td>
<td>0.11</td>
<td>0.53</td>
</tr>
<tr>
<td>$\beta_{1, lei}$</td>
<td>0.00</td>
<td>0.13</td>
<td>0.01</td>
</tr>
<tr>
<td>$\beta_{1, utl}$</td>
<td>0.02</td>
<td>0.14</td>
<td>0.16</td>
</tr>
<tr>
<td>$\beta_{1, htc}$</td>
<td>0.27</td>
<td>0.16</td>
<td>1.70</td>
</tr>
<tr>
<td>$\beta_{1, hea}$</td>
<td>0.10</td>
<td>0.16</td>
<td>0.63</td>
</tr>
<tr>
<td>$\beta_{3, IG}$</td>
<td>0.28</td>
<td>0.22</td>
<td>1.27</td>
</tr>
<tr>
<td>$\beta_{3, BB}$</td>
<td>0.14</td>
<td>0.15</td>
<td>0.95</td>
</tr>
<tr>
<td>$\beta_{3, CCC}$</td>
<td>-0.32</td>
<td>0.12</td>
<td>2.72</td>
</tr>
<tr>
<td>$\gamma_{1, IG}$</td>
<td>0.76</td>
<td>0.17</td>
<td>4.59</td>
</tr>
<tr>
<td>$\gamma_{1, BB}$</td>
<td>0.73</td>
<td>0.12</td>
<td>5.97</td>
</tr>
<tr>
<td>$\gamma_{1, B}$</td>
<td>0.44</td>
<td>0.06</td>
<td>7.14</td>
</tr>
<tr>
<td>$\gamma_{1, CCC}$</td>
<td>0.42</td>
<td>0.06</td>
<td>6.74</td>
</tr>
<tr>
<td>$\gamma_{2, IG}$</td>
<td>0.28</td>
<td>0.16</td>
<td>1.74</td>
</tr>
<tr>
<td>$\gamma_{2, BB}$</td>
<td>-0.05</td>
<td>0.12</td>
<td>0.43</td>
</tr>
<tr>
<td>$\gamma_{2, B}$</td>
<td>0.21</td>
<td>0.06</td>
<td>3.77</td>
</tr>
<tr>
<td>$\gamma_{2, CCC}$</td>
<td>-0.02</td>
<td>0.06</td>
<td>0.26</td>
</tr>
<tr>
<td>LogLik</td>
<td>-2994.74</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The table reports improvements in forecast error statistics associated with one-year ahead out-of-sample forecasts of time-varying (conditional) default probabilities. We report mean absolute error (MAE) and root mean square error (RMSE) statistics separately for investment grade (IG) and speculative grade (SG) firms. Columns 5 to 8 report reductions in MAE with respect to two benchmark models, i.e., M0a (no factors, unconditional probabilities) and M0b (three observed descriptors capture time variation in default rates). Models M1, M2, and M3 contain factors $F_t$, $f_{it}^{unc}$, and both $F_t$, $f_{it}^{unc}$, respectively. Model M4 provides an infeasible best case by using factor and parameter estimates from the full sample for forecasting. Columns 7 and 8 refer to the year 2000 forecasts of recession year 2001 default rates. Columns 9 to 15 report the forecast error statistics for a given year, where averaging occurs over industry groups.

<table>
<thead>
<tr>
<th></th>
<th>MAE</th>
<th>% M0a</th>
<th>% M0b</th>
<th>% M0a (2001)</th>
<th>% M0b (2001)</th>
<th>1997</th>
<th>1998</th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
</tr>
</thead>
<tbody>
<tr>
<td>M0a:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>MAE</td>
<td>IG</td>
<td>0.59</td>
<td>0.18</td>
<td>0.24</td>
<td>0.29</td>
<td>0.84</td>
<td>1.06</td>
<td>1.32</td>
<td>0.46</td>
<td>0.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>SG</td>
<td>3.55</td>
<td>3.05</td>
<td>2.08</td>
<td>2.24</td>
<td>3.49</td>
<td>7.18</td>
<td>4.47</td>
<td>2.12</td>
<td>3.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>IG</td>
<td>0.72</td>
<td>0.18</td>
<td>0.27</td>
<td>0.46</td>
<td>1.03</td>
<td>1.22</td>
<td>1.78</td>
<td>0.51</td>
<td>0.35</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>SG</td>
<td>4.07</td>
<td>3.65</td>
<td>2.50</td>
<td>3.14</td>
<td>4.05</td>
<td>7.93</td>
<td>4.83</td>
<td>2.48</td>
<td>3.93</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M0b:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>subset</td>
<td>MAE</td>
<td>IG</td>
<td>0.58</td>
<td>0.18</td>
<td>0.24</td>
<td>0.29</td>
<td>0.85</td>
<td>1.03</td>
<td>1.31</td>
<td>0.47</td>
<td>0.27</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>SG</td>
<td>3.48</td>
<td>3.18</td>
<td>1.98</td>
<td>2.25</td>
<td>3.74</td>
<td>7.03</td>
<td>4.24</td>
<td>2.24</td>
<td>3.22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X_{t}^{full}$</td>
<td>MAE</td>
<td>IG</td>
<td>0.72</td>
<td>0.21</td>
<td>0.26</td>
<td>0.46</td>
<td>1.04</td>
<td>1.19</td>
<td>1.76</td>
<td>0.50</td>
<td>0.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>SG</td>
<td>3.94</td>
<td>3.81</td>
<td>2.32</td>
<td>3.11</td>
<td>4.28</td>
<td>7.76</td>
<td>4.30</td>
<td>2.60</td>
<td>3.39</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M1:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$F_t$</td>
<td>MAE</td>
<td>IG</td>
<td>0.57</td>
<td>0.16</td>
<td>0.22</td>
<td>0.28</td>
<td>0.88</td>
<td>1.01</td>
<td>1.32</td>
<td>0.47</td>
<td>0.26</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>SG</td>
<td>3.32</td>
<td>2.72</td>
<td>1.86</td>
<td>2.37</td>
<td>4.13</td>
<td>6.76</td>
<td>4.00</td>
<td>2.28</td>
<td>2.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>IG</td>
<td>0.72</td>
<td>0.17</td>
<td>0.27</td>
<td>0.48</td>
<td>1.07</td>
<td>1.17</td>
<td>1.79</td>
<td>0.52</td>
<td>0.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>SG</td>
<td>3.78</td>
<td>3.23</td>
<td>2.12</td>
<td>3.37</td>
<td>4.68</td>
<td>7.43</td>
<td>4.09</td>
<td>2.67</td>
<td>2.66</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M2:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$f_{it}^{unc}$</td>
<td>MAE</td>
<td>IG</td>
<td>0.53</td>
<td>0.15</td>
<td>0.20</td>
<td>0.29</td>
<td>0.85</td>
<td>0.90</td>
<td>1.11</td>
<td>0.45</td>
<td>0.26</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>SG</td>
<td>3.04</td>
<td>2.41</td>
<td>1.27</td>
<td>2.19</td>
<td>3.51</td>
<td>5.75</td>
<td>4.67</td>
<td>2.07</td>
<td>2.44</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>IG</td>
<td>0.66</td>
<td>0.17</td>
<td>0.28</td>
<td>0.46</td>
<td>1.04</td>
<td>1.07</td>
<td>1.42</td>
<td>0.50</td>
<td>0.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>SG</td>
<td>3.52</td>
<td>2.66</td>
<td>1.39</td>
<td>3.06</td>
<td>4.07</td>
<td>6.35</td>
<td>5.59</td>
<td>2.45</td>
<td>2.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M3:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$f_{it}^{unc}$, $F_t$</td>
<td>MAE</td>
<td>IG</td>
<td>0.52</td>
<td>0.13</td>
<td>0.19</td>
<td>0.28</td>
<td>0.87</td>
<td>0.65</td>
<td>1.30</td>
<td>0.48</td>
<td>0.26</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>SG</td>
<td>2.94</td>
<td>2.12</td>
<td>1.29</td>
<td>2.79</td>
<td>4.03</td>
<td>5.24</td>
<td>3.97</td>
<td>2.19</td>
<td>1.88</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>IG</td>
<td>0.68</td>
<td>0.17</td>
<td>0.28</td>
<td>0.49</td>
<td>1.06</td>
<td>0.82</td>
<td>1.76</td>
<td>0.54</td>
<td>0.34</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>SG</td>
<td>3.33</td>
<td>2.32</td>
<td>1.45</td>
<td>3.85</td>
<td>4.57</td>
<td>5.64</td>
<td>4.18</td>
<td>2.53</td>
<td>2.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M4:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$f_{it}^{unc}$, $F_t$ (full sample)</td>
<td>MAE</td>
<td>IG</td>
<td>0.43</td>
<td>0.13</td>
<td>0.21</td>
<td>0.30</td>
<td>0.62</td>
<td>0.34</td>
<td>1.17</td>
<td>0.45</td>
<td>0.26</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>SG</td>
<td>2.60</td>
<td>2.40</td>
<td>1.38</td>
<td>2.53</td>
<td>2.00</td>
<td>3.64</td>
<td>4.48</td>
<td>3.04</td>
<td>1.34</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>IG</td>
<td>0.56</td>
<td>0.18</td>
<td>0.27</td>
<td>0.46</td>
<td>0.80</td>
<td>0.43</td>
<td>1.52</td>
<td>0.48</td>
<td>0.35</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>SG</td>
<td>3.26</td>
<td>2.68</td>
<td>1.66</td>
<td>3.19</td>
<td>2.54</td>
<td>5.27</td>
<td>5.88</td>
<td>3.33</td>
<td>1.52</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 1: Simulation Analysis

Graphs 1 and 2 contain the sampling distributions of R-squared goodness-of-fit statistics in regressions of \( \hat{F} \) on simulated factors \( F \), and conditional mean estimates \( \hat{E}[f_{uc}|y] \) on the true \( f_{uc} \), respectively. Graphs 3 to 6 plot the sampling distributions of key parameters \( \phi_{uc}, \beta, \lambda_0, \) and \( \gamma_1 \). The bottom panel plots two empirical distribution functions of the t-statistics associated with testing \( H_0: \gamma_1 = 0 \). In each simulation either \( F \) or \( \hat{F} \) are used to obtain Monte Carlo maximum likelihood parameter and standard error estimates. Distribution plots are based on 1000 simulations.
Figure 2: Aggregated Default Data and Default Fractions

The figure shows time series plots of (i) the total default counts $\sum_j y_{jt}$ aggregated to a univariate series, (ii) total number of firms $\sum_j k_{jt}$ in the database, and (iii) aggregate default fractions $\sum_j y_{jt} / \sum_j k_{jt}$ over time.
Figure 3: Default Fractions Scatterplots

The figure plots disaggregated default fractions $y_{jt}/k_{jt}$ over time for four rating groups AAA – BBB, BB, B, and CCC – C.
The first four principal components are calculated from iteratively balanced quarterly macro and financial time series data (N=1,...,120) using the EM algorithm of Stock and Watson (2002b).
Figure 5: Smoothed Default Signals for Investment-Grade Firms
The figure plots smoothed default signals $\theta_{jt}$ over time for investment grade firms. It further indicates how the time variation of the total signal can be decomposed into variation in estimated factors $\hat{f}_{uc}^t$, $\hat{F}_{1,t}$, and $\hat{F}_{2,t}$. All factors are scaled by their respective factor loadings/standard deviations.
Figure 6: Model fit to observed US quarterly default fractions
The figure compares quarterly aggregate default fractions as implied by different model specifications with the realized quarterly US default rate for S&P rated firms. The model-implied frequencies refer to models with (i) no covariates or factors, (ii) three observed covariates, i.e. filtered US unemployment rate, changes in filtered unemployment, and Baa corporate yield spread over 10y treasuries, (iii) the first two principal components from macro data, and (iv) two principal components and a latent frailty factor.