Online Appendix to Dynamic factor models with macro, frailty, and industry effects for U.S. default counts: the credit crisis of 2008

Siem Jan Koopman[a] *André Lucas*[a,b] *Bernd Schwaab*[c]

[a] VU University Amsterdam and Tinbergen Institute

[d] Duisenberg school of finance

[e] European Central Bank, Financial Research Division

December 22, 2011

Appendix A1: estimation via importance sampling

An analytical expression for the the maximum likelihood (ML) estimate of parameter vector ψ for the MiMe DFM is not available. A feasible approach to the ML estimation of ψ is the maximization of the likelihood function that is evaluated via Monte Carlo methods such as importance sampling. A short description of this approach is given below. A full treatment is presented by Durbin and Koopman (2001, Part II).

The observation density function of $y = (y_1', \ldots, y_T')'$ can be expressed by the joint density of y and $f = (f_1', \ldots, f_T')'$ where f is integrated out, that is

$$p(y; \psi) = \int p(y, f; \psi)\mathrm{d}f = \int p(y|f; \psi)p(f; \psi)\mathrm{d}f, \qquad \text{(A.1)}$$

where $p(y|f; \psi)$ is the density of y conditional on f and $p(f; \psi)$ is the density of f. A Monte Carlo estimator of $p(y; \psi)$ can be obtained by

$$\hat{p}(y; \psi) = M^{-1} \sum_{k=1}^{M} p(y|f^{(k)}; \psi), \qquad f^{(k)} \sim p(f; \psi),$$

for some large integer M. The estimator $\hat{p}(y; \psi)$ is however numerically inefficient since most draws $f^{(k)}$ will not contribute substantially to $p(y|f; \psi)$ for any ψ and $k = 1, \ldots, K$. Importance sampling improves the Monte Carlo estimation of $p(y; \psi)$ by sampling f from the Gaussian importance density $g(f|y; \psi)$. We can express the observation density function $p(y; \psi)$ by

$$p(y; \psi) = \int \frac{p(y, f; \psi)}{g(f|y; \psi)} g(f|y; \psi)\mathrm{d}f = g(y; \psi) \int \frac{p(y|f; \psi)}{g(y|f; \psi)} g(f|y; \psi)\mathrm{d}f. \qquad \text{(A.2)}$$

Since f is from a Gaussian density, we have $g(f; \psi) = p(f; \psi)$ and $g(y; \psi) = g(y, f; \psi) \;/\; g(f|y; \psi)$. In case $g(f|y; \psi)$ is close to $p(f|y; \psi)$ and in case simulation from $g(f|y; \psi)$ is feasible, the Monte Carlo estimator of the likelihood function is given by

$$\tilde{p}(y; \psi) = g(y; \psi) M^{-1} \sum_{k=1}^{M} \frac{p(y|f^{(k)}; \psi)}{g(y|f^{(k)}; \psi)}, \qquad f^{(k)} \sim g(f|y; \psi), \qquad \text{(A.3)}$$

is numerically much more efficient, see Kloek and van Dijk (1978), Geweke (1989) and Durbin and Koopman (2001).

The importance density $g(f|y;\psi)$ is based on an approximating, linear Gaussian state space model based on an observation equation for each y_{jt} in (1) and given by

$$y_{jt} = c_{jt} + \theta_{jt} + \varepsilon_{jt}, \qquad \varepsilon_{jt} \sim \mathrm{N}(0, h_{jt}), \tag{A.4}$$

where c_{jt} is a known mean, θ_{jt} is the unobserved signal and h_{jt} is a known variance, for $j = 1, \ldots, J+N$. For the normal variables y_{jt}, the signal θ_{jt} is equal to μ_{jt} of (5) and the variables $c_{jt} = 0$ and $h_{jt} = \sigma_j^2$ are known with $j = J+1, \ldots, J+N$. For the default counts y_{jt} in the approximating model, we let the signal θ_{jt} be equal to π^*_{jt} of (4), with $j = 1, \ldots, J$. The variables c_{jt} and h_{jt} for the default counts are determined such that the modes of $p(f|y;\psi)$ and $g(f|y;\psi)$ are equal, see Shephard and Pitt (1997), Durbin and Koopman (1997), and Durbin and Koopman (p. 191–195, 2001) for the details. The values for c_{jt} and h_{jt} are found iteratively and by means of the Kalman filter and an associated smoothing method.

To simulate values from the resulting importance density $g(f|y;\psi)$ based on the approximating model (A.4), the simulation smoothing method of Durbin and Koopman (2002) can be used. For a set of M draws $f^{(1)}, \ldots, f^{(M)}$ from $g(f|y;\psi)$, the evaluation of the likelihood function (A.3) via importance sampling relies on the computation of $p(y|f;\psi)$, $g(y|f;\psi)$, with $f = f^{(k)}$, and $g(y;\psi)$ for $k = 1, \ldots, M$. Density $p(y|f;\psi)$ is based on the model specifications in (3). Density $g(y|f;\psi)$ is based on the approximating, linear Gaussian model (A.4). Density $g(y;\psi)$ is effectively the likelihood function of the approximating model (A.4) and can be computed via the Kalman filter, see Durbin and Koopman (2001). Testing the assumptions underlying the application of importance sampling can be carried out using the procedures proposed by e.g. Koopman, Shephard, and Creal (2009).

Appendix A2: estimation of latent factors

Inference on the latent factors can also be based on importance sampling. In particular, it can be shown that

$$\mathrm{E}(f|y;\psi) = \int f \cdot p(f|y;\psi)\mathrm{d}f = \frac{\int f \cdot w(y,f;\psi)g(f|y;\psi)\mathrm{d}f}{\int w(y,f;\psi)g(f|y;\psi)\mathrm{d}f},$$

where $w(y,f;\psi) = p(y|f;\psi)/g(y|f;\psi)$. The estimation of $\mathrm{E}(f|y;\psi)$ via importance sampling can be achieved by

$$\tilde{f} = \sum_{k=1}^{M} w_k \cdot f^{(k)} \Big/ \sum_{k=1}^{M} w_k,$$

with $w_k = p(y|f^{(k)};\psi)/g(y|f^{(k)};\psi)$ and where $f^{(k)} \sim g(f|y;\psi)$ is obtained by simulation smoothing. The standard error of $\tilde{f}_i$, the ith element of $\tilde{f}$, is denoted by s_i and is computed by

$$s_i^2 = \left(\sum_{k=1}^{M} w_k \cdot (f_i^{(k)})^2 \Big/ \sum_{k=1}^{M} w_k \right) - \tilde{f}_i^2,$$

where $f_i^{(k)}$ is the ith element of $f^{(k)}$.

Appendix A3: macro data listing and time series plots

Table 1 and Figure 1 contain a listing and time series plots, respectively, of the macro data that is used for the empirical part of our analysis.

Appendix A4: information criteria

Table 2 reports likelihood-based information criteria. The standard AIC and BIC refer to the whole model which includes the default and macro data parts. The Bai and Ng (2002) panel criteria refer to the fit of macro factors f_t^m to the macro data y_{jt} for $j = J+1, \ldots, J+N$.

Table 1: Macroeconomic Time Series Data

The table gives a full listing of included macroeconomic time series data x_t and binary indicators b_t. All time series are obtained from the St. Louis Fed online database, http://research.stlouisfed.org/fred2/.

Category	Summary of time series in category	Shortname	Total no
(a) Macro indicators, and business cycle conditions	Industrial production index	indpro	
	Disposable personal income	dspi	
	ISM Manufacturing index	napm	5
	Uni Michigan consumer sentiment	umich	
	New housing permits	permit	
(b) Labour market conditions	Civilian unemployment rate	unrate	
	Median duration of unemployment	uempmed	
	Average weekly hours index	AWHI	4
	Total non-farm payrolls	payems	
(c) Monetary policy and financing conditions	Government bond term structure spread	gs10	
	Federal funds rate	fedfunds	
	Moody's seasoned Baa corporate bond yield	baa	
	Mortgage rates, 30 year	mortg	6
	10 year treasury rate, constant maturity	tssprd	
	Credit spread corporates over treasuries	credtsprd	
(d) Bank lending	Total Consumer Credit Outstanding	totalsl	
	Total Real Estate Loans, all banks	realln	2
(e) Cost of resources	PPI Fuels and related Energy	ppieng	
	PPI Finished Goods	ppifgs	3
	Trade-weighted U.S. dollar exchange rate	twexbmth	
(f) Stock market returns	S&P 500 yearly returns	s_p500	
	S&P 500 return volatility	vola	2
			22

Figure 1: Macroeconomic and financial time series data

We present time series of yearly growth rates in macroeconomic and financial data. For a listing of the data we refer to Table 1.

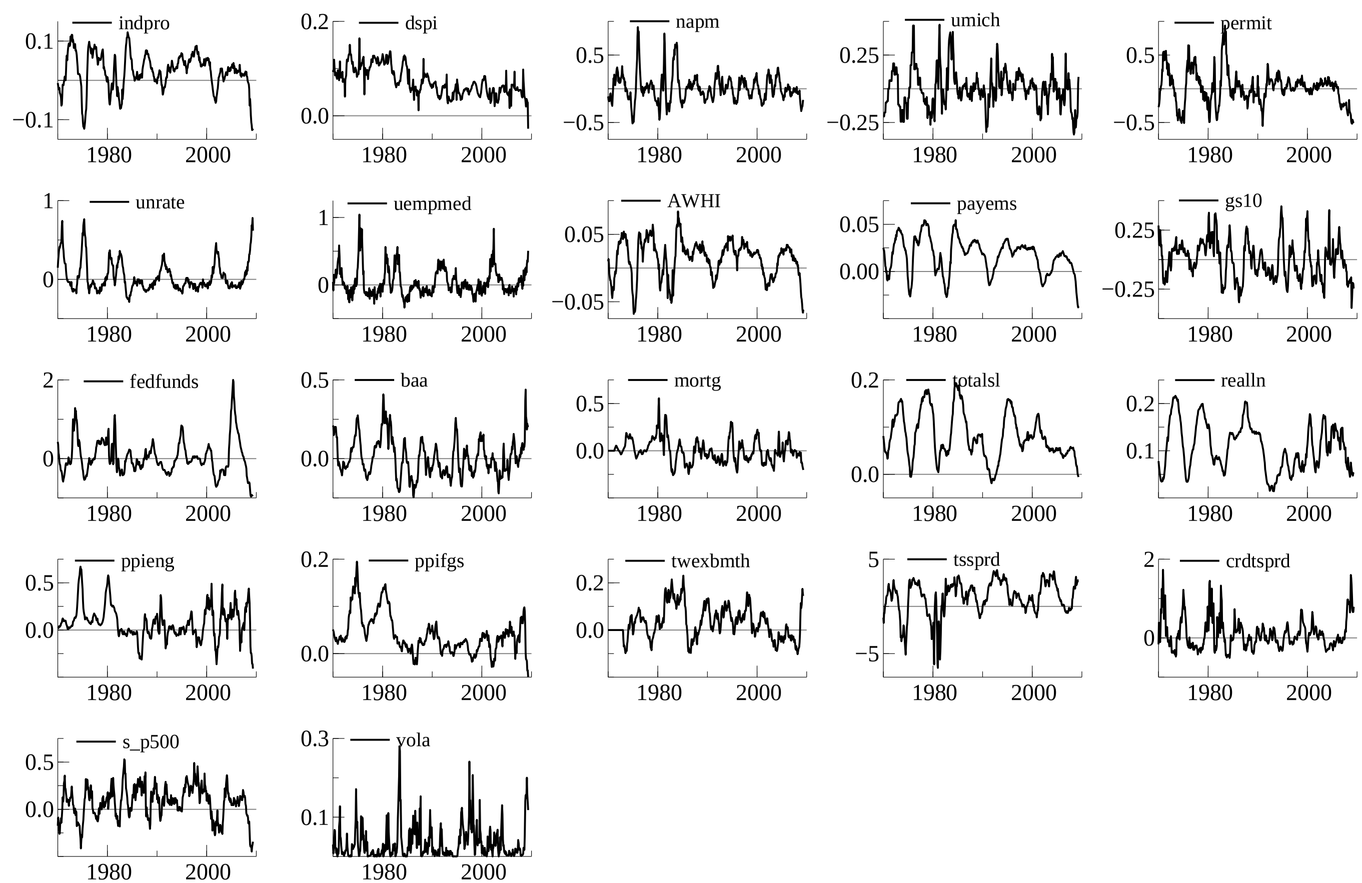


Table 2: Information criteria

We report likelihood-based information criteria (IC) to guide our model selection. The estimation sample is from 1971Q1 to 2009Q1. Row minimum values are printed in bold.

	F1	F2	F3	F4
loglik	-6530.2	-6277.3	-6182.5	-6133.7
#par	37	64	91	118
AIC	13134.5	12683.2	12548.2	**12505.4**
BIC	13415.0	**13168.1**	13237.2	13398.4
Bai Ng IC_1	-0.246	-0.454	**-0.523**	-0.474
Bai Ng IC_2	-0.239	-0.440	**-0.502**	-0.446
Bai Ng IC_3	-0.259	-0.481	**-0.563**	-0.527

Appendix A5: macroeconomic risk factor estimates

The top panel of Figure 2 presents the estimated risk factors f_t^m as defined in (4) and (5). We plot the estimated conditional mean of the factors, along with approximate standard error bands at a 95% confidence level. The factors are ordered row-wise from top-left to bottom-right according to their share of explained variation for the macro and financial data.

The bottom panel of Figure 2 presents the shares of variation in each macroeconomic time series that can be attributed to the common macroeconomic factors. The first two macroeconomic factors load mostly on labor market, production, and interest rate data. The last two factors displayed in the top panel of Figure 2 load mostly on survey sentiment data and changes in price level indicators. The macroeconomic factors capture 27.2%, 21.3%, 11.7%, and 8.3% of the total variation in the macro data panel, respectively (68.6% in total). The estimated factor loadings reveal that all four common factors f_t^m tend to load more on default probabilities of firms rated investment grade rather than speculative grade.

Figure 2: Macroeconomic risk factor estimates

The four panels present the estimated risk factors f_t^m as defined in (4) and (5). We present the estimated conditional mean of the factors, along with approximate standard error bands at a 95% confidence level. Details on the estimation and signal extraction methodology are available in A1 of this web appendix. Factors f^m are common to the (continuous) macro and financial as well as the (discrete) default count data.

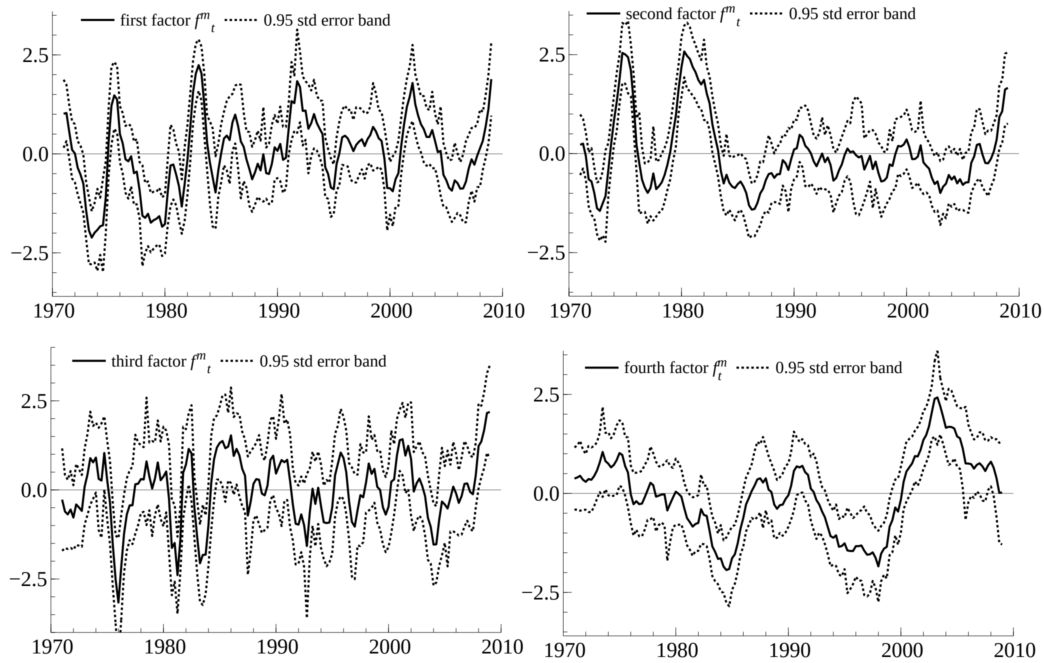
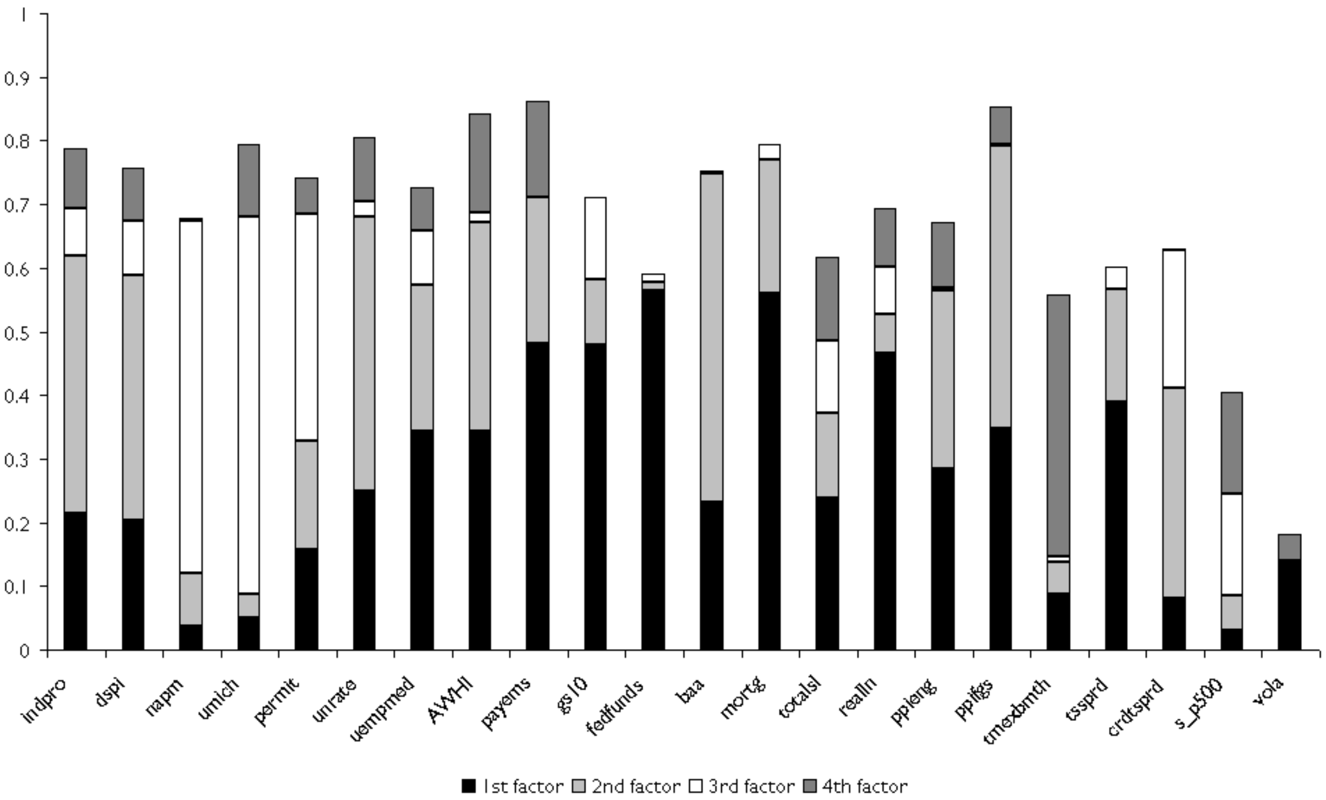


Figure 3: Macroeconomic risk factor loadings

The panel indicates which share of the variation in each macro-financial time series listed in Table 1 and plotted in Figure 1 can be attributed to each factor in f^m as plotted in Figure 2.

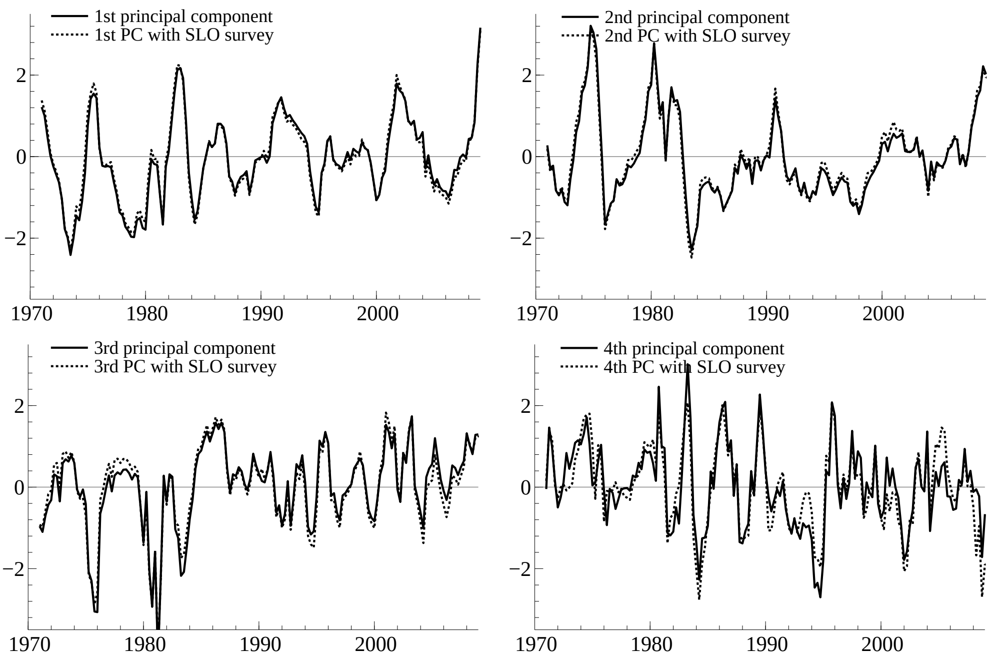


Appendix A6: macro principal components

Figure 4 plots the first four principal components from our macro data. Missing values are set to zero after standardization. The principal components change little when SLO bank lending standards are added to the panel as an additional explanatory variable.

Figure 4: Principal components of macro data

We plot the first four principal components from the macro data listed in Table 1. We also plot the principal components for the case when bank lending standards are added to the panel.

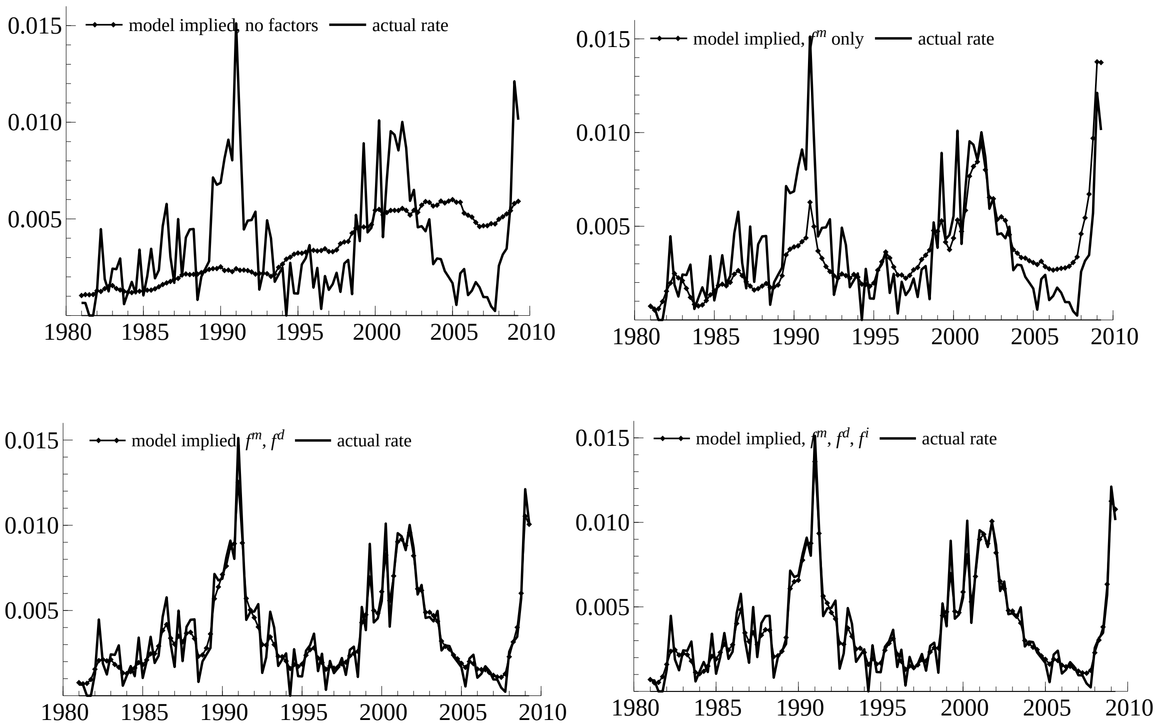


Appendix A7: graphs to illustrate model fit

Figure 5 presents the model-implied economy-wide default rate against the aggregate observed rates. We distinguish four specifications with (a) no factors, (b) f_t^m only, (c) f_t^m, f_t^d, and (d) all factors f_t^m, f_t^d, f_t^i. Based on these specifications, we assess the goodness of fit achieved at the aggregate level when adding latent factors. The static model fails to capture the observed default clustering around recession periods. The changes in the default rate for the static model are due to changes in the composition and quality of the rated universe. Such changes are captured by the rating and industry specific intercepts in the model. The upper-right panel indicates that the inclusion of macro variables helps to explain default rate variation. The latent frailty dynamics given by f_t^d, however, are clearly required for a good model fit. This holds both in low default periods such as 2002-2007, as well as in high default periods such as 1991. The bottom graphs of Figure 5 indicate that industry-specific developments cancel out in the cross-section to some extent and can thus be diversified. As a result, they may matter less from a (fully diversified) portfolio perspective.

Figure 5: Model fit to observed aggregate default rate

Each panel plots the observed quarterly default rate for all rated firms against the default rate implied by different model specifications. The models feature either (a) no factors, (b) only macro factors f^m, (c) macro factors and a frailty component f^m, f^d, and (d) all factors f^m, f^d, f^i, respectively.



Appendix A8: prediction error diagnostics

We report residual diagnostics for one step ahead prediction errors that pertain to the default count panel data. We define a time series of prediction errors as

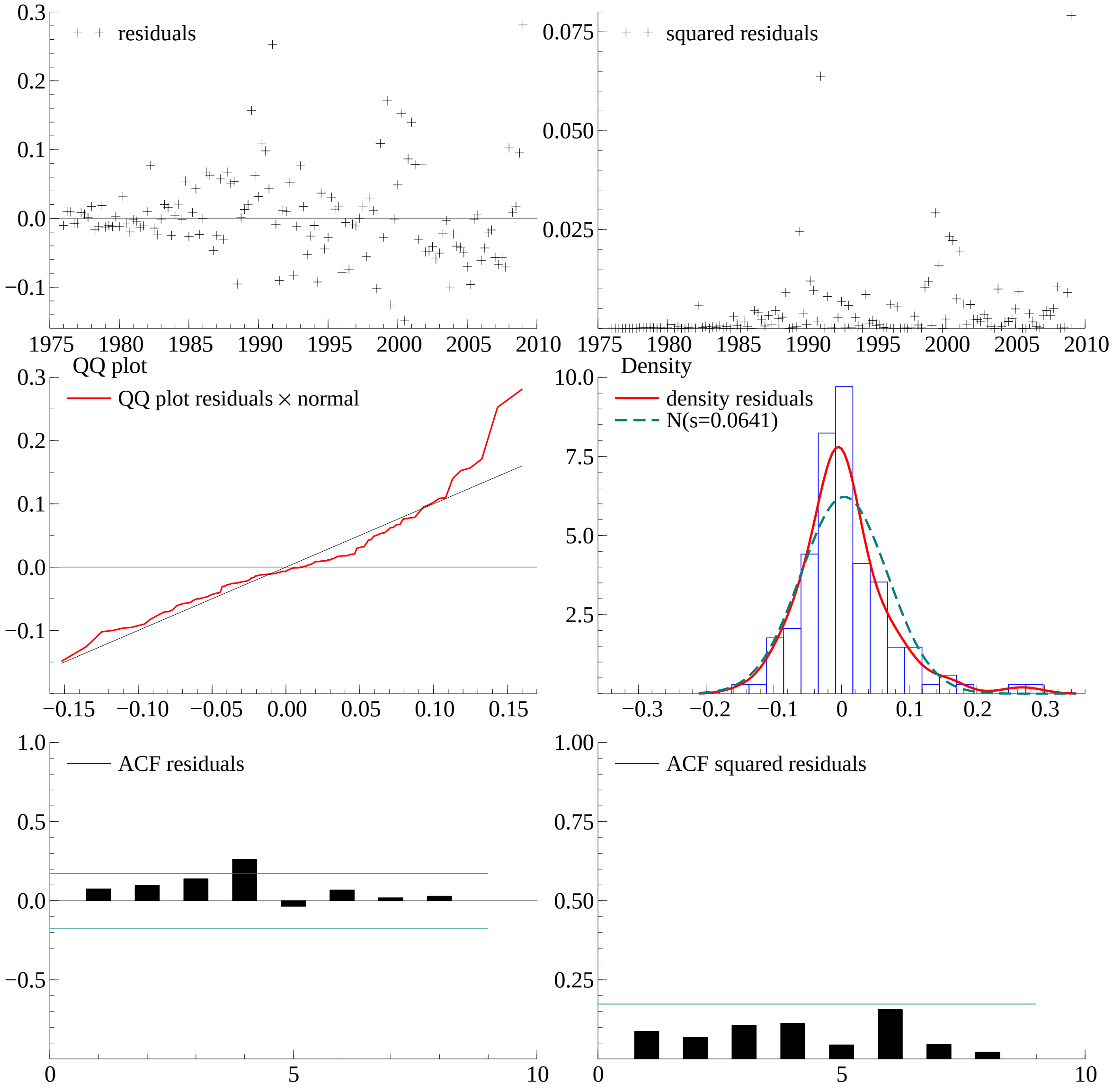
$$\hat{r}_{t|t-1} = \left(\sum_{j=1}^{J} \delta_{jt} \right)^{-1} \sum_{j=1}^{J} \delta_{jt} \left(y_{jt} - \hat{\pi}_{jt|t-1} k_{jt} \right), \tag{A.5}$$

where y_{jt} are the quarterly default counts, k_{jt} are the respective number of firms at risk at the beginning of quarter t, $j = 1, \ldots, J$, the indicator function δ_{jt} is equal to one if $k_{jt} > 0$ and zero otherwise, and $\hat{\pi}_{jt|t-1}$ is the one step ahead predicted pd for firms in cross section j. For prediction, parameter estimates are kept at full sample values for computational reasons, while risk factor estimates are obtained from an expanding window that contains data up to and including $t - 1$.

Figure 6 reports (a) mean prediction errors $\hat{r}_{t|t-1}$ over time, (b) squared prediction errors $\hat{r}^2_{t|t-1}$, (c) a QQ plot of the prediction errors against the normal, (d) an error histogram and associated density kernel estimate, as well as (e) the autocorrelation function (ACF) pertaining to errors and squared errors, respectively. Overall, the errors are zero on average and roughly standard normally distributed. Some larger deviations of observed from predicted values occur in the recession periods of 1991 and 2008-09. We note some leftover autocorrelation at the fourth lag in the error ACF. Overall, however, error autocorrelation does not seem to be a major issue. The autocorrelation at the fourth lag, along with some larger residuals in 1991 and 2008-09, both disappear if we base our diagnostics on deviations implied by smoothed (full sample) risk factor estimates.

Figure 6: Residual diagnostics

We report (a) prediction errors $\hat{r}_{t|t-1}$ over time, (b) squared prediction errors $\hat{r}^2_{t|t-1}$, (c) a QQ plot of the prediction errors against the normal, (d) an error histogram and associated density kernel estimate, (e) an estimate of the prediction error autocorrelation function, and (f) the autocorrelation function for squared errors.

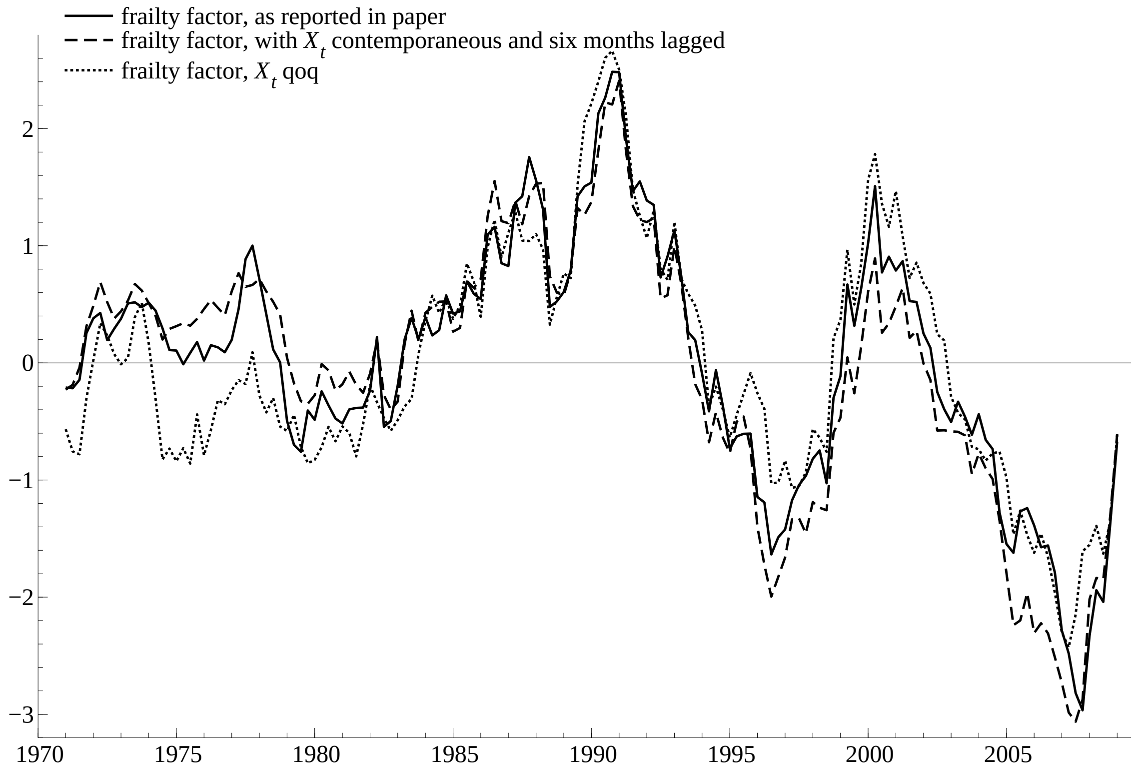


Appendix A9: frailty effects and macro data

Figure 7 compares three different estimates of the frailty factor f_t^d. The estimates are based on the same econometric specification but take in different macro data: (i) the original panel data as described in the paper, (ii) the original data stacked with its six months lagged values, doubling its cross sectional dimension, and (iii) macro data that replaces annual with quarterly growth rates. The frailty factor estimate is fairly robust to such changes and leading/lagging the macro data. This suggests that frailty is not caused by such timing effects. The reason why timing is not very important may be that such timing effects are captured indirectly in a static factor structure, see Stock and Watson (2002). In a linear Gaussian factor model, the static factors can be interpreted as a rotated version of current and lagged structural factors.

Figure 7: Frailty factor estimates for different macro data panels

We plot three conditional mean estimates of the frailty factor based on our favorite specification. Each model, however, now takes in a different transformation of the macro data: the original macro panel (22 covariates), the original panel stacked with its six months lagged values (44 covariates), and the original panel that replaces annual growth rates with quarterly rates.



Appendix A10: (un)conditional loss distributions

Section 4 considers a portfolio of short-term (rolling) loans to all Moody's rated U.S. firms. Loans are extended at the beginning of each quarter during 1981Q1 and 2008Q4 at no interest. A non-defaulting loan is re-extended after three months. The loan exposure to each firm at time t is given by the inverse of the total number of firms at that time, that is $(\sum_j k_{jt})^{-1}$. This implies that the total credit portfolio value is 1$ at all times. For simplicity, we assume a stressed loss-given-default of 80%. This example portfolio is stylized in many regards. Nevertheless, it allows us to investigate the importance of macroeconomic, frailty, and industry-specific dynamics for the risk measurement of a diversified loan or bond portfolio.

It is straightforward to simulate the portfolio credit loss distribution and associated risk measures for arbitrary credit portfolios in such a setting. First, the exposures k_{jt} are chosen to correspond to the portfolio exposures. Second, one uses the estimation methods detailed in this appendix to estimate the current position of the latent systematic risk factors. Third, one can use the transition equation (2) directly to simulate future risk factor realizations. Finally, conditional on the risk factor path, the defaults can be simulated by combining (3) and (4). Term structures of default rates can easily be obtained by combining model-implied quarterly probabilities over time.

References

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

Durbin, J. and S. J. Koopman (1997). Monte Carlo Maximum Likelihood estimation for non-Gaussian State Space Models. *Biometrika 84(3)*, 669–684.

Durbin, J. and S. J. Koopman (2001). *Time Series Analysis by State Space Methods*. Oxford: Oxford University Press.

Durbin, J. and S. J. Koopman (2002). A simple and efficient simulation smoother for state space time series analysis. *Biometrika 89(3)*, 603–616.

Geweke, J. (1989). Bayesian inference in econometric models using Monte Carlo integration. *Econometrica 57*, 1317–39.

Kloek, T. and H. K. van Dijk (1978). Bayesian estimates of equation system parameters: an application of integration by Monte Carlo. *Econometrica 46*, 1–20.

Koopman, S. J., N. Shephard, and D. Creal (2009). Testing the assumptions behind importance sampling. *Journal of Econometrics 149(1)*, 2–10.

Shephard, N. and M. K. Pitt (1997). Likelihood analysis of non-Gaussian measurement time series. *Biometrika 84*, 653–67.

Stock, J. and M. Watson (2002). Forecasting Using Principal Components From a Large Number of Predictors. *Journal of the American Statistical Association 97(460)*, 1167–1179.